School of Mathematical, Physical and Computational Sciences
Department of Mathematics and Statistics

Ergodic properties and response theory for a stochastic two-layer model of geophysical fluid dynamics.

Submitted for the degree of
Doctor of Philosophy in Mathematics

Giulia Carigi

June 2021

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Giulia Carigi

Abstract

In this work, ergodic properties of a stochastic medium complexity model for atmosphere and ocean dynamics are analysed. Specifically, we study a two–layer quasi–geostrophic (2LQG) model with the upper layer perturbed by additive noise for geophysical flows. This model is popular in the geosciences, for instance to study the effects of a random wind forcing on the ocean. Yet it is less studied in mathematics, especially if the stochastic perturbation is acting only on one of the layers. In this case the noise is effectively spatially degenerate, posing a significant challenge to the analysis.

After showing the model well–posedness, we focus on its long time average behaviour and ergodic properties: existence and uniqueness of an invariant measure (ergodicity), exponential convergence of solutions laws to the invariant measure (exponential stability or spectral gap), differentiability or only Hölder continuity of the invariant measure with respect to system parameters (linear or fractional response).

Existence of an invariant measure is shown with classic techniques. Its uniqueness is established using a recent technique from stochastic analysis called asymptotic coupling, to account for the noise spatial degeneracy. This is proved provided a certain passivity condition on the second layer holds. Under the same condition, exponential stability is shown by blending different recent approaches like the asymptotic coupling.

An important application of spectral gaps is response theory. The only result on linear response applicable to a large class of SPDEs is the work by Hairer and Madja (2010). We modify their approach treating a class of less regular observables. In particular we give a toolkit for linear and fractional response for SPDEs with moderately degenerate noise using the strength of a deterministic forcing as parameter. We apply such a framework to the 2D stochastic Navier-Stokes equation as test model, and finally to the stochastic 2LQG model.

Keywords: Ergodicity, Stochastic geophysical flow models, linear response, random wind forcing, spectral gap.

Acknowledgments

First and foremost I would like to wholeheartedly thank Jochen Broecker and Tobias Kuna for your guidance and support, which, even in the last very challenging 15 months, never failed. Through our endless discussions over the past years I have learnt enormously about the matters in this thesis and more importantly about research. Thanks for keeping on believing in me and my work, even when I may have forgotten to do so myself.

Thanks to Valerio Lucarini for your sincere encouragement and honest advice throughout the years. I much appreciated your “reality checks” over my work, both for the advancement of the scientific subject and my professional development.

I would like to thank Benedetta Ferrario for the insightful discussions, especially on unique ergodicity in Reading in early 2019. Work on this part of my thesis was also positively influenced by my time at the Institut Henri Poincaré during the thematic trimester on *The Mathematics of Climate and the Environment* in the Autumn of 2019. I particularly owe a thanks to Franco Flandoli for pointing me towards some enlightening literature on the topic during our stimulating conversations in Paris. I also acknowledge the organisers of the trimester for the financial support which made my participation possible.

Thanks to the MPE CDT, staff and students, present and past, for the great environment which fostered my curiosity and gave me endless opportunities to learn new skills, as well as for the social events and coffee breaks (particular shout to the Delta’s at Reading to whom I owe innumerable cups of fresh strong coffee!).

Outside of academia, a big *grazie* to my parents for supporting me in any possible way, for teaching me not to be afraid to dream big as well as for keeping me rooted. Thanks to Silvia, Gary, Gabriel for always believing in me and being there, as well as to my sisters from different mothers, Alessandra and Camilla, I hope one day we will manage to all live in the same country again. Above all, thanks Simone for your infallible loving support in all the sunny days and the rainy days we shared.

This work has been accomplished with the financial support of the Engineering and Physical Sciences Research Council Centre for Doctoral Training in the Mathematics of Planet Earth.

Contents

Introduction

> The atmosphere and oceanic sciences are sometimes thought of as not being 'beautiful' in the same way as some branches of theoretical physics. Yet surely quasi–geostrophic theory, and the quasi–geostrophic potential vorticity equation, are quite beautiful, both for their austerity of description and richness of behaviour.
>
> *G. Vallis [71]*

In this work we study the long time average behaviour of a stochastic version of an important model for large-scale atmosphere and ocean dynamics, the two–layer quasi–geostrophic model. In particular we will prove under which conditions the model is ergodic, namely when the time averages of an observable, for large times, can be approximated by spatial averages, and exponential mixing, meaning that this happens with exponential rate. Among other aspects, ergodicity provides the foundations to time series analysis as it allows to infer general properties of the system from long time averages of a single solution. In the context of climate science, ergodicity ensures that long time averages of a single realisation describe typical properties of the whole system independently of the initialization of climate models. This is different from weather forecast where prediction skill is the result of choosing specific initial conditions that approximate the current state of the atmosphere well.

Furthermore we study how the long term statistical properties are affected by changes in the parameters of the system, namely if the statistics of observables under the current set of parameters will still be valid under small perturbations of the parameters and how the perturbed and unperturbed statistics are related. By studying the response to perturbations in the parameters, namely *response theory*, for models in geophysical fluid dynamics, we give a mathematical insight into whether statistical properties derived under current conditions will be valid under future climates. Moreover we develop a framework for response theory suitable for dissipative stochastic partial differential equations with moderately degenerate noise, extending the applicability of the famous approach by Hairer and Madja [40].

Models for geophysical fluid dynamics. The term geophysical fluid dynamic (GFD) is most commonly used to refer to the general and formal treatment of atmosphere and ocean flows. These are nonlinear chaotic systems where the deep interaction between spatial and temporal scales creates the weather we all experience and, on long time scales, the climate of the Earth. In addition to being chaotic, these systems are also very high dimensional, for example it can be estimated that the atmosphere alone has roughly 10^{27} degrees of freedom. Therefore approximations best suitable to capture and describe macroscopic phenomena at different time and space scales have been introduced over the years. One significant example is for large scales dynamics (e.g. 1000 km for the atmosphere and 100 km for the ocean) at the mid-latitudes. Here the balance between the pressure gradient and the Coriolis force, the so-called geostrophic balance, and that between the pressure gradient and gravity, the so-called hydrostatic balance, are the main features of the fluid dynamic. These balances can be used to simplify the original set of equations for atmosphere and ocean dynamics, like the rotating shallow water equations, and the resulting model is the quasi–geostrophic (QG) approximation.

Already present in the literature from the late thirties of the twentieth century, the quasi–geostrophic model was systematically derived by Charney in 1948 in [13]. Used in early operational numerical weather forecasts, the QG model is still used extensively in research as it is strikes a balance between simplicity of formulation and range of the phenomenology it can model. Quasi–geostrophic models with several layers in particular are able to represent density stratification and provide insights into, for instance, atmosphere–ocean coupling or baroclinic instabilities. This type of instabilities is extremely common in the atmosphere and ocean and is at the origin of large scale weather phenomena like cyclones. The two–layer quasi–geostrophic (2LQG) model is one of the simplest models where the baroclinic instabilities arise.

The introduction of stochasticity in GFD model raises interesting questions both from a physical and mathematical point of view and there are several reasons one may want to consider stochastic perturbations. A classic example is Hasselmann's 1976 work [45] where slow components of the climate system are driven by the weather, which can be interpreted as random perturbation on shorter time scales. Stochastic terms can also account for the uncertainty generated by parametrization and approximations of the numerical simulation (see for example [16, 6]). Or the noise can account for the inevitable lack of precise measurements or knowledge of the processes involved. In this work we will look at the two–layer quasi–geostrophic model with a forcing on the top layer, to account for example of the wind forcing on the upper ocean, composed of a deterministic and a stochastic part, which is white in time and coloured in space.

The effect of a stochastic wind forcing on QG models has been a topic of research in meteorology and oceanography for at least forty years, in the single layer case e.g. [38, 58, 66, 67] and the multi–layer, see e.g. [70, 5, 14, 23, 60] as well as in the continuously stratified case [34, 59]. In the mathematical literature the stochastic single-layer QG model, with either additive or multiplicative noise, received surely more attention (see e.g. [9, 26, 27, 73]) than its multi–layer version. In fact, we can expect results for the single layer to extend to the multi–layer case when the random terms appears in all layers. Less studied, though, is the action of a stochastic forcing acting only on one of the layers and its consequent effects on the other layers and the whole dynamics.

To the best of this author's knowledge, the only reference for a mathematical investigation of a two–layer quasi–geostrophic model with a forcing on the top layer, prior to our work, is [15]. There the authors studied the long time dynamics of the model using the method of determining functionals for random dynamical systems. This method gives a way to parametrize the system global attractor by means of a finite number of functionals. Furthermore, under some conditions on the parameters of the system, it is shown that functionals depending only on the top layer suffice to describe the attractor. However, this approach does not give information on the statistics of the model as we will do here.

In the deterministic context, it was shown in [8] that the multi–layer quasi–geostrophic model is well posed and it admits a global attractor. This is done, as for other dissipative partial differential equations, by means of the Galerkin approximation and showing appropriate *a priori* bounds. In order to show the existence and uniqueness of the solutions for the stochastic version we will study an associated random equation (i.e. a partial differential equation with random coefficients). The well-posedness of this random PDE has not been explicitely verified until now, and it is also stated in [15].

Finally note that the model described so far must not be confused with the quasi-geostrophic model considered in other mathematical studies like [22, Section 13.13] and [63]. This model is, in fact, referred to as *surface* quasi-geostrophic in the applications and it describes a different physical scenario to the one analysed here.

Ergodic properties. In the context of climate science, a frequent underlying assumption is that long time averages of historic time series are relevant to describe properties of the whole system. This is often unavoidable as we obviously have a *single* realisation of the Earth climate system. In mathematical terms we express this assumption by requiring the climate system to be *ergodic*. In fact we talk of an ergodic system if the long time averages of an observable can be approximated by its averages with respect to a measure invariant for the dynamics.

Let $\{X_t, t \geq 0\}$ be the stochastic process of interest, in our case the solution of the stochastic two–layer quasi–geostrophic model. Since in this case $\{X_t\}$ is a Markov process, a sufficient condition for ergodicity is the existence of a *unique* invariant measure for the associated transition probabilities $\{P_t, t \geq 0\}$. To establish this it is also sufficient to show *asymptotic stability*, namely that $P_t(x)$ converges asymptotically in time to the unique invariant measure, independently of the initial data x. If the convergence is exponentially fast we talk of *exponential stability*.

There is a large body of literature [21, 39] on ergodic theory for infinite dimensional systems arising from SPDEs, in particular on the unique ergodicity and exponential stability of two–dimensional stochastic Navier–Stokes equations. Existence of an invariant measure for such model is established (see [18, 32]) by means of the Krylov-Bogoliubov theorem. This classic approach relies on two conditions: Feller property and tightness. The first corresponds to continuity with respect to initial condition of the solution. The second property describes the fact that the mass of transition probabilities does not "escape to infinity". Tightness is in fact shown exploiting the dissipativity properties of the solutions.

Ensuring that the invariant measure is unique is usually a harder task and several approaches are now available. A classic approach to establish unique ergodicity of a Markov process is to show that the process is irreducible and strong Feller [21]. Several results on the uniqueness of the invariant measure were shown since the 90's using this approach, for example for the stochastic two–dimensional Navier–Stokes equation see [33, 30], for the one layer QG model see [26]. However for SPDEs the strong Feller property fails to hold in cases where the noise is spatially degenerate, namely acting only on a subset of degrees of freedom. Eventually Hairer and Mattingly [41] provided a comprehensive approach for the 2D Navier–Stokes equations forced only in a four dimensional subspace, introducing the novel concept of asymptotic strong Feller. This seminal work is highly technical, by for example requiring Malliavin calculus and the Hörmander condition.

In this work we are considering a spatially degenerate noise as the noise appears only on the first layer. However as the noise acts on all modes of the first layer, meaning on an infinite dimensional subset of degrees of freedom, it is not as highly degenerate as in [41]. Therefore we can use a different approach to establish unique ergodicity, the asymptotic [44] (or generalised

[50]) coupling method. The recent work [37], using ideas from [44, Section 2], provides a compact account of the use of the asymptotic coupling method to establish unique ergodicity for several nonlinear SPDEs.

The main idea of the asymptotic coupling approach is to add a control to the stochastic forcing to synchronize solutions with different initial data. In finite dimensions, Girsanov's theorem ensures that the controlled equation and the original equation generate equivalent distributions. In the infinite dimensional case, an appropriate finite dimensional control on the unstable degrees of freedom is often sufficient to ensure synchronisation and permits application of Girsanov's theorem.

We will use this technique in a similar way to [37] on the stochastic two–layer quasi–geostrophic model to show the uniqueness of its invariant measure. In particular we will need to impose a condition involving the viscosity, bottom friction and the intensity of the noise to ensure the result. We may think of the imposed parameter condition as requiring the bottom layer (the one without noise) to be sufficiently dissipative to be determined by the top one.

Regarding exponential stability, i.e. the exponential convergence of the transition probabilities to the invariant measure, typically Harris' theorem provides conditions under which this holds in the total variation norm. A major difficulty with applying this theorem in the context of SPDEs is that the transition probabilities might be singular for different initial conditions in violation of a crucial condition in Harris' theorem. In [44, Section 4] a new framework is introduced to retrieve a version of Harris' theorem in the infinite dimensional context, which gives exponential rate of convergence in a Wasserstein-like distance, rather than in total variation.

In [50] the authors show how the asymptotic coupling method can be used not only for the uniqueness of the invariant measure, but also to establish convergence of transition probabilities (without a specified rate), taking a step forward towards unifying Section 2 and Section 4 of [44]. More recently [12] provided a set of verifiable conditions which, by means of the asymptotic coupling approach, provides exponential and sub-exponential rate of convergence to the invariant measure, improving [11] and [50]. These conditions are particularly suitable for SPDEs, including Navier–Stokes type equations like the quasi–geostrophic approximation.

However from the results in [12] it is not immediately clear how such toolbox for SPDEs gives the general Harris' theorem in [44]. It is challenging to retrieve the precise formulation of the Wasserstein-like distance with respect to which there is exponential convergence of transition probabilities. This will require us to establish a fine comparison and blending of the results in [12] and [44] as well as [11]. Next, applying this methodology to the two–layer quasi–geostrophic model we will prove that, under the same condition given for the uniqueness of the invariant measure, there is exponential convergence of transition probabilities.

An important immediate consequence of this result is a *spectral gap* property for the Markov semigroup with respect to the Wasserstein-like distance. This property means that the eigenvalues different from one of the Markov semigroup are concentrated in a disk of radius strictly smaller than one. Moreover, given the formulation of the Wasserstein-like distance, we will see that the spectral gap is in a space of Hölder continuous functions. We are particularly interested in obtaining a spectral gap for the semigroup operator, since this is a crucial ingredient to study the response of the model to changes in the parameters, especially whether the invariant measure is differentiable with respect to the parameters (*linear response theory*).

Response theory investigates the change of invariant objects under changes of a system parameter. In particular, we talk about linear response when the invariant measure is differentiable in the parameter. The derivative of the invariant measure with respect to the parameter at a certain point (the "current climate") allows to quantify the response of the system to perturbations ("climate change"), see [53]. In fact, even though the invariant measure is often a very singular object, it can nonetheless change smoothly with respect to changes in the parameters, at least in a weak sense. In particular one aims at a *response formula* which expresses the derivative of the invariant measure exclusively in terms of the unperturbed dynamics. In the applications this would mean having a way to compute statistics of the perturbed dynamics from those of the unperturbed. For more on the relevance of linear response theory in geophysics see for example applications like [1, 55, 53] or the recent review paper [36].

In the case of finite dimensional systems, there exists a large body of mathematical literature on linear response. For hyperbolic systems, in absence of stochasticity, the pioneering work of Ruelle [65] ensured the differentiability of invariant measures, in particular of SRB measures which carry crucial physical interpretation. The result has been extended also to partially hyperbolic systems in [24] but little is known for other classes of deterministic systems, finite or infinite dimensional. In particular the existence of SRB measures for Navier-Stokes seems entirely open as equations of fluid dynamic appears out of scope to be treated with these techniques. For a review on linear response theory in deterministic systems see the survey article [3].

In terms of stochastic systems, the impact of stochastic perturbations on Ruelle's linear response has been investigated for example in [52]. Recent works [35] and [2] pioneered linear response in finite dimensional random dynamical systems. However less is known for infinite dimensional systems associated to stochastic partial differential equations. To the best of this author's knowledge, the only result general enough to apply to dynamical systems associated to a large class of stochastic partial differential equations is the work of Hairer and Madja [40]. The response is studied in a weak sense, namely it is shown that the averages of a class of observables are differentiable with respect to the parameters of the systems, and moreover a formula for the derivative is given. One of the ingredients for this results is showing that the semigroup exhibits a spectral gap when acting on the desired class of observables. In [40] the authors consider the closure of the space of smooth observables with respect to a weighted C^1–norm. Such observables were already used in [42] to prove the spectral gap property for the 2D stochastic Navier–Stokes equation with highly degenerate noise as in [41].

Here we will reformulate the framework of [40] highlighting the necessary conditions for a general space of observables. In particular, we will show linear response for less regular observables than [40]. In fact, as discussed above, [44] provides sufficient conditions for the spectral gap property in a space of Hölder continuous functions, so less regular than those used in [40] and [42]. Using the strength of a deterministic forcing as the parameter with respect to which response is considered, we give a series of conditions to show linear response for dissipative SPDEs, like stochastic Navier–Stokes equation or the stochastic two–layer quasi–geostrophic model. This methodology is simpler to verify than the approach from [42, 40] as it does not require the use of Malliavin calculus for example, and can be extended more easily to models other than stochastic 2D Navier–Stokes. It is however less comprehensive as it cannot deal with highly degenerate noises and we need to impose stronger conditions on the nature of the perturbations to the dynamics we study the response for.

We focus on dissipative nonlinear equations with a deterministic forcing and the dependence of their invariant measure on the forcing strength. In particular we obtain differentiability and a linear response formula for forcings which are finite dimensional and in the range of the noise. For forcings not satisfying such conditions we can nevertheless show weak Hölder continuity of the invariant measure, often referred to as fractional response. This result does not provide a linear approximation of the perturbed dynamics in terms of the unperturbed one, but ensures that a small change in the intensity of the forcing does not cause a discontinuity in the long time average behaviour of observables that are at least Hölder continuous.

Outline of the thesis. In chapter 1 we recall the formulation of the two–dimensional stochastic Navier–Stokes (SNS) equations as we will use it as test model for the novel techniques, and present the deterministic two–layer quasi–geostrophic (2LQG) model. We close the first chapter by introducing the stochastic 2LQG model and its random version after [15] which we will later use to show existence and uniqueness of solutions and the existence of the invariant measure.

In chapter 2 we first show that there exists a unique solution of the stochastic 2LQG model, both in a weak and strong sense. These results, already known for the deterministic two-layer QG model, are now established for its stochastic version by means of classic techniques for random systems. In the second half of the chapter, namely section 2.4 and section 2.5, we consider respectively the stochastic 2LQG and SNS model and study the dependence of the solutions on the strength of the deterministic forcing (for 2LQG acting only on the first layer). We will show for both models that the solutions are locally Lipschitz and locally differentiable with respect to it. These technical results, not explicitly in the literature but straightforward to derive, will be crucial when studying the dependence of the invariant measure by the same parameter in chapter 5.

Then in chapter 3 we show existence and uniqueness of the invariant measure for the stochastic 2LQG model. Both these new results require a precise application of techniques available in the relevant literature pointed out above. In particular we first show the existence adapting the approach used for the stochastic Navier–Stokes equation in [32]. In the second part we show its uniqueness, conditional to the bottom friction being sufficiently large, using the asymptotic coupling method to deal with the nondegeneracy of the noise.

In chapter 4 we first merge the results from [44] and [12] providing a clear set-up to show exponential stability and a spectral gap property for nonlinear dissipative SPDEs. This set-up, which not only summarises the results in [44] and [12] but we believe is more transparent and easier to use, is a key contribution of this thesis. This methodology permits us for the first time to establish exponential stability and spectral gap for stochastic 2LQG model with moderately degenerate noise as discussed above. Furthermore we illustrate the methodology applying it to the stochastic Navier–Stokes equation.

Finally we dedicate chapter 5 to the dependence on the parameters of the invariant measure and in particular linear response. We develop a new methodology for SPDEs with moderate degenerate noise for observables that are less regular than in [40]. In particular we study response with respect to changes in the intensity of the forcing, as introduced in chapter 2. In section 5.2 we also provide a new toolkit for such SPDEs to ensure Hölder continuity of the invariant measure (fractional response). Both set-ups for linear and fractional response are a major contribution of this thesis. Finally in section 5.3 and section 5.4 we apply these results respectively to the stochastic Navier–Stokes equations and the stochastic 2LQG model showing linear and fractional response. This is the first mathematical result on linear response for the 2LQG model. With respect to the stochastic Navier-Stokes the contribution of this work is a detailed proof of response theory for moderate degenerate noise and in particular extending it to Hölder continuous observables.

Chapter 1

Intermediate complexity models for geophysical fluid dynamics

An important class of models for atmosphere and ocean dynamics are *quasi–geostrophic* models. Such medium complexity models, used in early operational numerical weather forecasts, capture the large scale features of atmosphere dynamics in the mid-latitudes and are still used extensively in research. Quasi–geostrophic models with several layers in particular are able to represent density stratification and provide insights into, for instance, atmosphere-ocean coupling or baroclinic instabilities [46, Chapter 8]. From a mathematical point of view we will see that these models are systems of several 2D Navier–Stokes equations in vorticity formulation coupled to each other. Therefore, while the centre of our investigation will be the stochastic two–layer quasi–geostrophic equation, we will also present the 2D Navier–Stokes model with additive noise.

We first recall some key concepts of stochastic analysis in Hilbert spaces (referring mainly to [22]) and in section 1.2 we briefly present the 2D stochastic Navier–Stokes equation (referring mainly to [69] and [62]). Then in section 1.3, after an introduction on the derivation of the quasi–geostrophic approximation, we lay down the mathematical framework and notations for the two–layer model. Finally in section 1.4 we introduce a stochastic perturbation on the top layer and present the corresponding stochastic two–layer quasi–geostrophic model.

1.1 Stochastic analysis in Hilbert spaces

Let U be a Hilbert space with scalar product $\langle \cdot\,,\,\cdot\rangle$ and associated norm $|\cdot|$. Given the probability space $(\Omega,\mathcal{F},\mathbb{P})$, let $(W(t))_{t\geq 0}$ be an infinite dimensional Wiener process with covariance operator $Q : U \to U$ and natural filtration $\{\mathcal{F}_t\,,\,t \geq 0\}$. The operator Q is assumed to be a nonnegative symmetric operator which is *trace class*, i.e. given $\{e_k\}_{k\in\mathbb{N}}$, a complete orthonormal basis of U

$$\operatorname{Tr} Q := \sum_{k=1}^{\infty}\langle Qe_k, e_k\rangle < \infty.$$

We denote by $L_1(U)$ the space of trace class operators on the space U. Equivalently if Q is trace class then

$$\left(\sum_{k\in\mathbb{N}} |Q^{1/2}e_k|^2\right)^{1/2} < \infty, \tag{1.1}$$

namely $Q^{1/2}$ is a *Hilbert-Schmidt* operator. We denote by $L_2(U)$ the space of Hilbert-Schmidt operators on U endowed with the norm

$$\|T\|_{L_2(U)} = \left(\sum_{k\in\mathbb{N}} |Te_k|^2\right)^{1/2},$$

and by $L_2^0(U)$ the space of all Hilbert-Schmidt operators from $Q^{1/2}(U)$ into U.

It can be shown (see e.g. [22, Proposition 4.3]) that a Q-Wiener process $W(t)$ can be written as a series of real valued Wiener processes more precisely

$$W(t) = \sum_{k=1}^{\infty}\sqrt{\sigma_k}\beta_k(t)g_k, \tag{1.2}$$

where σ_k are eigenvalues of Q with $\{g_k\}_{k\in\mathbb{N}}$ a corresponding orthonormal system of eigenfunctions, and $\beta_k(t)$, $k \in \mathbb{N}$, are independent real Brownian motions on $(\Omega,\mathcal{F},\mathbb{P})$.

A process $\phi(t)$ with values in U is called *stochastically integrable* if it is L_2^0-predictable and is such that

$$\int_0^T \|\phi(s)\|^2_{L_2^0}\,ds = \int_0^T \|\phi(s)Q^{1/2}\|^2_{L_2}\,ds < \infty \quad \mathbb{P}\text{-a.s.} \tag{1.3}$$

Last, we define the *quadratic variation* of the process

$$M(t) = \int_0^t \phi(s)\,dW(s) \tag{1.4}$$

as
$$\langle M \rangle_t = \int_0^t \|\phi(s)\|^2_{L_2^0}\, ds.$$

We are now ready to present the first stochastic partial differential equation we will deal with, the two–dimensional stochastic Navier–Stokes equations.

1.2 Stochastic Navier–Stokes equations

In this section we present the classic mathematical set up for the two-dimensional Navier-Stokes equation taking the main elements from the presentation for its deterministic version. We follow [69, Chapter III, Section 2] and [62, Chapter 9] but the interested reader can also refer to [68], one of the most complete references on the subject. For the stochastic version [22, Section 13.11] provides a brief presentation and an exhaustive list of references.

Let $\mathcal{D} = [0, L] \times [0, L] \subset \mathbb{R}^2$ with $L > 0$ and consider the two–dimensional (2D) stochastic Navier–Stokes equation on $\mathcal{D}$

$$\begin{aligned} &du + \nu \Delta u\, dt - (u \cdot \nabla) u\, dt = \nabla p\, dt + f\, dt + dW \\ &\operatorname{div} u = 0 \\ &u(0, x) = u_0. \end{aligned} \tag{1.5}$$

Here $u = u(t, x)$ is the velocity of an incompressible fluid, ν is the viscosity, $p(t, x)$ the pressure of the fluid, $f(x)$ is a time-independent deterministic forcing and W is a Q-Wiener process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the covariance operator Q nonnegative, symmetric and trace class. We consider (1.5) with periodic boundary conditions and we assume that the average flow vanishes, namely
$$\int_{\mathcal{D}} u(t, x)\, dx = 0 \quad \text{for all } t \geq 0.$$

Let $L^2(\mathcal{D})$ and $H^k(\mathcal{D})$, $k \in \mathbb{N}$, be the Sobolev spaces of L-periodic functions with norms respectively

$$|u|^2 := \int_{\mathcal{D}} |u(x)|^2\, dx \quad \text{and} \quad \|u\|_k^2 := \sum_{0 \leq |\alpha| \leq k} |D^\alpha u|^2, \quad \alpha \in \mathbb{N}^2. \tag{1.6}$$

and H^{-k}, $k \in \mathbb{N}$ the dual space of H^k. Furthermore denote by $\dot{L}^2(\mathcal{D})$ and $\dot{H}^k(\mathcal{D})$, $k \in \mathbb{N}$, the space of functions u in $L^2(\mathcal{D})$ and $H^k(\mathcal{D})$ such that

$$\int u(x)\, dx = 0.$$

As the velocity $u = u(t,x)$ is two–dimensional it is natural to introduce the following product spaces

$$\mathbf{L}^2(\mathcal{D}) = \left[L^2(\mathcal{D})\right]^2 \quad \text{and} \quad \mathbf{H}^k(\mathcal{D}) = \left[H^k(\mathcal{D})\right]^2, \tag{1.7}$$

where for $u = (u_1, u_2), v = (v_1, v_2) \in \mathbf{L}^2(\mathcal{D})$

$$(u,v)_{\mathbf{L}^2(\mathcal{D})} = (u_1, v_1) + (u_2, v_2), \tag{1.8}$$

$$|u|^2 := |u|^2_{\mathbf{L}^2(\mathcal{D})} = |u_1|^2 + |u_2|^2 \tag{1.9}$$

and for $u = (u_1, u_2), v = (v_1, v_2) \in \mathbf{H}^k(\mathcal{D})$

$$(\!(u,v)\!)_{\mathbf{H}^k(\mathcal{D})} = (\!(u_1, v_1)\!) + (\!(u_2, v_2)\!) \tag{1.10}$$

$$\|u\|_k^2 := \|u\|^2_{\mathbf{H}^k(\mathcal{D})} = \|u_1\|^2 + \|u_2\|^2. \tag{1.11}$$

We now consider the Hilbert spaces

$$\begin{aligned} \mathcal{H} &= \{u \in \dot{\mathbf{L}}^2(\mathcal{D}) \,:\, \operatorname{div} u = 0 \ \text{ in } \ \mathcal{D}\} \\ \mathcal{V} &= \{u \in \dot{\mathbf{H}}^1(\mathcal{D}) \,:\, \operatorname{div} u = 0 \ \text{ in } \ \mathcal{D}\} \end{aligned} \tag{1.12}$$

respectively with norms $|\cdot|$ and $\|\cdot\| := \|\cdot\|_1$. Elements of $\mathcal{H}$ and $\mathcal{V}$ then satisfy the divergence free condition and the boundary conditions by definition.

Let A denote the Stokes operator namely A is a linear operator on $\mathcal{H}$ such that

$$(Au, v) = (\!(u,v)\!) \quad \text{for all } \ u, v \in \mathcal{V}.$$

Its domain in $\mathcal{H}$, i.e.

$$D(A) = \{u \in \mathcal{V} \,:\, Au \in \mathcal{H}\}$$

can be shown to be

$$D(A) = \{u \in \dot{\mathbf{H}}^2(\mathcal{D}) \,:\, \operatorname{div} u = 0 \ \text{ in } \ \mathcal{D}\} = \dot{\mathbf{H}}^2(\mathcal{D}) \cap \mathcal{V},$$

so that $\mathcal{V} = D(A^{1/2})$ and $\|u\| = |A^{1/2}u|$.

Since we consider periodic boundary conditions, we have that $Au = -\Delta u$ for all $u \in D(A)$. Moreover, the operator A is a self-adjoint positive operator on $\mathcal{H}$, and we denote by $\{\lambda_k\}$ its eigenvalues and by $\{e_k\}$ a corresponding complete orthonormal system of eigenvectors.

Denote by $\mathcal{V}^*$ the dual of $\mathcal{V}$, then we have

$$D(A) \subset \mathcal{V} \subset \mathcal{H} \subset \mathcal{V}^*$$

where the inclusions are continuous and each space is dense in the following one.

The weak formulation of the Navier–Stokes equation is obtained by classical arguments (see e.g. [68, 62]), namely by taking the $\mathcal{H}$ scalar product of the equation with a test function $v \in \mathcal{V}$. In particular the term involving the pressure disappears since the periodic boundary conditions and the fact that $\operatorname{div} v = 0$ imply $(\nabla p, v) = 0$. Therefore given the trilinear form b

$$b(u,v,w) := \int_{\mathcal{D}} w(x) \cdot (u(x) \cdot \nabla) v(x)\, dx,$$

the original equation (1.5) becomes

$$(du, v) + \nu(Au, v)\, dt + b(u,u,v)\, dt = (f, v)\, dt + (v, dW) \tag{1.13}$$

Alternatively, given the bilinear operator $B(u,v) : \mathcal{V} \times \mathcal{V} \to \mathcal{V}^*$ defined as

$$\langle B(u,v), w \rangle = b(u,v,w),$$

we can rewrite (1.13) as

$$du + (\nu Au + B(u,u))\ dt = f\, dt + dW \quad u(0,\omega) = u_0(\omega) \tag{1.14}$$

which we expect to hold as an equality in $\mathcal{V}^*$ for $f \in \mathcal{V}^*$.

Crucial part of the study of the Navier–Stokes equations is the treatment of the trilinear form b and it is also where the main differences between the two and three dimensional versions lay. Thanks to Ladyzhenskaya's inequalities it is known that the trilinear form b satisfies the following identities and estimates:

Lemma 1.2.1 ([62, Propositions 9.1, 9.2])**.** *Consider the two–dimensional case $\mathcal{D} \subset \mathbb{R}^2$. Then*

$$b(u,v,w) = -b(u,w,v) \quad \text{for all } u \in \mathcal{H},\ v,w \in \mathcal{V} \tag{1.15}$$

hence the orthogonality relation $b(u,v,v) = 0$.

For periodic boundary conditions

$$b(u,u,Au) = 0 \quad \text{for all } u \in D(A). \tag{1.16}$$

Furthermore the following bounds hold: for all $u \in L^\infty$, $v \in \mathcal{V}$, $w \in \mathcal{H}$

$$|b(u,v,w)| \le \|u\|_\infty \|v\| |w|. \tag{1.17}$$

If u, v, $w \in \mathcal{V}$

$$|b(u,v,w)| \le k_B \|u\|^{1/2} |u|^{1/2} \|v\| \|w\|^{1/2} |w|^{1/2},$$

and if $u \in \mathcal{V}$, $v \in D(A)$, $w \in \mathcal{H}$

$$|b(u,v,w)| \le k_B \|u\|^{1/2} |u|^{1/2} \|v\|^{1/2} |Av|^{1/2} |w|. \tag{1.18}$$

for an appropriate positive constant k_B.

1.3 The two–layer quasi–geostrophic (2LQG) model

We are interested in what Vallis in his most comprehensive book on atmosphere and ocean dynamic, [71], described as "perhaps the most widely used set of equations for theoretical studies of atmosphere and ocean", namely the *shallow water quasi–geostrophic* equations or simply *quasi–geostrophic* equations. This model is ultimately an approximation of three dimensional Navier–Stokes equations on a rotating coordinate frame which best models the large-scale features, yet not planetary, of the atmosphere or the ocean at mid latitudes. It is at this scale that the main agents of the weather phenomena, like (anti)cyclons and oceanic currents like the Gulf stream, act and the quasi–geostrophic model is an intermediate system between planetary scales and smaller scales where convection enters predominantly.

We give a brief description of how the model can be derived but we refer to [71, Section 5.3] for a detailed presentation, or [54, Section 4.6] for a more rigorous mathematical derivation. As other GFD models, the quasi–geostrophic equations are determined by means of appropriate scaling and approximation by linear expansion around a small parameter. Consider the momentum equation of the rotating shallow water equation ([71, pg.127])

$$\partial_t u + u \cdot \nabla u + f \times u = -g \nabla \eta$$

where u is the horizontal velocity, f is the Coriolis force, g is the gravitational acceleration and η the height of the free surface. We consider the so-called β-plane approximation (see [71, Section 2.3.2]): this accounts for the fact that the vertical component of the rotation changes with the latitude y by writing the Coriolis parameter $f = f(y)$ as

$$f(y) = f_0 + \beta y, \tag{1.19}$$

with f_0 and β assigned constants.

By assuming the hydrostatic balance on the vertical direction, namely that the pressure is proportional to the height η, the right hand side of the rotating shallow water equations is nothing but the pressure gradient. By scale analysis it can be deduced that the time derivative has the same scale of the advection and we are left balancing the pressure gradient against the rotation and the advection terms. The so-called Rossby number *Ro* is a constant giving information on the relation between the advection and the rotation terms: the smaller is the Rossby number the larger is the effect of the rotation on the dynamic with respect to that of the advection.

Apart from the hydrostatic balance on the vertical component, large-scale flows are described in the zonal (horizontal) component by the geostrophic balance, namely when the rotation and the pressure gradient are in equilibrium with each other. As the term itself suggest, in the *quasi–geostrophic* equations these are instead *almost* in equilibrium as the Rossby number is small but not identically zero. The idea is then to use Ro as parameter for asymptotic expansion of the variables u and η in the rotating shallow water equations and derive the limit equations when $Ro \to 0$. This procedure defines the quasi–geostrophic equations.

In absence of dissipation or external forcing the one layer quasi–geostrophic model is nothing but the material conservation of a scalar quantity $q(x, y, t)$ i.e.

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \partial_t q + u \cdot \nabla q = 0, \tag{1.20}$$

where q, called quasi–geostrophic potential vorticity, is defined as

$$q := \Delta\psi + \beta y - \frac{f_0}{h}\eta.$$

Here η is the height of the free surface as above, h is the mean thickness of the layer, $f_0 + \beta y$ is the Coriolis parameter by the β-plane approximation, and ψ is the streamfunction of the fluid, namely $u = \nabla^\perp\psi$.

We consider the top to have a flat surface. This choice provides a simpler model but at the same time describes a case still relevant for example for the ocean where the variation of η about its mean position is very small compared to the mean depth of the upper ocean and with good approximation the surface can be considered flat. Then the quasi–geostrophic potential vorticity reduces to

$$q := \Delta\psi + \beta y.$$

and we can write (1.20) as

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \partial_t q + J(\psi, q) = 0 \tag{1.21}$$

with J the Jacobian operator $J(a, b) = \nabla^\perp a \cdot \nabla b$.

Let us now consider two layers of fluid one on top of each other with mean height h_1 for the top layer and h_2 for the bottom one, assuming (without loss of generality) that $h_1 \leq h_2$, with density respectively ρ_1 and ρ_2 with $\rho_1 < \rho_2$ and streamfunctions ψ_1 and ψ_2. This is a natural intermediate step between the one layer case and the continuously stratified model which account for a smooth density gradient on the vertical component, which allows us to study more features of the ocean and/or the atmosphere within a still rather simple framework. For example the two–layer QG model is one of the simplest models where crucial features of the weather system, like baroclinic instabilities, arise. Such a model can be derived in the same way as described above from the two–layer rotating shallow water equations (see e.g. [71, Section 5.3.2]) so that given the streamfunction $\boldsymbol{\psi} = (\psi_1, \psi_2)$ we have

$$\begin{aligned} \frac{\mathrm{d}q_1}{\mathrm{d}t} &= \partial_t q_1 + J(\psi_1, q_1) = 0 \\ \frac{\mathrm{d}q_2}{\mathrm{d}t} &= \partial_t q_2 + J(\psi_2, q_2) = 0, \end{aligned} \tag{1.22}$$

where, if we consider for simplicity flat bottom and top surface, the QG potential vorticites for the layers are defined as

$$\begin{aligned} q_1 &:= \Delta\psi_1 + F_1(\psi_2 - \psi_1) + \beta y \\ q_2 &:= \Delta\psi_2 + F_2(\psi_1 - \psi_2) + \beta y. \end{aligned}$$

Here F_1, F_2 are positive constants defined as

$$F_i := \frac{f_0^2}{g' h_i}, \tag{1.23}$$

with g' the reduced gravity $g' = g(\rho_2 - \rho_1)/\rho_0$ (see [61, pg. 419]) where ρ_0 is the characteristic value for the density. Further, denote

$$h_1 F_1 = h_2 F_2 = \frac{f_0^2}{g'} =: p. \tag{1.24}$$

Equivalently we can define q_1, q_2 as

$$\begin{aligned} q_1 &= \Delta\psi_1 + F_1(\psi_2 - \psi_1) \\ q_2 &= \Delta\psi_2 + F_2(\psi_1 - \psi_2), \end{aligned} \tag{1.25}$$

and (1.22) becomes

$$\begin{aligned} \partial_t q_1 + J(\psi_1, q_1 + \beta y) &= 0 \\ \partial_t q_2 + J(\psi_2, q_2 + \beta y) &= 0. \end{aligned}$$

In this work we will consider the case when the dynamic is affected by external forcing and dissipation. On both layer we take into account the effect of the eddy viscosity, on the top layer we consider a deterministic forcing which accounts for example for the effect of the wind on the upper ocean, and on the second layer the bottom friction r. In the next section we lay down its precise mathematical formulation and the main notations following closely the set up described in [8] and [15].

1.3.1 Mathematical setup

Let $\mathcal{D}$ be a squared domain $\mathcal{D} = [0, L] \times [0, L] \subset \mathbb{R}^2$ with $L > 0$ typical length scale of the dynamic (e.g. 10^6m for the atmosphere and 10^5m for the ocean) and consider the following equations

$$\begin{aligned} &\partial_t q_1 + J(\psi_1, q_1 + \beta y) = \nu\Delta^2\psi_1 + f \\ &\partial_t q_2 + J(\psi_2, q_2 + \beta y) = \nu\Delta^2\psi_2 - r\Delta\psi_2 \\ &\mathbf{q}(0, \mathbf{x}) = \mathbf{q}_0(\mathbf{x}) \end{aligned} \tag{1.26}$$

where $\mathbf{x} = (x_1, x_2) \in \mathcal{D}$, $\mathbf{q}(t, \mathbf{x}) = (q_1(t, \mathbf{x}), q_2(t, \mathbf{x}))^t$ is the so-called quasi–geostrophic potential vorticity, defined in (1.25) via the streamfunction of the fluid, $\boldsymbol{\psi}(t, \mathbf{x}) = (\psi_1(t, \mathbf{x}), \psi_2(t, \mathbf{x}))^t$, or in vectorial formulation

$$\mathbf{q} = (\Delta + M)\boldsymbol{\psi} \quad \text{with } M = \begin{pmatrix} -F_1 & F_1 \\ F_2 & -F_2 \end{pmatrix} \tag{1.27}$$

where $\Delta\boldsymbol{\psi} = (\Delta\psi_1, \Delta\psi_2)^t$ and F_1, F_2 are positive constants as in (1.23). The model (1.26) includes dissipation generated by the eddy viscosity ν on both layers and friction r on the bottom layer as well as a deterministic forcing on the top layer $f = f(t, \mathbf{x})$ with zero spatial averages, i.e.

$$\int_{\mathcal{D}} f(t, \mathbf{x})\, d\mathbf{x} = 0 \quad \text{for all } \ t \geq 0,$$

that accounts for the wind shear on the surface of the ocean. Furthermore we assume periodic boundary conditions for $\boldsymbol{\psi}$ in both directions with period L and we impose that

$$\int_{\mathcal{D}} \boldsymbol{\psi}(t, \mathbf{x})\, d\mathbf{x} = 0. \tag{1.28}$$

Next we set the notations for the two–layer quasi–geostrophic model used in this work. Let $L^2(\mathcal{D})$, $H^k(\mathcal{D})$, $k \in \mathbb{R}$ be the standard Sobolev spaces of L-periodic functions satisfying (1.28) with respectively L^2 norm defined as

$$\|u\|_0^2 := \int_{\mathcal{D}} |u(\mathbf{x})|^2\, d\mathbf{x}, \tag{1.29}$$

and the norm in $H^k(\mathcal{D})$ defined as

$$\|u\|_k^2 := \int_{\mathcal{D}} |(-\Delta)^{k/2} u(\mathbf{x})|^2\, d\mathbf{x}, \tag{1.30}$$

so that

$$\|\nabla u\|_k^2 = \|u\|_{k+1}^2 \quad \text{and} \quad \|\Delta u\|_k^2 = \|u\|_{k+2}^2.$$

We also introduce appropriate norms on the product spaces to deal with our coupled system. Given $\mathbf{u}$ and $\mathbf{v}$ elements of $H^k \times H^k$, $k > 0$ or $L^2 \times L^2$ for $k = 0$, define

$$\|\mathbf{u}\|_{p,k} = \left(\|\sqrt{h_1}u_1\|_k^p + \|\sqrt{h_2}u_2\|_k^p\right)^{1/p} \tag{1.31}$$

and in particular for $p = 2$ we write

$$\|\mathbf{u}\|_k^2 := \|\mathbf{u}\|_{2,k}^2 = h_1\|u_1\|_k^2 + h_2\|u_2\|_k^2 \tag{1.32}$$

$$(\mathbf{u}, \mathbf{v})_k := h_1(u_1, v_1)_k + h_2(u_2, v_2)_k. \tag{1.33}$$

Then we define

$$\mathbf{L}^2 = \left\{\mathbf{u} \in L^2 \times L^2 \,:\, \|\mathbf{u}\|_0^2 < \infty\right\} \tag{1.34}$$

$$\mathbf{H}^k = \left\{\mathbf{u} \in H^k \times H^k \,:\, \|\mathbf{u}\|_k^2 < \infty\right\}, \quad k > 0 \tag{1.35}$$

and we denote with $\mathbf{H}^{-k}$ the dual space of $\mathbf{H}^k$, $k > 0$.

Then Poincaré inequality in $\mathbf{H}^k$ reads as

$$\|\mathbf{u}\|_k \leq \lambda_1^{-1/2}\|\mathbf{u}\|_{k+1}, \quad k \geq 0, \tag{1.36}$$

where λ_1 is the smallest eigenvalue of the operator $-\Delta$.

Define the operator $\tilde{A} : \mathbf{H}^{k+2} \to \mathbf{H}^k$, $k \in \mathbb{R}$ as

$$\tilde{A}\mathbf{v} = -(\Delta + M)\mathbf{v} \quad \mathbf{v} \in \mathbf{H}^{k+2}. \tag{1.37}$$

It is easy to see that $\tilde{A}$ is self-adjoint, since $-\Delta$ is self-adjoint and for M we have

$$\begin{aligned}(M\mathbf{u}, \mathbf{v}) &= -h_1 v_1 F_1(u_1 - u_2) + h_2 v_2 F_2(u_1 - u_2)\\ &= -p(v_1 - v_2)(u_1 - u_2) = (\mathbf{u}, M\mathbf{v}).\end{aligned}$$

Furthermore its range

$$R_k(\tilde{A}) = \{\tilde{A}\mathbf{v} \,:\, \mathbf{v} \in \mathbf{H}^{k+2}\} \subset \mathbf{H}^k$$

is closed and $\tilde{A}^{-1}$ is a self-adjoint continuous operator in $\mathbf{H}^k$. Then we define

$$\mathbb{L}^2 := \mathbf{L}^2\big|_{R_0(\tilde{A})} \quad \text{and} \quad \mathbb{H}^k := \mathbf{H}^k\big|_{R_k(\tilde{A})} \quad k \in \mathbb{R}, k \neq 0.$$

Remark 1.3.1. Since the $\mathbf{L}^2$ and $\mathbf{H}^1$ norms and the $\mathbf{L}^2$ scalar product are the most used throughout all the chapters we denote them as follows

$$\begin{aligned}|v| := \|v\|_0 \quad &\text{and} \quad \|v\| := \|v\|_1\\ |\mathbf{v}| = h_1|v_1| + h_2|v_2| := \|\mathbf{v}\|_0 \quad &\text{and} \quad \|\mathbf{v}\| = h_1\|v_1\| + h_2\|v_2\| := \|\mathbf{v}\|_1\\ (u, v) := (u, v)_0 \quad &\text{and} \quad (\mathbf{u}, \mathbf{v}) = h_1(u_1, v_1) + h_2(u_2, v_2)\end{aligned}$$

also in accordance with what is typically used for Navier–Stokes equation.

Finally we introduce two new norms in $\mathbb{L}^2$ and $\mathbb{H}^{-1}$. For $\mathbf{u} \in \mathbb{H}^{-1}$ by definition there exists $\mathbf{v} \in \mathbf{H}^1$ such that $\mathbf{u} = \tilde{A}\mathbf{v}$, and we can define the norm on $\mathbb{H}^{-1}$

$$|||\mathbf{u}|||_{-1}^2 := \|\mathbf{v}\|^2 + p|v_1 - v_2|^2, \tag{1.38}$$

and, for $\mathbf{u} \in \mathbb{L}^2$ with $\mathbf{v} \in \mathbf{H}^2$ so that $\tilde{A}\mathbf{v} = \mathbf{u}$, define the norm on $\mathbb{L}^2$

$$|||\mathbf{u}|||_0^2 := \|\mathbf{v}\|_2^2 + p\|v_1 - v_2\|^2. \tag{1.39}$$

We set

$$\mathcal{H} = \left(\mathbb{H}^{-1}, |||\cdot|||_{-1}\right) \quad \text{and} \quad \mathcal{V} = \left(\mathbb{L}^2, |||\cdot|||_0\right). \tag{1.40}$$

Note that by Poincaré inequality $|||\mathbf{q}(t)|||_{-1}^2 \leq \lambda_1^{-1}|||\mathbf{q}(t)|||_0^2$ as by definition we have

$$\begin{aligned} |||\mathbf{q}(t)|||_{-1}^2 &= \|\boldsymbol{\psi}\|^2 + p|\psi_1 - \psi_2|^2 \\ &\leq \lambda_1^{-1}\left(|\Delta\boldsymbol{\psi}|^2 + p\|\psi_1 - \psi_2\|^2\right) = \lambda_1^{-1}|||\mathbf{q}(t)|||_0^2. \end{aligned} \tag{1.41}$$

Furthermore these norms are equivalent respectively to $\|\cdot\|_{-1}$ and $\|\cdot\|_0$, and have a series of useful properties.

Lemma 1.3.2. *Consider* $\mathbf{u} \in \mathbb{H}^{-1}$ *and* $\mathbf{v} \in \mathbf{H}^1$ *such that* $\tilde{A}\mathbf{v} = \mathbf{u}$. *Then the following relations hold:*

$$(\mathbf{u}, \mathbf{v}) = (\tilde{A}\mathbf{v}, \mathbf{v}) = |\tilde{A}^{1/2}\mathbf{v}|^2 = |||\mathbf{u}|||_{-1}^2 \tag{1.42}$$

$$\|\mathbf{v}\|^2 \leq |||\mathbf{u}|||_{-1}^2 \leq a_0\|\mathbf{v}\|^2 \tag{1.43}$$

for some $a_0 > 0$.

For $\mathbf{u} \in \mathbb{L}^2$ *and* $\mathbf{v} \in \mathbf{H}^2$ *such that* $\tilde{A}\mathbf{v} = \mathbf{u}$, *we have:*

$$(\mathbf{u}, \Delta\mathbf{v}) = -|||\mathbf{u}|||_0^2 \tag{1.44}$$

$$(|\Delta\mathbf{v}|^2 =)\|\mathbf{v}\|_2^2 \leq |||\mathbf{u}|||_0^2 \leq a_0\|\mathbf{v}\|_2^2. \tag{1.45}$$

Proof. We start by showing (1.42). By definition (1.37) of $\tilde{A}$ we have

$$(\mathbf{u}, \mathbf{v}) = (\tilde{A}\mathbf{v}, \mathbf{v}) = -(\Delta\mathbf{v}, \mathbf{v}) - (M\mathbf{v}, \mathbf{v})$$

then, by Green's theorem

$$-(\Delta\mathbf{v}, \mathbf{v}) = \|\mathbf{v}\|^2$$

and by (1.24), namely $F_1h_1 = F_2h_2 = p$ we have

$$-(M\mathbf{v}, \mathbf{v}) = F_1h_1(v_1 - v_2, v_1) + F_2h_2(v_2 - v_1, v_2) = p|v_1 - v_2|^2.$$

In a similar way we can show (1.44). By definition of $\tilde{A}$ we have

$$(\mathbf{u}, \Delta\mathbf{v}) = -(\Delta\mathbf{v}, \Delta\mathbf{v}) - (M\mathbf{v}, \Delta\mathbf{v})$$

then by definition of M (1.27), the relation (1.24) and Green's theorem

$$\begin{aligned} -(M\mathbf{v}, \Delta\mathbf{v}) &= h_1F_1(v_1 - v_2, \Delta v_1) + h_2F_2(v_2 - v_1, \Delta v_2) \\ &= -p((-\Delta)^{1/2}(v_1 - v_2), (-\Delta)^{1/2}v_1) + p((-\Delta)^{1/2}(v_1 - v_2), (-\Delta)^{1/2}v_2). \end{aligned}$$

Therefore we have

$$(\mathbf{u}, \Delta\mathbf{v}) = -|\Delta\mathbf{v}|^2 - p\|v_1 - v_2\|^2 = -|||\mathbf{u}|||_0^2.$$

Moving on to (1.43), the lower bound follows by definition of $|||\cdot|||_{-1}$. The upper bound can be computed thanks to Poincaré inequality (1.36) and the parallelogram law. Indeed we have

$$\begin{aligned} p|v_1 - v_2|^2 &\leq p\lambda_1^{-1}\|v_1 - v_2\|^2 \\ &\leq \frac{2p}{\lambda_1 \min(h_1, h_2)}(h_1\|v_1\|^2 + h_2\|v_2\|^2) \\ &= 2\lambda_1^{-1}\max(F_1, F_2)\|\mathbf{v}\|^2. \end{aligned}$$

Then since we assume $h_1 \leq h_2$, from (1.24) it follows that $F_1 \geq F_2$, and setting a_0 to be $\max(1, 2F_1/\lambda_1)$ we have the desired result

$$|||\mathbf{u}|||_{-1}^2 = \|\mathbf{v}\|^2 + p|v_1 - v_2|^2 \leq a_0\|\mathbf{v}\|^2.$$

The same calculation using the definition of $|||\cdot|||_0$ and the Poincaré inequality (1.36) for $k = 1$ gives (1.45). □

Similarly to the Navier–Stokes equation, taking the $\mathbf{L}^2$ scalar product of (1.26) with $\mathbf{v} \in \mathbf{H}^2$ we obtain the weak formulation

$$\begin{aligned} (\partial_t q_1, h_1v_1) + (J(\psi_1, q_1 + \beta y), h_1v_1) &= \nu(\Delta^2\psi_1, h_1v_1) + (f, h_1v_1) \\ (\partial_t q_2, h_2v_2) + (J(\psi_2, q_2 + \beta y), h_2v_2) &= \nu(\Delta^2\psi_2, h_2v_2) - r(\Delta\psi_2, h_2v_2). \end{aligned} \tag{1.46}$$

We can write (1.46) in vectorial formulation introducing the bilinear operator $B(\mathbf{u}, \mathbf{v})$ from $\mathbf{H}^2 \times \mathbf{H}^2$ into $\mathbf{H}^{-2}$ defined as

$$B(\mathbf{u}, \mathbf{v}) = \begin{pmatrix} J(u_1, \Delta v_1) + F_1J(u_1, v_2) \\ J(u_2, \Delta v_2) + F_2J(u_2, v_1) \end{pmatrix}, \tag{1.47}$$

for $\mathbf{u} = (u_1, u_2)$, $\mathbf{v} = (v_1, v_2) \in \mathbf{H}^2$. Then using the fact that $J(u,u) = \nabla^\perp u \cdot \nabla u = 0$ we can write (1.46) as

$$(\partial_t \mathbf{q}, \mathbf{v}) + (B(\boldsymbol{\psi}, \boldsymbol{\psi}), \mathbf{v}) + \beta\,(\partial_1 \boldsymbol{\psi}, \mathbf{v}) = \nu\left(\Delta^2 \boldsymbol{\psi}, \mathbf{v}\right) + \begin{pmatrix} h_1\langle f, v_1\rangle \\ -rh_2(\Delta\psi_2, v_2) \end{pmatrix}. \tag{1.48}$$

where ∂_1 denotes the derivative with respect to x_1.

Alternatively, regarding the variables $\mathbf{q}(t,\mathbf{x})$ and $\boldsymbol{\psi}(t,\mathbf{x})$ as trajectories in $\mathbb{H}^{-1}$ and $\mathbf{H}^1$ respectively

$$[\mathbf{q}(t)](\mathbf{x}) = \mathbf{q}(t,\mathbf{x}) \quad \text{and} \quad [\boldsymbol{\psi}(t)](\mathbf{x}) = \boldsymbol{\psi}(t,\mathbf{x}),$$

we can rewrite (1.48) as

$$\begin{aligned} &\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} + B(\boldsymbol{\psi}, \boldsymbol{\psi}) + \beta\partial_1\boldsymbol{\psi} = \nu\Delta^2\boldsymbol{\psi} + \begin{pmatrix} f \\ -r\Delta\psi_2 \end{pmatrix} \\ &\mathbf{q} = -\tilde{A}\boldsymbol{\psi}. \end{aligned} \tag{1.49}$$

Note that we expect the first equation to hold as equality in $\mathbb{H}^{-2}$ for $f \in H^{-2}$ and the relation between $\mathbf{q}$ and $\boldsymbol{\psi}$ as equality in $\mathbb{H}^{-1}$.

Table 1.1 contains a summary of the spaces and relative norms used throughout this work when working with the quasi–geostrophic model.

Space	Norm
$\mathbf{L}^2$	$\lvert\mathbf{v}\rvert^2 = h_1\lvert v_1\rvert^2 + h_2\lvert v_2\rvert^2$
$\mathbf{H}^1$	$\lVert\mathbf{v}\rVert^2 = h_1\lVert v_1\rVert^2 + h_2\lVert v_2\rVert^2$
$\mathbf{H}^k\ k \neq 0,1$	$\lVert\mathbf{v}\rVert_k^2 = h_1\lVert v_1\rVert_k^2 + h_2\lVert v_2\rVert_k^2$
$\mathcal{V} = \mathbb{L}^2 = \mathbf{L}^2\vert_{R(\tilde{A})}$	$\lVert\lvert\mathbf{u}\rvert\rVert_0^2 = \lVert\lvert -\tilde{A}\mathbf{v}\rvert\rVert_0^2 = \lVert\mathbf{v}\rVert_2^2 + p\lVert v_1 - v_2\rVert^2$
$\mathcal{H} = \mathbb{H}^{-1} = \mathbf{H}^{-1}\vert_{R(\tilde{A})}$	$\lVert\lvert\mathbf{u}\rvert\rVert_{-1}^2 = \lVert\lvert -\tilde{A}\mathbf{v}\rvert\rVert_{-1}^2 = \lVert\mathbf{v}\rVert^2 + p\lvert v_1 - v_2\rvert^2$

Table 1.1: Notations for the two–layer quasi–geostrophic model. The first block will be mainly used for the streamfunctions $\boldsymbol{\psi}$, while the second block is specific for the potential vorticities $\mathbf{q}$, i.e. functions in the range of the operator $\tilde{A}$.

1.3.2 Properties of the nonlinearity

The Jacobian operator

$$J(u,v) = \nabla^\perp u \cdot \nabla v,$$

has a series of useful properties and estimates that will be extensively used in this work. They are presented in [15, Lemma 3.1] which we report here to ease reference.

Lemma 1.3.3. *The Jacobian operator verifies the following properties*

$$J(u,v) = -J(v,u), \quad J(u,u) = 0 \quad u,v \in H^1 \tag{1.50}$$
$$(J(u,v),w) = (J(v,w),u), \quad (J(u,v),v) = 0 \quad u,v,w \in H^2 \tag{1.51}$$

and the following estimates hold

$$|(J(u,v),\Delta u)| \le c_0|\Delta v|\|u\||\Delta u|, \quad u,v \in H^2 \tag{1.52}$$
$$|(J(u,v),w)| \le c_1|\Delta u||\Delta v||w|, \quad u,v \in H^2, w \in L^2 \tag{1.53}$$
$$|(J(u,v),w)| \le c_1\|u\||\Delta v|\|w\|, \quad u,v,w \in H^2. \tag{1.54}$$

Here above $c_0 = 2+(\sqrt{2}\cdot\pi)^{-1}$ *and* $c_1 = c_0\lambda_1^{-\frac{1}{2}}$*, with* λ_1 *the smallest eigenvalue of* $-\Delta$*.*

We stress that the form of these estimates is specific to the fact that we have periodic boundary conditions. These results translate into the properties of the bilinear operator B (1.47):

Lemma 1.3.4. *Let* B *be the bilinear operator* (1.47)*, then for* $\mathbf{u}, \mathbf{w} \in \mathbf{H}^1$ *and* $\mathbf{v} \in \mathbf{H}^2$

$$(B(\mathbf{u},\mathbf{v}),\mathbf{w}) = -(B(\mathbf{w},\mathbf{v}),\mathbf{u}), \tag{1.55}$$
$$(B(\mathbf{u},\mathbf{v}),\mathbf{u}) = 0. \tag{1.56}$$

Moreover there exists positive constant k_0 *and* k_B *such that the following estimates hold for* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{H}^2$

$$|(B(\mathbf{u},\mathbf{u}),\mathbf{v})| \le k_0\|\mathbf{u}\||\Delta\mathbf{u}||\Delta\mathbf{v}| \tag{1.57}$$
$$|(B(\mathbf{u},\mathbf{v}),\mathbf{v})| \le \tfrac{\nu}{2}|\Delta\mathbf{v}|^2 + k_B|\Delta\mathbf{u}|^2\|\mathbf{v}\|^2 \tag{1.58}$$
$$|(B(\mathbf{u},\mathbf{v}),\mathbf{v})| \le \tfrac{\nu}{2}|\Delta\mathbf{u}|^2 + k_B|\Delta\mathbf{v}|^2\|\mathbf{v}\|^2. \tag{1.59}$$

Proof. The relation (1.56) follows from (1.55). The latter is a direct consequence of the definition of the operator B (1.47) and of the properties of the Jacobian $(J(u,v),w) = (J(v,w),u)$ and $J(u,v) = -J(v,u)$.

By the definition of B we have

$$\begin{aligned}|(B(\mathbf{u},\mathbf{u}),\mathbf{v})| = |&(J(u_1,\Delta u_1),h_1v_1) + F_1(J(u_1,u_2),h_1v_1)\\ &+(J(u_2,\Delta u_2),h_2v_2) + F_2(J(u_2,u_1),h_2v_2)|\,.\end{aligned}$$

Thanks to the estimates (1.53) and (1.54) we have

$$|(B(\mathbf{u},\mathbf{u}),\mathbf{v})| \le c_0 h_1 \|u_1\||\Delta u_1||\Delta v_1| + c_1 p\|u_1\|\|u_2\||\Delta v_1| \\ + c_0 h_2 \|u_2\||\Delta u_2||\Delta v_2| + c_1 p\|u_1\|\|u_2\||\Delta v_2|$$

Using Poincaré inequality (1.36), the fact that $c_1 = c_0 \lambda_1^{-1/2}$ and $p = F_1 h_1 = F_2 h_2$ it is easy to see that

$$\le c_0 \sqrt{h_1}\|u_1\||\Delta v_1| \left(\sqrt{h_1}|\Delta u_1| + \tfrac{\sqrt{pF_2}}{\lambda_1}\sqrt{h_2}|\Delta u_2|\right) \\ + c_0\sqrt{h_2}\|u_2\||\Delta v_2|\left(\tfrac{\sqrt{pF_1}}{\lambda_1}\sqrt{h_1}|\Delta u_1| + \sqrt{h_2}|\Delta u_2|\right).$$

Since we assumed without loss of generality that $h_1 \le h_2$ and $F_1 \ge F_2$, rearranging appropriately we get

$$|(B(\mathbf{u},\mathbf{u}),\mathbf{v})| \le \tfrac{c_0}{\sqrt{h_1}} \max(1, \tfrac{\sqrt{pF_1}}{\lambda_1}) \left(\sqrt{h_1}|\Delta u_1| + \sqrt{h_2}|\Delta u_2|\right) \cdot \\ \cdot \left(h_1\|u_1\||\Delta v_1| + h_2\|u_2\||\Delta v_2|\right).$$

As $a + b \le \sqrt{2(a^2+b^2)}$ we can derive the desired estimate (1.57)

$$|(B(\mathbf{u},\mathbf{u}),\mathbf{v})| \le k_0 |\Delta \mathbf{u}|\|\mathbf{u}\||\Delta \mathbf{v}|$$

where

$$k_0 = \tfrac{2c_0\sqrt{2}}{\sqrt{h_1}} \max(1, \tfrac{\sqrt{pF_1}}{\lambda_1}).$$

Next, both inequalities (1.58) and (1.59) derive from the same initial estimate we are about to derive. By definition of B and triangular inequality we have

$$|(B(\mathbf{u},\mathbf{v}),\mathbf{v})| \le h_1|(J(v_1,u_1),\Delta v_1)| + F_1 h_1 |(J(v_2,u_1),v_1)| \\ + h_2|(J(v_2,u_2),\Delta v_2)| + F_2 h_2|(J(v_1,u_2),v_2)|$$

so, by the estimates (1.52) and (1.54),

$$|(B(\mathbf{u},\mathbf{v}),\mathbf{v})| \le c_0 h_1 |\Delta u_1||\Delta v_1|\|v_1\| + c_1 F_1 h_1 \|v_1\|\|v_2\||\Delta u_1| \\ + c_0 h_2|\Delta u_2||\Delta v_2|\|v_2\| + c_1 F_2 h_2 \|v_1\|\|v_2\||\Delta u_2|. \quad (1.60)$$

Now we use Young's inequality to get first (1.58). Indeed

$$c_0 h_i |\Delta u_i||\Delta v_i|\|v_i\| \le \tfrac{c_0^2 h_i}{\nu}|\Delta u_i|^2\|v_i\|^2 + \tfrac{\nu h_i}{4}|\Delta v_i|^2,$$
$$c_1 F_i h_i \|v_i\|\|v_j\||\Delta u_i| \le \tfrac{c_1^2 F_i^2 h_i}{\nu\lambda_1}\|\Delta u_i\|^2\|v_j\|^2 + \tfrac{\nu\lambda_1 h_i}{4}\|v_i\|^2,$$

so that

$$|(B(\mathbf{u},\mathbf{v}),\mathbf{v})| \le \tfrac{c_0^2}{\nu}\left(h_1|\Delta u_1|^2\|v_1\|^2 + h_2|\Delta u_2|^2\|v_2\|^2\right) + \tfrac{\nu}{4}|\Delta \mathbf{v}|^2 \\ + \tfrac{c_1^2F_1^2h_1}{\nu\lambda_1}|\Delta u_1|^2\|v_2\|^2 + \tfrac{c_1^2F_2^2h_2}{\nu\lambda_1}|\Delta v_2|^2\|v_1\|^2 + \tfrac{\nu\lambda_1}{4}\|\mathbf{v}\|^2.$$

Then by Poincaré's inequality, and since $h_1 \le h_2$ and $F_1 \ge F_2$,

$$|(B(\mathbf{u},\mathbf{v}),\mathbf{v})| \le \tfrac{c_0^2}{\nu h_1}\left(h_1^2|\Delta u_1|^2\|v_1\|^2 + h_2^2|\Delta u_2|^2\|v_2\|^2\right) + \tfrac{\nu}{2}|\Delta \mathbf{v}|^2 \\ + \tfrac{c_1^2F_1^2}{\nu h_1\lambda_1}\left(h_1h_2|\Delta v_1|^2\|v_2\|^2 + h_1h_2|\Delta v_2|^2\|v_1\|^2\right).$$

Finally setting

$$k_B = \max\left(\tfrac{c_0^2}{\nu h_1}, \tfrac{c_1^2F_1^2}{\nu h_1\lambda_1}\right), \tag{1.61}$$

we have

$$|(B(\mathbf{u},\mathbf{v}),\mathbf{v})| \le \tfrac{\nu}{2}|\Delta\mathbf{v}|^2 + k_B|\Delta\mathbf{u}|^2\|\mathbf{v}\|^2.$$

To prove (1.59) we use Young's inequality in a different way in (1.60), namely

$$c_0h_i|\Delta u_i||\Delta v_i|\|v_i\| \le \frac{c_0^2h_i}{\nu}|\Delta v_i|^2\|v_i\|^2 + \frac{\nu h_i}{4}|\Delta u_i|^2$$
$$c_1F_ih_i\|v_i\|\|v_j\||\Delta u_i| \le \frac{c_1^2F_i^2h_i}{\nu}\|v_i\|^2\|v_j\|^2 + \frac{\nu h_i}{4}|\Delta u_i|^2.$$

Therefore we have

$$|(B(\mathbf{u},\mathbf{v}),\mathbf{v})| \le \tfrac{c_0^2}{\nu}\left(h_1|\Delta v_1|^2\|v_1\|^2 + h_2|\Delta v_2|^2\|v_2\|^2\right) + \tfrac{\nu}{2}|\Delta \mathbf{u}|^2 \\ + \tfrac{c_1^2F_1^2h_1}{\nu}\|v_1\|^2\|v_2\|^2 + \tfrac{c_1^2F_2^2h_2}{\nu}\|v_2\|^2\|v_1\|^2.$$

Poincaré's inequality gives

$$\tfrac{c_1^2F_1^2}{\nu}\|v_1\|^2\|v_2\|^2 + \tfrac{c_1^2F_2^2}{\nu}\|v_2\|^2\|v_1\|^2 \le \tfrac{c_1^2F_1^2h_1}{\nu\lambda_1}|\Delta v_1|^2\|v_2\|^2 + \tfrac{c_1^2F_2^2h_2}{\nu\lambda_1}|\Delta v_2|^2\|v_1\|^2,$$

so that

$$|(B(\mathbf{u},\mathbf{v}),\mathbf{v})| \le \tfrac{c_0^2}{h_1\nu}\left(h_1^2|\Delta v_1|^2\|v_1\|^2 + h_2^2|\Delta v_2|^2\|v_2\|^2\right) + \tfrac{\nu}{2}|\Delta \mathbf{u}|^2 \\ \tfrac{c_1^2F_1^2}{\nu h_1\lambda_1}\left(h_1h_2|\Delta v_1|^2\|v_2\|^2 + h_1h_2|\Delta v_2|^2\|v_1\|^2\right).$$

Recalling the definition of k_B we have the desired result

$$|(B(\mathbf{u},\mathbf{v}),\mathbf{v})| \le \tfrac{\nu}{2}|\Delta\mathbf{u}|^2 + k_B|\Delta\mathbf{v}|^2\|\mathbf{v}\|^2.$$

□

We conclude this section studying the derivative of the operator B as it will be useful when studying the differentiability of the solutions with respect to parameters in section 2.4.3.

Lemma 1.3.5. *Let B be the bilinear operator as defined* (1.47) *and consider its restriction on the diagonal space $\mathbb{D} = \{(\mathbf{u}, \mathbf{u}) \ : \ \mathbf{u} \in \mathbf{H}^2\}$. Then the operator $B : \mathbb{D} \to \mathbf{H}^{-2}$ satisfies the bound*

$$\|B(\mathbf{u}, \mathbf{u})\|_{-2} \leq k_0 \|\mathbf{u}\| |\Delta \mathbf{u}|. \tag{1.62}$$

Furthermore its derivative is well defined and is

$$DB(\mathbf{u}, \mathbf{u})\mathbf{v} = B(\mathbf{v}, \mathbf{u}) + B(\mathbf{u}, \mathbf{v}). \tag{1.63}$$

Proof. By definition

$$\|B(\mathbf{u}, \mathbf{u})\|_{-2} = \sup_{\|\mathbf{v}\|_2 \neq 0} \frac{|(B(\mathbf{u}, \mathbf{u}), \mathbf{v})|}{\|\mathbf{v}\|_2}.$$

Using the estimate (1.57) we have

$$\|B(\mathbf{u}, \mathbf{u})\|_{-2} \leq \sup_{\|\mathbf{v}\|_2 \neq 0} \frac{k_0 \|\mathbf{u}\| |\Delta \mathbf{u}| |\Delta \mathbf{v}|}{\|\mathbf{v}\|_2} \tag{1.64}$$

hence since $\|\mathbf{v}\|_2 = |\Delta \mathbf{v}|$ we get the desired estimate

$$\|B(\mathbf{u}, \mathbf{u})\|_{-2} \leq k_0 \|\mathbf{u}\| |\Delta \mathbf{u}|.$$

Next, by definition $DB(\mathbf{u}, \mathbf{u}) \in L(\mathbf{H}^2, \mathbf{H}^{-2})$ is the derivative of B at $\mathbf{u}$ if

$$\lim_{\|\mathbf{w}\| \to 0} \frac{\|B(\mathbf{u} + \mathbf{w}, \mathbf{u} + \mathbf{w}) - B(\mathbf{u}, \mathbf{u}) - DB(\mathbf{u}, \mathbf{u})\mathbf{w}\|_{-2}}{\|\mathbf{w}\|_2} = 0.$$

Now, by bilinearity one has

$$B(\mathbf{u} + \mathbf{w}, \mathbf{u} + \mathbf{w}) - B(\mathbf{u}, \mathbf{u}) = B(\mathbf{u}, \mathbf{w}) + B(\mathbf{w}, \mathbf{u}) + B(\mathbf{w}, \mathbf{w}),$$

therefore

$$\begin{aligned} \frac{\|B(\mathbf{u} + \mathbf{w}, \mathbf{u} + \mathbf{w}) - B(\mathbf{u}, \mathbf{u}) - DB(\mathbf{u}, \mathbf{u})\mathbf{w}\|_{-2}}{\|\mathbf{w}\|_2} &\leq \frac{\|B(\mathbf{w}, \mathbf{w})\|_{-2}}{\|\mathbf{w}\|_2} \\ &+ \frac{\|B(\mathbf{u}, \mathbf{w}) + B(\mathbf{w}, \mathbf{u}) - DB(\mathbf{u}, \mathbf{u})\mathbf{w}\|_{-2}}{\|\mathbf{w}\|_2}. \end{aligned}$$

To prove (1.63) we need to show that the first term on the right hand side converges to zero as $\|\mathbf{w}\|$ goes to zero. Using the estimate found in the first part, (1.62) we have

$$\frac{\|B(\mathbf{w},\mathbf{w})\|_{-2}}{\|\mathbf{w}\|_2} \leq k_0\|\mathbf{w}\| \leq k_0\lambda_1^{-1/2}\|\mathbf{w}\|_2.$$

Therefore, for all $\mathbf{w} \in \mathbf{H}^2$, $\frac{\|B(\mathbf{w},\mathbf{w})\|_{-2}}{\|\mathbf{w}\|_2} \to 0$ as $\|\mathbf{w}\|_2 \to 0$ and

$$DB(\mathbf{w},\mathbf{w})\mathbf{v} = B(\mathbf{v},\mathbf{w}) + B(\mathbf{w},\mathbf{v}).$$

□

Now we are ready to introduce the stochastic version of the two–layer quasi–geostrophic model we will study in this work.

1.4 A stochastic two–layer quasi–geostrophic model

Stochastically forced QG models have been considered in applications (see e.g. [38, 23]) and mathematical studies (see e.g. [9, 25, 26, 27]). Regarding the case of noise acting only on the first layer though, the only mathematical reference seems to be [15]. In section 1.4.1 we will present the mathematical set up for a such model following closely [15]. In section 1.4.2 we recall some technical results on the Ornstein-Uhlenbeck process which will be particularly relevant in chapter 2 and section 3.1.

1.4.1 Stochastic and Random model

Consider the framework described in section 1.3 with $\mathcal{H} = (\mathbb{H}^{-1}, |||\cdot|||_{-1})$ and $\mathcal{V} = (\mathbb{L}^2, |||\cdot|||_0)$ and perturb the top layer of the deterministic model (1.26) with a Q-Wiener process with values in L^2 to get

$$\begin{aligned} &dq_1 + J(\psi_1, q_1 + \beta y)\,dt = \left(\nu\Delta^2\psi_1 + f\right)dt + dW \\ &\partial_t q_2 + J(\psi_2, q_2 + \beta y) = \nu\Delta^2\psi_2 - r\Delta\psi_2 \\ &\mathbf{q} = -\tilde{A}\boldsymbol{\psi} = (\Delta + M)\boldsymbol{\psi}, \end{aligned} \tag{1.65}$$

and in vectorial formulation

$$d\mathbf{q} + (B(\boldsymbol{\psi},\boldsymbol{\psi}) + \beta\partial_1\boldsymbol{\psi})\,dt = \nu\Delta^2\boldsymbol{\psi}\,dt + \begin{pmatrix} f \\ -r\Delta\psi_1 \end{pmatrix} dt + d\mathbf{W} \tag{1.66a}$$

$$\mathbf{q} = -\tilde{A}\boldsymbol{\psi} = (\Delta + M)\boldsymbol{\psi} \tag{1.66b}$$

with initial condition $\mathbf{q}(0) = \mathbf{q}_0 \in \mathcal{H}$, where $\mathbf{W}$ the Wiener process on $\mathbf{L}^2$ defined as $\mathbf{W} = (W, 0)^t$.

To study the existence of solutions as well as the existence of invariant measure, concept that we will introduce precisely in chapter 3, we will use an approach most common in the study of random dynamical systems associated to SPDEs: transforming the stochastic equation in a partial differential equation with random coefficient by means of a change of variable (see for example [17]). One of its early uses was the 1973 work of Bensoussan and Temam on the stochastic Navier–Stokes equation [4] and this approach can now be considered classic in the specific literature.

Let $\eta(t, \mathbf{x}, \omega)$ be an auxiliary Ornstein-Uhlenbeck process defined by the solution of the linear equation

$$d\eta(t) - \nu(\alpha + 1)\Delta\eta(t)\, dt = dW(t), \quad \eta(0) = \eta_0 \in L^2 \tag{1.67}$$

with periodic boundary conditions, same realization of the noise, and $\alpha > 0$ a free control parameter. Now transform (1.66) via the change of variables

$$\tilde{q}_1 := q_1 - \eta; \; \tilde{q}_2 = q_2, \tag{1.68}$$

and use (1.67) to get

$$\begin{aligned} &d\tilde{\mathbf{q}} + (B(\boldsymbol{\psi}, \boldsymbol{\psi}) + \beta\partial_1\boldsymbol{\psi})\, dt = \nu\Delta^2\boldsymbol{\psi}\, dt + \begin{pmatrix} f - \nu(\alpha + 1)\Delta\eta(t) \\ -r\Delta\psi_1 \end{pmatrix} dt \\ &\tilde{\mathbf{q}} = (\Delta + M)\boldsymbol{\psi} - \begin{pmatrix} \eta \\ 0 \end{pmatrix}. \end{aligned} \tag{1.69}$$

To ensure the new variable $\tilde{\mathbf{q}}$ is in the range of $\tilde{A}$, namely it has a formulation as in (1.66b), we introduce the process $\boldsymbol{\xi} = (\xi_1, \xi_2)$, solution of the linear elliptic equation

$$(\Delta + M)\boldsymbol{\xi} = -\begin{pmatrix} \eta \\ 0 \end{pmatrix}. \tag{1.70}$$

Then define the new variable $\tilde{\boldsymbol{\psi}} = \boldsymbol{\psi} + \boldsymbol{\xi}$ so that we can write (1.69) as

$$\begin{aligned} &\frac{\mathrm{d}\tilde{\mathbf{q}}}{\mathrm{d}t} + B(\tilde{\boldsymbol{\psi}} - \boldsymbol{\xi}, \tilde{\boldsymbol{\psi}} - \boldsymbol{\xi}) + \beta\partial_1(\tilde{\boldsymbol{\psi}} - \boldsymbol{\xi}) = \nu\Delta^2(\tilde{\boldsymbol{\psi}} - \boldsymbol{\xi})\, dt + \begin{pmatrix} f - \nu(\alpha + 1)\Delta\eta \\ -r\Delta(\tilde{\psi}_2 - \xi_2) \end{pmatrix} \\ &\tilde{\mathbf{q}} = -\tilde{A}\tilde{\boldsymbol{\psi}} = (\Delta + M)\tilde{\boldsymbol{\psi}}. \end{aligned}$$

Finally, using (1.70), we get the following random version of the stochastically forced two–layer quasigeostrophic system in $\mathcal{H}$

$$\begin{aligned}
&\frac{\mathrm{d}\tilde{\mathbf{q}}}{\mathrm{d}t} + B(\tilde{\boldsymbol{\psi}} - \boldsymbol{\xi}, \tilde{\boldsymbol{\psi}} - \boldsymbol{\xi}) + \beta\partial_1(\tilde{\boldsymbol{\psi}} - \boldsymbol{\xi}) = \nu\Delta(\Delta\tilde{\boldsymbol{\psi}} + M\boldsymbol{\xi})\,dt + \begin{pmatrix} f - \nu\alpha\Delta\eta \\ -r\Delta(\tilde{\psi}_2 - \xi_2) \end{pmatrix} \\
&\tilde{\mathbf{q}} = (\Delta + M)\tilde{\boldsymbol{\psi}} \\
&\tilde{\mathbf{q}}(0) = \mathbf{q}_0 - \begin{pmatrix} \eta_0 \\ 0 \end{pmatrix}.
\end{aligned} \tag{1.71}$$

Last, note that given the definition of $\boldsymbol{\xi}$ (1.70), it is clear that the processes ξ_1, ξ_2 are more regular than the process η in the spatial variable, in particular if η has values in H^k then

$$\|\xi_i\|_{k+2} \le \|\eta\|_k, \quad i = 1, 2, \quad \text{and} \quad \|\xi_1 - \xi_2\|_{k+2} \le \|\eta\|_k \quad k \in \mathbb{R}. \tag{1.72}$$

1.4.2 Some properties of the Ornstein-Uhlenbeck process

Set $U = (H^k, \|\cdot\|)$ and let A be the Stokes operator $Au = -\Delta u$ so that $U = D(A^{k/2})$ and let e^{-tA} be the bounded semigroup generated by $-A$ in U (see e.g. [22, Appendix A.4]). Given a Wiener process W with covariance operator Q commuting with A, we consider the Stokes equation

$$d\eta + \alpha_1 A\eta\, dt = dW, \quad \eta(t_0) = \eta_0 \in U. \tag{1.73}$$

with constant $\alpha_1 = \nu(\alpha + 1) > 0$. By the classic theory of linear equations with additive noise there exists a unique solution η which we can expressed as

$$\eta(t) = \eta_0 e^{-(t-t_0)\alpha_1 A} + W_A(t) \qquad \text{with} \quad W_A(t) = \int_{t_0}^{t} e^{-\alpha_1(t-s)A}\, dW_s.$$

The process W_A is called *stochastic convolution* and has the following properties.

Lemma 1.4.1 ([22, Theorem 5.2])**.** *Let W be a Q-Wiener process on U. Consider the process*

$$W_A(t) = \int_0^t e^{-(t-s)\alpha_1 A}\, dW(s)$$

and assume that

$$\int_0^T \|e^{-t\alpha_1 A}\|^2_{L_2^0(U)}\, dt < \infty. \tag{1.74}$$

Then

(i) W_A *is Gaussian, continuous in mean square i.e.* $\mathbb{E}\|X(t)\|^2$ *is continuous in* t*, and has a predictable version;*

(ii) the trajectories of W_A *are* $\mathbb{P}$*-a.s. square integrable and* Law W_A *is a symmetric Gaussian measure on* $L^2(0,T;U)$*.*

Furthermore when Q is a trace class operator the process W_A admits a continuous version. It will be evident from the proof that this result holds true also for a less regular covariance operator but for sake of simplicity and consistency with the rest of this work we will always consider Q at least trace class.

Theorem 1.4.2. *Let* W *be a* Q*-Wiener process on* U *with covariance operator trace class* $Q \in L_1(U)$*. Then there exists a continuous version of* W_A *with values in* U*.*

Proof. Theorem 5.11 in [22] ensures that if there exists an $\delta \in (0, 1/2)$ such that

$$\int_0^t s^{-2\delta} \|e^{-s\alpha_1 A}\|^2_{L_2^0(U)}\, ds < \infty, \tag{1.75}$$

then there exists a continuous version of W_A with values in U. Let us then prove that (1.75) holds. Let ε be an arbitrary small positive constant, then

$$\begin{aligned}\int_0^t s^{-2\delta} \|e^{-s\alpha_1 A}\|^2_{L_2^0(U)}\, ds &= \int_0^t s^{-2\delta} \|e^{-s\alpha_1 A} Q^{1/2}\|^2_{L_2(U)}\, ds \\ &\le \|Q^{1/2} A^{\varepsilon - 1/2}\|^2_{L_2(U)} \int_0^t s^{-2\delta} \|A^{1/2-\varepsilon} e^{-s\alpha_1 A}\|^2_{L(U)}\end{aligned}$$

where $L(U)$ is the space of bounded linear operators on U. Furthermore it is known that (see e.g. [22, eqn 5.23]) there exists $c > 0$ such that

$$\|A^\beta e^{-t\alpha_1 A}\|_{L(U)} \le c t^{-\beta}$$

so that

$$\int_0^t s^{-2\delta} \|e^{-s\alpha_1 A}\|^2_{L_2^0(U)}\, ds \le \|Q^{1/2} A^{\varepsilon-1/2}\|^2_{L_2(U)} \int_0^t \frac{c^2}{s^{1-2\varepsilon+2\delta}}\, ds.$$

Then for any $0 < \delta < \varepsilon$ the integral on the right hand side is well defined and we have the desired result as long as

$$\|Q^{1/2} A^{\varepsilon-1/2}\|^2_{L_2(U)} < \infty$$

We can see that for such a condition to hold it is enough for Q to be trace class in U. In fact, if $\{e_k\}_{k\in\mathbb{N}}$ is a complete basis of U made of eigenfunctions of A we have

$$\|Q^{1/2}A^{\varepsilon-1/2}\|^2_{L_2(U)} = \sum_{k\in\mathbb{N}} \|Q^{1/2}A^{\varepsilon-1/2}e_k\|^2 = \sum_{k\in\mathbb{N}} \lambda_k^{2\varepsilon-1}\|Q^{1/2}e_k\|^2.$$

Since the eigenvalues of A form an increasing sequence in $\mathbb{R}^+$, for all $\varepsilon < 1/2$ we have

$$\sum_{k\in\mathbb{N}} \lambda_k^{2\varepsilon-1}\|Q^{1/2}e_k\|^2 \le \lambda_1^{2\varepsilon-1}\sum_{k\in\mathbb{N}} \|Q^{1/2}e_k\|^2 = \lambda_1^{2\varepsilon-1}\|Q^{1/2}\|^2_{L_2(U)}$$

so that it is enough to have $\|Q^{1/2}\|^2_{L_2(U)} < \infty$. □

In addition we can find an explicit bound for the mean square norm of η which can be made arbitrarily small picking a sufficiently large value of the parameter α_1, i.e. of the control parameter α. This result will later prove crucial in section 3.1.

Lemma 1.4.3. *Let η be the solution of* (1.73) *with initial data $\eta(t_0) = \eta_0 \in H^k$ and covariance Q trace class in H^k, i.e.* $\operatorname{Tr}_k Q = \operatorname{Tr}_0 A^k Q < \infty$ *for $k \ge 0$. Then*

$$\mathbb{E}\,\|\eta(t)\|_k^2 \le \|\eta(t_0)\|_k^2 e^{-2\alpha_1\lambda_1(t-t_0)} + \frac{\operatorname{Tr}_k Q}{2\alpha_1\lambda_1} \tag{1.76}$$

for all $t, t_0 \in \mathbb{R}$, $t_0 \le t$.

Proof. By Lemma A.3, since the function $f(t) = \mathbb{E}\|\eta(t)\|_k^2$ is continuous and non-negative, if we show that

$$f(s) - f(r) \le -\gamma \int_r^s f(\tau)\,d\tau + K(s-r) \quad \text{for all } r < s \tag{1.77}$$

then we have the desired result, i.e.

$$f(t) \le f(t_0)e^{-\gamma(t-t_0)} + K/\gamma \quad \text{for all } t \ge 0. \tag{1.78}$$

Take the H^k scalar product of (1.73) with η itself and by Itô formula we get

$$\|\eta(t)\|_k^2 = \|\eta(t_0)\|_k^2 - 2\alpha_1\int_{t_0}^t \|\eta(\tau)\|_{k+1}^2\,d\tau + (t-t_0)\operatorname{Tr}_k Q + 2\int_{t_0}^t (\eta, dW)_k$$

and in expectation

$$\mathbb{E}\,\|\eta(t)\|_k^2 = \mathbb{E}\,\|\eta(t_0)\|_k^2 - 2\alpha_1\mathbb{E}\int_{t_0}^t \|\eta(\tau)\|_{k+1}^2\,d\tau + (t-t_0)\,\mathrm{Tr}_k\,Q.$$

By Poincaré inequality $\|\eta\|_k^2 \le \lambda_1^{-1}\|\eta\|_{k+1}^2$ we have the desired expression

$$\mathbb{E}\,\|\eta(t)\|_k^2 \le \mathbb{E}\,\|\eta(t_0)\|_k^2 - 2\alpha_1\lambda_1\int_0^t \mathbb{E}\,\|\eta(\tau)\|_k^2\,d\tau + (t-t_0)\,\mathrm{Tr}_k\,Q$$

so that, setting $K = \mathrm{Tr}_U\,Q$ and $\gamma = 2\alpha_1\lambda_1$, Lemma A.3 gives

$$\mathbb{E}\,\|\eta(t)\|_k^2 \le \|\eta(t_0)\|_k^2 e^{-2\alpha_1\lambda_1(t-t_0)} + \frac{\mathrm{Tr}_k\,Q}{2\alpha_1\lambda_1}.$$

□

Furthermore, as we will see in an application of the results of chapter 4, (Example 4.1.14) the Ornestein Uhlenbeck process solution of (1.73) is ergodic, i.e. for all summable $f : U \to \mathbb{R}$,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t f(\eta(s))\,ds = \mathbb{E}f(\eta(0)) \quad \mathbb{P}\text{-a.s.}.$$

Summary and remarks

In this first chapter we set the mathematical framework for the models we will work on. After recalling the stochastic Navier–Stokes equation, we focused on the two–layer quasi–geostrophic model, with and without additive stochastic forcing. In particular, we introduced the Hilbert spaces essential in the analysis to come, see Table 1.1, and estimates on the Jacobian (Lemma 1.3.3 and Lemma 1.3.4) which will be used extensively.

Going forward, it is useful to keep in mind the relation between the velocity u, the streamfunction ψ and the potential vorticity q: the velocity is the orthogonal gradient of the streamfunction, so always one order of differentiability lower than the streamfunction, while the potential vorticity is comparable with the laplacian of the streamfunction, two orders of differentiability lower than it. This, together with classic results for Navier-Stokes, will give us an intuition on the spaces over which the stochastic two–layer QG model is well posed, topic central to the next chapter.

Chapter 2

Solutions and their properties

Given the similarities between the stochastic Navier–Stokes equation and the two–layer quasi–geostrophic model, it is natural to first look at the results for the former to prove similar ones for the latter. A possible approach to prove the existence of solutions is the one presented in [4, 32, 31] which we will briefly treat in section 2.1. This is by no means the only way possible (see e.g. [19]) but we selected it as it will prove extremely useful to show the existence of an invariant measure in the next chapter. Moreover this approach provides a direct proof of Theorem 3.1 [15] i.e. the well-posedness of the equation with random coefficients (1.71). In this approach the regularity of the Ornstein-Uhlenbeck process and that of the solutions of the random equation are combined to derive information on the solution of the stochastic equation. Afterwards, in section 2.4 and section 2.5, we will study how the solutions of the stochastic quasi–geostrophic model and the stochastic Navier–Stokes model depend on the intensity of the external forcing f. These results will later play a crucial role in chapter 5 where we will study the dependence of the long time average behaviour of these models with respect to this parameter.

2.1 Methodology and main results

We will explain the methodology first with the stochastic 2D Navier–Stokes equation as an example.

Consider the setup outlined in section 1.2 with the Hilbert spaces $(\mathcal{H}, |\cdot|)$ and $(\mathcal{V}, \|\cdot\|)$ as in (1.12) and the stochastic Navier–Stokes equation

$$du + \nu Au\, dt + B(u, u)\ dt = f\, dt + dW \quad u(0) = u_0 \tag{2.1}$$

where $u_0 \in \mathcal{H}$, W is a Q-Wiener process in $\mathcal{H}$ with Q trace class operator.

The first concept of solution we introduce is that of weak solution (see for example [22, Chapter 7]).

Definition 2.1.1. An $\mathcal{H}$-valued process $u(t)$, $t \in [0,T]$ is said to be a *weak* solution of (2.1) if the trajectories of u are $\mathbb{P}$-a.s. integrable,

$$u \in C([0,T];\mathcal{H}) \cap L^2(0,T;\mathcal{V}) \quad \mathbb{P}\text{-a.s.}$$

and if for all $\varphi \in D(A)$ and all $t \in [0,T]$ we have

$$(u(t),\varphi) + \int_0^t (u(s),A\varphi)\,ds + \int_0^t (B(u,u),\varphi)\,ds = (u_0,\varphi) + \int_0^t (f(s),\varphi)\,ds + (W(t),\varphi) \quad \mathbb{P}\text{-a.s.} \tag{2.2}$$

As observed in [31], a sufficiently regular noise would ensure the desired regularity of the solution and its uniqueness:

Theorem 2.1.2. *Let $u_0 \in \mathcal{H}$, $f \in L^2(0,T;\mathcal{V}^*)$. If Q is trace class in $\mathcal{V}$ then there exists a unique weak solution of* (2.1)

$$u \in C([0,T];\mathcal{H}) \cap L^2(0,T;\mathcal{V}) \quad \mathbb{P}\text{-}a.s.$$

Proof. The solution is derived thanks to the classic change of variables already introduced for the quasigeostrophic model in section 1.4.1. We consider the linear equation

$$dz + \nu Az\,dt = dW, \quad z(0) = 0 \tag{2.3}$$

with same realisation of the noise of (2.1), so that $v := u - z$ satisfies the random equation

$$\frac{\mathrm{d}v}{\mathrm{d}t} + \nu Av + B(v+z, v+z) = f, \quad v(0) = u_0. \tag{2.4}$$

It is known (e.g. [32, Proposition 4.1]) that there exists a unique solution $v \in C([0,T];\mathcal{H}) \cap L^2(0,T;\mathcal{V})$. Furthermore given the regularity of the covariance operator, the linear equation (2.3) has solution $z \in C([0,T];\mathcal{V})$ (Theorem 1.4.2). Then

$$u = v + z \in C([0,T];\mathcal{H}) \cap L^2(0,T;\mathcal{V})$$

as desired.

Finally we see that the solution is pathwise unique. Let u_1 and u_2 be two distinct solutions with same noise realisation, so that $U = u_1 - u_2$ satisfies

$$\frac{\mathrm{d}U}{\mathrm{d}t} + \nu AU + B(U,u_1) + B(u_2,U) = 0, \quad U(0) = u_1(0) - u_2(0). \tag{2.5}$$

Taking the $\mathcal{H}$ scalar product of (2.5) with U and using the properties of the bilinearity, Lemma 1.2.1, we have

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}|U|^2}{\mathrm{d}t} + \nu\|U\|^2 &\leq k_B\|U\|\|u_1\||U|\\ \text{(by Young inequality)} \qquad &\leq \frac{\nu\|U\|^2}{2} + \frac{k_B^2\|u_1\|^2}{2\nu}|U|^2.\end{aligned}$$

Hence

$$\frac{\mathrm{d}|U|^2}{\mathrm{d}t} + \nu\|U\|^2 \leq \frac{k_B^2\|u_1\|^2}{\nu}|U|^2,$$

and by the integral Gronwall's lemma Lemma A.2

$$|U(t)|^2 \leq |U(0)|^2 \exp\left(\frac{k_B}{\nu}\int_0^t \|u_1(s)\|^2\,ds\right).$$

Since $u \in L^2(0,T;\mathcal{V})$, the integral on the right hand side is well defined and we deduce that $u_1 = u_2$ whenever they have the same initial condition. □

Remark 2.1.3 (Generalised solution). In [32] and [31] the concept of solution used is that of *generalised solution* for (2.1). This is introduced to make possible working with a noise as irregular as possible, in particular for Q not necessarily trace class but with values at least in $D(A^{\frac{1}{4}+\varepsilon})$, $\varepsilon > 0$:

Definition 2.1.4 ([32, Definition 3.1], [31, Definition 3.1]). A stochastic process $u(t,\omega)$ is called *generalised solution* of (2.1) in $[0,T]$ if

$$u \in C([0,T];\mathcal{H}) \cap L^2(0,T;D(A^{1/4})) \quad \mathbb{P}\text{-a.s.},$$

it is progressively measurable in these topologies and (2.1) is satisfied $\mathbb{P}$-a.s. in the integral sense, i.e.

$$\begin{aligned}(u(t),\varphi) + \int_0^t (u(s),A\varphi)\,ds - \int_0^t (B(u,\varphi),u)\,ds &= (u_0,\varphi)\\ &+ \int_0^t (f(s),\varphi)\,ds + (W(t),\varphi)\end{aligned}$$

for all $t \in [0,T]$ and all $\varphi \in D(A)$.

Note that this definition is different from the one of weak solution due to the lower regularity of the solution. However if the process η is continuous with values in $\mathcal{V}$ the generalised solution is a weak solution itself. Let us finally remark that, as seen in the proof of Lemma 1.4.3, to ensure that η has a version continuous with values in $\mathcal{V}$ the covariance operator does not have necessarily to be trace class in principle. However it is not our goal to stretch the boundaries of regularity for the noise and we will simply require trace class type of noise within the most suitable space. For a clear concise presentation of the link between the noise and the type and regularity of the solutions we refer again the reader to [31].

Naturally, increasing the regularity of the initial condition and of the forcing, the solutions will be smoother and we talk about strong solution (e.g. [31, Theorem 3.3], [22, pg 122]). In fact in this case the equation does not hold as equality in H^{-1} but in L^2.

Definition 2.1.5. A stochastic process $u(t,\omega)$ is called *strong solution* of (2.1) in $[0,T]$ if

$$u \in C([0,T];\mathcal{V}) \cap L^2(0,T;D(A)) \quad \mathbb{P}\text{-a.s.},$$

and (2.1) is satisfied $\mathbb{P}$-a.s. as a pathwise identity in $\mathcal{H}$, i.e.

$$u(t) + \int_0^t Au(s)\, ds + \int_0^t B(u,u)\, ds = u_0 + \int_0^t f(s)\, ds + W(t)$$

for all $t \in [0,T]$.

Note that every strong solution is also a weak solution.
Then the following result holds:

Theorem 2.1.6 ([31, Theorem 3.3])**.** *If the initial condition u_0 is in $\mathcal{V}$, $f \in L^2(0,T;\mathcal{H})$ and Q is trace class in $D(A)$, then there exists a unique strong solution $u \in C([0,T];\mathcal{V}) \cap L^2(0,T;D(A))$ $\mathbb{P}$–almost surely.*

Proof. In this case z is continuous in time with values in $D(A)$ and it can be shown that $v \in C([0,T];\mathcal{V}) \cap L^2(0,T;D(A))$ (e.g. [32, Proposition 4.1]). □

In summary we see that the solution of the random Navier–Stokes equation, v, has the same regularity of the deterministic two–dimensional equation and consequently the regularity of the solution of the stochastic Navier–Stokes equation, u, is given by the combined regularity of the process z and v.

For the stochastic two–layer quasi–geostrophic (QG) model we will follow the same approach as outlined for the Navier–Stokes equation. Consider now the spaces $\mathcal{H}$ and $\mathcal{V}$ defined as in (1.40) i.e.

$$\mathcal{H} = (\mathbb{H}^{-1}, |||\cdot|||_{-1}) \quad \text{and} \quad \mathcal{V} = (\mathbb{L}^2, |||\cdot|||_0)$$

and equation (1.66) i.e.

$$
\begin{aligned}
&d\mathbf{q} + (B(\boldsymbol{\psi},\boldsymbol{\psi}) + \beta\partial_1\boldsymbol{\psi})dt = \nu\Delta^2\boldsymbol{\psi}\; dt + \begin{pmatrix} f \\ -r\Delta\psi_2 \end{pmatrix} dt + d\mathbf{W} \\
&\mathbf{q} = -\tilde{A}\boldsymbol{\psi} = (\Delta + M)\boldsymbol{\psi}
\end{aligned}
\tag{2.6}
$$

with initial condition $\mathbf{q}(0) = \mathbf{q}_0 \in \mathcal{H}$ and define the solution as follows:

Definition 2.1.7. A process $\mathbf{q}$ is a *weak solution* of (2.6) on $[0,T]$ if $\mathbb{P}$–almost surely $\mathbf{q} \in C([0,T];\mathcal{H}) \cap L^2(0,T;\mathcal{V})$, and it satisfies (2.6) in the integral sense:

$$
\begin{aligned}
(\mathbf{q}(t),\varphi) + \int_0^t (B(\boldsymbol{\psi},\boldsymbol{\psi}),\varphi) + \beta\,(\partial_1\boldsymbol{\psi},\varphi)\; ds = (\mathbf{q}_0,\varphi) + \nu\int_0^t (\Delta\boldsymbol{\psi},\Delta\varphi)\; ds \\
+ \int_0^t h_1(f,\varphi_1) - rh_2(\Delta\psi_2,\varphi_2)\; ds + h_1\,(W(t),\varphi_1)
\end{aligned}
$$

for all $\varphi = (\varphi_1,\varphi_2)^t \in \mathbf{H}^2$ and $t \in [0,T]$ where $(\cdot,\cdot)$ denotes the $\mathbf{L}^2$ scalar product.

As seen in section 1.3 the variables of interest $\mathbf{q}$ and $\boldsymbol{\psi}$, represent respectively the QG potential vorticity and the streamfunction. Recall that the velocity of the fluid is defined as the orthogonal gradient of the streamfunction. Therefore, as the velocity is continuous in time with values in L^2 and square integrable with values in H^1, we expect the streamfunctions and the potential vorticities to have the same regularity in time but with values in $\mathbf{H}^1$ and $\mathbf{H}^2$, and $\mathbb{H}^{-1}$ and $\mathbb{L}^2$ respectively. Then as in Navier–Stokes we required the noise on the level of velocities to be with values in H^1 in order to ensure existence of a unique solution, as the noise is now on the level of potential vorticites we expect to be sufficient to have W with values in L^2 to have a unique weak solution. We will show that this is indeed the case by verifying in section 2.2 the following result:

Theorem 2.1.8 (Weak solutions)**.** *We are given the initial condition* $\mathbf{q}_0 \in \mathcal{H}$*, the deterministic forcing* $f \in L^2(0,T;H^{-2})$ *and the covariance operator* Q *trace class in* L^2*. Then there exists a unique weak solution to* (2.6)

$$
\mathbf{q} \in C([0,T];\mathcal{H}) \cap L^2(0,T;\mathcal{V}) \quad \mathbb{P}\text{-}a.s.
$$

and as a consequence

$$
\boldsymbol{\psi} \in C([0,T];\mathbf{H}^1) \cap L^2(0,T;\mathbf{H}^2) \quad \mathbb{P}\text{-}a.s.
$$

As seen for the stochastic Navier–Stokes equation, increasing the regularity of the initial condition and of the forcing the solutions will be smoother and we talk about strong solution. In fact in this case the equation does not hold as equality in $\mathbb{H}^{-2}$ but in $\mathcal{H} = \mathbb{H}^{-1}$.

Definition 2.1.9. A stochastic process $\mathbf{q}$ is called *strong solution* of (2.6) in $[0, T]$ if

$$\mathbf{q} \in C([0,T];\mathcal{V}) \cap L^2(0,T;\mathbb{H}^1) \quad \mathbb{P}\text{-a.s.},$$

and (2.6) is satisfied $\mathbb{P}$-a.s. as a pathwise identity in $\mathcal{H}$ for all $t \in [0, T]$.

Theorem 2.1.10 (Strong solutions)**.** *We are given the initial condition* $\mathbf{q}_0 \in \mathcal{V}$*, the deterministic forcing* $f \in L^2(0,T;H^{-1})$ *and the covariance operator* Q *trace class in* H^1*. Then there exists a unique strong solution*

$$\mathbf{q} \in C([0,T];\mathcal{V}) \cap L^2(0,T;\mathbb{H}^1) \quad \mathbb{P}\text{-}a.s.$$

and as a consequence

$$\boldsymbol{\psi} =\in C([0,T];\mathbf{H}^2) \cap L^2(0,T;\mathbf{H}^3) \quad \mathbb{P}\text{-}a.s.$$

We will prove this result in section 2.3.

2.2 Weak solutions for stochastic 2LQG

Consider the linear auxiliary equation introduced in section 1.4.1, namely

$$d\eta + \alpha_1 A\eta \, dt = dW, \quad \eta(0) = \eta_0 \in L^2 \tag{2.7}$$

with W Wiener process with covariance matrix Q trace class operator in L^2. By Theorem 1.4.2 we know there exists a solution of (2.7) with continuous trajectories taking values in L^2, that is $\eta \in C(0,T;L^2)$. Then if we can show that the random differential equation (1.71) for $\tilde{\mathbf{q}} = \mathbf{q} - (\eta,\, 0)^t$, namely

$$\begin{aligned}
&\frac{\mathrm{d}\tilde{\mathbf{q}}}{\mathrm{d}t} + B(\tilde{\boldsymbol{\psi}} - \boldsymbol{\xi}, \tilde{\boldsymbol{\psi}} - \boldsymbol{\xi}) + \beta\partial_1(\tilde{\boldsymbol{\psi}} - \boldsymbol{\xi}) = \nu\Delta(\Delta\tilde{\boldsymbol{\psi}} + M\boldsymbol{\xi})\, dt + \begin{pmatrix} f - \nu\alpha\Delta\eta \\ -r\Delta(\tilde{\psi}_2 - \xi_2) \end{pmatrix} \\
&\tilde{\mathbf{q}} = (\Delta + M)\tilde{\boldsymbol{\psi}}
\end{aligned} \tag{2.8}$$

has a unique solution with trajectories continues with values in $\mathcal{H}$ and square integrable in $\mathcal{V}$, the process

$$\mathbf{q} := \tilde{\mathbf{q}} + \begin{pmatrix} \eta \\ 0 \end{pmatrix} \in C\left([0,T];\mathcal{H}\right) \cap L^2\left(0,T;\mathcal{V}\right) \tag{2.9}$$

is a solution of the original stochastic equation (2.6). Therefore we will dedicate most of the section to showing that the random equation (1.71) has a solution with the desired regularity. We will close with proving the pathwise uniqueness of the solutions corresponding to a given realisation of the noise.

The proof of the existence and uniqueness of the solution for the random equation has the same structure as the one for the deterministic equation, but we will have to take care of the new terms in η. The original deterministic model (1.49) has been treated in detail in [8]. There, for a k-layer model with Dirichlet boundary conditions the work shows existence, uniqueness and continuous dependence on initial condition of the solution. Furthermore the existence of a global attractor, bounded in $\mathbb{L}^2$, compact and connected in $\mathbb{H}^{-1}$ is established. We will adapt the original argument in [8] to the random equation (1.71), also drawing from some of the estimates already showed in [15]. Moreover we will consider deterministic forcing f independent of time as this will be the framework used in the rest of the chapters, although the following result can be extended without too much effort to time dependent forcings $f \in L^2(0,T;H^{-2})$.

Theorem 2.2.1. *Let $T > 0$, $\tilde{\mathbf{q}}_0 \in \mathcal{H}$ and $f \in H^{-2}(\mathcal{D})$. Then for $\mathbb{P}$-a.e. $\omega \in \Omega$, the system* (2.8) *has a unique solution $\tilde{\mathbf{q}}$ such that*

$$\tilde{\mathbf{q}} \in C\left([0,T];\mathcal{H}\right) \cap L^2\left(0,T;\mathcal{V}\right),$$

and consequently

$$\tilde{\boldsymbol{\psi}} \in C\left([0,T];\mathbf{H}^1\right) \cap L^2\left(0,T;\mathbf{H}^2\right).$$

Moreover the solution depends continuously on the initial condition $\tilde{\mathbf{q}}_0 \in \mathcal{H}$.

Proof. Note that to ease the notation of this proof we drop the tilde for the random equation (1.71).

Step 1: *A priori estimate in* $\mathcal{H} = (\mathbb{H}^{-1}, |||\cdot|||_{-1})$.
First of all, using (1.42), i.e. $(\mathbf{q},\boldsymbol{\psi}) = -|||\mathbf{q}|||_{-1}^2$, we have

$$\begin{aligned}\left(\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t},\boldsymbol{\psi}\right) &= \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{q},\boldsymbol{\psi}) - \left(\mathbf{q},\frac{\mathrm{d}\boldsymbol{\psi}}{\mathrm{d}t}\right)\\ &= -\frac{\mathrm{d}|||\mathbf{q}|||_{-1}^2}{\mathrm{d}t} - \left(-\tilde{A}\boldsymbol{\psi},\frac{\mathrm{d}\boldsymbol{\psi}}{\mathrm{d}t}\right)\end{aligned}$$

and since the operator $\tilde{A}$ is self-adjoint

$$\begin{aligned}\left(\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t},\boldsymbol{\psi}\right) &= -\frac{\mathrm{d}|||\mathbf{q}|||_{-1}^2}{\mathrm{d}t} - \left(\boldsymbol{\psi},\frac{\mathrm{d}(-\tilde{A}\boldsymbol{\psi})}{\mathrm{d}t}\right)\\ 2\left(\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t},\boldsymbol{\psi}\right) &= -\frac{\mathrm{d}|||\mathbf{q}|||_{-1}^2}{\mathrm{d}t}.\end{aligned}$$

Then, taking the $\mathbf{L}^2$ scalar product of (2.8) with $\boldsymbol{\psi}$ we have

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}|||\mathbf{q}|||_{-1}^2}{\mathrm{d}t} + \nu|\Delta\boldsymbol{\psi}|^2 + rh_2\|\psi_2\|^2 &= (B(\boldsymbol{\psi}-\boldsymbol{\xi},\boldsymbol{\psi}-\boldsymbol{\xi}),\boldsymbol{\psi})\\ &- \beta\,(\partial_1\boldsymbol{\psi},\boldsymbol{\psi}) - \beta(\partial_1\boldsymbol{\xi},\boldsymbol{\psi}) + \nu(\Delta M\boldsymbol{\xi},\boldsymbol{\psi}) - (f,h_1\psi_1)\\ &+ \nu\alpha(\Delta\eta,h_1\psi_1) - r(\xi_2,h_2\psi_2). \quad (2.10)\end{aligned}$$

From Equation (2.10), using the properties of B (Lemma 1.3.4) as well as the estimate (1.72) it can be shown (see Lemma 4.2 in [15]) that

$$\frac{\mathrm{d}}{\mathrm{d}t}|||\mathbf{q}(t)|||_{-1}^2 + \nu|\Delta\boldsymbol{\psi}(t)|^2 \leq d_0|\eta(t)|^2\|\boldsymbol{\psi}\|^2 + m(t) \qquad (2.11)$$

with

$$m(t,\omega) = d_1|\eta(t,\omega)|^4 + d_2|\eta(t,\omega)|^2 + d_3 \qquad (2.12)$$

and d_0, d_1, d_2, d_3 constants depending on the parameters of the system. The only difference in the computation carried out in [15], where f is assumed in H^{-1} rather than in H^{-2}, is in treating the term (f,ψ_1). There the authors have

$$-2h_1(f,\psi_1) \leq 2h_1\|f\|_{-1}\|\psi_1\| \leq \frac{9h_1}{\nu\lambda_1}\|f\|_{-1}^2 + \frac{\nu h_1}{9}|\Delta\psi_1|^2$$

and so $d_3 = 9h_1\|f\|_{-1}^2/\nu\lambda_1$, whereas we consider $f \in H^{-2}$ so

$$-2h_1(f,\psi_1) \leq 2h_1\|f\|_{-2}|\Delta\psi_1| \leq \frac{9h_1}{\nu}\|f\|_{-2}^2 + \frac{\nu h_1}{9}|\Delta\psi_1|^2,$$

and $d_3 = 9h_1\|f\|_{-2}^2/\nu$.

By the estimate (1.43) and Poincaré inequality (1.36) it is easy to see that

$$\|\boldsymbol{\psi}\|^2 \leq |||\mathbf{q}|||_{-1}^2 \leq \tfrac{a_0}{\lambda_1}|\Delta\boldsymbol{\psi}|^2$$

and as a consequence (2.11) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}|||\mathbf{q}(t)|||_{-1}^2 \leq \left(d_0|\eta(t)|^2 - \tfrac{\nu\lambda_1}{a_0}\right)|||\mathbf{q}(t)|||_{-1}^2 + m(t). \qquad (2.13)$$

Therefore, by Gronwall's inequality Lemma A.2, we get a pointwise bound

$$\begin{aligned}|||\mathbf{q}(t)|||_{-1}^2 \leq |||\mathbf{q}(0)|||_{-1}^2 \exp\Big(-\tfrac{\nu\lambda_1}{a_0}t + \int_0^t d_0|\eta(\tau)|^2\, d\tau\Big) \\ + \int_0^t \exp\Big(-\tfrac{\nu\lambda_1}{a_0}(t-s) + \int_s^t d_0|\eta(\tau)|^2 d\tau\Big) m(s)\, ds. \quad (2.14)\end{aligned}$$

Note that the integral

$$\int_0^t d_0|\eta(\tau)|^2\, d\tau$$

is well defined as η has continuous trajectories with values in $L^2(\mathcal{D})$.

Step 2: *Integral bound in* $\mathcal{V} = (\mathbb{L}^2, |||\cdot|||_0)$.
From (2.11) can also derive an integral bound in $\mathcal{V} = (\mathbb{L}^2, |||\cdot|||_0)$. Integrating (2.11) over $[s,t] \subset [0,T]$ we have

$$|||\mathbf{q}(t)|||_{-1}^2 - |||\mathbf{q}(s)|||_{-1}^2 + \nu \int_s^t |\Delta\boldsymbol{\psi}(\tau)|^2\, d\tau \leq d_0 \int_s^t \big(|\eta(\tau)|^2 \|\boldsymbol{\psi}\|^2 + m(\tau)\big)\, d\tau,$$

then, using the estimate (1.43) and dropping $|||\mathbf{q}(t)|||_{-1}^2$, we get

$$\nu \int_s^t |\Delta\boldsymbol{\psi}(\tau)|^2\, d\tau \leq |||\mathbf{q}(s)|||_{-1}^2 + d_0 \int_s^t \big(|\eta(\tau)|^2 |||\mathbf{q}(\tau)|||_{-1}^2 + m(\tau)\big)\, d\tau. \quad (2.15)$$

Note that

$$\begin{aligned}|||\mathbf{q}|||_0^2 &= |\Delta\boldsymbol{\psi}|^2 + p\|\psi_1 - \psi_2\|^2 \\ &\leq |\Delta\boldsymbol{\psi}|^2 + 2F_1\left(h_1\|\psi_1\|^2 + h_2\|\psi_2\|^2\right) \\ &\leq |\Delta\boldsymbol{\psi}|^2 + 2F_1|||\mathbf{q}|||_{-1}^2.\end{aligned}$$

It follows that

$$\begin{aligned}\int_0^T |||\mathbf{q}(\tau)|||_0^2\, d\tau &= \int_0^T |\Delta\boldsymbol{\psi}(\tau)|^2\, d\tau + 2F_1 \int_0^T |||\mathbf{q}(\tau)|||_{-1}^2\, d\tau \\ &\leq \nu^{-1}|||\mathbf{q}(0)|||_{-1}^2 + \int_0^T \nu^{-1} d_0 m(\tau)\, d\tau \\ &\quad + \int_0^T \left(\nu^{-1}d_0|\eta(\tau)|^2 + 2F_1\right) |||\mathbf{q}(\tau)|||_{-1}^2\, d\tau.\end{aligned} \quad (2.16)$$

Thanks to equation (2.14) we know that $|||\mathbf{q}(t)|||_{-1}^2$ is bounded, so $\mathbf{q} \in L^2(0,T;\mathcal{V})$.

Step 3: *Galerkin approximation.*
The results presented so far are only formal and to ensure existence of a solution we use the classic Galerkin method considering the truncated eigenfunction expansions of the variable of interest $\mathbf{q}$. Given an orthonormal basis $(e_k)_{k\in\mathbb{N}}$ of $\mathcal{H}$ made of eigenfunctions of $-\Delta$ define

$$\mathbf{q}^{(n)} := \left(\sum_{k=1}^{n}(q_1, e_k)e_k, \sum_{k=1}^{n}(q_2, e_k)e_k\right)^t .$$

Equivalently we write $\mathbf{q}^{(n)} = \Pi_n\mathbf{q}$, where Π_n is the projection operator in $\mathbf{H}^1$ onto the space spanned by the first n eigenfunctions of $-\Delta$. First note that the linear operator $(\Delta + M)$ commutes with the projection Π_n so that

$$\mathbf{q}^{(n)} = \Pi_n\mathbf{q} = \Pi_n(\Delta + M)\boldsymbol{\psi} = (\Delta + M)\boldsymbol{\psi}^{(n)} \tag{2.17}$$

where $\boldsymbol{\psi}^{(n)} := \Pi_n\boldsymbol{\psi}$. Consider the equation for the n-dimensional approximation $\{\mathbf{q}^{(n)}, \boldsymbol{\psi}^{(n)}\}$

$$\begin{aligned}\frac{\mathrm{d}\mathbf{q}^{(n)}}{\mathrm{d}t} + \Pi_n B(\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}, \boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}) + \Pi_n\partial_1(\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}) = \\ \nu\Pi_n\Delta(\Delta\boldsymbol{\psi}^{(n)} + M\boldsymbol{\xi}) + \begin{pmatrix}\Pi_n f - \nu\alpha\Pi_n\Delta\eta \\ -r\Pi_n\Delta(\psi_2^{(n)} - \xi_2)\end{pmatrix}\end{aligned} \tag{2.18}$$

with initial condition

$$\mathbf{q}^{(n)}(0) = \Pi_n\mathbf{q}_0 = \mathbf{q}_0^{(n)} \quad \text{with} \quad \mathbf{q}_0^{(n)} \to \mathbf{q}_0 \;\text{ for } n \to \infty.$$

We want to find a bound for $|||\mathbf{q}^{(n)}|||_{-1}$ uniform in n. Take the $\mathbf{L}^2$ inner product of (2.18) with $\boldsymbol{\psi}^{(n)}$ and note that

$$\begin{aligned}\left(\Pi_n B(\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}, \boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}), \boldsymbol{\psi}^{(n)}\right) &= \left(B(\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}, \boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}), \boldsymbol{\psi}^{(n)}\right), \\ \left(\Pi_n\Delta\eta, \boldsymbol{\psi}^{(n)}\right) = \left(\Delta\eta, \boldsymbol{\psi}^{(n)}\right) \quad &\text{and} \quad \left(\Pi_n\Delta\xi, \boldsymbol{\psi}^{(n)}\right) = \left(\Delta\xi, \boldsymbol{\psi}^{(n)}\right).\end{aligned}$$

Hence we get

$$\begin{aligned}\frac{\mathrm{d}|||\mathbf{q}^{(n)}|||_{-1}^2}{\mathrm{d}t} + \nu|\Delta\boldsymbol{\psi}^{(n)}|^2 + rh_2\|\psi_2^{(n)}\|^2 &= \left(B(\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}, \boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}), \boldsymbol{\psi}^{(n)}\right) \\ - \left(\partial_1\boldsymbol{\psi}^{(n)}, \boldsymbol{\psi}^{(n)}\right) - (\partial_1\boldsymbol{\xi}, \boldsymbol{\psi}^{(n)}) &+ \nu(\Delta M\boldsymbol{\xi}, \boldsymbol{\psi}^{(n)}) - h_1(f, \psi_1^{(n)}) \\ &+ \nu\alpha(\Delta\eta, \psi_1^{(n)}) - r(\xi_2, \psi_2^{(n)}).\end{aligned}$$

Then the formal bounds (2.14), (2.16) derived for (2.10) can be shown for $\mathbf{q}^{(n)}$ as in Step 1 and 2. Most important is the fact that such bounds are uniform in

n as it implies that $\mathbf{q}^{(n)}$ is bounded uniformly in n in $L^\infty(0,T;\mathcal{H}) \cap L^2(0,T;\mathcal{V})$. Therefore we can find a subsequence $(\mathbf{q}^{(k_n)})_{n\in\mathbb{N}}$ such that

$$\mathbf{q}^{(k_n)} \rightharpoonup \mathbf{q} \in L^\infty(0,T;\mathcal{H}) \cap L^2(0,T;\mathcal{V}) \tag{2.19}$$

and as a consequence

$$\boldsymbol{\psi}^{(k_n)} \rightharpoonup \boldsymbol{\psi} \in L^\infty(0,T;\mathbf{H}^1) \cap L^2(0,T;\mathbf{H}^2). \tag{2.20}$$

Now the limit $\mathbf{q}$ is the candidate solution and to close the argument one should ensure that it satisfies (2.8), by showing the appropriate convergence in n of each term in (2.18). Similarly to the classic approach for Navier–Stokes equation (see [62]), the result (2.20) for $\boldsymbol{\psi}^{(k_n)}$ is sufficient to show convergence of the linear terms, whereas for the time derivative and the bilinearity B extra arguments are necessary. The convergence of the bilinearity, that is

$$B(\boldsymbol{\psi}^{(n)}, \boldsymbol{\psi}^{(n)}) \overset{\star}{\rightharpoonup} B(\boldsymbol{\psi}, \boldsymbol{\psi}) \tag{2.21}$$

in $L^2(0,T;\mathbf{H}^{-2})$, follows from the same arguments for the deterministic model so we will omit them here (see [7] and [8]). We will instead show that $d\mathbf{q}^{(n)}/dt$ is bounded in $L^2(0,T;\mathbf{H}^{-2})$ uniformly on n so to extract a converging subsequence. Note that the limit will then be precisely $d\mathbf{q}/dt$ by definition of weak derivative.

To prove that $d\mathbf{q}^{(n)}/dt$ is bounded in $L^2(0,T;\mathbf{H}^{-2})$ uniformly in n we show that each of the terms on the right hand side of

$$\begin{aligned}\frac{\mathrm{d}\mathbf{q}^{(n)}}{\mathrm{d}t} &= -\Pi_n B(\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}, \boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}) - \partial_1(\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}) \\ &\quad + \nu\Delta(\Delta\boldsymbol{\psi}^{(n)} + M\boldsymbol{\xi}) + \begin{pmatrix} \Pi_n f - \nu\alpha\Pi_n\Delta\eta \\ -r\Pi_n\Delta(\psi_2^{(n)} - \xi_2) \end{pmatrix}\end{aligned} \tag{2.22}$$

is bounded in such a space uniformly in n. For the viscosity term $\nu\Delta^2\boldsymbol{\psi}$ we have

$$\begin{aligned}\int_0^T \|\Delta^2\boldsymbol{\psi}^{(n)}\|_{-2}^2\, ds &\le \int_0^T \|\Delta\|^2_{L(\mathbf{L}^2,\mathbf{H}^{-2})} |\Delta\boldsymbol{\psi}^{(n)}|^2\, ds \\ &\le \|\Delta\|^2_{L(\mathbf{L}^2,\mathbf{H}^{-2})} \int_0^T \left|\!\left|\!\left|\mathbf{q}^{(n)}\right|\!\right|\!\right|_0^2 ds\end{aligned}$$

and since Δ is a continuous linear operator from $\mathbf{L}^2$ into $\mathbf{H}^{-2}$ and $\mathbf{q}^{(n)} \in L^2(0,T;\mathcal{V})$ we have that $\Delta^2\boldsymbol{\psi}^{(n)}$ is bounded in $L^2(0,T;\mathbf{H}^{-2})$ uniformly in n. Next, the *a priori* bound (2.14) of $\left|\!\left|\!\left|\mathbf{q}^{(n)}\right|\!\right|\!\right|_{-1}^2$ implies that $\|\boldsymbol{\psi}^{(n)}\|^2$, and in particular $\partial_1 \boldsymbol{\psi}^{(n)}$, is bounded in $L^2(0,T;\mathbf{H}^{-2})$ uniformly in n. Since f is assumed to be in H^{-2} and $\langle \Pi_n f, v\rangle := \langle f, v^{(n)}\rangle$, it follows that $\Pi_n f$ is bounded in $L^2(0,T;\mathbf{H}^{-2})$ uniformly in n.

The main difference with the deterministic model is surely in dealing with the random coefficients, those involving the processes η and $\boldsymbol{\xi}$, but we see next that their boundedness is ensured by the regularity of the Ornstein-Uhlenbeck process η. By the definition of M, namely

$$M = \begin{pmatrix} -F_1 & F_1 \\ F_2 & -F_2 \end{pmatrix},$$

and the estimate (1.72) for $|\Delta\xi_i|$, $i = 1, 2$ we have

$$\begin{aligned} |\Delta M\boldsymbol{\xi}|^2 &= p(F_1 + F_2)|\Delta\xi_1 - \Delta\xi_2|^2 \\ &\le p(F_1 + F_2)|\eta|^2 \end{aligned}$$

and in particular then

$$\int_0^T \|\Delta M\boldsymbol{\xi}\|_{-2}^2\, ds \le \int_0^T |\Delta M\boldsymbol{\xi}|\, ds \le p(F_1 + F_2)\int_0^T |\eta|^2\, ds.$$

Since η has continuous trajectories with values in L^2, $M\Delta\boldsymbol{\xi}$ is bounded in $L^2(0,T;\mathbf{H}^{-2})$. With the same argument also $\Delta\xi_2$ and $\partial_1\boldsymbol{\xi}$ can be shown to be bounded as desired. Finally notice that

$$\begin{aligned} \int_0^T \|\Pi_n \Delta\eta\|_{-2}^2\, ds &= \int_0^T \|\Delta\eta^{(n)}\|_{-2}^2\, ds \\ \le \int_0^T \|\Delta\|_{L(L^2,H^{-2})}^2 |\eta^{(n)}|^2\, ds &\le \|\Delta\|_{L(L^2,H^{-2})}^2 \int_0^T |\eta|^2\, ds, \end{aligned}$$

so that the regularity of random process η gives the desired result once more.

It is left to show that the bilinearity B is bounded in $L^2(0,T;\mathbf{H}^{-2})$. Indeed

$$\|\Pi_n B(\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}, \boldsymbol{\psi}^{(n)} - \boldsymbol{\xi})\|_{-2} \le \|B(\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}, \boldsymbol{\psi}^{(n)} - \boldsymbol{\xi})\|_{-2}$$

and by the estimate (1.62) for $\|B(u,u)\|_{-2}$ we have

$$\|B(\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}, \boldsymbol{\psi}^{(n)} - \boldsymbol{\xi})\|_{-2} \le k_0 \|\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}\| |\Delta\boldsymbol{\psi}^{(n)} - \Delta\boldsymbol{\xi}|.$$

Therefore

$$\int_0^T \|B(\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}, \boldsymbol{\psi}^{(n)} - \boldsymbol{\xi})\|_{-2}^2\, dt \le k_0^2 \sup_{t\in[0,T]} \|\boldsymbol{\psi}^{(n)} - \boldsymbol{\xi}\|^2 \int_0^T |\Delta\boldsymbol{\psi}^{(n)} - \Delta\boldsymbol{\xi}|^2\, dt$$

and by (1.72)

$$\leq 4k_0^2 \sup_{t\in[0,T]} \left(\|\psi^{(n)}\|^2 + \|\eta\|_{-1}^2\right) \int_0^T |\Delta\psi^{(n)}|^2 + |\eta|^2 \, dt.$$

To conclude that $\Pi_n B(\psi^{(n)} - \xi, \psi^{(n)} - \xi)$ is uniformly bounded in $L^2(0,T;\mathbf{H}^{-2})$ it is enough to recall fact that η has continuous trajectories in L^2 as well as the results from the previous step, namely that $\psi^{(n)}$ is bounded in $L^\infty(0,T;\mathbf{H}^1)$ and in $L^2(0,T;\mathbf{H}^2)$ uniformly in n.

Finally all terms in (2.18) converge in $L^2(0,T;\mathbf{H}^{-2})$ and the limit function $\mathbf{q}$ in (2.19) is a solution of (2.8) with $\mathbf{q} \in L^\infty(0,T;\mathbb{H}^{-1}) \cap L^2(0,T;\mathbb{L}^2)$.

We actually have a better result, namely $\mathbf{q} \in C(0,T;\mathbb{H}^{-1}) \cap L^2(0,T;\mathbb{L}^2)$ thanks to the following result:

Theorem 2.2.2 ([62, Theorem 7.2])**.** *Let $(H, |\cdot|)$ and $(V, \|\cdot\|)$ be Hilbert spaces such that $V \subset\subset H$. Suppose that a function $u \in L^2(0,T;V)$ is such that $du/dt \in L^2(0,T;V^*)$. Then u is continuous from $[0,T]$ to H with*

$$\sup_{t\in[0,T]} |u(t)| \leq C \left(\|u\|_{L^2(0,T;V)} + \|du/dt\|_{L^2(0,T;V^*)}\right).$$

Step 4: *Uniqueness and continuous dependence on initial condition.* Let $\mathbf{q}$ and $\mathbf{p}$ be two solutions of (2.8) with respective streamfunctions ψ and ϕ and initial conditions $\mathbf{q}_0 \neq \mathbf{p}_0$. Then the differences $\mathbf{u} = \mathbf{q} - \mathbf{p}$ and $\mathbf{v} = \psi - \phi$ satisfy

$$\begin{aligned} &\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} + B(\psi - \xi, \mathbf{v}) + B(\mathbf{v}, \phi - \xi) + \beta\partial_1\mathbf{v} = \nu\Delta^2\mathbf{v} + \begin{pmatrix} 0 \\ -r\Delta v_2 \end{pmatrix} \\ &\mathbf{u} = (\Delta + M)\mathbf{v} \\ &\mathbf{u}(0) = \mathbf{q}_0 - \mathbf{p}_0 \end{aligned} \tag{2.23}$$

with periodic boundary condition, where we have used that $B(u,u) - B(v,v) = B(u-v,u) + B(v,u-v)$.

Now take the $\mathbf{L}^2$ scalar product of (2.23) with $\mathbf{v}$ to get

$$-\frac{1}{2}\frac{\mathrm{d}|||\mathbf{u}|||_{-1}^2}{\mathrm{d}t} + (B(\psi - \xi, \mathbf{v}), \mathbf{v}) + \beta(\partial_1\mathbf{v}, \mathbf{v}) = \nu|\Delta\mathbf{v}|^2 + rh_2\|v_2\|^2$$

where we have used again (1.42) and the property (1.56) of the bilinearity. The periodic boundary conditions give that $(\partial_1\mathbf{v}, \mathbf{v}) = 0$ so we are left with

$$\frac{1}{2}\frac{\mathrm{d}|||\mathbf{u}|||_{-1}^2}{\mathrm{d}t} + \nu|\Delta\mathbf{v}|^2 + rh_2\|v_2\|^2 = (B(\psi - \xi, \mathbf{v}), \mathbf{v}).$$

We can bound the right hand side using the estimate (1.58) for the bilinearity

$$
\begin{aligned}
|(B(\boldsymbol{\psi}-\boldsymbol{\xi},\mathbf{v}),\mathbf{v})| &\leq \tfrac{1}{2}k_B|\Delta\boldsymbol{\psi}-\Delta\boldsymbol{\xi}|^2\|\mathbf{v}\|^2+\tfrac{\nu}{2}|\Delta\mathbf{v}|^2\\
&\leq \left(k_B|\Delta\boldsymbol{\psi}|^2+k_Bh_2|\eta|^2\right)\|\mathbf{v}\|^2+\tfrac{\nu}{2}|\Delta\mathbf{v}|^2
\end{aligned}
$$

It follows that, also using the fact that $\|\mathbf{v}\|$ is bounded by $|||\mathbf{u}|||_{-1}$,

$$
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}|||\mathbf{u}|||_{-1}^2+\frac{\nu}{2}|\Delta\mathbf{v}|^2+rh_2\|v_2\|^2\leq\left(k_B|\Delta\boldsymbol{\psi}|^2+k_Bh_2|\eta|^2\right)|||\mathbf{u}|||_{-1}^2
$$

and by Gronwall's lemma Lemma A.1

$$
|||\mathbf{u}(t)|||_{-1}^2\leq|||\mathbf{u}(0)|||_{-1}^2\exp\left(2k_B\int_0^t|\Delta\boldsymbol{\psi}|^2\,ds+2k_Bh_2\int_0^t|\eta|^2\,ds\right). \tag{2.24}
$$

Clearly, by (2.15) and the regularity of η, the right hand side is well defined and this bound gives continuous dependence on initial conditions and the desired uniqueness when $\mathbf{q}_0=\mathbf{p}_0$. □

Consequently we can show that the solution $\mathbf{q}:=\tilde{\mathbf{q}}+\boldsymbol{\eta}$ as in (2.9) is unique and continuous with respect to the initial condition.

Proposition 2.2.3. *The solution of the stochastic equation* (2.6) *depends continuously on the initial condition. In particular the solution is pathwise unique.*

Proof. Let $\mathbf{p}$ be a solution of the stochastic equation (2.6) with initial condition $\mathbf{p}_0\neq\mathbf{q}_0$, streamfunction $\boldsymbol{\phi}=(-\tilde{A})^{-1}\mathbf{p}$ and same realization of the noise. Then the differences $\mathbf{u}=\mathbf{q}-\mathbf{p}$ and $\mathbf{v}=\boldsymbol{\psi}-\boldsymbol{\phi}$ satisfy the following (integral) equation

$$
\begin{aligned}
&\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}+B(\boldsymbol{\psi},\mathbf{v})+B(\mathbf{v},\boldsymbol{\phi})+\beta\partial_1\mathbf{v}=\nu\Delta^2\mathbf{v}+\begin{pmatrix}0\\-r\Delta v_2\end{pmatrix}\\
&\mathbf{u}=(\Delta+M)\,\mathbf{v}\\
&\mathbf{u}(0)=\mathbf{q}_0-\mathbf{p}_0
\end{aligned} \tag{2.25}
$$

To bound $|||\mathbf{u}|||_{-1}$ we take the $\mathbf{L}^2$ product with $\mathbf{v}$, to get

$$
-\frac{1}{2}\frac{\mathrm{d}|||\mathbf{u}|||_{-1}^2}{\mathrm{d}t}+(B(\boldsymbol{\psi},\mathbf{v}),\mathbf{v})+(B(\mathbf{v},\boldsymbol{\phi}),\mathbf{v})+\beta(\partial_1\mathbf{v},\mathbf{v})=
\nu(\Delta^2\mathbf{v},\mathbf{v})-rh_2(\Delta v_2,v_2).
$$

Using the property (1.56) of the bilinearity and the fact that $(\partial_1 \mathbf{v}, \mathbf{v})$

$$\frac{1}{2}\frac{\mathrm{d}|||\mathbf{u}|||_{-1}^2}{\mathrm{d}t} + \nu|\Delta \mathbf{v}|^2 + rh_2\|v_2\|^2 = (B(\boldsymbol{\psi}, \mathbf{v}), \mathbf{v})$$

Again by the bound on the bilinearity (1.58) and Young inequality we have

$$\frac{\mathrm{d}|||\mathbf{u}|||_{-1}^2}{\mathrm{d}t} + \nu|\Delta \mathbf{v}|^2 + 2rh_2\|v_2\|^2 \leq 2k_B|\Delta\boldsymbol{\psi}|^2|||\mathbf{u}|||_{-1}^2.$$

and by integral Gronwall lemma Lemma A.2 we have that for almost all times

$$|||\mathbf{u}(t)|||_{-1}^2 \leq |||\mathbf{u}(0)|||_{-1}^2 \exp\left(2k_B\int_0^t |\Delta\boldsymbol{\psi}|^2\, ds\right).$$

Since $\mathbf{q} \in L^2(0,T;\mathbb{L}^2)$ and that $\mathbf{q} = -\tilde{A}\boldsymbol{\psi} = (\Delta + M)\boldsymbol{\psi}$ we have that $\Delta\boldsymbol{\psi} \in L^2(0,T;\mathbf{L}^2)$ so that the exponential factor is well defined and we obtain uniqueness when we consider the same initial condition $\mathbf{q}_0 = \mathbf{p}_0$.

□

2.3 Strong solutions for stochastic 2LQG

In this section we will prove Theorem 2.3.1, namely the existence and uniqueness of strong solutions for the stochastic two–layer quasi–geostrophic model. Now the covariance operator Q is assumed to be trace class operator on H^1 so that $\eta \in C(0,T;H^1)$, the forcing is $f \in H^{-1}$ and the initial condition $\mathbf{q}_0$ is in $\mathcal{V}$. Therefore if the random equation (2.8) has a unique solution $\tilde{\mathbf{q}}$ with trajectories continues with values in $\mathcal{V}$ and square integrable in $\mathbb{H}^1$ the process

$$\mathbf{q} := \tilde{\mathbf{q}} + \begin{pmatrix}\eta\\0\end{pmatrix} \in C\left([0,T];\mathcal{V}\right) \cap L^2\left(0,T;\mathbb{H}^1\right) \tag{2.26}$$

is a solution of the original stochastic equation (2.6) and it can be shown to be unique. As every strong solution is in particular a weak solution, the uniqueness of strong solutions follows from that of weak solutions which we established in the previous section. We dedicate the rest of the section to proving the following theorem for the random equation (2.8)

Theorem 2.3.1. *If* $\tilde{\mathbf{q}}_0 \in \mathcal{V} = (\mathbb{L}^2, |||\cdot|||_0)$ *and* $f \in H^{-1}(\mathcal{D})$ *then there is a unique solution of* (2.8) *that satisfies*

$$\tilde{\mathbf{q}} \in C([0,T];\mathbb{L}^2) \cap L^2(0,T;\mathbf{H}^1)$$

and as a consequence

$$\tilde{\boldsymbol{\psi}} \in C([0,T];\mathbf{H}^2) \cap L^2(0,T;\mathbf{H}^3).$$

Furthermore the solutions depend continuously on the initial condition $\tilde{\mathbf{q}}_0$.

Proof. The proof has the same structure used for Theorem 2.2.1 for the existence and uniqueness of the weak solutions and for the deterministic case in [8]. Here we will only show Step 1 and Step 2, namely the appropriate bounds for the strong solutions. We omit the Galerkin approximation argument which will bring little novelty at this point. In the presentation we drop the tilde to ease the notation.

Step 1: *A priori bound in* $(\mathbb{L}^2, \lVert\!|\cdot|\!\rVert_0)$. We want to obtain an estimate of the form

$$\frac{\mathrm{d}}{\mathrm{d}t} \lVert\!|\mathbf{q}|\!\rVert_0^2 \leq g_1(t) \lVert\!|\mathbf{q}|\!\rVert_0^2 + g_2(t) \tag{2.27}$$

in order to recover a bound like

$$\lVert\!|\mathbf{q}(t)|\!\rVert_0^2 \leq e^{\int_0^t g_1(s)\,ds} \lVert\!|\mathbf{q}_0|\!\rVert_0^2 + \int_0^t e^{\int_0^s g_1(\tau) d\tau} g_2(s)\,ds, \tag{2.28}$$

thanks to Gronwall lemma Lemma A.1. At the end of this step we will find such estimate where g_1 and g_2 are random functions dependent on η and all the parameters of the system.

Let us start by taking the $\mathbf{L}^2$ inner product of (2.8) with $\Delta\boldsymbol{\psi}$

$$\left(\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t}, \Delta\boldsymbol{\psi}\right) + (B(\boldsymbol{\psi}-\boldsymbol{\xi}, \boldsymbol{\psi}-\boldsymbol{\xi}), \Delta\boldsymbol{\psi}) + \beta(\partial_1(\boldsymbol{\psi}-\boldsymbol{\xi}), \Delta\boldsymbol{\psi}) = \nu(\Delta^2\boldsymbol{\psi}, \Delta\boldsymbol{\psi})$$
$$+ \nu(M\Delta\boldsymbol{\xi}, \Delta\boldsymbol{\psi}) + h_1(f - \nu\alpha\Delta\eta, \Delta\psi_1) - h_2 r(\Delta\psi_2 - \Delta\xi_2, \Delta\psi_2). \tag{2.29}$$

First of all by (1.44) we have that

$$\left(\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t}, \Delta\boldsymbol{\psi}\right) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} \lVert\!|\mathbf{q}|\!\rVert_0^2. \tag{2.30}$$

Next, observe that by Green's theorem and the periodic boundary conditions, equation (2.29) can be written as

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} \lVert\!|\mathbf{q}|\!\rVert_0^2 + \nu\|\Delta\boldsymbol{\psi}\|^2 + (B(\boldsymbol{\psi}-\boldsymbol{\xi}, \boldsymbol{\psi}-\boldsymbol{\xi}), \Delta\boldsymbol{\psi}) + \beta(\partial_1(\boldsymbol{\psi}-\boldsymbol{\xi}), \Delta\boldsymbol{\psi}) =$$
$$\nu(M\Delta\boldsymbol{\xi}, \Delta\boldsymbol{\psi}) + h_1(f - \nu\alpha\Delta\eta, \Delta\psi_1) - h_2 r|\Delta\psi_2|^2 + h_2 r(\Delta\xi_2, \Delta\psi_2). \tag{2.31}$$

Using Cauchy-Schwarz and Young inequalities we have

$$\beta|(\partial_1(\boldsymbol{\psi}-\boldsymbol{\xi}), \Delta\boldsymbol{\psi})| \leq \tfrac{\beta^2}{2}\|\boldsymbol{\psi}\|^2 + \tfrac{1}{2}|\Delta\boldsymbol{\psi}|^2 + \tfrac{\beta^2}{2}\|\boldsymbol{\psi}\|^2 + \tfrac{1}{2}|\Delta\boldsymbol{\xi}|^2$$

and by the estimate (1.72) on $\boldsymbol{\xi}$

$$\beta|(\partial_1(\boldsymbol{\psi}-\boldsymbol{\xi}),\Delta\boldsymbol{\psi})| \leq \beta^2\|\boldsymbol{\psi}\|^2 + \tfrac{1}{2}|\Delta\boldsymbol{\psi}|^2 + h_2|\eta|^2. \tag{2.32}$$

Secondly, using the definition of $\boldsymbol{\xi}$ (1.70) and the definition of the matrix M (1.27) we derive

$$\begin{aligned}\nu(\Delta M\boldsymbol{\xi},\Delta\boldsymbol{\psi}) &= \nu p\,(\Delta\xi_1-\Delta\xi_2,\Delta\psi_2-\Delta\psi_1)\\ &\leq \nu p|\Delta\xi_1-\Delta\xi_2||\Delta\psi_1-\Delta\psi_2|\end{aligned}$$

and, using Young's inequality and the estimates (1.72) for the process $\boldsymbol{\xi}$, we have

$$\nu(\Delta M\boldsymbol{\xi},\Delta\boldsymbol{\psi}) \leq \tfrac{\nu^2 p}{2}|\eta|^2 + \tfrac{F_1}{2}|\Delta\boldsymbol{\psi}|^2.$$

Next,

$$|h_1(f-\nu\alpha\Delta\eta,\Delta\psi_1)| \leq \tfrac{2h_1}{\nu}\|f\|_{-1}^2 + 2\nu h_1\alpha^2\|\eta\|^2 + \tfrac{h_1\nu}{4}\|\Delta\psi_1\|^2.$$

Finally, (1.72) gives

$$h_2 r\,(\Delta\xi_2,\Delta\psi_2) \leq \frac{rh_2}{2}\left(|\Delta\xi_2|^2+|\Delta\psi_2|^2\right) \leq \frac{rh_2}{2}\left(|\eta|^2+|\Delta\psi_2|^2\right).$$

Therefore pulling these estimates together in (2.31) we get

$$\begin{aligned}\tfrac{1}{2}\tfrac{d}{dt}|||\mathbf{q}|||_0^2 + \tfrac{3\nu}{4}\|\Delta\boldsymbol{\psi}\|^2 + \tfrac{rh_2}{2}|\Delta\psi_2|^2 &\leq \beta^2\|\boldsymbol{\psi}\|^2 + \tfrac{1+F_1}{2}|\Delta\boldsymbol{\psi}|^2\\ &+|\eta|^2\left(h_2+\tfrac{rh_2}{2}+\tfrac{\nu^2 p}{2}\right)+\tfrac{2h_1}{\nu}\|f\|_{-1}^2\\ &+2\nu h_1\alpha^2\|\eta\|^2 + (B(\boldsymbol{\psi}-\boldsymbol{\xi},\boldsymbol{\psi}-\boldsymbol{\xi}),\Delta\boldsymbol{\psi}).\end{aligned} \tag{2.33}$$

Finally we treat the bilinearity. By definition of B (1.47) we have

$$\begin{aligned}&(B(\boldsymbol{\psi}-\boldsymbol{\xi},\boldsymbol{\psi}-\boldsymbol{\xi}),\Delta\boldsymbol{\psi}) =\\ &\quad h_1(J(\psi_1-\xi_1,\Delta\psi_1-\Delta\xi_1),\Delta\psi_1) + p(J(\psi_1-\xi_1,\psi_2-\xi_2),\Delta\psi_1)\\ &\quad +h_2(J(\psi_2-\xi_2,\Delta\psi_2-\Delta\xi_2),\Delta\psi_2) + p(J(\psi_2-\xi_2,\psi_1-\xi_1),\Delta\psi_2)\end{aligned}$$

which, thanks to the properties of J in Lemma 1.3.3, reduces to

$$\begin{aligned}&= -h_1(J(\psi_1-\xi_1,\Delta\xi_1+F_1\xi_2),\Delta\psi_1) + p(J(\psi_1-\xi_1,\psi_2),\Delta\psi_1)\\ &\quad -h_2(J(\psi_2-\xi_2,\Delta\xi_2+F_2\xi_1),\Delta\psi_2) + p(J(\psi_2-\xi_2,\psi_1),\Delta\psi_2)\end{aligned}$$

Now recall the definition of ξ_1,ξ_2 in (1.70) namely

$$\begin{aligned}\Delta\xi_1 - F_1(\xi_1-\xi_2) &= -\eta\\ \Delta\xi_2 - F_2(\xi_2-\xi_1) &= 0.\end{aligned}$$

With simple manipulations and using the fact that $J(v,v)=0$, we see that

$$\begin{aligned}(B(\boldsymbol{\psi}-\boldsymbol{\xi},\boldsymbol{\psi}-\boldsymbol{\xi}),\Delta\boldsymbol{\psi}) &= h_1(J(\psi_1-\xi_1,\eta),\Delta\psi_1)-p(J(\psi_1,\xi_1),\Delta\psi_1)\\ &+p(J(\psi_1-\xi_1,\psi_2),\Delta\psi_1)-p(J(\psi_2,\xi_2),\Delta\psi_2)+p(J(\psi_2-\xi_2,\psi_1),\Delta\psi_2)\end{aligned}$$

so finally

$$\begin{aligned}&= h_1(J(\psi_1-\xi_1,\eta),\Delta\psi_1)-p(J(\psi_1,\xi_1),\Delta\psi_1)-p(J(\psi_2,\xi_2),\Delta\psi_2)\\ &+p(J(\psi_1,\psi_2),\Delta\psi_1-\Delta\psi_2)-p(J(\xi_1,\psi_2),\Delta\psi_1)-p(J(\xi_2,\psi_1),\Delta\psi_2).\end{aligned} \tag{2.34}$$

Then we are left to bound each of these terms. By the estimates on the Jacobian (1.53), (1.54), and Young inequality we have

$$\begin{aligned}h_1|(J(\psi_1-\xi_1,\eta),\Delta\psi_1)| &\leq h_1c_1\|\Delta\psi_1\||\Delta\psi_1|\|\eta\|+h_1c_1|\Delta\xi_1|\|\Delta\psi_1\|\|\eta\|\\ &\leq \tfrac{h_1\nu}{4}\|\Delta\psi_1\|^2+\tfrac{4h_1c_1^2}{\nu}\|\eta\|^2|\Delta\psi_1|^2+\tfrac{4h_1c_1^2}{\nu}|\Delta\xi_1|^2\|\eta\|^2\end{aligned}$$

and using once more the relation (1.72) between $\boldsymbol{\xi}$ and η and Poincaré inequality

$$\leq \tfrac{h_1\nu}{4}\|\Delta\psi_1\|^2+\tfrac{4h_1c_1^2}{\nu}\|\eta\|^2|\Delta\psi_1|^2+\tfrac{4h_1c_1^2}{\nu\lambda_1}\|\eta\|^4. \tag{2.35}$$

Secondly, by means of estimate (1.52) for the Jacobian, we derive for $i=1,2$

$$p|(J(\psi_i,\xi_i),\Delta\psi_i)| \leq c_0p|\Delta\xi_i|\|\psi_i\||\Delta\psi_i| \leq \tfrac{c_0h_iF_i}{2}\left(|\Delta\psi_i|^2+|\Delta\xi_i|^2\|\psi_i\|^2\right).$$

By Poincaré inequality and estimating the terms in ξ by η as above, we get

$$p|(J(\psi_1,\xi_1),\Delta\psi_1)+(J(\psi_2,\xi_2),\Delta\psi_2)| \leq \left(\tfrac{c_0F_1}{2}+\tfrac{c_0F_1}{2\lambda_1}|\eta|^2\right)|\Delta\boldsymbol{\psi}|^2 \tag{2.36}$$

Moving on to the next term in (2.34), using the estimate (1.54) and Young inequality we have

$$\begin{aligned}p|(J(\psi_1,\psi_2),\Delta\psi_1-\Delta\psi_2)| &\leq c_1p\|\psi_1\||\Delta\psi_2|\|\Delta\psi_1-\Delta\psi_2\|\\ &\leq \tfrac{c_1^2pF_1}{\nu}\|\psi_1\|^2|\Delta\psi_2|^2+\tfrac{\nu}{4}\|\Delta\boldsymbol{\psi}\|^2.\end{aligned} \tag{2.37}$$

Finally, the estimates (1.53) and (1.72) give

$$\begin{aligned}p(J(\xi_1,\psi_2),\Delta\psi_1)-p(J(\xi_2,\psi_1),\Delta\psi_2) &\leq c_1p\left(|\Delta\xi_1|+|\Delta\xi_2|\right)|\Delta\psi_1||\Delta\psi_2|\\ &\leq 2c_1|\eta|\left(\tfrac{F_1h_1}{2}|\Delta\psi_1|^2+\tfrac{F_2h_2}{2}|\Delta\psi_2|\right)\\ &\leq c_1F_1|\eta||\Delta\boldsymbol{\psi}|^2.\end{aligned} \tag{2.38}$$

Therefore using the estimates (2.35), (2.36), (2.37) and (2.38) just obtained in (2.33) we get

$$\frac{1}{2}\frac{d}{dt}|||\mathbf{q}|||_0^2 + \frac{\nu}{4}\|\Delta\boldsymbol{\psi}\|^2 + \frac{rh_2}{2}|\Delta\psi_2|^2 \le \beta^2\|\boldsymbol{\psi}\|^2 + |\eta|^2\left(h_2 + \frac{rh_2}{2} + \frac{\nu^2 p}{2}\right)$$
$$+\left(\frac{1+F_1+c_0F_1}{2} + c_1F_1|\eta| + \frac{c_0F_1}{2\lambda_1}|\eta|^2 + \frac{4h_1c_1^2}{\nu}\|\eta\|^2\right)|\Delta\boldsymbol{\psi}|^2 + \frac{2h_1}{\nu}\|f\|_{-1}^2 + 2\nu h_1\alpha^2\|\eta\|^2$$
$$+\frac{4h_1c_1^2}{\nu\lambda_1}\|\eta\|^4 + \frac{c_1^2pF_1}{\nu}\|\psi_1\|^2|\Delta\psi_2|^2.$$

Using Poincaré inequality on $|\eta|$ and the definition of the $\mathbb{H}^{-1}$ norm for $\mathbf{q}$ we have the following more compact estimate

$$\frac{d}{dt}|||\mathbf{q}|||_0^2 + \frac{\nu}{2}\|\Delta\boldsymbol{\psi}\|^2 + rh_2|\Delta\psi_2|^2 \le c_4\|\eta\|^2 + c_5\|\eta\|^4 + 2\beta^2|||\mathbf{q}|||_{-1}^2$$
$$+\frac{4h_1}{\nu}\|f\|_{-1}^2 + \left(c_2 + 2c_1F_1|\eta| + c_6|\eta|^2 + c_5\lambda_1\|\eta\|^2 + c_3|||\mathbf{q}|||_{-1}^2\right)|||\mathbf{q}|||_0^2, \quad (2.39)$$

where

$$c_2 = 1 + F_1 + c_0F_1, \quad c_3 = \frac{c_1^2F_2F_1}{\nu h_1}$$
$$c_4 = 2\frac{h_2}{\lambda_1} + \frac{rh_2}{\lambda_1} + \frac{\nu^2 p}{\lambda_1} + 4\nu h_1\alpha^2, \quad c_5 = \frac{8h_1c_1^2}{\nu\lambda_1}, \quad c_6 = \frac{c_0F_1}{\lambda_1}.$$

Then, by Gronwall inequality we have that

$$|||\mathbf{q}(t)|||_0^2 \le |||\mathbf{q}(0)|||_0^2 \exp\left(\int_0^t g_1(s)\,ds\right) + \int_0^t g_2(s)\exp\left(\int_s^t g_1(\tau)\,d\tau\right)ds, \quad (2.40)$$

with

$$g_1(t) := c_2 + 2c_1F_1|\eta(t)| + c_6|\eta|^2 + c_5\lambda_1\|\eta(t)\|^2 + c_3|||\mathbf{q}(t)|||_{-1}^2, \quad (2.41)$$
$$g_2(t) := c_4\|\eta(t)\|^2 + c_5\|\eta(t)\|^4 + 2\beta^2|||\mathbf{q}(t)|||_{-1}^2 + \frac{4h_1}{\nu}\|f\|_{-1}^2. \quad (2.42)$$

Note that (2.40) is well defined for any fixed t thanks to the pointwise bound on $|||\mathbf{q}(s)|||_{-1}^2$ given by the *a priori* bound (2.14) in the $\mathcal{H}$ norm and since the process η has continuous trajectories in $H^1(\mathcal{D})$.

Step 2: *Integral bound in* $\mathbf{H}^1$. First of all we have that

$$\|\mathbf{q}\|^2 = \|(\Delta + M)\boldsymbol{\psi}\|^2 \le 2\|\Delta\boldsymbol{\psi}\|^2 + 2\|M\boldsymbol{\psi}\|^2$$
$$= 2\|\Delta\boldsymbol{\psi}\|^2 + 2p(F_1 + F_2)\|\psi_1 - \psi_2\|^2$$

so that by definition of $|||\cdot|||_0$ we have

$$\int_0^t \|\mathbf{q}(s)\|^2\,ds \le 2\int_0^t \|\Delta\boldsymbol{\psi}(s)\|^2\,ds + 2p(F_1 + F_2)\int_0^t |||\mathbf{q}(s)|||_0^2\,ds.$$

We can bound the second term on the right hand side by means of the *a priori* bound (2.40) derived in the previous step. For the first term we integrate over time the energy estimate (2.39) to get

$$\nu \int_0^t \|\Delta \boldsymbol{\psi}(s)\|^2 \, ds \leq \|\!|\mathbf{q}_0|\!\|_0^2 + \int_0^t g_1(s) \|\!|\mathbf{q}(s)|\!\|_0^2 \, ds + c_4 \int_0^t \|\eta(s)\|^2 \, ds \\ + c_5 \int_0^t \|\eta(s)\|^4 \, ds + 2\beta^2 \int_0^t \|\!|\mathbf{q}(s)|\!\|_{-1}^2 \, ds + \tfrac{4h_1}{\nu} \|f\|_{-1}^2 t.$$

For the terms containing $\|\!|\mathbf{q}(s)|\!\|_{-1}^2$ and $\|\!|\mathbf{q}(s)|\!\|_0^2$ we use the *a priori* bounds (2.14) and (2.40), respectively. For the terms in η, the continuity in H^1 is sufficient to ensure they are well defined.

□

2.4 Regularity with respect to the forcing for stochastic 2LQG

Consider the system (1.66) with covariance operator Q trace class in L^2 and the forcing $f \in H^{-2}$ with intensity modulated by a multiplicative constant $a > 0$,

$$\begin{aligned} & d\mathbf{q} + (B(\boldsymbol{\psi}, \boldsymbol{\psi}) + \beta \partial_1 \boldsymbol{\psi}) \, dt = \nu \Delta^2 \boldsymbol{\psi} \, dt + \begin{pmatrix} af \\ -r\Delta\psi_2 \end{pmatrix} dt + d\mathbf{W} \\ & \mathbf{q} = (\Delta + M)\boldsymbol{\psi} \\ & \mathbf{q}(0) = \mathbf{q}_0 \in \mathcal{H}. \end{aligned} \tag{2.43}$$

In this section we will study the dependence of the solution $\mathbf{q}(t, \omega; \mathbf{q}_0, a)$ with respect to the parameter a showing that it is locally Lipschitz continuous but also differentiable. These results will prove crucial in section 5.4 to study the dependence on a of the associated Markov semigroup $\mathcal{P}_t^a$ and its invariant measure (see chapter 3 for a precise definition of this concepts).

Let us first introduce a series of results which will be useful throughout the rest of the document.

2.4.1 Energy estimate for stochastic 2LQG

The following lemma is a classic result of stochastic analysis which will be used extensively when proving energy estimates for the stochastic 2LQG model or stochastic Navier–Stokes equations.

Lemma 2.4.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration $\{\mathcal{F}_t,\ t \in [0,T]\}$ and let $\phi(t)$ be a process adapted to $\{\mathcal{F}_t\}_{t\le T}$ with*

$$\mathbb{E}\int_0^T \phi^2(s)\,ds < \infty.$$

Further, $\{W(t)\}$ is a $\{\mathcal{F}_t\}_{t\le T}$ Brownian motion. Let X_t, $t \in [0,T]$ be a continuous martingale defined as

$$X_t := \int_0^t \phi(s)\,dW(s),$$

then for all $\gamma > 0$ and $R > 0$

$$\mathbb{P}\left(\sup_{t\in[0,T]} X_t - \frac{\gamma}{2}\langle X\rangle_t \ge R\right) \le e^{-\gamma R}. \tag{2.44}$$

Furthermore we have

$$\mathbb{E}\exp\left(K\sup_{t\in[0,T]}\left(X_t - \frac{\gamma}{2}\langle X\rangle_t\right)\right) < \infty \tag{2.45}$$

for all $0 < K < \gamma$.

Proof. First of all note that

$$\mathbb{P}\left(\sup_{t\in[0,T]} X_t - \frac{\gamma}{2}\langle X\rangle_t \ge R\right) = \mathbb{P}\left(\sup_{t\in[0,T]} \exp\left(\gamma X_t - \frac{\gamma^2}{2}\langle X\rangle_t\right) \ge e^{\gamma R}\right).$$

The exponential process

$$Z_t^\gamma := \exp\left(\gamma X_t - \frac{\gamma^2}{2}\langle X\rangle_t\right)$$

is a right continuous supermartingale (as it is a positive local martingale). Then the submartingale inequality, see e.g. [47, Theorem 1.3.8(ii)], gives

$$\mathbb{P}\left(\sup_{t\in[0,T]} Z_t^\gamma \ge e^{\gamma R}\right) \le \frac{\mathbb{E}Z_0^\gamma + \mathbb{E}\left[(Z_T^\gamma)^-\right]}{e^{\gamma R}}$$

where $x^- = \max\{-x, 0\}$. Since Z^γ is a positive process and $Z_0^\gamma = 1$ we get

$$\mathbb{P}\left(\sup_{t\in[0,T]} X_t - \frac{\gamma}{2}\langle X\rangle_t \ge R\right) \le e^{-\gamma R}.$$

In order to show (2.45) it is enough to note that for the random variable $Y = \sup_{t\in[0,T]} \left(X_t - \frac{\gamma}{2}\langle X\rangle_t\right)$ and any $K > 0$

$$\begin{aligned}\mathbb{E}\exp(KY) &= \int_{-\infty}^{\infty} K\exp(Ky)\mathbb{P}\left(Y \geq y\right)\, dy \\ &\leq 1 + \int_0^{\infty} K\exp(Ky)\mathbb{P}\left(Y \geq y\right)\, dy.\end{aligned}$$

Therefore, by (2.44)

$$\mathbb{E}\exp\left(K\sup_{t\in[0,T]}\left(X_t - \frac{\gamma}{2}\langle X\rangle_t\right)\right) \leq 1 + \int_0^{\infty} K\exp(-y(\gamma - K))\, dy \tag{2.46}$$

which stays finite for all $K < \gamma$. □

Furthermore in the context of Lemma 2.4.1, observe that the random variable

$$\Xi_\gamma := \sup_{t\in[0,T]} X_t - \gamma\langle X\rangle_t$$

is almost surely finite.

For the quasi–geostrophic model (2.43) we define

$$X_t^a := \int_0^t (\psi_1(s,a), dW(s)) \tag{2.47}$$

so that its quadratic variation is

$$\langle X^a\rangle_t := \int_0^t \|(\psi_1(s,a),\cdot)\|^2_{L_2^0}\, ds. \tag{2.48}$$

By the definition of L_2^0 norm we have

$$\|(\psi_1,\cdot)\|^2_{L_2^0} = \sum_{k\in\mathbb{N}} |(\psi_1, Q^{1/2}e_k)|^2$$

so by Cauchy-Schwartz inequality

$$\leq \sum_{k\in\mathbb{N}} |\psi_1|^2|Q^{1/2}e_k|^2 = |\psi_1|^2\sum_{k\in\mathbb{N}} |Q^{1/2}e_k|^2$$

and by the definition of trace we conclude that

$$\|(\psi_1,\cdot)\|^2_{L_2^0} \leq \operatorname{Tr} Q\, |\psi_1|^2. \tag{2.49}$$

Consequently we can bound the quadratic variation (2.48) as follows

$$\langle X^a \rangle_t \leq \operatorname{Tr} Q \int_0^t |\psi_1(s,a)|^2 \, ds. \tag{2.50}$$

Therefore Lemma 2.4.1 holds for

$$\Xi_\gamma^a := \sup_{t\geq 0} X_t^a - \gamma \langle X^a \rangle \tag{2.51}$$

giving

$$\mathbb{E} \exp\big(K \Xi_\gamma^a\big) < \infty \quad \text{for all } \ K < 2\gamma.$$

In particular, from (2.46) we can compute the explicit bound

$$\mathbb{E} \exp\big(K \Xi_\gamma^a\big) < \frac{K}{2\gamma - K} \quad \text{for all } \ a > 0. \tag{2.52}$$

The next theorem provides a crucial estimate for the model of interest by means of Lemma 2.4.1.

Theorem 2.4.2. *Consider the stochastic 2LQG model* (2.43) *with initial condition* $\mathbf{q}(0) = \mathbf{q}_0 \in \mathcal{H}$ *and* $f \in H^{-2}(\mathcal{D})$*. Let* $\mathbf{q}$ *be its unique solution and set*

$$T_Q := \operatorname{Tr}\left[(Q^{1/2})^* \tilde{A}^{-1} Q^{1/2}\right]. \tag{2.53}$$

Then, given Ξ_γ^a *as in* (2.51)*, for all* $0 < \delta < \nu$ *there exists* $\gamma \geq 0$ *such that*

$$|||\mathbf{q}(t)|||_{-1}^2 + \delta \int_0^t |\Delta \boldsymbol{\psi}|^2 \, ds \leq |||\mathbf{q}_0|||_{-1}^2 + t\left(\tfrac{h_1}{\nu}\|af\|_{-2}^2 + T_Q\right) + 2h_1 \Xi_\gamma^a. \tag{2.54}$$

Proof. Let us apply Itô formula (see e.g. [22, Theorem 4.32]) to compute $d|||\mathbf{q}|||_{-1}^2$. Since $\mathbf{q} = -\tilde{A}\boldsymbol{\psi}$ we have

$$-d|||\mathbf{q}|||_{-1}^2 = d(\mathbf{q}, \boldsymbol{\psi}) = -d\left(\mathbf{q}, \tilde{A}^{-1}\mathbf{q}\right).$$

By Itô formula for the transformation $F(t,\mathbf{q}) = \left(\mathbf{q}, \tilde{A}^{-1}\mathbf{q}\right)$ and since $\tilde{A}$ is self-adjoint we have

$$d\left(\mathbf{q}, \tilde{A}^{-1}\mathbf{q}\right) = 2(d\mathbf{q}, \tilde{A}^{-1}\mathbf{q}) + \operatorname{Tr}\left[(Q^{1/2})^* \tilde{A}^{-1} Q^{1/2}\right] dt$$

(see e.g. [51, pg. 129] for a detailed computation). Therefore, given (2.53), we have that $-d|||\mathbf{q}|||_{-1}^2 = 2(d\mathbf{q}, \boldsymbol{\psi}) - T_Q \, dt$, which gives

$$-d|||\mathbf{q}|||_{-1}^2 = 2\left(\nu(\Delta^2 \boldsymbol{\psi}, \boldsymbol{\psi}) + h_1(af, \psi_1) - r(\Delta\psi_2, \psi_2) - \tfrac{1}{2}T_Q\right) dt + 2(\boldsymbol{\psi}, d\mathbf{W})$$

where we have used that $(B(\boldsymbol{\psi},\boldsymbol{\psi}),\boldsymbol{\psi}) = 0$ and $(\partial_1\boldsymbol{\psi},\boldsymbol{\psi}) = 0$. By Green's theorem and the definition of $\mathbf{W}$

$$d|||\mathbf{q}|||_{-1}^2 = -2\left(\nu|\Delta\boldsymbol{\psi}|^2 + h_1(af,\psi_1) + rh_2\|\psi_2\|^2 - \tfrac{1}{2}T_Q\right)\,dt - 2h_1\left(\psi_1, dW\right). \tag{2.55}$$

Next, using Cauchy-Schwartz, Young's and Poincaré inequalities we can bound the deterministic forcing term as usual

$$-2(af,h_1\psi_1) \leq 2|(af,h_1\psi_1)| \leq \tfrac{h_1}{\nu}\|af\|_{-2}^2 + \nu h_1|\Delta\psi_1|^2$$

and using this estimate in (2.55) we have

$$|||\mathbf{q}(t)|||_{-1}^2 - |||\mathbf{q}_0|||_{-1}^2 + \nu\int_0^t|\Delta\boldsymbol{\psi}|^2\,ds + 2rh_2\int_0^t\|\psi_2\|^2\,ds \leq \kappa(a)t + 2h_1X_t \tag{2.56}$$

where $\kappa(a) = \frac{h_1}{\nu}\|af\|_{-2}^2 + T_Q$ and X_t is as in (2.47). Then with a simple manipulation we get

$$|||\mathbf{q}(t)|||_{-1}^2 - |||\mathbf{q}_0|||_{-1}^2 + \nu\int_0^t|\Delta\boldsymbol{\psi}|^2\,ds + 2rh_2\int_0^t\|\psi_2\|^2\,ds \leq \\ \kappa t + 2h_1\Xi_\gamma^a + 2h_1\gamma\langle X^a\rangle_t. \tag{2.57}$$

where Ξ_γ^a is as in (2.51).

Using Poincaré inequality twice in (2.50) we have

$$\langle X\rangle_t \leq \frac{\operatorname{Tr}Q}{\lambda_1^2}\int_0^t|\Delta\psi_1|^2\,ds. \tag{2.58}$$

which in (2.57) gives

$$|||\mathbf{q}(t)|||_{-1}^2 - |||\mathbf{q}_0|||_{-1}^2 + \left(\nu - \frac{2\gamma\operatorname{Tr}Q}{\lambda_1^2}\right)\int_0^t|\Delta\boldsymbol{\psi}|^2\,ds + 2rh_2\int_0^t\|\psi_2\|^2\,ds \\ \leq \kappa(a)t + 2h_1\Xi_\gamma^a.$$

Setting

$$\delta := \nu - \frac{2\gamma\operatorname{Tr}Q}{\lambda_1^2}$$

so that, for all $0 \leq \gamma \leq \nu - (\lambda_1^2)/(2\operatorname{Tr}Q)$ we have $0 \leq \delta \leq \nu$ and

$$|||\mathbf{q}(t)|||_{-1}^2 + \delta\int_0^t|\Delta\boldsymbol{\psi}|^2\,ds \leq |||\mathbf{q}_0|||_{-1}^2 + \kappa(a)t + 2h_1\Xi_\gamma^a. \tag{2.59}$$

□

2.4.2 Lipschitz continuity with respect to the forcing

We will show that the solution of (2.43) is locally Lipschitz in the parameter a with respect to the $\mathcal{H}$-norm $|||\cdot|||_{-1}$.

Theorem 2.4.3. *For $a \in \mathbb{R}$, let $\mathbf{q}(t,a)$ be the solution of* (2.43) *with initial condition $\mathbf{q}_0 \in \mathcal{H}$. Then for any $a_0 \in \mathbb{R}$, $T \geq 0$ and a.a. $\omega \in \Omega$ there is finite $C = C(a_0, t, \omega)$ such that*

$$|||\mathbf{q}(t,a) - \mathbf{q}(t,a_0)|||_{-1}^2 \leq C|a-a_0|^2 \quad \textit{for all } a \in \mathbb{R},\, t \in [0,T]. \tag{2.60}$$

Proof. Let $\mathbf{q}(t)$ and $\tilde{\mathbf{q}}(t)$ be the unique solutions of (2.43) respectively with parameter a and a_0, and same realization of the noise. Then the difference $\mathbf{u} := \mathbf{q} - \tilde{\mathbf{q}}$, with corresponding streamfunction $\boldsymbol{\phi} := \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}$, satisfies the following equation

$$\begin{aligned} &\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} + B(\boldsymbol{\phi}, \boldsymbol{\psi}) + B(\tilde{\boldsymbol{\psi}}, \boldsymbol{\phi}) + \beta\partial_1\boldsymbol{\phi} = \nu\Delta^2\boldsymbol{\phi} + \begin{pmatrix} (a-a_0)f \\ -r\Delta\phi_2 \end{pmatrix} \\ &\mathbf{u} = (\Delta + M)\,\boldsymbol{\phi}, \\ &\mathbf{u}(0) = 0 \end{aligned} \tag{2.61}$$

that should be interpreted in the integral sense.

To bound $|||\mathbf{u}|||_{-1}$ we take, as in the proof of Theorem 2.2.1, the $\mathbf{L}^2$ product with $\boldsymbol{\phi}$: by the properties of the nonlinearity and Green theorem we get

$$\frac{1}{2}\frac{\mathrm{d}|||\mathbf{u}|||_{-1}^2}{\mathrm{d}t} + \nu|\Delta\boldsymbol{\phi}|^2 + rh_2\|\phi_2\|^2 = (B(\tilde{\boldsymbol{\psi}}, \boldsymbol{\phi}), \boldsymbol{\phi}) - h_1(a-a_0)(f, \phi_1) \tag{2.62}$$

By the bound on the bilinearity (1.58) we have

$$|(B(\tilde{\boldsymbol{\psi}}, \boldsymbol{\phi}), \boldsymbol{\phi})| \leq k_B|\Delta\tilde{\boldsymbol{\psi}}|^2\|\boldsymbol{\phi}\|^2 + \tfrac{\nu}{2}|\Delta\boldsymbol{\phi}|^2$$

and by Young's inequality

$$h_1(a-a_0)|(f,\phi_1)| \leq |a-a_0|^2\frac{h_1\|f\|_{-2}^2}{\nu} + \frac{\nu h_1}{4}|\Delta\phi_1|^2$$

so that

$$\begin{aligned} \frac{1}{2}\frac{\mathrm{d}|||\mathbf{u}|||_{-1}^2}{\mathrm{d}t} + \frac{\nu}{2}|\Delta\boldsymbol{\phi}|^2 + rh_2\|\phi_2\|^2 &\leq k_B|\Delta\tilde{\boldsymbol{\psi}}|^2\|\boldsymbol{\phi}\|^2 \\ &\quad + |a-a_0|^2\frac{h_1\|f\|_{-2}^2}{\nu} + \frac{\nu h_1}{4}|\Delta\phi_1|^2. \end{aligned}$$

Finally, rearranging and using (1.43) we have

$$\frac{\mathrm{d}|||\mathbf{u}|||_{-1}^2}{\mathrm{d}t}+\frac{\nu}{2}|\Delta\boldsymbol{\phi}|^2+2rh_2\|\phi_2\|^2 \leq 2k_B|\Delta\tilde{\boldsymbol{\psi}}|^2|||\mathbf{u}|||_{-1}^2+|a-a_0|^2\frac{h_1\|f\|_{-2}^2}{\nu}. \quad (2.63)$$

and by Gronwall's lemma Lemma A.1 we have that for almost all times

$$|||\mathbf{u}(t)|||_{-1}^2 \leq |a-a_0|^2\tfrac{h_1\|f\|_{-2}^2}{\nu}\int_0^t \exp\left(2k_B\int_0^t |\Delta\tilde{\boldsymbol{\psi}}|^2\,d\tau\right)ds. \quad (2.64)$$

Theorem 2.4.2 ensures that the integral of $|\Delta\tilde{\boldsymbol{\psi}}|^2$ is almost surely finite and in particular gives the following estimate when using (2.54)

$$|||\mathbf{u}(t)|||_{-1}^2 \leq |a-a_0|^2 h_1\nu^{-1}\|f\|_{-2}^2\, t\, \exp\left(\tfrac{2k_B}{\delta}\left(|||\mathbf{q}_0|||_{-1}^2+\kappa(a_0)t+\Xi_\gamma^{a_0}\right)\right) \quad (2.65)$$

where $\kappa(a_0) = T_Q + h_1\nu^{-1}\|a_0 f\|_{-2}^2$. □

As a consequence of Theorem 2.4.3 we can also show the following integral property:

Corollary 2.4.4. *For $a \in \mathbb{R}$, let $\mathbf{q}(t,a)$ be the solution of* (2.43) *with initial condition $\mathbf{q}_0 \in \mathcal{H}$ and corresponding streamfunction $\boldsymbol{\psi}(t,a)$. Then for any $a_0 \in \mathbb{R}$, $T \geq 0$ and a.a. $\omega \in \Omega$ there is finite $K = K(a_0,t,\omega)$ such that*

$$\int_0^t |\Delta\boldsymbol{\psi}(s,a)-\Delta\boldsymbol{\psi}(s,a_0)|^2\,ds \leq K|a-a_0|^2 \quad \textit{for all } a\in\mathbb{R},\ t\in[0,T].$$

Proof. Consider again (2.63) and integrate both sides between 0 and t to find,

$$\frac{\nu}{2}\int_0^t |\Delta\boldsymbol{\phi}|^2\,ds \leq 2k_B\int_0^t |\Delta\tilde{\boldsymbol{\psi}}|^2|||\mathbf{u}|||_{-1}^2\,ds + |a-a_0|^2\frac{th_1\|f\|_{-2}^2}{\nu}$$

and, by Theorem 2.4.3, we have

$$\int_0^t |\Delta\boldsymbol{\phi}|^2\,ds \leq |a-a_0|^2\left(\frac{4k_B}{\nu}\int_0^t C(s)|\Delta\tilde{\boldsymbol{\psi}}|^2\,ds + \frac{2th_1\|f\|_{-2}^2}{\nu^2}\right). \quad (2.66)$$

By Theorem 2.4.2 the integral on the right hand side is well defined for almost all ω, concluding the proof. □

2.4.3 Differentiability with respect to the forcing

The solution of (2.43) is not only locally Lipschitz continuous with respect to the parameter a, but we now show it to be differentiable. Formally the derivative $D_a\mathbf{q}(t,a)|_{a=a_0}$ would satisfy the following equation

$$\frac{\mathrm{d}(D_a\mathbf{q})}{\mathrm{d}t} + DB(\boldsymbol{\psi},\boldsymbol{\psi})(D_a\boldsymbol{\psi}) + \beta\partial_1(D_a\boldsymbol{\psi}) = \nu\Delta^2(D_a\boldsymbol{\psi}) + \begin{pmatrix} f \\ -r\Delta(D_a\psi_2)\end{pmatrix} \tag{2.67}$$

where $\boldsymbol{\psi} = \boldsymbol{\psi}(t,a_0)$, since the additive noise is not depending on a. The result is divided in two parts, first in Theorem 2.4.5 we ensure that (2.67) has a unique solution. Then in Theorem 2.4.6 we show that such a solution is the desired derivative.

Consider the system for the variable $\mathbf{v}\in\mathcal{H}$ on the time interval $[0,T]$

$$\begin{aligned} &\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} + DB(\boldsymbol{\psi},\boldsymbol{\psi})\mathbf{g} + \beta\partial_1\mathbf{g} = \nu\Delta^2\mathbf{g} + \begin{pmatrix} f \\ -r\Delta g_2\end{pmatrix} \\ &\mathbf{v} = (\Delta + M)\mathbf{g}, \\ &\mathbf{v}(0) = 0 \end{aligned} \tag{2.68}$$

where $\boldsymbol{\psi} = \boldsymbol{\psi}(t,a_0)$ is the streamfunction solution of (2.43) when $a = a_0$.

Theorem 2.4.5. *Let $T>0$ and $f\in H^{-2}(\mathcal{D})$. Then the system* (2.68) *has a unique solution* $\mathbf{v}$ *such that*

$$\mathbf{v}\in C(0,T;\mathcal{H})\cap L^2(0,T;\mathcal{V}) \quad \mathbb{P}\text{-}a.s.$$

so that

$$\mathbf{g}\in C(0,T;\mathbf{H}^1)\cap L^2(0,T;\mathbf{H}^2) \quad \mathbb{P}\text{-}a.s..$$

Proof. Consider the n-dimensional Galerkin approximation of (2.68) with Π_n the projection onto the space spanned by the first n eigenvalues of $-\Delta$. The operators $(\Delta+M)$ and Δ are linear bounded operators which commute with Π_n so that

$$\mathbf{v}^{(n)} = \Pi_n(\Delta+M)g = (\Delta+M)\Pi_n\mathbf{g} = (\Delta+M)\mathbf{g}^{(n)}$$

and

$$\frac{\mathrm{d}\mathbf{v}^{(n)}}{\mathrm{d}t} + \Pi_n B(\boldsymbol{\psi},\mathbf{g}^{(n)}) + \Pi_n B(\mathbf{g}^{(n)},\boldsymbol{\psi}) + \beta\partial_1\mathbf{g}^{(n)} = \nu\Delta^2\mathbf{g}^{(n)} + \begin{pmatrix} \Pi_n f \\ -r\Delta\mathbf{g}_2^{(n)}\end{pmatrix} \tag{2.69}$$

where we have used Lemma 1.3.5 to express the action of the operator $DB(\boldsymbol{\psi},\boldsymbol{\psi})$. We want to find a bound for $\left|\left|\left|\mathbf{v}^{(n)}\right|\right|\right|_{-1}$ which is uniform in n. As in the previous results, we take the $\mathbf{L}^2$ product with $\mathbf{g}^{(n)} \in \Pi_n \mathbf{H}^1$ and get

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left|\left|\left|\mathbf{v}^{(n)}\right|\right|\right|_{-1}^2 + \nu|\Delta \mathbf{g}^{(n)}|^2 + rh_2\|g_2^{(n)}\|^2 = (B(\boldsymbol{\psi},\mathbf{g}^{(n)}),\mathbf{g}^{(n)}) - h_1(\Pi_n f, g_1^{(n)}).$$

As usual we apply Cauchy-Schwartz, Young and Poincaré inequalities to the deterministic forcing term to get

$$h_1(f, g_1^{(n)}) \leq \frac{h_1\|f\|_{-2}^2}{\nu} + \frac{\nu h_1}{4}|\Delta g_1^{(n)}|$$

and use the estimate (1.58) for the bilinearity and the fact that $\|\mathbf{g}^{(n)}\|^2 \leq \left|\left|\left|\mathbf{v}^{(n)}\right|\right|\right|_{-1}^2$ to get

$$\frac{\mathrm{d}\left|\left|\left|\mathbf{v}^{(n)}\right|\right|\right|_{-1}^2}{\mathrm{d}t} + \frac{\nu}{2}|\Delta \mathbf{g}^{(n)}|^2 + 2rh_2\|g_2^{(n)}\|^2 \leq \tfrac{2h_1}{\nu}\|f\|_{-2}^2 + 2k_B|\Delta\boldsymbol{\psi}|^2\left|\left|\left|\mathbf{v}^{(n)}\right|\right|\right|_{-1}^2. \tag{2.70}$$

By Gronwall's inequality we have

$$\left|\left|\left|\mathbf{v}^{(n)}(t)\right|\right|\right|_{-1}^2 \leq \tfrac{2h_1}{\nu}\|f\|_{-2}^2 \int_0^t \exp\left(2k_B\int_s^t |\Delta\boldsymbol{\psi}|^2\, d\tau\right) ds. \tag{2.71}$$

Again Theorem 2.4.2 ensures the integrating factor is well defined for almost all ω and as a consequence $\left|\left|\left|\mathbf{v}^{(n)}(t)\right|\right|\right|_{-1}^2$ is bounded uniformly in n in $L^\infty(0,T;\mathcal{H})$ for a.a. $\omega \in \Omega$.

Integrating (2.70) over time we have

$$\tfrac{\nu}{2}\int_0^t |\Delta\mathbf{g}^{(n)}|^2 \leq \tfrac{2h_1\|f\|_{-2}^2}{\nu}\, t + \int_0^t 2k_B|\Delta\boldsymbol{\psi}|^2\left|\left|\left|\mathbf{v}^{(n)}\right|\right|\right|_{-1}^2 ds, \tag{2.72}$$

so that by the definition of the $|||\cdot|||_0$ norm (1.39)

$$\int_0^t |||\mathbf{v}|||_0^2\, ds \leq \tfrac{2h_1\|f\|_{-2}^2}{\nu}\, t + \int_0^t 2k_B|\Delta\boldsymbol{\psi}|^2\left|\left|\left|\mathbf{v}^{(n)}\right|\right|\right|_{-1}^2 ds.$$

Using the *a priori* bound (2.71) and again Theorem 2.4.2, we can conclude that $\mathbf{v} \in L^2(0,T;\mathcal{V})$.

With arguments similar to the proof of Theorem 2.2.1 we can show that $d\mathbf{v}^{(n)}/dt$ is bounded uniformly in n in $L^2(0,T;\mathcal{V}^*)$ so that we can extract convergent subsequences

$$\mathbf{v}_n \rightharpoonup \mathbf{v} \quad \text{in } L^\infty(0,T;\mathcal{H}) \cap L^2(0,T;\mathcal{V}) \tag{2.73}$$

$$\frac{\mathrm{d}\mathbf{v}_n}{\mathrm{d}t} \rightharpoonup \mathbf{v} \quad \text{in } L^2(0,T;\mathbf{H}^{-2}). \tag{2.74}$$

Then, with arguments that require little novelty at this point there exists a solution of (2.68) such that

$$\mathbf{v} \in C(0,T;\mathcal{H}) \cap L^2(0,T;\mathcal{V})$$

and one can ensure as well that the solution is unique and continuous with respect to the initial condition.

□

Next we have to ensure that such a unique solution $\mathbf{v}$ is exactly the derivative of the potential vorticity $\mathbf{q}$ with respect to the parameter a:

Theorem 2.4.6. *Let* $\mathbf{q}$ *be the solution of* (2.43) *and* $\mathbf{v}$ *the corresponding solution of* (2.68). *Then*

$$\lim_{a\to a_0} \left|\left|\left| \frac{\mathbf{q}(t,a)-\mathbf{q}(t,a_0)}{a-a_0} - \mathbf{v}(t,a_0) \right|\right|\right|_{-1} = 0 \quad \textit{for all } t \in [0,T], \mathbb{P}\textit{-a.s.}.$$

Proof. For brevity of notations let us introduce the following variables

$$\begin{aligned} \boldsymbol{\phi}(t) &= \boldsymbol{\psi}(t,a) - \tilde{\boldsymbol{\psi}}(t,a_0) \\ \mathbf{u}(t) &= \mathbf{q}(t,a) - \tilde{\mathbf{q}}(t,a_0) = (\Delta + M)\boldsymbol{\phi} \\ \boldsymbol{\gamma}(t) &= \frac{\boldsymbol{\phi}(t)}{a-a_0} - \mathbf{g}(t,a_0) \\ \boldsymbol{\theta}(t) &= \frac{\mathbf{u}(t)}{a-a_0} - \mathbf{v}(t,a_0) = (\Delta + M)\boldsymbol{\gamma}, \end{aligned} \tag{2.75}$$

so that, given (2.61) and (2.68), the variables of interest $\{\boldsymbol{\theta},\boldsymbol{\gamma}\}$ satisfy

$$\frac{\mathrm{d}\boldsymbol{\theta}}{\mathrm{d}t} + \frac{B(\boldsymbol{\phi},\tilde{\boldsymbol{\psi}}) + B(\boldsymbol{\psi},\boldsymbol{\phi})}{a-a_0} - DB(\tilde{\boldsymbol{\psi}},\tilde{\boldsymbol{\psi}})\mathbf{g} + \beta\partial_1\boldsymbol{\gamma} = \nu\Delta^2\boldsymbol{\gamma} + \begin{pmatrix} 0 \\ -r\Delta\gamma_2 \end{pmatrix}.$$

Using the formula (1.63) for the derivative of the bilinearity

$$\begin{aligned} \frac{B(\boldsymbol{\phi},\tilde{\boldsymbol{\psi}}) + B(\boldsymbol{\psi},\boldsymbol{\phi})}{a-a_0} - DB(\tilde{\boldsymbol{\psi}},\tilde{\boldsymbol{\psi}})\mathbf{g} &= B\left(\frac{\boldsymbol{\phi}}{a-a_0},\tilde{\boldsymbol{\psi}}\right) + B\left(\boldsymbol{\psi},\frac{\boldsymbol{\phi}}{a-a_0}\right) \\ &\quad - B(\tilde{\boldsymbol{\psi}},\mathbf{g}) - B(\mathbf{g},\tilde{\boldsymbol{\psi}}), \end{aligned}$$

so, introducing an appropriate term, we have

$$\begin{aligned} \frac{B(\boldsymbol{\phi},\tilde{\boldsymbol{\psi}}) + B(\boldsymbol{\psi},\boldsymbol{\phi})}{a-a_0} - DB(\tilde{\boldsymbol{\psi}},\tilde{\boldsymbol{\psi}})\mathbf{g} &= B(\boldsymbol{\gamma},\tilde{\boldsymbol{\psi}}) + B\left(\boldsymbol{\psi},\frac{\boldsymbol{\phi}}{a-a_0}\right) \\ &\quad - B\left(\tilde{\boldsymbol{\psi}},\frac{\boldsymbol{\phi}}{a-a_0}\right) + B\left(\tilde{\boldsymbol{\psi}},\frac{\boldsymbol{\phi}}{a-a_0}\right) - B(\tilde{\boldsymbol{\psi}},\mathbf{g}). \end{aligned}$$

Therefore $\boldsymbol{\theta}$ satisfies the following equation with a now familiar structure

$$\frac{\mathrm{d}\boldsymbol{\theta}}{\mathrm{d}t} + B(\boldsymbol{\gamma}, \tilde{\boldsymbol{\psi}}) + B(\tilde{\boldsymbol{\psi}}, \boldsymbol{\gamma}) + \frac{B(\boldsymbol{\phi}, \boldsymbol{\phi})}{a - a_0} + \beta\partial_1\boldsymbol{\gamma} = \nu\Delta^2\boldsymbol{\gamma} + \begin{pmatrix} 0 \\ -r\Delta\gamma_2 \end{pmatrix}.$$

Take the $\mathbf{L}^2$-scalar product of both sides with $\boldsymbol{\gamma}$ so that

$$\frac{1}{2}\frac{\mathrm{d}|||\boldsymbol{\theta}|||_{-1}^2}{\mathrm{d}t} + \nu|\Delta\boldsymbol{\gamma}|^2 + rh_2\|\gamma_2\|^2 = -(B(\tilde{\boldsymbol{\psi}}, \boldsymbol{\gamma}), \boldsymbol{\gamma}) - \frac{(B(\boldsymbol{\phi}, \boldsymbol{\phi}), \boldsymbol{\gamma})}{a - a_0}$$

where we have used the properties of the nonlinearity (1.51) and (1.56).

By (1.58) and (1.59) we have respectively

$$|(B(\tilde{\boldsymbol{\psi}}, \boldsymbol{\gamma}), \boldsymbol{\gamma})| \leq \tfrac{\nu}{2}|\Delta\boldsymbol{\gamma}|^2 + k_B|\Delta\tilde{\boldsymbol{\psi}}|^2\|\boldsymbol{\gamma}\|^2, \tag{2.76}$$

$$\frac{|(B(\boldsymbol{\phi}, \boldsymbol{\phi}), \boldsymbol{\gamma})|}{a - a_0} \leq \tfrac{\nu}{2}|\Delta\boldsymbol{\gamma}|^2 + \frac{k_B|\Delta\boldsymbol{\phi}|^2\|\boldsymbol{\phi}\|^2}{(a - a_0)^2}, \tag{2.77}$$

which provide the following energy estimate for the variable $\boldsymbol{\theta}$

$$\frac{\mathrm{d}|||\boldsymbol{\theta}|||_{-1}^2}{\mathrm{d}t} \leq \frac{2k_B|\Delta\boldsymbol{\phi}|^2\|\boldsymbol{\phi}\|^2}{(a - a_0)^2} + 2k_B|\Delta\tilde{\boldsymbol{\psi}}|^2\|\boldsymbol{\gamma}\|^2.$$

Recall that $\|\boldsymbol{\gamma}\|^2 \leq |||\boldsymbol{\theta}|||_{-1}^2$ so that, by Gronwall's inequality, we have

$$|||\boldsymbol{\theta}(t)|||_{-1}^2 \leq \int_0^t \frac{2k_B|\Delta\boldsymbol{\phi}|^2\|\boldsymbol{\phi}\|^2}{(a - a_0)^2} \exp\left(2k_B\int_s^t |\Delta\tilde{\boldsymbol{\psi}}|^2\, d\tau\right)\, ds.$$

We have shown in Theorem 2.4.3 that $\mathbf{q}(t, a)$ is locally Lipschitz with respect to a and in particular

$$\|\boldsymbol{\phi}(t)\|^2 \leq |||\mathbf{u}(t)|||_{-1}^2 \leq |a - a_0|^2\tfrac{h_1\|f\|_{-2}^2}{\nu}\int_0^t \exp\left(2k_B\int_s^t |\Delta\tilde{\boldsymbol{\psi}}|^2\, d\tau\right) ds \tag{2.78}$$

so that

$$|||\boldsymbol{\theta}(t)|||_{-1}^2 \leq \tfrac{2k_Bh_1\|f\|_{-2}^2}{\nu}\exp\left(4k_B\int_0^t |\Delta\tilde{\boldsymbol{\psi}}|^2\, d\tau\right)\int_0^t |\Delta\tilde{\boldsymbol{\phi}}|^2\, ds.$$

The Lipschitz continuity of $\boldsymbol{\psi}$ in $L^2(0, T; \mathbf{H}^2)$ given by Corollary 2.4.4, in particular estimate (2.66), gives

$$|||\boldsymbol{\theta}|||_{-1}^2 \leq \tfrac{2k_Bh_1\|f\|_{-2}^2}{\nu}\exp\left(4k_B\int_0^t |\Delta\tilde{\boldsymbol{\psi}}|^2\, d\tau\right)\cdot$$

$$\cdot\left((a - a_0)^2\tfrac{2th_1\|f\|_{-1}^2}{\nu^2\lambda_1} + \tfrac{4k_B}{\nu}\int_0^t |||\mathbf{u}|||_{-1}^2|\Delta\tilde{\boldsymbol{\psi}}|^2\, ds\right).$$

Thanks to (2.78) we have

$$\int_0^t \|\|\mathbf{u}\|\|_{-1}^2 |\Delta\tilde{\boldsymbol{\psi}}|^2\, ds \le (a-a_0)^2 \frac{th_1\|f\|_{-2}^2}{\nu} \exp\left(2k_B \int_0^t |\Delta\tilde{\boldsymbol{\psi}}|^2\, d\tau\right) \int_0^t |\Delta\tilde{\boldsymbol{\psi}}|^2\, d\tau$$

hence

$$\|\|\boldsymbol{\theta}(t)\|\|_{-1}^2 \le \tfrac{2k_B h_1\|f\|_{-2}^2}{\nu} \exp\left(4k_B \int_0^t |\Delta\tilde{\boldsymbol{\psi}}|^2\, d\tau\right)(a-a_0)^2 \cdot$$
$$\cdot \left(\frac{2th_1\|f\|_{-1}^2}{\nu^2\lambda_1} + \frac{4tk_B h_1\|f\|_{-2}^2}{\nu^2} \exp\left(2k_B \int_0^t |\Delta\tilde{\boldsymbol{\psi}}|^2\, d\tau\right) \int_0^t |\Delta\tilde{\boldsymbol{\psi}}|^2\, d\tau\right).$$

Rearranging appropriately we find a well defined function $R(t)$ such that

$$\|\|\boldsymbol{\theta}(t)\|\|_{-1}^2 \le (a-a_0)^2 R(t) \exp\left(\int_0^t 7k_B |\Delta\tilde{\boldsymbol{\psi}}(a_0)|^2\, ds\right). \tag{2.79}$$

Since $\tilde{\boldsymbol{\psi}} \in L^2(0,T;\mathbf{H}^2)$, the exponential is well defined and therefore we can conclude that

$$\|\|\boldsymbol{\theta}(t)\|\|_{-1} = \left\|\left\|\frac{\mathbf{u}(t)}{a-a_0} - \mathbf{v}(t,0;a_0)\right\|\right\|_{-1}$$

vanishes on the limit of a to a_0 as desired. □

2.5 Regularity with respect to the forcing for stochastic Navier–Stokes

We conclude this chapter with an investigation of stochastic Navier–Stokes model as presented in section 1.2 where $\mathcal{H} = (\mathbf{L}^2, |\cdot|)$ and $\mathcal{V} = (\mathbf{H}^1, \|\cdot\|)$. Consider (1.5) with the deterministic forcing $f \in H^{-1}$ modulated by a constant $a > 0$, i.e.

$$du + (\nu Au + B(u,u))\, dt = af\, dt + dW \quad u(0) = u_0. \tag{2.80}$$

Here W is a Q-Wiener process with Q trace class in H^1 so that the solution is

$$u(t) \in C(0,T;\mathcal{H}) \cap L^2([0,T];\mathcal{V}).$$

As we did for the stochastic two–layer quasi–geostrophic model, we will show that the solution $u(t,a)$ is locally Lipschitz continuous and differentiable with respect to the parameter a, after proving some useful bounds.

2.5.1 Energy estimate for stochastic Navier–Stokes

Similarly to what done in (2.47)-(2.51) let us now set

$$X_t^a := \int_0^t (u(s,a), dW(s)) \tag{2.81}$$

so that its quadratic variation is

$$\langle X^a \rangle_t := \int_0^t \|(u(s,a),\cdot)\|^2_{L_2^0}\, ds$$

and, given the estimate (2.49) for the integrand, we have

$$\langle X^a \rangle_t \leq \operatorname{Tr} Q \int_0^t |u(s,a)|^2\, ds. \tag{2.82}$$

Furthermore Lemma 2.4.1 holds for

$$\Xi_\gamma^a := \sup_{t\geq 0} \left(X_t^a - \gamma \langle X^a \rangle_t \right) \tag{2.83}$$

and we have the following result parallel to Theorem 2.4.2.

Theorem 2.5.1. *Consider the stochastic Navier-Stokes model* (2.80) *with initial condition* $u(0) = u_0 \in \mathcal{H}$, $f \in H^{-1}$ *and* $q := \operatorname{Tr} Q$. *Then, given* Ξ_γ^a *as in* (2.83) *for all* $0 < \delta < \nu$ *there exists* $\gamma \geq 0$ *such that*

$$|u(t)|^2 + \delta \int_0^t \|u(s)\|^2\, ds \leq |u_0|^2 + t \left(\frac{a^2}{\nu} \|f\|_{-1}^2 + q \right) + 2\Xi_\gamma^a. \tag{2.84}$$

Proof. Consider (2.80) and take the $\mathcal{H}$-scalar product with u itself to get, by use of Itô lemma,

$$|u(t)|^2 - |u_0|^2 + 2\nu \int_0^t \|u(s)\|^2\, ds = \int_0^t 2\langle af, u\rangle\, ds + t \operatorname{Tr} Q + 2 \int_0^t \langle u, \cdot \rangle\, dW_s. \tag{2.85}$$

where we have used that $(B(u,u),u) = 0$ and that $(Au,u) = \|u\|^2$.

Using Poincaré's inequality in (2.82) we get

$$\langle X^a \rangle_t \leq q \lambda_1^{-1} \int_0^t \|u\|^2\, ds.$$

We bound the deterministic forcing term as follows

$$2\langle af, u\rangle \leq \frac{\|af\|_{-1}^2}{\nu} + \nu \|u\|^2 \tag{2.86}$$

so that, using the estimate for the quadratic variation in (2.85) we have

$$|u(t)|^2 + (\nu - \gamma q \lambda_1^{-1}) \int_0^t \|u\|^2 \, ds \le |u_0|^2 + t \left(\frac{a^2}{\nu} \|f\ |_{-1}^2 + q \right) + 2\Xi_\gamma^a.$$

For $0 < \gamma < \nu\lambda_1/q$, define

$$\delta = \nu - \gamma q \lambda_1^{-1}$$

to get the desired statement.

□

2.5.2 Lipschitz continuity with respect to the forcing

Thanks to Theorem 2.5.1 we can show that the solution of (2.80) is locally Lipschitz continuous with respect to the strength a of the forcing.

Theorem 2.5.2. *For $a \in \mathbb{R}$, let $u(t,a)$ be the solution of* (2.80) *with initial condition $u_0 \in \mathcal{H}$. Then for any $a_0 \in \mathbb{R}$, $T \ge 0$ and a.a. $\omega \in \Omega$ there is finite $C = C(a_0, t, \omega)$ such that*

$$|u(t,a) - u(t,a_0)|^2 \le C|a - a_0|^2 \quad \textit{for all } a \in \mathbb{R},\ t \le T. \tag{2.87}$$

Proof. Set $u(t) := u(t,\omega; u_0, a_0)$, $v(t) := u(t,\omega; u_0, a)$ and $w := u - v$. Then it is easy to see that w must satisfy the following equation

$$\frac{dw}{dt} + \nu A w + B(w,u) + B(v,w) = (a - a_0) f, \quad w(0) = 0. \tag{2.88}$$

As usual we compute an energy estimate taking the $\mathcal{H}$ scalar product of (2.80) with w

$$\frac{1}{2}\frac{d|w|^2}{dt} + \nu\|w\|^2 + (B(w,u), w) = (a - a_0)(f, w),$$

where we have used that $(B(v,w), w) = 0$. Using the estimates for the trilinear form (Lemma 1.2.1), Cauchy-Schwartz we get

$$\frac{1}{2}\frac{d|w|^2}{dt} + \nu\|w\|^2 \le k_B |w| \|w\| \|u\| + |a - a_0| \|f\|_{-1} \|w\|, \tag{2.89}$$

and by Young inequality

$$\frac{1}{2}\frac{d|w|^2}{dt} + \frac{\nu}{2}\|w\|^2 \le \frac{k_B^2}{\nu}\|u\|^2 |w|^2 + \frac{|a - a_0|^2}{\nu}\|f\|_{-1}^2. \tag{2.90}$$

By Gronwall's inequality we get

$$|w(t,\omega)|^2 \leq |a-a_0|^2 \frac{2\|f\|_{-1}^2}{\nu} \int_0^t \exp\left(\frac{2k_B^2}{\nu}\int_s^t \|u(\tau)\|^2\, d\tau\right)\, ds. \tag{2.91}$$

Thanks to Theorem 2.5.1 part (i) i.e. (2.84) we have that

$$|w(t,\omega)|^2 \leq |a-a_0|^2 \tfrac{2\|f\|_{-1}^2}{\nu}\, t\, \exp\left(\tfrac{2k_B^2}{\nu\delta}\left(|u_0|^2 + t\left(\tfrac{\|a_0 f\|^2}{\nu} + q\right) + 2\Xi_\gamma^{a_0}\right)\right) \tag{2.92}$$

□

Next we show another immediate consequence of Theorem 2.5.2.

Corollary 2.5.3. *For $a \in \mathbb{R}$, let $u(t,a)$ be the solution of* (2.80) *with initial condition $u_0 \in \mathcal{H}$. Then for any $a_0 \in \mathbb{R}$, $T \geq 0$ and a.a. $\omega \in \Omega$ there exists $K = K(a_0,t,\omega)$ such that*

$$\int_0^t \|u(t,a) - u(t,a_0)\|^2\, ds \leq K|a-a_0|^2 \quad \textit{for all } a \in \mathbb{R},\ t \in [0,T].$$

namely the solution is locally Lipschitz continuous with respect to a in the $L^2(0,t;\mathcal{V})$ norm.

Proof. To show (2.5.3) holds we integrate in time (2.90) to get

$$\int_0^t \|w\|^2\, ds \leq \frac{2k_B^2}{\nu^2}\int_0^t \|u\|^2|w|^2\, ds + |a-a_0|^2 \frac{t\|f\|_{-1}^2}{\nu^2}. \tag{2.93}$$

Then using (2.87) and the fact that $u \in L^2(0,T;\mathcal{V})$ almost surely, we get local Lipschitz continuity of u in the $L^2(0,T;\mathcal{V})$ norm. □

2.5.3 Differentiability with respect to the forcing

Finally we show that the solution of (2.80) is differentiable with respect to the parameter a. Taking *formally* the derivative of (2.80) we have

$$\frac{d(D_a u)}{dt} + \nu A(D_a u) + DB(u,u)(D_a u) = f, \quad (D_a u)(0) = 0.$$

As done for the bilinear operator of the 2LQG model in Lemma 1.3.5, it can be showed that the operator $DB(u,u) \in L(\mathcal{V},\mathcal{V}^*)$ has explicit formulation

$$DB(u,u)v = B(u,v) + B(v,u). \tag{2.94}$$

Then as done for the quasi–geostrophic model, first we show that this equation has a unique solution, second we ensure that this solution is precisely the desired derivative.

Theorem 2.5.4. *Let* $u = u(t, a_0)$ *be the solution of* (2.80)*. Then the equation over the time interval* $[0,T]$

$$\frac{d\xi}{dt} + \nu A\xi + DB(u,u)\xi = f, \quad \xi(0) = 0 \tag{2.95}$$

has a unique solution with values in $C([0,T];\mathcal{H}) \cap L^2(0,T;\mathcal{V})$ *for almost all* ω *which is continuous with respect to the initial condition.*

Proof. Existence. We take the n-dimensional Galerkin approximation

$$\frac{d\xi^{(n)}}{dt} + \nu A\xi^{(n)} + \Pi_n DB(u,u)\xi^{(n)} = \Pi_n f \quad \xi^{(n)}(0) = 0. \tag{2.96}$$

and show that $(\xi^{(n)})_{n\in\mathbb{N}}$ is bounded in $L^\infty(0,T;\mathcal{H})$ and $L^2(0,T;\mathcal{V})$ uniformly in n. Taking the inner product with $\xi^{(n)}$ we get

$$\frac{1}{2}\frac{d|\xi^{(n)}|^2}{dt} + \nu\|\xi^{(n)}\|^2 + (DB(u,u)\xi^{(n)}, \Pi_n\xi^{(n)}) = (f, \Pi_n\xi^{(n)}).$$

Using the relation (2.94) and the fact that $(B(u,v),v) = 0$ we obtain

$$\frac{1}{2}\frac{d|\xi^{(n)}|^2}{dt} + \nu\|\xi^{(n)}\|^2 + (B(\xi^{(n)},u),\xi^{(n)}) = (f, \Pi_n\xi^{(n)}).$$

Therefore, via the trilinear form estimates (Lemma 1.2.1), as well as Cauchy Schwartz and Young's inequalities, we have

$$\frac{1}{2}\frac{d|\xi^{(n)}|^2}{dt} + \frac{\nu}{2}\|\xi^{(n)}\|^2 \le \frac{k_B^2}{\nu}\|u\|^2|\xi^{(n)}|^2 + \frac{\|f\|_{-1}^2}{\nu}. \tag{2.97}$$

Finally, via Gronwall's inequality,

$$|\xi^{(n)}(t)|^2 \le \int_0^t \tfrac{2\|f\|_{-1}^2}{\nu} \exp\left(\tfrac{2k^2}{\nu}\int_s^t \|u\|^2 d\tau\right) ds. \tag{2.98}$$

Since $u \in L^2(0,T;\mathcal{V})$ for almost every ω, we have that $\xi^{(n)}$ is bounded uniformly in n in $L^\infty(0,T;\mathcal{H})$ for a.a. $\omega \in \Omega$.

Integrating (2.97) over time we have

$$\int_0^t \|\xi^{(n)}\|^2 ds \le \int_0^t \|u\|^2|\xi^{(n)}|^2 ds + \frac{t\|f\|_{-1}^2}{2\nu}$$

and the *a priori* bound (2.98) gives us that the sequence $\xi^{(n)}$ is bounded in $L^2(0,T;\mathcal{V})$ uniformly in n for almost every ω.

Now, we know we can extract a subsequence converging weakly both in $L^\infty(0,T;\mathcal{H})$ and $L^2(0,T;\mathcal{V})$. By further arguments that do not require any novelty one could show its limit ξ is precisely a solution of the desired equation and $\xi \in C([0,T];\mathcal{H}) \cap L^2(0,T;\mathcal{V})$.

Uniqueness. If $\eta = \xi - \overline{\xi}$ is the difference between two solutions then it satisfies

$$\frac{d\eta}{dt} + \nu A\eta + DB(u,u)\eta = 0 \quad \eta_0 = \xi_0 - \overline{\xi}_0$$

and, treating as usual the derivative of the nonlinearity, the following energy estimate holds

$$\frac{1}{2}\frac{d|\eta|^2}{dt} + \frac{\nu}{2}\|\eta\|^2 \leq \frac{k_B^2}{2\nu}\|u\|^2|\eta|^2.$$

Therefore

$$|\eta(t)|^2 \leq |\eta(0)|^2 \exp\left(\tfrac{k_B^2}{\nu}\int_0^t \|u\|^2 ds\right).$$

and, by the fact that $u(\omega) \in L^2(0,T;\mathcal{V})$ for almost all ω, we have the result. □

We close this section by showing that ξ is indeed $D_a u(t)$, derivative of the solution u with respect to the parameter a .

Theorem 2.5.5. *Let $u(t,a)$ be the solution of* (2.80) *and ξ the solution of* (2.95)*. Then for all $t \in [0,T]$*

$$\lim_{a\to a_0}\left|\frac{u(t,a)-u(t,a_0)}{a-a_0} - \xi(t,a_0)\right| = 0 \quad \mathbb{P}\text{-}a.s. \tag{2.99}$$

Proof. Let $u(t)$ and $v(t)$ be solutions of (2.80) respectively for the parameter a and a_0 and set $w(t) = u(t) - v(t)$ and

$$\theta(t) = \frac{w(t)}{a-a_0} - \xi(t,a_0), \tag{2.100}$$

then θ satisfies the following equation

$$\frac{d\theta}{dt} + \nu A\theta + DB(v,v)\theta = DB(v,v)\frac{w(t)}{a-a_0} - \frac{B(u,u)-B(v,v)}{a-a_0}.$$

Given the fact that

$$DB(v,v)\theta = B(v,\theta) + B(\theta,v)$$

and the bilinearity of B we have

$$\frac{d\theta}{dt} + \nu A\theta + B(v,\theta) + B(\theta,v) = \frac{B(v,w)+B(w,v)}{a-a_0} - \\ - \frac{B(u,w)+B(w,v)}{a-a_0},$$

hence

$$\frac{d\theta}{dt} + \nu A\theta + B(v,\theta) + B(\theta,v) = -\frac{B(w,w)}{a-a_0}. \tag{2.101}$$

Taking the $\mathcal{H}$ product of (2.101) with θ and using the properties of the bilinarity we get

$$\frac{1}{2}\frac{d|\theta|^2}{dt} + \nu\|\theta\|^2 \leq k_B\|v\||\theta|\|\theta\| + \frac{k_B|w|\|w\|\|\theta\|}{a-a_0},$$

and by Young's inequality

$$\frac{1}{2}\frac{d|\theta|^2}{dt} + \frac{\nu}{2}\|\theta\|^2 \leq \frac{k_B^2\|v\|^2}{\nu}|\theta|^2 + \frac{k_B^2|w|^2\|w\|^2}{\nu(a-a_0)^2}. \tag{2.102}$$

It is then enough to have local Lipschitz continuity in $\mathcal{H}$ and in $L^2(0,T;\mathcal{V})$ of the solution $u(t,a)$ with respect to the parameter. In fact, using Gronwall's inequality, we have

$$|\theta(t)|^2 \leq \tfrac{2k^2}{\nu}\int_0^t \exp\left(\int_s^t \tfrac{2k^2}{\nu}\|v\|^2\, d\tau\right)\frac{|w|^2}{(a-a_0)^2}\|w\|^2\, ds. \tag{2.103}$$

Thanks to the local Lipschitz continuity of $u(t)$ with respect to the parameter a in the $\mathcal{H}$-norm (2.91) we have

$$\begin{aligned}|\theta(t)|^2 &\leq \tfrac{2k^2}{\nu}e^{2k^2\nu^{-1}\int_0^t \|u_\tau(a_0)\|^2\, d\tau}\int_0^t \tfrac{2}{\nu}\|f\|_{-1}^2 e^{2k^2\nu^{-1}\int_0^s \|u_\tau(a_0)\|^2\, d\tau}\|w\|^2\, ds \\ &\leq \tfrac{4k^2}{\nu^2}\|f\|_{-1}^2 e^{\int_0^t 4k^2\nu^{-1}\|u_\tau(a_0)\|^2\, d\tau}\int_0^t \|w\|^2\, ds.\end{aligned}$$

Furthermore using the bound (2.93) for $\|w\|_{L^2(0,T;\mathcal{V})}$ we get

$$|\theta(t)|^2 \leq \tfrac{4k^2}{\nu^2}\|f\|_{-1}^2 e^{\int_0^t 4k^2\nu^{-1}\|u_\tau(a_0)\|^2\, d\tau}\left(\tfrac{2k^2}{\nu^2}\int_0^t \|u\|^2|w|^2\, ds + \right. \\ \left. + |a-a_0|^2\frac{t\|f\|_{-1}^2}{\nu^2}\right). \tag{2.104}$$

Using once more thanks to the local Lipschitz continuity (2.91), after some simple computations, we get

$$|\theta(t)|^2 \leq |a-a_0|^2 R(t) \exp\left(\int_0^t \tfrac{7k^2}{\nu} \|u_\tau(a_0)\|^2 \, d\tau\right) \tag{2.105}$$

with $R(t)$ depending on $\|f\|_{-1}$, the viscosity ν and k_B constant from the bound on the trilinear form which carries information on the domain $\mathcal{D}$. Since $u \in L^2(0,t;\mathcal{V})$ we can conclude that $|\theta(t)|^2 \to 0$ almost surely for $a \to a_0$ as desired. □

Summary and remarks

In this chapter we discussed two major topics, the solution theory for the stochastic two–layer quasi–geostrophic model (2.6), and the regularity with respect to the intensity of the forcing of both the stochastic two–layer quasi–geostrophic model and the stochastic Navier–Stokes equation.

We first introduced the concepts of weak and strong solution, drawing a parallel to Navier-Stokes, and then showed that such solutions exists and are pathwise unique. We achieved these results by means of the associated random equation (2.8). We ensured well–posedness of the random equation with methodology proper of the theory of partial differential equations. Some estimates established along the way, like the *a priori* bounds (2.40) and (2.14), will prove crucial in the next chapter when studying the existence of an invariant measure for the system.

In the second part of the chapter, we dealt with the regularity of the solutions with respect to the parameter of interest, showing almost sure local Lipschitz continuity as well as almost sure differentiability. These results are essential to study further the dependence of the long–term dynamics on the parameters, topic which will be the focus of chapter 5. Furthermore, the estimate (2.52), as well as energy bounds in Theorem 2.4.2 for the quasi–geostrophic model, and in Theorem 2.5.1 for the Navier-Stokes equation, will be extensively used in all the chapters that follow.

Chapter 3

Unique ergodicity

We are interested in studying the long time average behaviour of the solutions of the stochastic 2LQG model studied in the previous chapter. This type of information is the focus of ergodic theory and specifically we talk of an *ergodic system* if the long time averages of an observable can be approximated by its averages with respect to a measure that is *invariant* for the dynamics.

For Markovian systems we study these properties by means of the *Markov semigroup* associated to the stochastic process $\{X_t, t \geq 0\}$ of interest, in our case the solution of a stochastic partial differential equation (SPDE) modelling atmosphere or ocean dynamics. Let us introduce these key concepts precisely (we refer to [39], [22] and [21]).

The following concepts can be defined for more general spaces but here we are in particular interested in Hilbert spaces. Let $(\mathcal{H}, |\cdot|)$ be a Hilbert space with Borel σ-algebra $\mathcal{B}(\mathcal{H})$. We denote $B_b(\mathcal{H})$ the space of all real bounded Borel functions on $\mathcal{H}$, $\mathcal{M}(\mathcal{H})$ the space of signed measures on $\mathcal{H}$ and $\mathcal{M}_1(\mathcal{H})$ the space of probability measures on $\mathcal{H}$.

Definition 3.0.1 (Markov transition function)**.** The map $P_t(x, \Gamma)$, $x \in \mathcal{H}$, $\Gamma \in \mathcal{B}(\mathcal{H})$ is called Markov transition function on $\mathcal{H}$ if

(i) $P_t(x, \cdot)$ is a probability measure on $(\mathcal{H}, \mathcal{B}(\mathcal{H}))$ for each $t \geq 0$, $x \in \mathcal{H}$;

(ii) $P_t(\cdot, \Gamma)$ is a $\mathcal{B}(\mathcal{H})$-measurable function for all $t \geq 0$, $\Gamma \in \mathcal{B}(\mathcal{H})$;

(iii) $P_{t+s}(x, \Gamma) = \int P_s(y, \Gamma)\, P_t(x, dy)$ for each $t, s \geq 0$, $x \in \mathcal{H}$, $\Gamma \in \mathcal{B}(\mathcal{H})$;

(iv) $P_0(x, \Gamma) = \mathbb{1}_\Gamma(x)$ for each $x \in \mathcal{H}$, $\Gamma \in \mathcal{B}(\mathcal{H})$.

Definition 3.0.2 (Markov semigroup)**.** Let P_t, $t \geq 0$, be a Markov transition function, then we define the associated Markov semigroup as a family of linear operators $\{\mathcal{P}_t : t \geq 0\}$ on $B_b(\mathcal{H})$ by

$$\mathcal{P}_t\varphi(x) := \int \varphi(y)\, P_t(x, dy) \tag{3.1}$$

for all $t \geq 0$, $x \in \mathcal{H}$, $\varphi \in B_b(\mathcal{H})$, such that $\mathcal{P}_{t+s} = \mathcal{P}_t \circ \mathcal{P}_s$ for all nonnegative s, t. Furthermore for any $t \geq 0$ and $\mu \in \mathcal{M}_1(\mathcal{H})$ we define the dual operator $\mathcal{P}_t^*\mu$ as

$$\mathcal{P}_t^*\mu(\Gamma) := \int \mathcal{P}_t \mathbb{1}_\Gamma(x)\, \mu(dx) \tag{3.2}$$

for $t \geq 0$, $\Gamma \in \mathcal{B}(\mathcal{H})$.

Note in particular that then

$$\mathcal{P}_t^*\mu(\Gamma) = \int P_t(x, \Gamma)\, \mu(dx), \tag{3.3}$$

and for $\varphi \in B_b(\mathcal{H})$

$$\int \varphi(z)\, (\mathcal{P}_t^*\mu)(dz) = \int \mathcal{P}_t\varphi(y)\, \mu(dy) = \int\int \varphi(z) P_t(y, dz)\, \mu(dy). \tag{3.4}$$

Note that $\mathcal{P}_t$ for each $t \geq 0$ is itself a Markov operator according to:

Definition 3.0.3 (Markov operator)**.** A Markov operator over $\mathcal{H}$ is a bounded linear operator $\mathcal{P} : B_b(\mathcal{H}) \to B_b(\mathcal{H})$ such that

(i) $\mathcal{P}1 = 1$;

(ii) $\mathcal{P}\varphi \geq 0$ if $\varphi \geq 0$;

(iii) If a sequence $(\varphi_n)_{n\in\mathbb{N}}$ of functions in $B_b(\mathcal{H})$ converges pointwise to an element $\varphi \in B_b(\mathcal{H})$, then $\mathcal{P}\varphi_n$ converges pointwise to $\mathcal{P}\varphi$.

A measure on $\mathcal{H}$, $\mu \in \mathcal{M}(\mathcal{H})$ is called *invariant* with respect to $\mathcal{P}_t$ if $\mathcal{P}_t^*\mu = \mu$ for all $t \geq 0$, i.e.

$$\int_{\mathcal{H}} \varphi\, d\mu = \int_{\mathcal{H}} \mathcal{P}_t\varphi\, d\mu \quad \text{for all } \varphi \in B_b(\mathcal{H}),\, t \geq 0.$$

Furthermore an invariant measure μ is called *ergodic* if, any $\varphi \in L^2(\mathcal{H}, \mu)$ such that

$$\mathcal{P}_t\varphi = \varphi \quad \mu\text{-a.s.} \quad \text{for all } t > 0, \tag{3.5}$$

is μ-a.s. constant. Given the Markov semigroup $\mathcal{P}_t$, a sufficient condition to show ergodicity is the existence of a *unique* invariant measure, see for example

[21, Theorem 3.2.6]. This is the approach we will follow for the stochastic two–layer quasi–geostrophic equation by showing first the existence of an invariant measure and secondly, in section 3.2, its uniqueness. We close this introduction showing how to define the Markov semigroup associated to the solution of an autonomous stochastic differential equation.

Consider the following general setup to be later used to lay out the methodology in section 3.1.1 and section 3.2.1. Let $(\mathcal{V}, \|\cdot\|)$ be another Hilbert space with $\mathcal{V} \subset\subset \mathcal{H}$ and consider the stochastic equation on $[t_0, T]$

$$dX + AX\,dt = F(X)\,dt + dW, \quad X(t_0) = x \in \mathcal{H}, \tag{3.6}$$

where W is a Wiener process on $(\Omega, \mathcal{F}, \mathbb{P})$ with covariance operator Q trace class in $\mathcal{V}$, the operator $A : \mathcal{V} \to \mathcal{V}^*$ is invertible and nonnegative selfadjoint in $\mathcal{H}$. It follows that $A^{-1} : \mathcal{H} \to \mathcal{H}$ is compact with eigenvectors giving an orthonormal basis of $\mathcal{H}$. The mapping $F : \mathcal{V} \to \mathcal{V}^*$ is such that there exists a unique solution $X = X(t)$ for any initial condition $x \in \mathcal{H}$ and

(i) $X \in C([t_0, T]; \mathcal{H}) \cap L^2(t_0, T; \mathcal{V})$ $\mathbb{P}$-a.s. for all $T \geq 0$;

(ii) (3.6) holds as an equality in $L^2([t_0, T]; \mathcal{V}^*)$;

(iii) X is continuous with respect to the initial condition.

Since the solutions are defined over the time intervals $[t_0, T]$ for all T, by Kolmogorov extension theorem, it is possible to have the law of $\{X(t),\, t \geq 0\}$ well defined.

Given that X is continuous with respect to the initial condition, [22, Theorem 9.14] ensures that $\{X(t, t_0, x),\ t \geq t_0\}$ is a Markov process. In addition, since (3.6) is autonomous, the solution is time–homogeneous i.e.

$$\text{Law}\, X(t; t_0, x) = \text{Law}\, X(t - t_0; 0, x) \quad \text{for all } \ t_0 \leq t.$$

We can define the associated Markov transition function $P_t(x, \Gamma)$ as the probability for the process to end in the set Γ at time t when starting from x at time 0

$$P_t(x, \Gamma) = \mathbb{P}(X(t; 0, x) \in \Gamma),$$

namely the measure of Γ under the law of $X(t; 0, x)$. Furthermore since $\{X(t),\, t \geq 0\}$ is time–homogeneous

$$P_{t-t_0}(x, \Gamma) = \mathbb{P}(X(t - t_0; 0, x) \in \Gamma) = \mathbb{P}(X(t; t_0, x) \in \Gamma).$$

Then we define the associated Markov semigroup acting on $B_b(\mathcal{H})$ by

$$\mathcal{P}_t\varphi(x) := \mathbb{E}\,\varphi(X(t; 0, x)) \tag{3.7}$$

for all $t \geq 0$, $x \in \mathcal{H}$ and $\varphi \in B_b(\mathcal{H})$. By time-homogeneity we have

$$\mathbb{E}\varphi(X(t;t_0,x)) = \mathbb{E}\,\varphi(X(t-t_0;0,x)) = \mathcal{P}_{t-t_0}\varphi(x) \tag{3.8}$$

for all $t \geq t_0$, $x \in \mathcal{H}$ and $\varphi \in B_b(\mathcal{H})$. See for example to [22, Section 9] for further reference on the relations and definitions here introduced.

3.1 Existence of invariant measures

3.1.1 Methodology

A classic technique to show the existence of an invariant measure is the Krylov-Bogoliubov theorem for Markov semigroups, see [21, Section 3.1] or [39, Section 6] for reference. Its statement requires the following two concepts:

Definition 3.1.1. A Markov semigroup $\mathcal{P}_t$, $t \geq 0$ is called a *Feller* semigroup if for any $\varphi \in C_b(\mathcal{H})$ (the continuous bounded functions on $\mathcal{H}$) and any $t \geq 0$ one has $\mathcal{P}_t\varphi \in C_b(\mathcal{H})$.

Definition 3.1.2. A collection of probability measures $M \subset \mathcal{M}_1(\mathcal{H})$ is called *tight* if, for any $\varepsilon > 0$, there exists a compact set $K_\varepsilon \subset \mathcal{H}$ such that

$$\mu(K_\varepsilon) > 1-\varepsilon \quad \text{for all } \mu \in M.$$

Then the existence theorem reads as follows:

Theorem 3.1.3 (Krylov-Bogoliubov)**.** *Let $\mathcal{P}_t$, $t \geq 0$, be a Feller Markov semigroup over the space $\mathcal{H}$. Assume that there exists $\mu_0 \in \mathcal{M}_1(\mathcal{H})$ such that the sequence $\{\mathcal{P}_t^*\mu_0\}$ is tight. Then, there exists at least one invariant probability measure for $\mathcal{P}_t$.*

When the semigroup is defined by the unique solution of a stochastic equation as in (3.7), the Feller property is a direct consequence of the continuous dependence with respect to the initial conditions of such a solution:

Proposition 3.1.4. *Let $X(t,\omega;t_0,x)$ be the unique solution of* (3.6) *which depends continuously on the initial condition x. Then the associated Markov semigroup $\mathcal{P}_t$* (3.7) *is Feller.*

Proof. Let $\varphi \in C_b(\mathcal{H})$. The solution $X(t,\omega;t_0,x)$ is continuous with respect to the initial condition x then $\varphi(X(t,\omega;t_0,x))$ is continuous in x. Moreover as φ is bounded, by the bounded convergence theorem we have

$$\mathbb{E}\,\varphi(X(t,\cdot;t_0,x)) \to \mathbb{E}\,\varphi(X(t,\cdot;t_0,\tilde{x})) \quad \text{for } x \to \tilde{x}. \tag{3.9}$$

It follows that $\mathcal{P}_t\varphi$ is continuous at any $\tilde{x} \in \mathcal{H}$. □

For the two dimensional stochastic Navier–Stokes equation, the Krylov-Bogoliubov theorem is used in [32] to ensure the existence of an invariant measure when the Wiener process W is $D(A^{-\frac{1}{4}+\varepsilon})$-valued with $\varepsilon > 0$. We will mainly follow the presentation in [32], but also [20] and [21, Section 11.5] to study the stochastic two–layer quasi–geostrophic model (3.58) in section 3.1.2. For a general equation like (3.6) Theorem 3.1.5 will show the existence of an invariant measure following closely the proof of Theorem 3.3 in [32] for the stochastic Navier–Stokes equation. In order to do so we extend (3.6) to negative times.

Let $W(t)$ be a Q-Wiener process with Q trace class in $\mathcal{V}$ and let $V(t)$, $t \geq 0$ be a Wiener process with same law as $W(t)$ and independent of it. Then we define a Wiener process for $t \in \mathbb{R}$ as follows

$$\bar{W}(t) := \begin{cases} W(t) & t \geq 0 \\ V(-t) & t \leq 0 \end{cases} \tag{3.10}$$

with filtration

$$\bar{\mathcal{F}}_t = \sigma\left(\bar{W}(t_2) - \bar{W}(t_1) \,:\, t_1 \leq t_2 \leq t\right) \quad t \in \mathbb{R}. \tag{3.11}$$

Note that $\{\bar{W}(t)\}$ is *not* adapted to $\{\bar{\mathcal{F}}_t\}$ but for any $t_0 \in \mathbb{R}$, $\{\bar{W}(t) - \bar{W}(t_0),\, t \geq t_0\}$ is a martingale with respect to $\{\bar{\mathcal{F}}_t\}$.

Then, instead of (3.6), consider the general equation

$$dX + AX\,dt = F(X)\,dt + d\bar{W}, \quad X(t_0) = x \in \mathcal{H}, \tag{3.12}$$

for $t, t_0 \in \mathbb{R}$, $t_0 \leq t$.

Theorem 3.1.5. *Let X be the unique solution of* (3.12)*, $\mathcal{P}_t$, $t \in \mathbb{R}$ the associated Markov semigroup, and suppose X is time–homogeneous i.e. for all $x \in \mathcal{H}$, $t \in \mathbb{R}$*

$$\text{Law}\{X(t+s, \cdot; t, x),\, s \geq 0\} = \text{Law}\{X(s, \cdot; 0, x),\, s \geq 0\}. \tag{3.13}$$

Consider the family of solutions

$$\{X(0, \omega; -\tau, 0) \,:\, \tau \geq 0\} \tag{3.14}$$

with initial condition $X(-\tau) = 0$, and the associated family of laws $\{\vartheta_\tau \in \mathcal{M}_1 \,:\, \tau \geq 0\}$ i.e.

$$\vartheta_\tau(\Gamma) = \mathbb{P}(X(0, \omega; -\tau, 0) \in \Gamma) \tag{3.15}$$

for $\Gamma \in \mathcal{B}(\mathcal{H})$. If $\{\vartheta_\tau\}_{\tau \geq 0}$ is tight and $\mathcal{P}_t$ is Feller, then the system (3.12) *has at least an invariant measure.*

Proof. Define the family of measures $\{\mu_T\ ;\ T \geq 0\}$ as

$$\mu_T(A) = \frac{1}{T}\int_0^T \vartheta_\tau(A)\, d\tau \quad \text{for all } A \in \mathcal{B}(\mathcal{H}). \tag{3.16}$$

By the tightness of $\{\vartheta_\tau\}_{\tau\geq 0}$ for any ε there exists a compact set K_ε such that

$$\mu_T(K_\epsilon^c) = \frac{1}{T}\int_0^T \vartheta_\tau(K_\epsilon^c)\, d\tau \leq \varepsilon,$$

so that also $\{\mu_T\}_{T\geq 0}$ is tight. Then, by Prohorov theorem, there exists at least one accumulation point $\mu_* \in \mathcal{M}_1$ and a sequence μ_{t_n} converging weakly to μ_* for $t_n \to \infty$. We are left to ensure that μ_* is indeed invariant, i.e. $P_t^*\mu_* = \mu_*$ for all $t \geq 0$.

Observe that

$$\int \varphi\, d\vartheta_\tau = \int \varphi(X(0,\omega\,; -\tau, 0))\,\mathbb{P}(d\omega) = \mathbb{E}\varphi(X(0,\cdot\,; -\tau, 0))$$

and, by the homogeneity of X (3.13),

$$\mathbb{E}\varphi(X(0,\cdot\,; -\tau, 0)) = \mathbb{E}\,\varphi(X(\tau,\cdot\,; 0, 0))].$$

Moreover

$$\mathbb{E}\,\varphi(X(\tau,\cdot; 0,0)) = (\mathcal{P}_\tau\varphi)(0) = \int \mathcal{P}_\tau\varphi\, d\delta_0 = \int \varphi\, d(\mathcal{P}_\tau^*\delta_0).$$

It follows that $\vartheta_\tau = \mathcal{P}_\tau^*\delta_0$. Then by the semigroup property of $\mathcal{P}_t^*$ we have

$$\mathcal{P}_s^*\vartheta_\tau = \mathcal{P}_s^*\mathcal{P}_\tau^*\delta_0 = \mathcal{P}_{\tau+s}^*\delta_0 = \vartheta_{\tau+s}. \tag{3.17}$$

By the definition of μ_T and (3.17) we have

$$\mathcal{P}_s^*\mu_{t_n} = \frac{1}{t_n}\int_0^{t_n} \mathcal{P}_s^*\vartheta_\tau\, d\tau = \frac{1}{t_n}\int_0^{t_n} \vartheta_{\tau+s}\, d\tau,$$

so that

$$\mathcal{P}_s^*\mu_{t_n} = \frac{1}{t_n}\int_0^{t_n} \vartheta_\tau\, d\tau + \frac{1}{t_n}\int_{t_n-s}^{t_n} \vartheta_{\tau+s} d\tau - \frac{1}{t_n}\int_{-s}^{0} \vartheta_{\tau+s} d\tau.$$

Taking the weak-$\star$ limit on both sides for $n \to \infty$ we get

$$\lim_{n\to\infty} \mathcal{P}_s^*\mu_{t_n} = \lim_{n\to\infty} \mu_{t_n} = \mu_*.$$

From the Feller property of $\mathcal{P}_t$, it follows that $\mathcal{P}_t^*$ is continuous in the weak-$\star$ topology of $\mathcal{M}_1$ i.e. if $\mu_n \stackrel{\star}{\to} \mu$ for $n \to \infty$ then $\mathcal{P}_t^* \mu_n \stackrel{\star}{\to} \mathcal{P}_t^* \mu$, meaning

$$\int \mathcal{P}_t \varphi \, d\mu_n \to \int \mathcal{P}_t \varphi \, d\mu_* \quad \forall \varphi \in C_b(\mathcal{H}).$$

Since $\mathcal{P}_t$ is Feller, $\mathcal{P}_t \varphi$ is also continuous and bounded, and the convergence in weak-$\star$ topology of μ_{t_n} against μ concludes the argument.

In conclusion,

$$\mathcal{P}_s^* \mu = \mathcal{P}_s^* (\lim_{n\to\infty} \mu_{t_n}) = \lim_{n\to\infty} \mathcal{P}_s^* \mu_{t_n} = \mu_*$$

as desired. □

Recall the setup of (3.12) which was introduced in connection with (3.6). In particular we have two Hilbert spaces $(\mathcal{V}, \|\cdot\|)$ and $(\mathcal{H}, |\cdot|)$ with $\mathcal{V}$ dense and compactly contained in $\mathcal{H}$ so that the bounded sets in $\mathcal{V}$ are a class of compact subsets of $\mathcal{H}$. We will exploit this useful fact in Lemma 3.1.6 which gives a verifiable sufficient condition to prove tightness of the family $\{\vartheta_\tau\}$.

Lemma 3.1.6. *Let X be the unique solution of* (3.12) *and $\{\vartheta_\tau \in \mathcal{M}_1 : \tau \geq 0\}$ the family of laws as in* (3.15). *If there exists an a.s. finite random variable $R(\omega)$ such that*

$$\sup_{\tau\geq 0} \|X(0,\omega;-\tau,0)\| \leq R(\omega), \tag{3.18}$$

then the family of measures $\{\vartheta_\tau\}_{\tau\geq 0}$ is tight in $\mathcal{H}$

Proof. Given $\varepsilon > 0$ and a positive constant r_ε define the set

$$K_\varepsilon = \{x \in \mathcal{V} \, : \, \|x\| \leq r_\varepsilon\}$$

which is bounded in $\mathcal{V}$, hence compact in $\mathcal{H}$ since $\mathcal{V}$ is assumed compactly contained in $\mathcal{H}$. Then

$$\vartheta_\tau(K_\varepsilon^c) = \mathbb{P}\left(\|X(0,\omega;-\tau,x)\| > r_\varepsilon \right) \leq \mathbb{P}\left(\sup_{\tau\geq 0} \|X(0,\omega;-\tau,x)\| > r_\varepsilon \right).$$

If (3.18) holds, then

$$\vartheta_\tau(K_\varepsilon^c) \leq \mathbb{P}(R(\omega) > r_\varepsilon).$$

As R is almost surely finite, for every ε we can find an r_ε large enough such that

$$\mathbb{P}(R(\omega) > r_\varepsilon) < \varepsilon,$$

hence the result. □

From the proof of Theorem 3.1.5 it is easy see that it would be enough to have the family of laws of $X(\tau,\omega;0,0)$ being tight rather than those of $X(0,\omega;-\tau,0)$ and from Lemma 3.1.6 it would be enough to show that there exists a random variable R such that

$$\sup_{\tau\geq 0} X(\tau,\omega;0,0) \leq R(\omega)$$

However it is not clear if this condition holds as in fact on the limit for large τ the solution is expected to approach a stationary process, if ergodic. On the other hand considering the final time fixed and pulling the initial condition to $-\infty$ the solution will converge to the value of the stationary process at such fixed final time.

In the rest of the section we will show that the quasi–geostrophic model satisfies the hypothesis of Theorem 3.1.5.

3.1.2 Stochastic two–layer quasi–geostrophic model

Given an arbitrary $t_0 \in \mathbb{R}$ and $t \geq t_0$ then $\mathbf{q}(t,\omega;t_0,\mathbf{q}_0)$ is the unique (weak) solution of

$$\begin{aligned} &d\mathbf{q} + B(\boldsymbol{\psi},\boldsymbol{\psi}) + \beta\partial_1\boldsymbol{\psi} = \nu\Delta^2\boldsymbol{\psi}\, dt + \begin{pmatrix} f \\ -r\Delta\psi_2 \end{pmatrix} + d\begin{pmatrix} \bar{W} \\ 0 \end{pmatrix} \\ &\mathbf{q} = (\Delta + M)\boldsymbol{\psi} \\ &\mathbf{q}(t_0) = \mathbf{q}_0 \in \mathcal{H} \end{aligned} \tag{3.19}$$

with deterministic and stochastic forcing now more regular, i.e. $f \in H^{-1}$ and $\bar{W}$ being a Q-Wiener process, as in (3.10), with Q trace class operator in H^1. Then we define the associated Markov semigroup acting on the family of bounded measurable functions $\varphi : \mathcal{H} \to \mathbb{R}$ as in (3.7) i.e.

$$\mathcal{P}_{t-t_0}\varphi(\mathbf{q}_0) = \mathbb{E}\,\varphi(\mathbf{q}(t;t_0,\mathbf{q}_0)).$$

Consider the auxiliary Ornstein-Uhlenbeck equation (1.73) now for all $t \in \mathbb{R}$

$$d\eta + \alpha_1 A\eta = d\bar{W}. \tag{3.20}$$

Let $\eta_\infty(t)$, $t \in \mathbb{R}$ be the stationary solution of (3.20) for which we have the following explicit formulation

$$\eta_\infty(t) = \int_{-\infty}^{t} \exp(-\alpha_1(t-s)A)\, d\bar{W}(s). \tag{3.21}$$

Given the properties seen in section 1.4.2, $\eta_\infty(t)$ is ergodic, has trajectories continuous with values in H^1 (Theorem 1.4.2) and, taking the limit of initial time $t_0 \to -\infty$ in (1.76), is such that

$$\mathbb{E}|\eta_\infty(t)|^2 \leq \frac{\operatorname{Tr}_0 Q}{2\alpha_1\lambda_1}. \tag{3.22}$$

With a careful adaptations of the proofs seen in section 2.2 it can be shown that $\tilde{\mathbf{q}} + (\eta_\infty, 0)^t$ is solution of (3.19) and has the same properties as seen in Theorem 2.2.1. Then the Feller property is a direct consequence of the continuity with respect to initial conditions of the solution $\mathbf{q}$.

Making use of Lemma 3.1.6 we will show that the family of solutions (3.14) is bounded uniformly in the initial time, so to ensure the family of their laws $\{\vartheta_\tau\}_\tau$ is tight. The statement is a parallel version of Theorem 3.2 in [32] adapted and proved for our model of interest instead of 2D Stochastic Navier Stokes.

Theorem 3.1.7. *Let* $\mathbf{q} = \mathbf{q}(0, \omega; t_0, 0)$ *be solution over* $[t_0, 0]$ *of* (3.19) *with initial condition* $\mathbf{q}(t_0) = 0$*. Then the family of their laws* $\{\vartheta_{t_0} : t_0 \leq 0\}$ *is tight in* $\mathcal{H}$*.*

Proof. Recall that given

$$\mathcal{H} = (\mathbb{H}^{-1}, |||\cdot|||_{-1}) \quad \text{and} \quad \mathcal{V} = (\mathbb{L}^2, |||\cdot|||_0)$$

we have in particular $\mathcal{V} \subset\subset \mathcal{H}$ so that a closed set in $\mathcal{V}$ is compact in $\mathcal{H}$. Therefore, given Lemma 3.1.6, we want to show that

$$\sup_{t_0 \leq 0} |||\mathbf{q}(0, \omega; t_0, 0)|||_0 \leq R(\omega). \tag{3.23}$$

for an appropriate $R(\omega)$ random variable almost surely finite. Given $\boldsymbol{\eta}_\infty = (\eta_\infty, 0)^t$, let $\tilde{\mathbf{q}} = \mathbf{q} - \boldsymbol{\eta}_\infty$ be the solution of the random system associated to (3.19) in the time interval $[t_0, 0]$ with initial condition $\tilde{\mathbf{q}}(t_0) = -\boldsymbol{\eta}_\infty(t_0)$. To simplify the notation we write $\tilde{\mathbf{q}}(t, t_0)$ for $\tilde{\mathbf{q}}(t, \omega; t_0, -\boldsymbol{\eta}_\infty(t_0))$.

Consider the *a priori* bound in $(\mathbb{L}^2, |||\cdot|||_0)$ as in (2.40) over the time interval $[r, 0]$, with $t_0 \leq r \leq 0$,

$$|||\tilde{\mathbf{q}}(0, t_0)|||_0^2 \leq |||\tilde{\mathbf{q}}(r, t_0)|||_0^2 \exp\left(\int_r^0 g_1(s, t_0)\, ds\right) \\ + \int_r^0 \exp\left(\int_t^0 g_1(s, t_0)\, ds\right) g_2(t, t_0) dt$$

with

$$g_1(t,t_0) := c_2 + 2c_1F_1|\eta_\infty(t)| + c_6|\eta_\infty(t)|^2 + c_5\lambda_1\|\eta_\infty(t)\|^2 + c_3|||\tilde{\mathbf{q}}(t,t_0)|||_{-1}^2, \tag{3.24}$$

$$g_2(t,t_0) := c_4\|\eta_\infty(t)\|^2 + c_5\|\eta_\infty(t)\|^4 + 2\beta^2|||\tilde{\mathbf{q}}(t,t_0)|||_{-1}^2 + \tfrac{4h_1}{\nu}\|f\|_{-1}^2. \tag{3.25}$$

Integrate r over $[-1,0]$ to get

$$|||\tilde{\mathbf{q}}(0,t_0)|||_0^2 \leq (A)\cdot(B) + (C) \tag{3.26}$$

with

$$(A) := \exp\left(\int_{-1}^{0} g_1(t,t_0)\,dt\right) \tag{3.27}$$

$$(B) := \int_{-1}^{0} |||\tilde{\mathbf{q}}(r,t_0)|||_0^2\,dr \tag{3.28}$$

$$(C) := \int_{-1}^{0} \exp\left(\int_{r}^{0} g_1(s,t_0)\,ds\right) g_2(r,t_0)\,dr. \tag{3.29}$$

We now want to bound the right hand side of (3.26) with an a.s. finite random variable uniformly in t_0 and we will do it in several steps, starting with the exponential term (A).

Bound for (A). By the definition of the function $g_1(t,t_0)$ we want to find a bound uniform in t_0 for

$$\exp\left(c_2 + \int_{-1}^{0} 2c_1F_1|\eta_\infty(t)| + c_6|\eta_\infty(t)|^2 + c_5\lambda_1\|\eta_\infty(t)\|^2 + c_3|||\tilde{\mathbf{q}}(t,t_0)|||_{-1}^2\,dt\right). \tag{3.30}$$

The terms involving η_∞, namely

$$2c_1F_1\int_{-1}^{0}|\eta_\infty(t)|\,dt + c_6\int_{-1}^{0}|\eta_\infty(t)|^2\,dt + c_5\lambda_1\int_{-1}^{0}\|\eta_\infty(t)\|^2\,dt$$

are independent of t_0 so we only have to know that they are well defined, which is true since η_∞ is continuous with values in H^1. Therefore we are left to show that there exists an almost surely finite random variable $R(\omega)$ such that

$$\int_{-1}^{0} |||\tilde{\mathbf{q}}(t,t_0)|||_{-1}^2\,dt < R \quad \text{for all } t_0 < 0.$$

Consider the *a priori* bound (2.14) for $\tilde{\mathbf{q}}$ in $\mathcal{H}$, for $t_0 \le t \le 0$, i.e.

$$\begin{aligned}|||\tilde{\mathbf{q}}(t,t_0)|||_{-1}^2 \le |||\tilde{\mathbf{q}}(t_0)|||_{-1}^2 \exp\Big(\int_{t_0}^t -\frac{\nu\lambda_1}{a_0} + d_0|\eta_\infty(u,\omega)|^2\,du\Big) \\ + \int_{t_0}^t \exp\Big(\int_s^t -\frac{\nu\lambda_1}{a_0} + d_0|\eta_\infty(u,\omega)|^2\,du\Big) m(s,\omega)\,ds. \quad (3.31)\end{aligned}$$

with $m(t,\omega)$ given by (2.12), namely

$$m(t,\omega) = d_1|\eta_\infty(t,\omega)|^4 + d_2|\eta_\infty(t,\omega)|^2 + d_3, \quad d_1,d_2,d_3 \ge 0,$$

and let us first focus on the integrating factor. We assumed the Ornstein-Uhlenbeck process η_∞ to be an ergodic stationary solution of (3.21) continuous with values in H^1, hence Birkhoff ergodic theorem ensures that

$$\lim_{s\to-\infty} \frac{1}{|s|}\int_s^0 |\eta_\infty(r,\omega)|^2\,dr = \mathbb{E}\,|\eta_\infty(0)|^2 \quad \mathbb{P}\text{-a.a. } \omega. \tag{3.32}$$

By the estimate (3.22) for $t=0$ we have

$$\mathbb{E}\,|\eta_\infty(0)|^2 \le \frac{\operatorname{Tr} Q}{2\alpha_1\lambda_1},$$

so that, for values of the control parameter α large enough, $\mathbb{E}|\eta_\infty(0)|^2$ can be made arbitrarily small and in particular, we pick α such that

$$d_0\mathbb{E}|\eta_\infty(0)|^2 \le \frac{\nu\lambda_1}{2a_0}.$$

Consequently, with such a choice of α, we get

$$\lim_{s\to-\infty} \frac{1}{|s|}\int_s^0 -\frac{\nu\lambda_1}{a_0} + d_0|\eta_\infty(r,\omega)|^2\,dr = -\frac{\nu\lambda_1}{a_0} + d_0\mathbb{E}\,|\eta_\infty(0)|^2 \le -\frac{\nu\lambda_1}{2a_0}$$

and in particular there exists a random time $\tau(\omega) < 0$ such that, for all $s < \tau(\omega)$

$$\int_s^0 -\frac{\nu\lambda_1}{a_0} + d_0|\eta_\infty(r,\omega)|^2\,dr \le -\frac{\nu\lambda_1|s|}{2a_0}. \tag{3.33}$$

By continuity of the trajectories of the process η_∞, we can find a bound for the remaining part of the interval, $[\tau(\omega),0]$, meaning that there is an a.s.-finite random variable $C(\omega)$ such that

$$\sup_{\tau(\omega)\le s\le 0} \int_s^0 -\frac{\nu\lambda_1}{a_0} + d_0|\eta_\infty(r,\omega)|^2\,dr \le C(\omega). \tag{3.34}$$

Next, given $\kappa = (\nu\lambda_1)/(2a_0)$, define the process

$$\zeta(s,\omega) = \begin{cases} -\kappa|s| & s < \tau(\omega) \\ C(\omega) & \tau(\omega) \le s \le 0 \end{cases} \tag{3.35}$$

so that for all $t \in [\tau(\omega), 0]$ we have the following estimate

$$\int_s^t -\frac{\nu\lambda_1}{a_0} + d_0|\eta_\infty(u,\omega)|^2\,du \le \zeta(s,\omega).$$

Using this estimate in (3.31), we have that for all $t \in [\tau(\omega), 0]$ and $t_0 \le t$,

$$|||\tilde{\mathbf{q}}(t)|||_{-1}^2 \le e^{\zeta(t_0,\omega)}|||\tilde{\mathbf{q}}(t_0)|||_{-1}^2 + \int_{t_0}^t e^{\zeta(s,\omega)} m(s,\omega)\,ds. \tag{3.36}$$

Given the initial condition $\tilde{\mathbf{q}}(t_0) = \boldsymbol{\eta}_\infty(t_0)$ and the definition of $\zeta(t_0)$ we have

$$e^{\zeta(t_0)}|||\tilde{\mathbf{q}}(t_0)|||_{-1}^2 \le \begin{cases} ke^{C(\omega)}\|\eta_\infty(t_0)\|_{-1}^2 & \tau \le t_0 \le 0 \\ ke^{-\kappa|t_0|}\|\eta_\infty(t_0)\|_{-1}^2 & t_0 < \tau \end{cases} \tag{3.37}$$

where we have used that $|||\cdot|||_{-1}$ is equivalent to $\|\cdot\|_{-1}$.

As η_∞ is a continuous stationary Gaussian process and $L^2 \subset H^{-1}$, the following lemma provides a useful bound.

Lemma 3.1.8 ([21, Lemma 15.4.4])**.** *Let Z be a continuous stationary Gaussian process on a separable Banach space U. Then for arbitrary $\delta > 0$ there exists a random variable Y_δ such that $\mathbb{P}$-a.s.*

$$\|Z(t)\|_U \le Y_\delta(1 + |t|^\delta) \tag{3.38}$$

for all $t \le 0$.

In particular there exists a random variable Y such that

$$|\eta_\infty(t_0)|^2 \le Y(1 + |t_0|)^2.$$

Then, using this result in (3.37), we see that

$$\sup_{t_0\le 0} e^{\zeta(t_0)}|||\tilde{\mathbf{q}}(t_0)|||_{-1}^2 \le Ye^C(1 + |\tau|) + Y\sup_{t_0\le\tau} e^{-\kappa|t_0|}(1 + |t_0|) =: M.$$

We have hence found a random variable almost surely finite that bounds the first term in (3.36), uniformly in the initial time t_0.

It is then left to ensure that also the integral on the right hand side of (3.36) can be bounded uniformly in t_0. Recall the definition of $m =$

$d_1|\eta_\infty|^4 + d_2|\eta_\infty|^2 + d_3$. Then by Lemma 3.1.8 there exists a random variable $Y(\omega)$ such that $\mathbb{P}$-a.s.

$$|\eta_\infty(t)| \le Y(1+|t|) \quad \forall t \le 0. \tag{3.39}$$

Hence relabelling Y appropriately

$$\int_{t_0}^{t} e^{\zeta(s,\omega)} d_2|\eta_\infty|^2 \, ds \le Y(\omega) \int_{t_0}^{t} e^{\zeta(s,\omega)} d_2(1+|s|)^2 \, ds \tag{3.40}$$

In the case $t_0 \ge \tau(\omega)$, then $\zeta(s,\omega) = C(\omega)$ for all $t_0 \le s \le t$ and so

$$\begin{aligned} \int_{t_0}^{t} e^{\zeta(s,\omega)}(1+|s|)^2 \, ds &= e^{C(\omega)} \int_{t_0}^{t} (1+|s|)^2 \, ds \\ &\le e^{C(\omega)} \int_{\tau(\omega)}^{0} (1+|s|)^2 \, ds =: R_1(\omega). \end{aligned}$$

Conversely, when $t_0 < \tau(\omega)$, for $s \le \tau(\omega)$ we have $\zeta(s,\omega) = \kappa s$ and therefore

$$\int_{t_0}^{t} e^{\zeta(s,\omega)}(1+|s|)^2 \, ds = \underbrace{\int_{t_0}^{\tau(\omega)} e^{\kappa s}(1+|s|)^2 \, ds}_{(I)} + \underbrace{e^{C(\omega)} \int_{\tau(\omega)}^{t} (1+|s|)^2 \, ds}_{(II)}.$$

Then since

$$\sup_{t_0 \le 0} \int_{t_0}^{\tau(\omega)} e^{\kappa s}(1+|s|)^2 \, ds < \infty$$

there exists an almost surely finite random variable R_2 such that

$$\sup_{t_0 \le 0} \int_{t_0}^{t} e^{\zeta(s,\omega)}(1+|s|)^2 \, ds < R_2(\omega).$$

Therefore, going back to (3.40), we have shown that there exist a random variable $R_3(\omega)$, a.s. finite, such that

$$\int_{t_0}^{t} e^{\zeta(s,\omega)} d_2|\eta_\infty|^2 \, ds \le R_3(\omega) \tag{3.41}$$

for all $t \in [\tau(\omega), 0]$ and all $t_0 \le t$.

Similar calculations, using again Lemma 3.1.8 to get

$$|\eta_\infty(t)|^4 \le Y_1(1+|t|)^4 \quad \forall t \le 0,$$

will give a bound also for $\int_{t_0}^{t} e^{\zeta(s,\omega)+C(\omega)} d_1|\eta_\infty|^4 \, ds$.

Finally we can conclude that there exists an a.s. finite random variable $R_4(\omega)$ such that

$$|||\tilde{\mathbf{q}}(t,\omega;t_0,\mathbf{q}_0)|||_{-1}^2 \leq R_4(\omega)$$

for all $t \in [\tau(\omega),0]$ and $t_0 \leq t$. Assume without loss of generality that $\tau(\omega) \leq -1$, then,

$$|||\tilde{\mathbf{q}}(t,\omega;t_0,\mathbf{q}_0)|||_{-1}^2 \leq R_4(\omega) \quad \text{for all } t \in [-1,0],\, t_0 \leq t. \tag{3.42}$$

Finally looking back at the entire term (A) in (3.30)

$$(A) = \exp\left(c_2 + \int_{-1}^0 2c_1F_1|\eta_\infty(t)| + c_6|\eta_\infty(t)|^2 + c_5\lambda_1\|\eta_\infty(t)\|^2 + c_3|||\tilde{\mathbf{q}}(t,t_0)|||_{-1}^2\,dt\right)$$

by the estimate (3.42) just obtained, it can be bounded as follows

$$\begin{aligned}(A) &\leq \exp\left(c_2 + \int_{-1}^0 2c_1F_1|\eta_\infty(t)| + c_6|\eta_\infty(t)|^2 + c_5\lambda_1\|\eta_\infty(t)\|^2\,dt + c_3R_4\right)\\ &=: R_5 \end{aligned} \tag{3.43}$$

Bound for (B)**.** Next to find a bound for (B) defined in (3.28) we need to show that

$$\int_{-1}^0 |||\tilde{\mathbf{q}}(r,t_0)|||_0^2\,dr$$

is bounded uniformly in t_0.

From the *a priori* bound (2.16) for $\tilde{\mathbf{q}}$ in $L^2(0,T;\mathcal{V})$ we have that

$$\begin{aligned}\int_{-1}^0 |||\tilde{\mathbf{q}}(r,t_0)|||_0^2\,dr &\leq \nu^{-1}|||\tilde{\mathbf{q}}(-1,t_0)|||_{-1}^2 + \int_{-1}^0 \nu^{-1}d_0m(r)\,dr\\ &\int_{-1}^0 \left(\nu^{-1}d_0|\eta_\infty(r)|^2 + 2F_1\right)|||\tilde{\mathbf{q}}(r,t_0)|||_{-1}^2\,dr.\end{aligned} \tag{3.44}$$

By the estimate (3.42) obtained for $|||\tilde{\mathbf{q}}(r,t_0)|||_{-1}^2$ in the previous step for all $r \in [-1,0]$, we have

$$(B) \leq \frac{R_4}{\nu} + \frac{R_4}{\nu}\int_{-1}^0 (d_0|\eta_\infty(r)|^2 + 2F_1)\,dr + \frac{d_0}{\nu}\int_{-1}^0 m(r)\,dr.$$

Bound for (C). Let (C) be as in (3.29) i.e.

$$(C) := \int_{-1}^0 \exp\left(\int_r^0 g_1(s,t_0)ds\right)g_2(r,t_0)\,dr$$

and g_2 as in (3.25), i.e.

$$c_4\|\eta_\infty(t)\|^2 + c_5\|\eta_\infty(t)\|^4 + 2\beta^2 |||\tilde{\mathbf{q}}(t,t_0)|||_{-1}^2 + \tfrac{4h_1}{\nu}\|f\|_{-1}^2.$$

Since $r \in [-1,0]$, by the estimate (3.43) we have that

$$\exp\left(\int_r^0 g_1(s,t_0)\,ds\right) \le \exp\left(\int_{-1}^0 g_1(s,t_0)\,ds\right) \le R_5$$

so that, again by the estimate (3.42) for $|||\tilde{\mathbf{q}}(t,t_0)|||_{-1}^2$

$$2\beta^2 \int_{-1}^0 \exp\left(\int_r^0 g_1(s,t_0)ds\right) |||\tilde{\mathbf{q}}(t,t_0)|||_{-1}^2\,dr \le 2\beta^2 R_5 R_4.$$

Then

$$(C) \le R_5 \int_{-1}^0 c_4\|\eta_\infty(t)\|^2 + c_5\|\eta_\infty(t)\|^4\,dt + 2\beta^2 R_5 R_4 + \tfrac{4h_1}{\nu}\|f\|_{-1}^2 R_5 =: R_6$$

and the right hand side is a well defined almost surely finite random variable since η_∞ is continuous with values in H^1.

Finally putting together the estimate for the expressions (A), (B) and (C) we have

$$\sup_{t_0\le 0} |||\tilde{\mathbf{q}}(0,\omega;t_0,\tilde{\mathbf{q}}_0)|||_0^2 \le R_7(\omega),$$

where $R_7(\omega)$ is an a.s. finite random variable. Since $\mathbf{q}(0,t_0) = \tilde{\mathbf{q}}(0,t_0) + \boldsymbol{\eta}_\infty(0)$ and $|||\boldsymbol{\eta}_\infty|||_0^2 \le |\boldsymbol{\eta}_\infty|^2 = h_1|\eta_\infty|^2$, we have

$$|||\mathbf{q}(0,t_0)|||_0^2 \le |||\tilde{\mathbf{q}}(0,t_0)|||_0^2 + h_1|\eta_\infty(0)|^2 \le R_7 + h_1|\eta_\infty(0)|^2 =: R_8 \qquad (3.45)$$

for all $t_0 \le 0$, as desired.

□

We have then shown that the family of measures on $\mathcal{H}$

$$\vartheta_{t_0} = \text{Law}(\mathbf{q}(0,\omega;t_0,0)), \quad t_0 \le 0$$

is tight, and consequently we have the following result.

Theorem 3.1.9. *The stochastic two–layer quasi–geostrophic model* (3.19) *has an invariant measure.*

Proof. First of all $\mathbf{q}(t)$ is time–homogeneous: in fact η_∞ is stationary and the forcing f is time–independent, hence $\tilde{\mathbf{q}}$ is time–homogeneous, and consequently $\tilde{\mathbf{q}} + \boldsymbol{\eta}_\infty$ is.

Next, Theorem 3.1.7 ensured tightness of the family of measures $\{\vartheta_{t_0}\}$. Finally, $\mathcal{P}_t$ is Feller as $\mathbf{q}$ is continuous with respect to its initial condition. Then Theorem 3.1.5 ensures the desired result. □

3.2 Uniqueness of the invariant measure

3.2.1 Methodology

A classic approach to establish unique ergodicity of a Markov process is based on *Doob's theorem* (see for example [21, Section 4.2]):

Theorem 3.2.1 (Doob)**.** *Let $\mathcal{P}_t$ be a stochastically continuous Markov semigroup with invariant measure μ. If $\mathcal{P}_t$ is t_0-regular for some $t_0 > 0$, i.e. all transition probabilities $P_{t_0}(x,\cdot)$ are mutually equivalent, then μ is unique, strongly mixing and equivalent to all measures $P_t(x,\cdot)$ for all $t > t_0$.*

In particular, if the Markov semigroup is *irreducible* at some time s_0, i.e. for any nonempty open set Γ and all $x \in \mathcal{H}$, $P_{s_0}(x,\Gamma) > 0$, and *strong Feller* at some time t_0, i.e. for any φ measurable bounded function, $\mathcal{P}_{t_0}\varphi$ is continuous and bounded, then $\mathcal{P}_t$ can be proved to be $(t_0 + s_0)$-regular (see [21]). This was the approach used for example in [33, 30] to prove the ergodicity of stochastic Navier–Stokes equations for non degenerate noise.

However for SPDEs the strong Feller property fails to hold in cases where the noise is spatially degenerate. To cope with this problem the concept of *asymptotic strong Feller* was introduced in [41], providing a comprehensive approach for the 2D Navier Stokes equations forced only in a four dimensional subspace.

Alternatively, for moderately degenerate noise, the asymptotic (or generalised) coupling method has proven successful. Introduced in the early 2000s [48, 10, 29], it was used under minimal assumptions to prove unique ergodicity for Markov operators on Polish spaces in [44] and applied to several nonlinear SPDEs in [37]. The main idea of this approach is to add a control in the stochastic forcing to synchronize solutions with different initial data. In finite dimensions, Girsanov's theorem ensures that the controlled equation and the original equation generate equivalent distributions. In the infinite dimensional case, an appropriate finite dimensional control on the unstable degrees of freedom is often sufficient to ensure synchronisation and permits application of the classical Girsanov theorem.

We present such a method following closely [37] and [44]. Let $\mathcal{H}$ be an Hilbert space with norm $|\cdot|$ and consider two probability measures μ_1 and μ_2 on it. A probability measure Γ on $\mathcal{H}\times\mathcal{H}$ is called *coupling* of μ_1, μ_2 if, given $\Pi_1(x,y) = x$, $\Pi_2(x,y) = y$ the projections of $\mathcal{H}\times\mathcal{H}$ onto its two components,

$$\Gamma\Pi_1^{-1} = \mu_1 \quad \text{and} \quad \Gamma\Pi_2^{-1} = \mu_2.$$

We denote the set of all such couplings as $\mathcal{C}(\mu_1,\mu_2)$. Equivalently we call a pair of random variables (ξ_1,ξ_2) a coupling of μ_1,μ_2 if $\operatorname{Law}\xi_1 = \mu_1$ and $\operatorname{Law}\xi_2 = \mu_2$.

A weaker notion of coupling is that of *equivalent coupling*, namely when we require the marginals to be only absolutely continuous with respect to the target measures rather than equal:

Definition 3.2.2. Given two measures μ_1, μ_2 on $\mathcal{H}$, a probability measure Γ on $\mathcal{H} \times \mathcal{H}$ is an *equivalent coupling* of μ_1 and μ_2 if $\Gamma\Pi_i^{-1} \ll \mu_i$, $i = 1, 2$.

Let $\mathcal{P}$ be a Markov operator on $\mathcal{H}$. We denote as $\mathcal{H}^{\mathbb{N}}$ the pathspace representing trajectories of $\mathcal{P}$ over $\mathcal{H}$. For any measure μ on $\mathcal{H}$, we denote as $\mu\mathcal{P}^{\mathbb{N}}$ the suspension of μ to $\mathcal{H}^{\mathbb{N}}$. Then it can be shown (see e.g. [21]) that μ being ergodic on $\mathcal{H}$ for $\mathcal{P}$ is equivalent to $\mu\mathcal{P}^{\mathbb{N}}$ being ergodic on $\mathcal{H}^{\mathbb{N}}$ for the classic shift map.

Consider now two measures μ_1, μ_2 on the pathspace $\mathcal{H}^{\mathbb{N}}$. Then a probability measure on $\mathcal{H}^{\mathbb{N}} \times \mathcal{H}^{\mathbb{N}}$ is called *asymptotic coupling* when it satisfies the following definition:

Definition 3.2.3. The diagonal at infinity D are the sequences elements of $\mathcal{H}^{\mathbb{N}}$ which converge to each other, i.e.

$$D := \{(u, v) \in \mathcal{H}^{\mathbb{N}} \times \mathcal{H}^{\mathbb{N}} : \lim_{n\to\infty} |u_n - v_n| = 0\}. \tag{3.46}$$

Then given two measures μ_1, μ_2 on $\mathcal{H}^{\mathbb{N}}$, a coupling $\Gamma \in \mathcal{C}(\mu_1, \mu_2)$ is called *asymptotic coupling* if $\Gamma(D) = 1$.

It can be proven ([44, Theorem 1.1]) that given two ergodic invariant measures μ_1, μ_2 for a Markov operator $\mathcal{P}$, they are equal if there exists an *asymptotic coupling* Γ of $\mu_1\mathcal{P}^{\mathbb{N}}$ and $\mu_2\mathcal{P}^{\mathbb{N}}$.

Actually a seemingly weaker condition is enough to give uniqueness of the invariant measure for $\mathcal{P}$, for which we use the notion of equivalent coupling instead.

Theorem 3.2.4 ([37, Corollary 2.2],[44, Theorem 1.1])**.** *Let D be the diagonal at infinity. If for any $u_0, v_0 \in \mathcal{H}$ there exists an equivalent coupling Γ of $\delta_{u_0}\mathcal{P}^{\mathbb{N}}$ and $\delta_{v_0}\mathcal{P}^{\mathbb{N}}$ such that $\Gamma(D) > 0$, then there exists at most one ergodic invariant measure for $\mathcal{P}$ on $\mathcal{H}$.*

We omit the proof of this result which can be found in the references above, and we continue focusing on how we can apply this method to stochastic differential equations in Hilbert spaces. As previously discussed, we will use a control to synchronize solutions with different initial condition, but then the controlled system loses its connection to the uncontrolled one. Girsanov theorem provides such a connection but only ensuring that the controlled process has at most *equivalent* distribution to the original. Nevertheless as we will see in Theorem 3.2.8, this equivalence will be enough to ensure uniqueness of the invariant measure thanks to Theorem 3.2.4. Let us first recall the precise formulation of Girsanov theorem. We refer for example to [47, Section 3.5] for a proof.

Theorem 3.2.5 (Girsanov)**.** *Given the probability space* $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$*, let* $(B(t))_{t\in[0,\infty)}$ *be a* n*-dimensional Wiener process with covariance matrix* Q *and* $(h(t))_{t\in[0,\infty)}$ *a* n*-dimensional stochastically integrable process (see section 1.1). Assume the process*

$$Z^h(t) := \exp\left(-\int_0^t Q^{-1}h(s)\, dB(s) - \tfrac{1}{2}\int_0^t |Q^{-1/2}h(s)|^2\, ds\right) \tag{3.47}$$

is a $\mathbb{P}$*-martingale, and for each* $T \geq 0$ *define the probability measure* $\tilde{\mathbb{P}}_T$ *on* $\mathcal{F}_T$ *by*

$$\tilde{\mathbb{P}}_T(A) := \mathbb{E}\left[\mathbb{1}_A Z^h(T)\right], \quad \text{for all } A \in \mathcal{F}_T.$$

Define the Itô process $\tilde{B}(t)$*,* $t \geq 0$*, with filtration* $\{\mathcal{F}_t\}_t$ *by*

$$\tilde{B}(t) := B(t) + \int_0^t h(s)\, ds, \quad t \in [0,\infty) \tag{3.48}$$

Then for each fixed $T \in [0,\infty)$*, the process* $(\tilde{B}(t))_{t\in[0,T]}$ *is a* n*-dimensional Wiener process on* $(\Omega, \mathcal{F}_T, \tilde{\mathbb{P}}_T)$*.*

A classical sufficient, but not necessary, condition for Theorem 3.2.5 to apply, namely for (3.47) to be a $\mathbb{P}$-martingale, is given by the following result:

Lemma 3.2.6 ([64, Theorem IV.37.8])**.** *Let* $(h(t))_{t\in[0,\infty)}$ *be a* n*-dimensional process adapted to the filtration generated by the Wiener process* B *with covariance matrix* Q*. Suppose that for each* $t \geq 0$ *there exists a constant* K_t *such that:*

$$\int_0^t |Q^{-1/2}h(s)|^2\, ds < K_t \quad \mathbb{P}\text{-a.s.} \tag{3.49}$$

Then the process

$$Z^h(t) := \exp\left(-\int_0^t Q^{-1}h(s)\, dB(s) - \tfrac{1}{2}\int_0^t |Q^{-1/2}h(s)|^2\, ds\right)$$

is a $\mathbb{P}$*-martingale.*

Therefore, under Theorem 3.2.5 the laws of B and $\tilde{B}$, denoted μ_B and $\mu_{\tilde{B}}$, are equivalent measures on $C([0,T],\mathbb{R}^n)$ for all compact intervals $[0,T]$, $T > 0$ and we have a precise formulation of the Radon-Nikodyn density, which is the exponential martingale $(Z_t^h)_{t\in[0,T]}$. However, in some of applications which follow, we would like the process $\tilde{B}$ to be defined *for all* $t \in [0,\infty)$. There exists a version of Girsanov theorem on $C([0,\infty),\mathbb{R}^n)$ for drifts h which are progressively measurable, see for example [47, Corollary 3.5.2]. Then without further conditions, the probabilities $\mathbb{P}$ and $\tilde{\mathbb{P}}$ are mutually absolutely continuous when restricted to $\mathcal{F}_T^B$, $T \in [0,\infty)$. On the other hand, when we do not require explicit knowledge of the density, the following weaker formulation of Girsanov theorem is enough to ensure absolute continuity of $\mu_{\tilde{B}}$ with respect to μ_B.

Theorem 3.2.7 ([51, Theorem 7.4])**.** *Given the probability space with filtration $(\Omega,\mathcal{F},\mathbb{P},\mathcal{F}_t)$, let $(B(t))_{t\in[0,\infty)}$ be a n-dimensional Wiener process with convariance matrix Q, and $(h(t))_{t\in[0,\infty)}$ a n-dimensional progressively measurable process. Let μ_B be the associated Wiener measure on $C([0,\infty),\mathbb{R}^d)$ and $\mu_{\tilde{B}}$ the law of the process $\tilde{B}$ in $C([0,\infty),\mathbb{R}^d)$ defined by*

$$\tilde{B}(t) := B(t) + \int_0^t h(s)\, ds, \quad t \in [0,\infty). \tag{3.50}$$

Then, if

$$\int_0^\infty |Q^{-1/2}h(s)|^2\, ds < \infty \quad \mathbb{P}\text{-a.s.} \tag{3.51}$$

the law of $\tilde{B}$ is a absolutely continuous with respect to the law of B, i.e. $\mu_{\tilde{B}} \ll \mu_B$ as measures on $C([0,\infty),\mathbb{R}^n)$.

Recall that the Q-Wiener process W in a Hilbert space $\mathcal{H}$ can be expressed as in (1.2), i.e.

$$W(t) = \sum_{k\in\mathbb{N}} \sqrt{\sigma_k}\beta_k(t)e_k \tag{3.52}$$

where σ_k and e_k are respectively the eigenvalues and eigenfuctions of the covariance operator Q, forming an orthonormal basis of $\mathcal{H}$, and $\beta_k(t)$ are mutually independent real valued Wiener processes.

Define

$$W_n = \sum_{k=1}^n \sqrt{\sigma_k}\beta_k(t)e_k, \quad \text{and} \quad W^n = \sum_{k=n+1}^\infty \sqrt{\sigma_k}\beta_k(t)e_k, \tag{3.53}$$

so that $W = W_n + W^n$. Notice that (see e.g. [22, Section 4.2.2]) W_n and W^n are independent Wiener processes with covariance matrix respectively

$$Q_n = \sum_{k=0}^n \sigma_k e_k \otimes e_k \quad \text{and} \quad Q^n = \sum_{k=n+1}^\infty \sigma_k e_k \otimes e_k. \tag{3.54}$$

Let $\tilde{W}_n$ be a n-dimensional Brownian motion independent from and equal in law to W_n. Then, as all β_k are mutually independent, $\tilde{W}(t) := \tilde{W}_n + W^n$ is a Q-Wiener process equal in law to W.

We are now ready to state the main theorem of this section. This is a detailed yet compact summary of the framework introduced in [37] for a series of nonlinear stochastic partial differential equations. This will give verifiable conditions to be then used for our model of interest.

Theorem 3.2.8. *Consider the system of equations*

$$dX(t) = AX\,dt + F(X)\,dt + dW(t), \quad X(0) = x \tag{3.55}$$

$$d\tilde{Y}(t) = A\tilde{Y}\,dt + F(\tilde{Y})\,dt + G(X,\tilde{Y})\mathbb{1}_{t\le\tau}\,dt + dW(t), \quad \tilde{Y}(0) = y \tag{3.56}$$

where $G : \mathcal{H} \times \mathcal{H} \to \mathcal{H}_n$ is such that $G(X,\tilde{Y})\mathbb{1}_{t\le\tau} \in \operatorname{range} Q$ and τ is a stopping time with respect to the filtration generated by $\{(X_s,\tilde{Y}_s) : s \le t\}$. Assume there exists $n > 0$ such that the control G and the stopping time τ satisfy the following conditions:

(i) the system (3.55)-(3.56) *has a global solution*

$$X \in C(0,\infty;\mathcal{H}) \cap L^2(0,\infty;\mathcal{V});$$

(ii) the following condition holds

$$\int_0^\infty |Q_n^{-1/2} G(X(t),\tilde{Y}(t))|^2 \mathbb{1}_{t\le\tau}\,dt < C \quad \mathbb{P}\text{-}a.s.; \tag{3.57}$$

(iii) $\mathbb{P}(\tau = \infty) > 0$*;*

(iv) $|X(t) - \tilde{Y}(t)| \to 0$ *for* $t \to \infty$ *on the event* $\{\tau = \infty\}$*.*

Then (3.55) *has at most one invariant measure.*

Proof. Let $\mathcal{P}_t$, $t \ge 0$ be the Markov semigroup (3.7) associated to (3.55). We want to apply Theorem 3.2.4, so, for some $t > 0$, we have to find a suitable asymptotically equivalent coupling Γ of $\delta_x \mathcal{P}_t^{\mathbb{N}}$ and $\delta_y \mathcal{P}_t^{\mathbb{N}}$ for any pair of initial conditions x, y.

Given condition (ii) we can apply Theorem 3.2.7, hence the n-dimensional process

$$\tilde{W}_n(t) := W_n(t) + \int_0^t G(X(t),\tilde{Y}(t))\mathbb{1}_{t\le\tau}\,ds \quad t \ge 0$$

is a Wiener process with law absolutely continuous with respect to the law of W_n. As a consequence the process $\tilde{W} = \tilde{W}_n + W^n$ is absolutely continuous with respect to the original Wiener process W and Law $\tilde{Y}$ is absolutely continuous with respect to Law Y with Y solution of

$$dY(t) = AY\,dt + F(Y)\,dt + dW(t), \quad Y(0) = y.$$

This means that the law of the pair $(X, \tilde{Y})$ has marginals which are absolutely continuous with respect to the solutions to (3.55) starting respectively from x and y. Therefore, for some $t > 0$, the law induced by $\{(X(nt), \tilde{Y}(nt)) : n \in \mathbb{N}\}$ on $\mathcal{H}^{\mathbb{N}} \times \mathcal{H}^{\mathbb{N}}$, which we denote by $\Gamma_{(X,\tilde{Y})}$, is an asymptotically equivalent coupling of $\delta_x \mathcal{P}_t^{\mathbb{N}}$ and $\delta_y \mathcal{P}_t^{\mathbb{N}}$.

Finally conditions (iii) and (iv) ensure that $\Gamma_{(X,\tilde{Y})}$ gives positive probability to the diagonal at infinity D, as

$$\mathbb{P}\left((X(nt), \tilde{Y}(nt)) \in D\right) = \mathbb{P}\left(\lim_{n\to\infty} |X(nt) - \tilde{Y}(nt)| = 0\right) > 0.$$

Therefore by Theorem 3.2.4 there exists at most one ergodic invariant measure for $\mathcal{P}_t$. □

In the rest of the section we will prove that the conditions of Theorem 3.2.8 are satisfied for the stochastic two–layer quasi–geostrophic model.

3.2.2 Stochastic two–layer quasi–geostrophic model

Consider again our model of interest i.e.

$$\begin{aligned} &d\mathbf{q} + B(\boldsymbol{\psi}, \boldsymbol{\psi}) + \beta\partial_1\boldsymbol{\psi} = \nu\Delta^2\boldsymbol{\psi}\,dt + \begin{pmatrix} f \\ -r\Delta\psi_1 \end{pmatrix} + \begin{pmatrix} dW \\ 0 \end{pmatrix}. \\ &\mathbf{q} = (\Delta + M)\boldsymbol{\psi} \end{aligned} \tag{3.58}$$

Recall that in this case the noise is straightforwardly spatially degenerate as it acts only on one of the two layers. Consequently the classic Doob's theorem does not apply. At the same time on the first layer we assume that the noise is full, so as a whole it is only moderately degenerate, allowing us to apply the asymptotic coupling method just presented.

The main result of this section is an application of the general Theorem 3.2.8 to our model of interest and its proof follows closely the arguments developed in [37] for stochastic Navier–Stokes equation, among other examples. We will find an appropriate control function so that all the conditions of Theorem 3.2.8 are met, but under a restriction on the parameters of the system, in particular when the bottom friction parameter r is sufficiently large. We will later discuss what the physical interpretation of such a result could be and if conditions on other parameters can be considered alternatively.

Theorem 3.2.9. *For set of parameters such that*

$$2r > \tfrac{h_1 k_B}{\nu^2}\|f\|_{-2}^2 + \tfrac{2k_B}{\nu}\operatorname{Tr}\left[(Q^{1/2})^*\tilde{A}^{-1}Q^{1/2}\right], \tag{3.59}$$

the system (3.58) *has at most one ergodic invariant measure.*

Proof. Consider the modified system

$$\begin{aligned}&\frac{\mathrm{d}\tilde{\mathbf{q}}}{\mathrm{d}t} + B(\tilde{\boldsymbol{\psi}},\tilde{\boldsymbol{\psi}}) + \beta\partial_1\tilde{\boldsymbol{\psi}} = \nu\Delta^2\tilde{\boldsymbol{\psi}}\,dt + \begin{pmatrix} f + G(\mathbf{q},\tilde{\mathbf{q}})\mathbb{1}_{t\le\tau} \\ -r\Delta\tilde{\psi}_1\end{pmatrix} + d\mathbf{W}.\\ &\tilde{\mathbf{q}} = (\Delta + M)\tilde{\boldsymbol{\psi}}\end{aligned} \tag{3.60}$$

with periodic boundary condition and initial condition $\tilde{\mathbf{q}}(0) = \tilde{\mathbf{q}}_0 \neq \mathbf{q}_0$, where $G(\mathbf{q},\tilde{\mathbf{q}}) : \mathcal{H}\times\mathcal{H} \to \mathcal{H}_N$ is a finite dimensional control function, which we will specify later, and τ a stopping time with respect to the natural filtration for $(\mathbf{q},\tilde{\mathbf{q}})$. To prove the uniqueness of the invariant measure μ for $\mathcal{P}_t$, we look for $N > 0$, G and τ satisfying the conditions in Theorem 3.2.4.

Given $\mathbf{q}$, solution of (3.58) with streamfunction $\boldsymbol{\psi}$, define the variables

$$\mathbf{u} = \mathbf{q} - \tilde{\mathbf{q}} \quad \text{and} \quad \mathbf{v} = \boldsymbol{\psi} - \tilde{\boldsymbol{\psi}}. \tag{3.61}$$

Then $\mathbf{u} = (\Delta + M)\mathbf{v}$ and it satisfies the integral equation

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} + B(\mathbf{v},\tilde{\boldsymbol{\psi}}) + B(\boldsymbol{\psi},\mathbf{v}) = \nu\Delta^2\mathbf{v} - \begin{pmatrix} G(\mathbf{q},\tilde{\mathbf{q}})\mathbb{1}_{t\le\tau} \\ r\Delta v_2\end{pmatrix} \tag{3.62}$$

with initial condition $\mathbf{u}_0 = \mathbf{q}_0 - \tilde{\mathbf{q}}_0 \neq 0$.

Let λ_k be an increasing sequence of eigenvalues of $-\Delta$ with corresponding eigenvectors e_k forming an orthonormal basis for $\mathcal{H}$. Consider the following finite dimensional control

$$G(\mathbf{q},\tilde{\mathbf{q}}) = \alpha\Pi_N(\Delta\psi_1 - \Delta\tilde{\psi}_1) = \alpha\Pi_N\Delta v_1 \tag{3.63}$$

where Π_N is the projection onto $\mathcal{H}_N = \operatorname{span}\{e_k,\ k = 1,\dots N\}$ and $\alpha > 0$ is a parameter to be found below.

Define the stopping time

$$\tau_K := \inf\left\{t \ge 0\ :\ \int_0^t \left|Q_N^{-1/2}\alpha\Pi_N\Delta v_1\right|^2 ds \ge K\right\} \tag{3.64}$$

so that

$$\int_0^\infty \left|Q_N^{-1/2}G(\mathbf{q},\tilde{\mathbf{q}})\mathbb{1}_{t\le\tau_K}\right|^2 dt \le K$$

and condition (ii) in Theorem 3.2.4 holds.

The existence of a unique solution for the controlled system, i.e condition (i), can be shown as in the last chapter but transforming the stochastic forcing with the use of Girsanov theorem (Theorem 3.2.5). Following [50, Remark 8] where this is done for the stochastic Navier–Stokes equation, we can rewrite the equation for the first component of (3.60) as

$$d\tilde{q}_1(t) + J(\tilde{\psi}_1, \tilde{q}_1 + \beta y)dt = (\nu\Delta - \alpha\mathbb{1}_{t\leq\tau}\Pi_N)\Delta\tilde{\psi}_1\, dt + f\, dt \\ + \alpha\Pi_N\Delta\psi_1\mathbb{1}_{t\leq\tau}\, dt + dW(t)$$

and define the process $\hat{W}$ by

$$\hat{W}(t) = W(t) + \int_0^t \alpha\Pi_N\Delta\psi_1(s)\mathbb{1}_{t\leq\tau}\, ds.$$

To ensure that we can use Girsanov theorem Theorem 3.2.5 we show that condition (3.49) holds. First observe that

$$\int_0^t \left|Q_N^{-1/2}\alpha\Pi_N\Delta\psi_1\mathbb{1}_{t\leq\tau}\right|^2 ds \leq \alpha h_1^{-1}\|Q_N^{-1/2}\|^2 \int_0^t |\Delta\boldsymbol{\psi}|^2\, ds$$

then by Theorem 2.4.2 we have a bound for the $L^2(0,T;L^2)$ norm of $\Delta\boldsymbol{\psi}$

$$\leq \frac{\alpha\|Q_N^{-1/2}\|^2}{h_1\delta}\left(\|\!|\mathbf{q}_0|\!\|_{-1}^2 + t\left(\tfrac{h_1}{2\nu\lambda_1}\|f\|_{-1}^2 + T_Q\right) + \Xi_\gamma\right)$$

where Ξ_γ is an almost surely finite random variable. Then by Lemma 3.2.6 Z^h is a martingale and $\hat{W}$ is a Brownian motion equivalent with respect to W for every finite interval $[0,T]$. Furthermore the term $-\alpha\Pi_N\Delta\tilde{\psi}_1\mathbb{1}_{t\leq\tau}\, dt$ will not affect substantially the estimates derived in the last chapter to prove the existence of a solution. In fact to provide an *a priori* bound in $\mathcal{H}$ we took the $\mathbf{L}^2$ scalar product with $\boldsymbol{\psi}$ of the random equation associated with the controlled system. Therefore this new term appears on the left hand side of (2.10) as

$$-(\alpha\Pi_N\Delta\tilde{\psi}_1, h_1\tilde{\psi}_1) = \alpha\|\Pi_N\tilde{\psi}_1\|^2.$$

and all the subsequent calculations can be easily adapted. Another difference with the previous case is that in this case the auxiliary Ornstein-Uhlenbeck process will satisfy

$$d\eta + \alpha_1 A\eta\, dt = d\hat{W}$$

where $\hat{W}$ is a Brownian motion with respect to a different probability measure but this will not affect the solution theory. Then the controlled equation

(3.60) with control (3.63) has a unique solution defined for all $[0,T]$, $T \geq 0$ and by Kolmogorov existence theorem we can extend it to $\mathbb{R}^+$.

Next we want to show condition (iv) holds, namely

$$\lim_{t\to\infty} |||\mathbf{u}(t)|||_{-1}^2 = 0 \quad \text{on } \{\tau_K = \infty\}.$$

Taking the $\mathbf{L}^2$ scalar product of (3.62) with $\mathbf{v}$, we obtain

$$\left(\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}, \mathbf{v}\right) + (B(\boldsymbol{\psi}, \mathbf{v}), \mathbf{v}) = \nu|\Delta\mathbf{v}|^2 - h_1(\alpha\Pi_N\Delta v_1 \mathbb{1}_{t\leq\tau}, v_1) + rh_2\|v_2\|^2,$$

where we have used the fact that $(B(a,b),a) = 0$ (1.56). Recall that by (1.42) we have

$$\left(\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t}, \mathbf{v}\right) = -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|||\mathbf{u}|||_{-1}^2$$

so that we have the following equation

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|||\mathbf{u}|||_{-1}^2 + \nu|\Delta\mathbf{v}|^2 + rh_2\|v_2\|^2 = (B(\boldsymbol{\psi}, \mathbf{v}), \mathbf{v}) + h_1(\alpha\Pi_N\Delta v_1\mathbb{1}_{t\leq\tau}, v_1). \quad (3.65)$$

Note that we can write G as

$$G(\mathbf{q}, \tilde{\mathbf{q}}) = \alpha\Delta v_1 - E_N$$

where, denoting the orthogonal complement of Π_N as $\Pi_N^\perp$, $E_N = \alpha\Pi_N^\perp\Delta v_1$. Recall (see e.g. [62, pg 366]) the generalised Poincaré inequalities for any $N \geq 1$ i.e.

$$\|\Pi_N v_1\|_{k+1}^2 \leq \lambda_N\|\Pi_N v_1\|_k^2 \quad \text{and} \quad \|\Pi_N^\perp v_1\|_k^2 \leq \lambda_N^{-1}\|\Pi_N^\perp v_1\|_{k+1}^2. \quad (3.66)$$

Then we have an appropriate bound for the norm of the error term E_N as

$$\begin{aligned}(G(\mathbf{q}, \tilde{\mathbf{q}})), v_1) &= \alpha(\Delta v_1 - \Pi_N^\perp\Delta v_1, v_1)\\ &= -\alpha h_1\|v_1\|^2 + \alpha h_1\|\Pi_N^\perp v_1\|^2\\ &\leq -\alpha h_1\|v_1\|^2 + \alpha h_1\lambda_N^{-1}|\Delta v_1|^2.\end{aligned}$$

Consequently using this estimates in (3.65) we get for $t \in [0, \tau_K)$

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|||\mathbf{u}|||_{-1}^2 + \nu|\Delta\mathbf{v}|^2 + rh_2\|v_2\|^2 \leq (B(\boldsymbol{\psi}, \mathbf{v}), \mathbf{v}) - \alpha h_1\|v_1\|^2 + \alpha h_1\lambda_N^{-1}|\Delta v_1|^2. \quad (3.67)$$

By (1.58) we have

$$(B(\boldsymbol{\psi}, \mathbf{v}), \mathbf{v}) \leq \tfrac{\nu}{2}|\Delta\mathbf{v}|^2 + k_B|\Delta\boldsymbol{\psi}|^2\|\mathbf{v}\|^2, \quad (3.68)$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}t}|||\mathbf{u}|||_{-1}^2+\nu|\Delta\mathbf{v}|^2+2\alpha h_1\|v_1\|^2+2rh_2\|v_2\|^2 \leq 2k_B|\Delta\boldsymbol{\psi}|^2\|\mathbf{v}\|^2+2\alpha h_1\lambda_N^{-1}|\Delta v_1|^2.$$

Therefore, setting $\alpha \geq r$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}|||\mathbf{u}|||_{-1}^2 + \left(\nu - 2\alpha\lambda_N^{-1}\right)|\Delta\mathbf{v}|^2 \leq \|\mathbf{v}\|^2\left(2k_B|\Delta\boldsymbol{\psi}|^2 - 2r\right), \tag{3.69}$$

and we choose N so that

$$\nu - 2\alpha\lambda_N^{-1} > 0. \tag{3.70}$$

Therefore Gronwall lemma (Lemma A.1) gives

$$|||\mathbf{u}(t)|||_{-1}^2 \leq |||\mathbf{u}(0)|||_{-1}^2 \exp\left(-2rt + 2k_B\int_0^t |\Delta\boldsymbol{\psi}|^2\,ds\right) \quad \text{for all } t \in [0, \tau_K). \tag{3.71}$$

It is only left to bound appropriately the $L^2(0,t;\mathbf{H}^2)$ norm of $\boldsymbol{\psi}$. Theorem 2.4.2 gave the estimate (2.54) i.e.

$$|||\mathbf{q}(t)|||_{-1}^2 - |||\mathbf{q}_0|||_{-1}^2 + \delta\int_0^t |\Delta\boldsymbol{\psi}|^2\,ds - t\left(\tfrac{h_1}{2\nu}\|f\|_{-2}^2 + T_Q\right) \leq 2h_1\Xi_\gamma$$

for $\delta = \nu - (2\gamma\,\mathrm{Tr}\,Q)/(\lambda_1^2) > 0$ with

$$\Xi_\gamma = \sup_{t\geq 0}(X_t - \gamma\langle X\rangle_t), \quad X_t := \int_0^t (\psi_1, dW).$$

Define the event $\mathcal{E}_R$ to be

$$\mathcal{E}_R := \Big\{|||\mathbf{q}(t)|||_{-1}^2 - |||\mathbf{q}(0)|||_{-1}^2 + \delta\int_0^t |\Delta\boldsymbol{\psi}|^2\,ds \\ -t\left(\tfrac{h_1}{2\nu}\|f\|_{-2}^2 + T_Q\right) \leq R, \quad \forall t \geq 0\Big\}. \tag{3.72}$$

By the exponential martingale estimate (Lemma 2.4.1) we have

$$\mathbb{P}(\mathcal{E}_R^c) \leq \mathbb{P}(2h_1\Xi_\gamma \geq R) \leq e^{-\frac{\gamma R}{2h_1}}$$

and as a consequence $\mathcal{E}_R$ is a set with positive probability.

Therefore on the set $\mathcal{E}_R$ we have

$$\delta\int_0^t |\Delta\boldsymbol{\psi}|^2\,ds \leq |||\mathbf{q}(0)|||_{-1}^2 + t\left(\tfrac{h_1}{2\nu}\|f\|_{-2}^2 + T_Q\right) + R$$

and so going back to (3.71) we find on the set $\mathcal{E}_R \cap \{\tau_K > t\}$ that

$$|||\mathbf{u}(t)|||_{-1}^2 \le |||\mathbf{u}(0)|||_{-1}^2 \, e^{\frac{2k_B}{\delta}(|||\mathbf{q}_0|||_{-1}^2+R)} \exp\left(-t\left(2r - \tfrac{2k_B}{\delta}T_Q - \tfrac{h_1 k_B}{\delta\nu}\|f\|_{-2}^2\right)\right).$$

Then, setting the arbitrary parameter $\delta \in (0,\nu)$ to be for example $\delta = \nu/2$, whenever

$$C_r := 2r - \tfrac{2k_B}{\nu}T_Q - \tfrac{h_1 k_B}{\nu^2}\|f\|_{-2}^2 > 0, \tag{3.73}$$

we have that $|||\mathbf{u}(t,\omega)|||_{-1}^2$ is exponentially decaying in time over $[0,\tau_K)$ and so

$$\lim_{t\to\infty} |||\mathbf{u}(t,\omega)|||_{-1}^2 = 0 \quad \text{for all } \omega \in \mathcal{E}_R \cap \{\tau_K = \infty\}.$$

Finally it is left to ensure that there exists a K so that $\mathcal{E}_R \cap \{\tau_K = \infty\}$ is non empty and has positive probability. We will actually show that there exists $K > 0$ such that $\mathcal{E}_R \subset \{\tau_K = \infty\}$.

Recall the definition of the stopping time (3.64) and observe that the following estimate holds

$$\begin{aligned}\int_0^t \left|Q_N^{-1/2}\alpha\Pi_N\Delta v_1\right|^2 ds &\le \alpha^2\|Q_N^{-1/2}\|^2 \int_0^t |\Pi_N\Delta v_1|^2\,ds \\ &\le \tfrac{\alpha^2}{h_1}\|Q_N^{-1/2}\|^2 \int_0^t |\Delta\mathbf{v}|^2\,ds.\end{aligned} \tag{3.74}$$

We can focus on showing the right hand side is exponentially decaying in time. Integrate (3.69) over $[0,t]$ with $t \le \tau_K$, and use again that $\|\mathbf{v}\| \le |||\mathbf{u}|||_{-1}$ to get

$$\left(\nu - 2\alpha h_1\lambda_N^{-1}\right)\int_0^t |\Delta\mathbf{v}|^2\,ds \le |||\mathbf{u}(0)|||_{-1}^2 + \int_0^t |||\mathbf{u}(s)|||_{-1}^2\left(-2r + 2k_B|\Delta\boldsymbol{\psi}|^2\right)ds$$

and by (3.71) we have the bound

$$\left(\nu - 2\alpha h_1\lambda_N^{-1}\right)\int_0^t |\Delta\mathbf{v}|^2\,ds \le |||\mathbf{u}(0)|||_{-1}^2 \exp\left(-2rt + 2k_B\int_0^t |\Delta\boldsymbol{\psi}|^2\,ds\right).$$

Since

$$\mathcal{E}_R \subset \left\{\delta\int_0^t |\Delta\boldsymbol{\psi}|^2\,ds - |||\mathbf{q}(0)|||_{-1}^2 - t\left(\tfrac{h_1}{2\nu}\|f\|_{-2}^2 + T_Q\right) \le R\right\}$$

we conclude that on $\mathcal{E}_R$ for $t \in [0,\tau_K)$

$$\int_0^t |\Delta\mathbf{v}|^2\,ds \le \frac{|||\mathbf{u}(0)|||_{-1}^2}{\nu - 2\alpha h_1\lambda_N^{-1}}\exp\left(\tfrac{2k_B}{\nu}(|||\mathbf{q}_0|||_{-1}^2 + R) - C_r t\right). \tag{3.75}$$

Then using (3.75) in (3.74) we have that

$$\int_0^t |Q_N^{-1/2}\alpha\Pi_N\Delta v_1|^2\,ds \le \frac{\alpha^2\|Q_N^{-1/2}\|^2 |||\mathbf{u}(0)|||_{-1}^2}{h_1(\nu-2\alpha h_1\lambda_N^{-1})}\exp\left(\tfrac{2k_B}{\nu}(|||\mathbf{q}_0|||_{-1}^2+R)-C_r t\right).$$

for all $\omega \in \mathcal{E}_R$ and $t \in [0,\tau_K(\omega)]$. Choosing the parameter K large enough we can conclude that

$$\int_0^t |Q_N^{-1/2}\alpha\Pi_N\Delta v_1|^2\,ds \le \frac{K}{2} \quad \text{for all } t<\tau_K \tag{3.76}$$

and by continuity this is true also for $t=\tau_K$. Now suppose ω is such that $\tau_K(\omega)$ is finite, by the definition of the stopping time we have that

$$\int_0^{\tau_K} |Q_N^{-1/2}\alpha\Pi_N\Delta v_1|^2\,ds = K.$$

This leads to a contradiction with (3.76) and proves that no such ω exists. Hence there exists $K>0$ such that $\mathcal{E}_R \subset \{\tau_K=\infty\}$. In conclusion we have found a set of positive probability over which $|||\mathbf{q}(t)-\tilde{\mathbf{q}}(t)|||_{-1}$ converges to zero on the infinite time horizon. □

Remark 3.2.10 (Finite dimensional noise). From the literature (e.g. [37]) we know that the asymptotic coupling method applies also when the noise acts only on finitely many modes, as long as enough of them are activated. This lower bound on the dimension of the noise arose also in the argument just presented, when we required the condition (3.70) to be satisfied by N. Therefore, with few modifications to the proof of Theorem 3.2.9, the ergodicity holds also for the model perturbed on the top layer only by a N dimensional noise as long as (3.70) holds.

It is interesting to highlight that the ergodicity result for the 2LQG model holds also under conditions different from (3.73). In fact with a simple modification in the proof of Theorem 3.2.9 we can retrieve a condition also, or solely, involving the viscosity. From (3.69), namely

$$\frac{\mathrm{d}}{\mathrm{d}t}|||\mathbf{u}|||_{-1}^2 + \left(\nu-2\alpha\lambda_N^{-1}\right)|\Delta\mathbf{v}|^2 \le \|\mathbf{v}\|^2\left(2k_B|\Delta\boldsymbol{\psi}|^2-2r\right),$$

where N is such that $\nu-2\alpha\lambda_N^{-1}>0$, by Poincaré inequality, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}|||\mathbf{u}|||_{-1}^2 + \lambda_1\left(\nu-2\alpha\lambda_N^{-1}\right)\|\mathbf{v}\|^2 \le \|\mathbf{v}\|^2\left(2k_B|\Delta\boldsymbol{\psi}|^2-2r\right).$$

Using (1.43) i.e. $\|\mathbf{v}\| \le |||\mathbf{u}|||_{-1} \le a_0\|\mathbf{v}\|$ with $a_0 = \max(1, 2F_1/\lambda_1)$, we derive

$$\frac{\mathrm{d}}{\mathrm{d}t}|||\mathbf{u}|||_{-1}^2 \le |||\mathbf{u}|||_{-1}^2 \left(2k_B|\Delta\boldsymbol{\psi}|^2 - 2r - \tfrac{\lambda_1}{a_0^2}\left(\nu - 2\alpha\lambda_N^{-1}\right)\right),$$

so that, thanks to Gronwall lemma,

$$|||\mathbf{u}(t)|||_{-1}^2 \le |||\mathbf{u}(0)|||_{-1}^2 \exp\left(-t(2r + \tfrac{\lambda_1}{a_0^2}\left(\nu - 2\alpha\lambda_N^{-1}\right)) + 2k_B\int_0^t |\Delta\boldsymbol{\psi}|^2\right).$$

With the same arguments developed in the proof of Theorem 3.2.9 we can conclude that

$$\lim_{t\to\infty} |||\mathbf{u}(t)|||_{-1}^2 = 0$$

on a set of positive probability, whenever

$$2r + \tfrac{\lambda_1}{a_0^2}\left(\nu - 2\alpha\lambda_N^{-1}\right) > \tfrac{2k_B}{\nu}T_Q + \tfrac{h_1k_B}{\nu^2}\|f\|_{-2}^2. \tag{3.77}$$

In particular this result provides ergodicity of the model also when $r = 0$ as long as the viscosity is large enough. The presence of a large viscosity would also imply that we can consider smaller values of N, namely more degenerate noise on the first layer. A similar result holds also for the stochastic Navier-Stokes equation. In fact in [56] ergodicity is ensured in a large viscosity scenario even with a finite dimensional stochastic forcing.

On the one hand, it is clear, from a physical point of view, that we can expect ergodicity when there is strong dissipation on both layers, for example, as just discussed, by means of a large viscosity. On the other hand, the imposed parameter condition (3.73) requires sufficient dissipation only on one of the two layers by requiring the bottom layer (the one without noise) to be enslaved by the top one, or to converge autonomously, by means of a minimum requirement for the friction. However, a natural question which arises in this context is whether the ergodicity result can be achieved even when no particular condition on the parameters is imposed. This is not clear directly from our analysis nor the available literature and it will be discussed further in the Conclusions.

We close the section on the ergodicity of the two–layer quasi–geostrophic model by discussing our result from a quantitative point of view. In Theorem 3.1.9 we required (3.73) to hold, i.e.

$$2r > \tfrac{2k_B}{\nu}T_Q + \tfrac{h_1k_B}{\nu^2}\|f\|_{-2}^2$$

in order to establish uniqueness of the invariant measure. Is this a quantitatively meaningful requirement or should it rather be interpreted as a qualitative statement? We will see that the latter is the case, as the typical range of values for r is on significantly smaller scales than $h_1 k_B \|f\|_{-2}^2/\nu^2$.

The constant k_B is given by (1.61) i.e.

$$k_B = \max\left(\frac{c_0^2}{\nu h_1}, \frac{c_1^2 F_1^2}{\nu h_1 \lambda_1}\right)$$

where c_0, c_1 are as in Lemma 1.3.3, namely $c_0 = 2+(\pi\sqrt{2})^{-1}$ and $c_1 = c_0\lambda_1^{-1/2}$. The smallest eigenvalue of $-\Delta$ is $\lambda_1^{1/2} = 2\pi/L$ where L is the side of the domain (see e.g. [62, Lemma 5.40]). Then, using the definition of F_1, (1.23)

$$k_B = \frac{(2+(\pi\sqrt{2})^{-1})^2}{\nu h_1} \max\left(1, \left(\frac{L^2 f_0^2 \rho_0}{4\pi^2 g h_1(\rho_2-\rho_1)}\right)^2\right).$$

Recall that the gravitational acceleration is $g = 9.81ms^{-2}$ and the average Coriolis force at mid–latitude is $f_0 = 8\times 10^{-5}s^{-1}$. Next, let us consider the values for the ocean as presented in [15]: the horizontal spatial scale L in the ocean is typically $1000km$, the depth of the layers is $500m$, the eddy viscosity ν is $50m^2s^{-1}$, the mean density ρ_0 is $1025kgm^{-3}$ and the density difference between the layers is $25kgm^{-3}$. By using the values listed above we have

$$\begin{aligned}\left(\frac{L^2 f_0^2 \rho_0}{4\pi^2 g h_1(\rho_2-\rho_1)}\right)^2 &= \left(\frac{(10^{12}m^2)(8\times 10^{-5}s^{-1})^2(1.025\times 10^3 kgm^{-3})}{4\pi^2(9.81ms^{-2})(5\times 10^2 m)(25kgm^{-3})}\right)^2 \\ &= 1.836\end{aligned}$$

so that

$$k_B = \frac{(2+(\pi\sqrt{2})^{-1})^2}{\nu h_1}\left(\frac{L^2 f_0^2 \rho_0}{4\pi^2 g h_1(\rho_2-\rho_1)}\right)^2 = 3.636\times 10^{-4} sm^{-3}.$$

Next, we consider the mean deterministic wind forcing f to be as in [15]

$$f(x,y) = \frac{2\pi\tau_0}{\rho_0 h_1 L}\sin\frac{2\pi y}{L}, \tag{3.78}$$

where τ_0 is the wind tension and is of order $0.1Nm^{-2}$. By explicitly computing $\|f\|_{-2}^2$, we get

$$\begin{aligned}\|f\|_{-2}^2 &= \int_0^L\int_0^L |\Delta^{-1} f|^2\,dxdy \\ &= \left(\frac{L}{2\pi}\right)^4 \int_0^L\int_0^L |f|^2\,dxdy = \left(\frac{L^2\tau_0}{2\pi\rho_0 h_1}\right)^2.\end{aligned}$$

As a consequence

$$\|f\|_{-2}^2 = \left(\frac{(10^{12}m^2)(10^{-1}Nm^{-2})}{2\pi(1.025 \times 10^3 kgm^{-3})(5 \times 10^2 m)} \right)^2 = 9.644 \times 10^8 m^6 s^{-4}.$$

Finally, pulling these measures together we see that

$$\frac{h_1 k_B}{\nu^2}\|f\|_{-2}^2 = \frac{(5 \times 10^2 m)(3.636 \times 10^{-4} sm^{-3})(9.644 \times 10^8 m^6 s^{-4})}{2.5 \times 10^3 m^4 s^{-2}} = 7{\times}10^4 s^{-1}.$$

However, according to the data reported in [15] the parameter r should be of the order $10^{-5}s^{-1}$. Therefore, the ergodicity result Theorem 3.1.9 cannot provide a realistic physical constraint on the bottom friction, but rather qualitative information on the long-time behaviour of the model.

Summary and remarks

This chapter includes a number of original results regarding the stochastic two–layer quasi–geostrophic model. In the first part we showed that there exists an invariant measure for the model. To do so, we particularly focused our efforts in proving Theorem 3.1.7, namely that the family of laws of the solutions indexed by the initial time is tight.

In the second part, we ensured that, when indeed there exists an invariant measure, this is unique. Given the spatial degeneracy of our noise, we used a recently developed technique, the asymptotic coupling method. The idea behind this method will be central also in the next chapter when we will show the exponential convergence of transition probabilities to the invariant measure. Moreover, Girsanov theorem, which we recalled here, will be an important tool when studying the dependence of the invariant measure with respect to parameters, as we will do in chapter 5.

In using the asymptotic coupling method, we introduced a control, in (3.63), which is solely dependent on the top layer. This choice led to the desired ergodicity, but also to it being conditional on parameters like the bottom friction being sufficiently large. As discussed, the result holds also with the following modifications: having a finite dimensional noise on the top layer instead of full noise, as long as (3.70) holds, or, for sufficiently large values of the viscosity, having a very degenerate noise and even no bottom friction. It is not clear though whether ergodicity holds without extra conditions on the friction or viscosity. Intuitively, it may be that a different coupling, carrying information of the whole dynamics rather than just of the top layer, would help achieve this result.

Chapter 4

Exponential stability and spectral gap

In this chapter we will use a general form of Harris' theorem developed in [44] to prove exponential convergence of the transition probabilities to the invariant measure, a result stronger than unique ergodicity, for our model of interest. In particular we will use the framework for SPDEs introduced in [12] and show sufficient conditions for the convergence.

In the first section we provide the details of the methodology from the literature giving streamlined proofs of the desired results. In section 4.2 and section 4.3 we apply these techniques respectively to the stochastic Navier–Stokes equations (section 1.2) and the stochastic two–layer quasi–geostrophic model (section 1.4).

4.1 Methodology

In the finite dimensional context Harris' theorem (see e.g. [44, Theorem 1.5], [43], [57]) provides conditions under which the transition probabilities of a Markov process converge to the invariant measure. In this case the convergence is with respect to the total variation distance which is defined as follows. Let $\mathcal{H}$ be a Hilbert space with scalar product $(\cdot, \cdot)$ and associated norm $|\cdot|$. Given a probability measure $\mu \in \mathcal{M}_1(\mathcal{H})$, the *total variation norm* of μ is (see e.g. [57, pg 315])

$$\begin{aligned}\|\mu\|_{TV} &= \sup\{|\langle\varphi,\mu\rangle| \,:\, \varphi:\mathcal{H}\to\mathbb{R} \text{ continuous}, |\varphi(x)|\le 1 \text{ for all } x\in\mathcal{H}\}\\ &= \sup_{A\in\mathcal{B}(\mathcal{H})} \mu(A)\end{aligned}$$

where we denoted

$$\langle \varphi, \mu \rangle := \int_{\mathcal{H}} \varphi(x)\, \mu(dx).$$

Then we define the *total variation distance* between two probability measures $\mu_1, \mu_2 \in \mathcal{M}_1$ as

$$d_{TV}(\mu_1, \mu_2) = \|\mu_1 - \mu_2\|_{TV}. \tag{4.1}$$

Harris' theorem ensures convergence of transition probabilities if there exists a so-called *small set* which is visited infinitely often by the process, and the speed of convergence is related to how fast the process return to such set. Precisely a small set for a Markov process $\mathcal{P}_t$ with transition probabilities P_t is defined as follows:

Definition 4.1.1 (small set)**.** A set $A \subset \mathcal{H}$ is *small* if there exists a time $t > 0$ and a constant $\varepsilon > 0$ such that

$$d_{TV}\left(P_t(x, \cdot), P_t(y, \cdot)\right) \leq 1 - \varepsilon \quad \text{for all } \ x, y \in A.$$

Then a major difficulty with applying Harris' theorem in the context of stochastic partial differential equations is that the transition probabilities $P_t(x, \cdot)$ and $P_t(y, \cdot)$ might be singular for different initial conditions $x \neq y$. In that case

$$d_{TV}\left(P_t(x, \cdot), P_t(y, \cdot)\right) = 1$$

and there are no small sets.

In [44] a weaker notion of small set is introduced and a version of Harris' theorem in infinite dimensions is proved, which gives exponential rate of convergence in a *Wasserstein-like distance*, rather than in total variation.

Let us define precisely such Wasserstein-like distance to state the result. We call a function $d : \mathcal{H} \times \mathcal{H} \to \mathbb{R}_+$ to be *distance-like* function, or equivalently a *semimetric*, when it is symmetric, lower semi-continuous and such that $d(x, y) = 0 \Leftrightarrow x = y$. When the symmetry fails, we refer to d as a *premetric*.

A semimetric d on $\mathcal{H}$ can be lifted to a semimetric on the level of probabilities called *Wasserstein semimetric* W_d: given two probability measures $\mu_1, \mu_2 \in \mathcal{M}_1(\mathcal{H})$ set

$$W_d(\mu_1, \mu_2) := \inf_{\Gamma \in \mathcal{C}(\mu_1, \mu_2)} \int d(x, y)\, \Gamma(dx, dy), \tag{4.2}$$

where $\mathcal{C}(\mu_1, \mu_2)$ is the set of all couplings of μ_1, μ_2 as introduced in section 3.2.1. The classic coupling lemma (e.g. [72, Theorem 4.1],[49, Proposition 4.3.3]) ensures that the infimum in this definition is always reached by some coupling, given the lower semi-continuity of d.

Lemma 4.1.2 (Coupling lemma). *Let $\mathbb{X}$ be a Polish space and the distance-like function d be lower semi-continuous. Then for any probability measures μ_1, μ_2 on $\mathbb{X}$, there exists a coupling $\Gamma^* \in \mathcal{C}(\mu_1, \mu_2)$ such that*

$$W_d(\mu_1, \mu_2) = \int d(x, y)\ \Gamma^*(dx, dy)$$

or equivalently there exists a couple of random variables (ξ_*, η_*) such that $\mathrm{Law}(\xi) = \mu_1$ *and* $\mathrm{Law}(\eta) = \mu_2$*,*

$$W_d(\mu_1, \mu_2) = \mathbb{E}\, d(\xi_*, \eta_*).$$

*We call any such Γ^* or (ξ_*, η_*) a d-*optimal coupling*.*

We can also use the coupling lemma to give an equivalent formulation of the total variation distance (4.1) which we will use often later on. The total variation distance has probabilistic representation by the discrete metric $l(x, y) = \mathbb{1}(x \neq y)$ i.e. given two measures μ_1, μ_2

$$d_{TV}(\mu_1, \mu_2) = 2W_l(\mu_1, \mu_2).$$

Then by the coupling lemma there exists a l-optimal coupling $(\xi_*, \eta_*) \in \mathcal{C}(\mu_1, \mu_2)$ such that

$$d_{TV}(\mu_1, \mu_2) = 2\mathbb{P}(\xi_* \neq \eta_*). \tag{4.3}$$

Given a distance-like function d, define the associated Lipschitz seminorm as

$$\|\varphi\|_d := \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{d(x, y)}. \tag{4.4}$$

It is immediate that $\|a\varphi\|_d = |a|\|\varphi\|_d$ for all $a \in \mathbb{R}$, and the triangular inequality holds as

$$\begin{aligned}\|\varphi_1 + \varphi_2\|_d &= \sup_{x \neq y} \frac{|\varphi_1(x) + \varphi_2(x) - \varphi_1(y) - \varphi_2(y)|}{d(x, y)} \\ &\leq \sup_{x \neq y} \frac{|\varphi_1(x) - \varphi_1(y)|}{d(x, y)} + \sup_{x \neq y} \frac{|\varphi_2(x) - \varphi_2(y)|}{d(x, y)}.\end{aligned}$$

On the other hand $\|\varphi\|_d = 0$ only implies that φ is a constant function as for all $C \in \mathbb{R}$, $\|\varphi - C\|_d = \|\varphi\|_d$.

When d is a distance function, the Wasserstein distance W_d has a useful equivalent formulation by means of the Kantorovich-Rubinstein formula (see e.g. [28]) i.e.

$$\sup_{\|\varphi\|_d \leq 1} |\langle \varphi, \mu_1 \rangle - \langle \varphi, \mu_2 \rangle| = W_d(\mu_1, \mu_2). \tag{4.5}$$

On the other hand when d is only a distance-like function, like in the framework we are about to work with, we retain the following weaker relation

$$\sup_{\|\varphi\|_d \le 1} |\langle \varphi, \mu_1 \rangle - \langle \varphi, \mu_2 \rangle| \le W_d(\mu_1, \mu_2). \tag{4.6}$$

In fact, for any coupling $\Gamma \in \mathcal{C}(\mu_1, \mu_2)$

$$|\langle \varphi, \mu_1 \rangle - \langle \varphi, \mu_2 \rangle| = \left| \int \varphi(x) - \varphi(y)\, \Gamma(dx, dy) \right| \le \|\varphi\|_d \int d(x, y)\, \Gamma(dx, dy).$$

By taking the infimum over Γ, we get

$$|\langle \varphi, \mu_1 \rangle - \langle \varphi, \mu_2 \rangle| \le \|\varphi\|_{\tilde{d}}\, W_d(\mu_1, \mu_2) \tag{4.7}$$

which in particular implies (4.6).

We are now ready to relax the definition of small set swapping the total variation distance for the Wasserstein semimetric associated to a distance-like function d.

Definition 4.1.3 ([44, Definition 4.4])**.** Let $d : \mathcal{H} \times \mathcal{H} \to [0, 1]$ be a distance-like function on $\mathcal{H}$ and $\mathcal{P}$ a Markov operator on $\mathcal{H}$. A set $K \subset \mathcal{H}$ is called *d-small* for $\mathcal{P}$ if there exists $\varepsilon > 0$ such that

$$W_d(P_t(x, \cdot), P_t(y, \cdot)) \le 1 - \varepsilon \tag{4.8}$$

for all $x, y \in K$.

As explained for example in [43], the proof of the original Harris' theorem relies on the existence of a small set which is visited infinitely often. Usually candidates for small sets are the level set of the so called *Lyapunov functions*:

Definition 4.1.4 ([44])**.** A measurable function $V : \mathcal{H} \to \mathbb{R}_+$ is called *Lyapunov function* for $\mathcal{P}_t$ if there exist positive constants C_V, γ, K_V such that

$$\mathcal{P}_t V(x) \le C_V e^{-\gamma t} V(x) + K_V \quad \text{for all } x \in \mathcal{H}, t \ge 0. \tag{4.9}$$

It can be shown (see [11, Lemma 4.1]) that Lyapunov functions are integrable with respect to the invariant measure μ_*. Then we also have an explicit upper bound for its integral by means of Definition 4.1.4: by the invariance of μ_* with respect to $\mathcal{P}_t$, $t \ge 0$ we have that

$$\langle V, \mu_* \rangle = \langle \mathcal{P}_t V, \mu_* \rangle \quad \text{for all } t \ge 0,$$

and by Definition 4.1.4

$$\langle V, \mu_* \rangle = \langle \mathcal{P}_t V, \mu_* \rangle \le C_V e^{-\gamma t} \langle V, \mu_* \rangle + K_V.$$

It follows that

$$\langle V,\, \mu_* \rangle \leq \frac{K_V}{(1 - C_V e^{-\gamma t})}, \tag{4.10}$$

for any fixed $t > \gamma^{-1} \ln C_V$.

As the authors in [44] observe, there is a last ingredient necessary to give a general form of Harris' theorem for Wasserstein-like distances: the semidistance is contracting for the semigroup $\{\mathcal{P}_t,\, t \geq 0\}$. In the original statement this condition is not made explicit as it is automatically verified by the total variation distance. In fact it is easy to see that the total variation distance is always contracting with respect to a Markov operator: given a Markov operator $\mathcal{P}$ and two probability measures μ_1, μ_2, by definition of the total variation distance (4.1), we have

$$\begin{aligned} d_{TV}(\mathcal{P}^*\mu_1, \mathcal{P}^*\mu_2) &= \sup_A |(\mathcal{P}^*\mu_1)(A) - (\mathcal{P}^*\mu_2)(A)| \\ &= \sup_A |\langle \mathcal{P}\mathbb{1}_A, \mu_1 \rangle - \langle \mathcal{P}\mathbb{1}_A, \mu_2 \rangle| \end{aligned}$$

and since $(\mathcal{P}\mathbb{1}_A)(x) = P(x, A) \in [0, 1]$ for any $A \in \mathcal{B}(\mathcal{H})$

$$\leq \sup_{f\,:\,|f| \leq 1} |\langle f, \mu_1 \rangle - \langle f, \mu_2 \rangle| = d_{TV}(\mu_1, \mu_2).$$

Therefore the price to pay for being able to use other distances-like functions and for a weaker concept of small set is to ask explicitly the distance-like function to be contracting for $\mathcal{P}_t$.

Definition 4.1.5 ([44, Definition 4.6])**.** Let $\mathcal{P}$ be a Markov operator on $\mathcal{H}$ with associated transition function $P(\cdot, \cdot)$. Given a time $t \geq 0$, a distance-like function $d : \mathcal{H} \times \mathcal{H} \to [0, 1]$ is called *contracting* for $\mathcal{P}$ if there exists $\alpha < 1$ such that for every pair $x, y \in \mathcal{H}$ with $d(x, y) < 1$

$$W_d(P(x, \cdot), P(y, \cdot)) \leq \alpha d(x, y). \tag{4.11}$$

Going through the proof of Theorem 4.8 in [44], namely the part dedicated to their general version of Harris' theorem, it is clear that first and foremost they demonstrate the following intermediate result:

Theorem 4.1.6 ([44, Theorem 4.8])**.** *Let $\mathcal{P}_t$, $t \geq 0$, be a Markov semigroup over $\mathcal{H}$ admitting a continuous Lyapunov function V. Suppose that there exists $T > 0$ and a distance-like function $d : \mathcal{H} \times \mathcal{H} \to [0, 1]$ which is contracting for $\mathcal{P}_T$, and such that the level set $\{x \in \mathcal{H} : V(x) \leq 4K_V\}$ is d-small for $\mathcal{P}_T$. Then $\{\mathcal{P}_t,\, t \geq 0\}$ has at most one invariant measure μ_*. Furthermore defining $\tilde{d}(x, y)^2 = d(x, y)(1 + V(x) + V(y))$, there exists $t_* > 0$ and $\rho < 1$ such that*

$$W_{\tilde{d}}(P_{t_*}(x, \cdot), P_{t_*}(y, \cdot)) \leq \rho \tilde{d}(x, y). \tag{4.12}$$

From it two important corollaries will follow, namely the exponential convergence of transition probabilities to the invariant measure, Corollary 4.1.7, and a spectral gap result, Corollary 4.1.8. The precise formulation of the rate of convergence in Corollary 4.1.7 will be particularly useful in section 5.3 and section 5.4 when we will show properties of the invariant measure of the stochastic Navier–Stokes equations and the stochastic two–layer QG model with respect to system parameters.

Corollary 4.1.7. *Theorem 4.1.6 implies that there exists $\gamma > 0$ such that given $\mu, \nu \in \mathcal{M}_1$*

$$W_{\tilde{d}}(\mathcal{P}_{kt_*}\mu, \mathcal{P}_{kt_*}\nu) \leq e^{-\gamma k t_*} W_{\tilde{d}}(\mu, \nu) \quad \textit{for all } x \in \mathcal{H},\ k \in \mathbb{N}. \tag{4.13}$$

Furthermore if μ_* is the invariant measure for $\mathcal{P}_t$, there exists $C > 0$ such that

$$W_{\tilde{d}}(P_{kt_*}(x, \cdot), \mu_*) \leq C(1 + V(x))e^{-\gamma k t_*} \quad \textit{for all } x \in \mathcal{H}, k \in \mathbb{N}. \tag{4.14}$$

Proof. First of all we can prove that, as a consequence of (4.12), we have

$$W_{\tilde{d}}(P_{kt_*}(x, \cdot), P_{kt_*}(y, \cdot)) \leq \rho^k \tilde{d}(x, y) \quad \text{for all } k \in \mathbb{N}. \tag{4.15}$$

By the semigroup property for all $t, s \geq 0$

$$P_{t+s}(x, dv) = \int_{\mathcal{H}} P_t(z, dv) P_s(x, dz) = \int_{\mathcal{H}} P_s(z, dv) P_t(x, dz).$$

For any $x, y \in \mathcal{H}$ and $k \in \mathbb{N}$, let $\Gamma^{(x,y)}_{kt_*}$ be the optimal coupling for $P_{kt_*}(x, \cdot)$ and $P_{kt_*}(y, \cdot)$, which we know exists by Lemma 4.1.2, and define the measure on $\mathcal{H} \times \mathcal{H}$

$$\pi^{(x,y)}_{kt_*}(dv_1, dv_2) := \iint_{\mathcal{H}\times\mathcal{H}} \Gamma^{(z_1,z_2)}_{t_*}(dv_1, dv_2) \Gamma^{(x,y)}_{(k-1)t_*}(dz_1, dz_2).$$

It is easy to see that $\pi^{(x,y)}_{kt_*}$ is a coupling for $P_{kt_*}(x, \cdot)$ and $P_{kt_*}(y, \cdot)$ and consequently

$$\begin{aligned} W_{\tilde{d}}(P_{kt_*}(x, \cdot), P_{kt_*}(y, \cdot)) &\leq \int \tilde{d}(v_1, v_2)\, \pi^{(x,y)}_{kt_*}(dv_1, dv_2) \\ &= \iint \tilde{d}(v_1, v_2)\, \Gamma^{(z_1,z_2)}_{t_*}(dv_1, dv_2)\, \Gamma^{(x,y)}_{(k-1)t_*}(dz_1, dz_2) \\ &= \int W_{\tilde{d}}(P_{t_*}(z_1, \cdot), P_{t_*}(z_2, \cdot))\, \Gamma^{(x,y)}_{(k-1)t_*}(dz_1, dz_2) \end{aligned}$$

since $\Gamma_{t_*}^{(z_1,z_2)}$ is the optimal coupling. Therefore using the estimate (4.12)

$$W_{\tilde{d}}(P_{kt_*}(x,\cdot),P_{kt_*}(y,\cdot)) \le \rho \int \tilde{d}(z_1,z_2)\Gamma_{(k-1)t}^{(x,y)}(dz_1,dz_2)$$
$$= \rho W_{\tilde{d}}(P_{(k-1)t_*}(x,\cdot),P_{(k-1)t_*}(y,\cdot)).$$

In particular for $k=2$ we have

$$W_{\tilde{d}}(P_{2t_*}(x,\cdot),P_{2t_*}(y,\cdot)) \le \rho W_{\tilde{d}}(P_{t_*}(x,\cdot),P_{t_*}(y,\cdot)) \le \rho^2 \tilde{d}(x,y)$$

and (4.15) follows iteratively.

Let us now show that (4.13) holds. As in the proof of [44, Theorem 4.8] we observe that, since $W_{\tilde{d}}$ is convex, by Jensen inequality

$$W_{\tilde{d}}(\mathcal{P}_{kt_*}\mu,\mathcal{P}_{kt_*}\nu) \le \int W_{\tilde{d}}(P_{kt_*}(x,\cdot),P_{kt_*}(y,\cdot))\,\Gamma(dx,dy),$$

for any $\mu,\nu \in \mathcal{M}_1$ and coupling $\Gamma \in \mathcal{C}(\mu,\nu)$. Then by (4.15) we have

$$W_{\tilde{d}}(\mathcal{P}_{kt_*}\mu,\mathcal{P}_{kt_*}\nu) \le \rho^k \int \tilde{d}(x,y)\,\Gamma(dx,dy),$$

for any $\Gamma \in \mathcal{C}(\mu,\nu)$, hence setting $\gamma := -t_*^{-1}\ln\rho > 0$ follows

$$W_{\tilde{d}}(\mathcal{P}_{kt_*}\mu,\mathcal{P}_{kt_*}\nu) \le \rho^k W_{\tilde{d}}(\mu,\nu) = e^{-\gamma k t_*} W_{\tilde{d}}(\mu,\nu).$$

Finally it is easy to see that (4.14) holds. In fact setting $\mu=\delta_x$, Dirac measure, and $\nu=\mu_*$, the invariant measure, (4.13) gives

$$W_{\tilde{d}}(P_{kt_*}(x,\cdot),\mu_*) \le e^{-\gamma k t_*} W_{\tilde{d}}(\delta_x,\mu_*)$$
$$\le e^{-\gamma k t_*}\int \tilde{d}(x,y)\,\mu_*(dy).$$

By the definition of $\tilde{d}$

$$\int \tilde{d}(x,y)\,\mu_*(dy) \le \int (1+V(x)+V(y))^{1/2}\,\mu_*(dy)$$
$$\le 1+V(x)+\int V(y)^{1/2}\,\mu_*(dy).$$

Given the bound (4.10) we have

$$\int \tilde{d}(x,y)\,\mu_*(dy) \le 1+V(x)+\left(\frac{K_V}{(1-C_V e^{-\gamma s})}\right)^{1/2},$$

for any $s \ge \gamma^{-1}\ln C_V$, hence there exists $C>0$ such that

$$W_{\tilde{d}}(P_{kt_*}(x,\cdot),\mu_*) \le C(1+V(x))e^{-\gamma k t_*}.$$

□

As pointed out in [44, Remark 4.10] if the assumptions of Theorem 4.1.6 hold uniformly in t_* belonging to an open interval of times, then Corollary 4.1.7 holds for all t and not just for integer multiples of t_*. In particular in [11], Theorem 2.4 shows that (4.14) can be extended to all positive times $t \geq 0$ i.e.

$$W_{\tilde{d}}(P_t(x,\cdot),\mu_*) \leq C(1+V(x))e^{-\gamma t} \quad \text{for all } \ x \in \mathcal{H}, t \geq 0. \tag{4.16}$$

From Theorem 4.1.6 we can derive a further property for the semigroup $\mathcal{P}_t$ which will prove crucial in the next chapter to establish regular dependence of the invariant measure with respect to the parameters. Recall that μ_* being invariant for $\mathcal{P}_t$ means that μ_* is eigenvector of $\mathcal{P}_t^*$ with eigenvalue 1. Then the next corollary implies that all other eigenvalues are contained in the disk of radius $e^{-\gamma t}$ around the origin, or we may say $\mathcal{P}_t^*$ exhibits a spectral gap.

Corollary 4.1.8 (Spectral gap)**.** *Suppose all conditions in Theorem 4.1.6 are met and let $\|\cdot\|_{\tilde{d}}$ be the Lipschitz seminorm* (4.4) *associated to the distance-like function $\tilde{d}$. Then there exists $\rho < 1$ and $t_* > 0$ such that*

$$\|\mathcal{P}_{t_*}\varphi - \langle\varphi,\mu_*\rangle\|_{\tilde{d}} \leq \rho\|\varphi - \langle\varphi,\mu_*\rangle\|_{\tilde{d}} \tag{4.17}$$

for all $\varphi : \mathcal{H} \to \mathbb{R}$ such that $\|\varphi\|_{\tilde{d}} < \infty$.

Proof. By definition of the Lipschitz norm (4.4) and of transition semigroup we have

$$\|\mathcal{P}_t\varphi - \langle\varphi,\mu_*\rangle\|_{\tilde{d}} = \sup_{x\neq y} \frac{|\langle\varphi, P_t(x,\cdot)\rangle - \langle\varphi, P_t(y,\cdot)|}{\tilde{d}(x,y)} \tag{4.18}$$

By (4.7), the weak version of the Kantorovich-Rubinstein formula, we have

$$\|\mathcal{P}_t\varphi - \langle\varphi,\mu_*\rangle\|_{\tilde{d}} \leq \|\varphi\|_{\tilde{d}} \sup_{x\neq y} \frac{W_{\tilde{d}}\left(P_t(x,\cdot), P_t(y,\cdot)\right)}{\tilde{d}(x,y)}. \tag{4.19}$$

Now Theorem 4.1.6 ensures that there exists $t_* > 0$ and $\rho < 1$ such that

$$W_{\tilde{d}}\left(P_{t_*}(x,\cdot), P_{t_*}(y,\cdot)\right) \leq \rho\tilde{d}(x,y)$$

and so

$$\|\mathcal{P}_{t_*}\varphi - \langle\varphi,\mu_*\rangle\|_{\tilde{d}} \leq \rho\|\varphi\|_{\tilde{d}} = \rho\|\varphi - \langle\varphi,\mu_*\rangle\|_{\tilde{d}} \tag{4.20}$$

as desired. □

In [44] the authors proceed to show the general form of Harris' theorem to hold for stochastic delay equations using the asymptotic coupling approach. Recently [12] provided a set of general verifiable conditions for nonlinear dissipative stochastic model to apply the asymptotic coupling approach not only to show the uniqueness of the invariant measure as seen in the previous chapter, but to have an explicit rate of convergence of the transition probabilities to it. More precisely the assumptions in [12] are shown to be sufficient to apply the main result in [11], i.e. subgeometric rate of convergence to the invariant measure, as well as the results from [44] presented so far, i.e. exponential rate of convergence. However the result in [12] may appear challenging to navigate when one wants to retrieve the formulation of the appropriate distance-like function d and see immediately how the conditions for SPDEs imply Theorem 4.1.6.

In the next section we will open this tool box for SPDEs provided by [12] and give a compact proof of how their conditions gives a contracting semimetric d and the existence of a Lyapunov function with level sets which are d-small.

4.1.1 Framework for SPDEs

Let $(\mathcal{H}, |\cdot|)$ and $(\mathcal{V}, \|\cdot\|)$ be Hilbert spaces with $\mathcal{V} \subset\subset \mathcal{H}$ and consider the stochastic equation

$$dX = (AX + F(X))\, dt + dW, \quad X(0) = x \tag{4.21}$$

where A is a nonnegative linear operator, W is a Wiener process on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathcal{H}$ and trace class covariance operator Q, and F a nonlinear function such that there exists a unique solution for any initial condition $X(0) = x$. As for Navier–Stokes equations and the two–layer quasi–geostrophic equations, we assume the solution to be in $C([0,T];\mathcal{H}) \cap L^2(0,T;\mathcal{V})$ for all $T > 0$ and a.a. $\omega \in \Omega$, and to be continuous with respect to initial condition. Then define the Markov semigroup $\mathcal{P}_t$ as

$$\mathcal{P}_t \varphi(x) = \mathbb{E}\, \varphi(X(t;x)) \tag{4.22}$$

for all $\varphi : \mathcal{H} \to \mathbb{R}$ bounded measurable functions.

As in the previous chapter consider $\tilde{Y}$ solution of

$$d\tilde{Y} = \left(A\tilde{Y} + F(\tilde{Y}) + G(X, \tilde{Y})\right) dt + dW, \quad Y(0) = y \tag{4.23}$$

with $G : \mathcal{H} \times \mathcal{H} \to \mathcal{H}_n$ where $\mathcal{H}_n$ is an n-dimensional subspace of $\mathcal{H}$. In particular given the expansion of the Wiener process

$$W(t) = W_n(t) + W^n(t) = \sum_{k=1}^{n} \sqrt{\sigma_k}\beta_k(t) e_k + \sum_{k=n+1}^{\infty} \sqrt{\sigma_k}\beta_k(t) e_k, \tag{4.24}$$

where W_n and W^n are Wiener processes with covariance matrices respectively Q_n and Q^n, as in (3.54), the expression $|Q_n^{-1/2}G(X(s),\tilde{Y}(s))|^2$ is well defined.

We are now ready to state the set of verifiable conditions laid out in [12] for SPDEs as (4.21).

Assumption A There exists a $n > 0$ and a finite dimensional control $G : \mathcal{H} \times \mathcal{H} \to \mathcal{H}_n$ such that X, solution of (4.21) and $\tilde{Y}$, solution of the controlled equation (4.23), satisfy the following conditions

A1 There exist $\kappa_0 > 0$ and $\kappa_1 \geq 0$ such that for all $t \geq 0$

$$|X(t) - \tilde{Y}(t)|^2 \leq |x-y|^2 \exp\left(-\kappa_0 t + \kappa_1 \int_0^t \|X(s)\|^2\, ds\right). \tag{4.25}$$

A2 There exists $\kappa_2 > 0$, $\kappa_3 \geq 0$ and a random variable Ξ_γ such that

$$|X(t)|^2 + \kappa_2 \int_0^t \|X(s)\|^2\, ds \leq |x|^2 + \kappa_3 t + \Xi_\gamma \quad t \geq 0 \tag{4.26}$$

with $\kappa_0 > \kappa_1\kappa_3/\kappa_2$ and

$$\mathbb{P}(\Xi_\gamma \geq R) \leq e^{-2\gamma R} \quad R \geq 0. \tag{4.27}$$

A3 There exists a positive constant $c > 0$ such that for each $t \geq 0$ and $s \in [0,t]$

$$|G(X(s),\tilde{Y}(s))|^2 \leq c|X(s) - \tilde{Y}(s)|^2. \tag{4.28}$$

A4 There exists a measurable function $V : \mathcal{H} \to \mathbb{R}_+$ such that for some $\gamma_1 > 0$, $K > 0$

$$\mathbb{E}V(X(t)) \leq \mathbb{E}V(X(s)) + \int_s^t (-\gamma_1 \mathbb{E}V(X(\tau)) + K)\, d\tau, \quad t \geq s \geq 0. \tag{4.29}$$

Furthermore for any $M > 0$ the function $x \mapsto |x|^2$ is bounded on the level sets $\{V \leq M\}$.

Given κ_1, κ_2 from **A1** and **A2** respectively define the premetric θ_α depending on a positive parameter α as

$$\theta_\alpha(x,y) := |x-y|^{2\alpha} e^{\alpha \upsilon |x|^2} \quad \text{with } \upsilon := \frac{\kappa_1}{\kappa_2} \tag{4.30}$$

and given $N \in \mathbb{N}$ define a distance-like function d_N as

$$d_N(x,y) := N\theta_\alpha(x,y) \wedge N\theta_\alpha(y,x) \wedge 1. \tag{4.31}$$

We will see in Theorem 4.1.10 that, thanks to Assumption A, there exists α_0 such that for all $\alpha \in (0, \alpha_0)$ the conditions of Theorem 4.1.6 are satisfied with respect to the associated distance-like function d_N for N large enough.

First though we show that thanks to Assumption A, the process $\tilde{Y}$ is absolutely continuous with respect to the solution of

$$dY = (AY + F(Y))\, dt + dW, \quad Y(0) = y \neq x. \tag{4.32}$$

A similar result has already be shown in the proof of Theorem 3.2.8 with the main difference that now we do not restrict the equivalence to the set $\{\tau = \infty\}$.

Proposition 4.1.9. *Let the process Y be the solution of* (4.32) *so that $P_t(y, \cdot) = \mathrm{Law}(Y(t))$. Then if Assumptions **A1-A3** hold, $\tilde{Y}$ solution of the controlled equation* (4.23) *is absolutely continuous with respect to Y.*

Proof. Given the Wiener process W and its decomposition (4.24) we define the process

$$\tilde{W}_n(t) = W_n(t) + \int_0^t G(X, \tilde{Y})\, ds. \tag{4.33}$$

Then, as seen in the previous chapter, $\tilde{W} = \tilde{W}_n + W^n$ is absolutely continuous with respect to $W = W_n + W^n$ if $\tilde{W}_n$ is absolutely continuous with respect to W_n, and Girsanov theorem helps showing the latter. To make use of Girsanov theorem, we have to ensure that the hypothesis of Lemma 3.2.6 is satisfied i.e. for all $T > 0$

$$\int_0^T \left| Q_n^{-1/2} G(X, \tilde{Y}) \right|^2 ds < \infty \quad \mathbb{P}\text{-a.s.} \tag{4.34}$$

By **A3** and **A1** we have that for any $T \geq 0$ and $t \in [0, T]$

$$|G(X(t), \tilde{Y}(t))|^2 \leq c|x - y|^2 \exp\Big(-\kappa_0 t + \kappa_1 \int_0^t \|X(s)\|^2\, ds \Big),$$

and by **A2**

$$|G(X(t), \tilde{Y}(t))|^2 \leq c|x - y|^2 \exp\Big(- \left(\kappa_0 - \tfrac{\kappa_1 \kappa_3}{\kappa_2}\right) t + \tfrac{\kappa_1}{\kappa_2} \left(|x|^2 + \Xi_\gamma\right) \Big).$$

It follows that for any $T \geq 0$

$$\begin{aligned} \int_0^T |Q_n^{-1/2} G(X(t), \tilde{Y}(t))|^2\, dt &\leq \|Q_n^{-1/2}\|^2 \int_0^T |G(X(t), Y(t))|^2\, dt \\ &\leq \tfrac{c\|Q_n^{-1/2}\|^2}{\chi} |x - y|^2 \exp\Big(\tfrac{\kappa_1}{\kappa_2} \left(|x|^2 + \Xi_\gamma\right) \Big) \left(1 - e^{-\chi T}\right) \end{aligned} \tag{4.35}$$

where $\chi = \kappa_0 - \frac{\kappa_1\kappa_3}{\kappa_2} > 0$. Note that, given the estimate (4.27), the random variable Ξ_γ is almost surely finite. Then (4.34) holds as desired and $\tilde{Y}$ is absolutely continuous with respect to Y on $[0,T]$ for all $T > 0$. □

We are ready to state the main theorem of this section which is a summary of Theorem 4.2 and Theorem 2.4 in [12]. Such a result combined with Theorem 4.1.6 provides a synthesis of the results in section 2.2 and section 4.2 in [12].

Theorem 4.1.10. *Suppose Assumption A holds and, given κ_1, κ_2 and γ as in **A1** and **A2**, set*

$$\upsilon = \frac{\kappa_1}{\kappa_2} \quad \text{and} \quad \alpha_0 = \frac{1}{2} \wedge \frac{2\gamma}{\upsilon + 2\gamma}. \tag{4.36}$$

*Then the function V given by **A4** is a Lyapunov function with $C_V = 1$ and $K_V = K/\gamma_1$. Moreover, for all $\alpha \in (0, \alpha_0)$, there exists $N_* \in \mathbb{N}$ and $t_* \in \mathbb{R}_+$ such that for all $t > t_*$ and $N \in \mathbb{N}$ such that $N > N_*$:*

- *(i) the distance-like function d_N (4.31) is contracting for $\mathcal{P}_t$;*
- *(ii) the level set $\{x \in \mathcal{H} : V(x) \le 4K_V\}$ is d_N-small.*

In order to prove this result we introduce first the following three lemmas. The first one is the so-called gluing lemma, see e.g. [72, pg.11] or [49, Lemma 4.3.2].

Lemma 4.1.11 (Gluing lemma)**.** *Let (ξ, η) and (ξ', η') be two pairs of random elements valued in a Borel measurable space $(X, \mathscr{X})$ such that η and ξ' have the same distribution. Then on a proper probability space, there exist three random elements $\zeta_1, \zeta_2, \zeta_3$ such that the law of (ζ_1, ζ_2) in $(X \times X, \mathscr{X} \otimes \mathscr{X})$ coincides with the law of (ξ, η) and the law of (ζ_2, ζ_3) coincides with the law of (ξ', η').*

The second lemma gives upper bounds for the total variation distance of two finite dimensional Wiener processes like W_n and $\tilde{W}_n$ as in (4.33).

Lemma 4.1.12 ([12, Theorem A.5])**.** *Let B be a n-dimensional Wiener process, $(h(t))_{t\ge 0}$ a progressively measurable n-dimensional process and define*

$$\tilde{B}(t) = \int_0^t h(s)\,ds + B(t) \quad t \ge 0.$$

Fix $t > 0$, then, if for some $\delta \in (0,1)$

$$M_\delta := \mathbb{E}\left(\int_0^t |h(s)|^2\,ds\right)^\delta < \infty \tag{4.37}$$

the following bounds hold:

$$d_{TV}(\mathrm{Law}(B(s))_{s\le t}, \mathrm{Law}(\tilde{B}(s))_{s\le t}) \le 2^{(1-\delta)/(1+\delta)} M_\delta^{1/(1+\delta)}; \tag{4.38}$$

$$d_{TV}(\mathrm{Law}(B(s))_{s\le t}, \mathrm{Law}(\tilde{B}(s))_{s\le t}) \le 1 - \frac{1}{6}\min\left(\frac{1}{8}, \exp\bigl(-(2^{2-\delta}M_\delta)^{1/\delta}\bigr)\right). \tag{4.39}$$

The last lemma gives an estimate for the distance of two transition probabilities with different initial condition which will be the first block in building the proof of Theorem 4.1.10.

Lemma 4.1.13. *Let* $X(t)$ *and* $Y(t)$ *be the solutions of* (4.21) *and* (4.32) *respectively and* $P_t(x,\cdot), P_t(y,\cdot)$ *their respective laws. If Assumption A holds, then there exists* $\alpha_0 > 0$ *such that for all* $\alpha \in (0,\alpha_0)$ *and* $N \in \mathbb{N}$ *there exists two positive constants* C_Ξ *and* γ *such that*

$$W_{d_N}(P_t(x,.), P_t(y,.)) \le \tfrac{1}{2} d_{TV}(\mathrm{Law}(W_n(t))_{s\le t}, \mathrm{Law}(\tilde{W}_n(t))_{s\le t}) + N C_\Xi \theta_\alpha(x,y) e^{-\chi\alpha t}, \tag{4.40}$$

for all $x, y \in \mathcal{H}$.

Proof. First of all we show that by means of the coupling lemma and the gluing lemma

$$W_{d_N}(P_t(x,.), P_t(y,.)) \le \tfrac{1}{2} d_{TV}(P_t(y,\cdot), \mathrm{Law}\,\tilde{Y}(t)) + \mathbb{E}[d_N(X(t), \tilde{Y}(t))].$$

By the coupling lemma we know that there exists a coupling $(Y_*, \tilde{Y}_*) \in \mathcal{C}(P_t(y,\cdot), \mathrm{Law}(\tilde{Y}(t)))$ such that

$$d_{TV}(P_t(y,\cdot), \mathrm{Law}(\tilde{Y}(t))) = 2\mathbb{P}(Y_* \ne \tilde{Y}_*). \tag{4.41}$$

Consider the two pairs $(X, \tilde{Y})$ and $(\tilde{Y}_*, Y_*)$, then since $\tilde{Y}_*$ and $\tilde{Y}$ have the same law, the gluing lemma Lemma 4.1.11 ensures that there exist three random variables Z_1, Z_2, Z_3 such that $\mathrm{Law}(Z_1, Z_2) = \mathrm{Law}(X, \tilde{Y})$ and $\mathrm{Law}(Z_2, Z_3) = \mathrm{Law}(\tilde{Y}_*, Y_*)$. Therefore

$$\mathrm{Law}(Z_2) = \mathrm{Law}(\tilde{Y}_*) = \mathrm{Law}(\tilde{Y}) \quad \text{and} \quad \mathrm{Law}(Z_3) = \mathrm{Law}(Y_*) = \mathrm{Law}(Y).$$

It follows that (Z_1, Z_3) is a coupling of $\mathrm{Law}(X)$, $\mathrm{Law}(Y)$ and as a consequence

$$W_{d_N}(P_t(x,\cdot), P_t(y,\cdot)) \le \mathbb{E}\, d_N(Z_1, Z_3).$$

Therefore, using the fact that $d_N \leq 1$,

$$\begin{aligned} W_{d_N}(P_t(x,\cdot),P_t(y,\cdot)) &\leq \mathbb{E}[d_N(Z_1,Z_3)\mathbb{1}(Z_2 \neq Z_3)] \\ &\quad + \mathbb{E}[d_N(Z_1,Z_3)\mathbb{1}(Z_2 = Z_3)] \\ &\leq \mathbb{E}[\mathbb{1}(Z_2 \neq Z_3)] + \mathbb{E}[d_N(Z_1,Z_2)]. \end{aligned} \tag{4.42}$$

Now, by the fact that $\text{Law}(Z_2,Z_3) = \text{Law}(\tilde{Y}_*,Y_*)$ and (4.41)

$$\mathbb{E}[\mathbb{1}(Z_2 \neq Z_3)] = \mathbb{E}[\mathbb{1}(\tilde{Y}_* \neq Y_*)] = \tfrac{1}{2} d_{TV}(P_t(y,\cdot),\text{Law}(\tilde{Y}(t))).$$

Back to Equation (4.42), since $\text{Law}(Z_1,Z_2) = \text{Law}(X,\tilde{Y})$, we have the desired result

$$W_{d_N}(P_t(x,.),P_t(y,.)) \leq \tfrac{1}{2} \underbrace{d_{TV}(P_t(y,\cdot),\text{Law}\,\tilde{Y}(t))}_{(I)} + \underbrace{\mathbb{E}[d_N(X(t),\tilde{Y}(t))]}_{(II)}. \tag{4.43}$$

Next we look at the two terms separately starting from (I). From the theory of stochastic differential equations the solutions of (4.32) and (4.23) can be seen as image via a measurable function Φ^y of their driving noise, W and $\tilde{W}$ respectively i.e. $Y(t) = \Phi^y((W(s))_{s\leq t})$ and $\tilde{Y} = \Phi^y((\tilde{W}(s))_{s\leq t})$. Then

$$d_{TV}(\text{Law}\,Y(t),\text{Law}\,\tilde{Y}(t)) \leq d_{TV}(\text{Law}(W(s))_{s\leq t},\text{Law}(\tilde{W}(s))_{s\leq t}). \tag{4.44}$$

By the coupling lemma there exists a coupling $(B_*,\tilde{B}_*)$ of $\text{Law}(W(t))_{t\geq 0}$ and $\text{Law}(\tilde{W}(t))_{t\geq 0}$ such that

$$d_{TV}(\text{Law}(W(t))_{t\geq 0},\text{Law}(\tilde{W}(t))_{t\geq 0}) = 2\mathbb{P}(B_* \neq \tilde{B}_*)$$

and a coupling $(\xi_*,\tilde{\xi}_*)$ of the finite dimensional Wiener processes $\text{Law}(W_n(t))_{t\geq 0}$ and $\text{Law}(\tilde{W}_n(t))_{t\geq 0}$ such that

$$d_{TV}(\text{Law}(W_n(t))_{t\geq 0},\text{Law}(\tilde{W}_n(t))_{t\geq 0}) = 2\mathbb{P}(\xi_* \neq \tilde{\xi}_*).$$

Since the Wiener processes $W = W_n + W^n$ and $\tilde{W} = \tilde{W}_n + W^n$ differ only for the their first n-dimensional part we have that for all couplings $\xi,\tilde{\xi} \in \mathcal{C}(\text{Law}(W_n(s))_{s\leq t},\text{Law}(\tilde{W}_n(s))_{s\leq t})$

$$\mathbb{P}(B_* \neq \tilde{B}_*) \leq \mathbb{P}(\xi \neq \tilde{\xi}).$$

Then, by the definition of optimal coupling

$$d_{TV}(\text{Law}(W(s))_{s\leq t},\text{Law}(\tilde{W}(s))_{s\leq t}) \leq d_{TV}(\text{Law}(W_n(s))_{s\leq t},\text{Law}(\tilde{W}_n(s))_{s\leq t})$$

giving the first part of (4.40).

Next we look at (II) in (4.43). By definition of d_N (4.31) we have

$$\mathbb{E}\, d_N(X(t), \tilde{Y}(t)) \leq N\mathbb{E}\,\theta_\alpha(X(t), \tilde{Y}(y)). \tag{4.45}$$

Combining Assumption **A1** and **A2** we have that

$$|X(t) - \tilde{Y}(t)|^2 \leq |x-y|^2 \exp\big(-\chi t + \upsilon(|x|^2 - |X(t)|^2 + \Xi_\gamma)\big) \tag{4.46}$$

namely, for all $\alpha > 0$

$$|X(t) - \tilde{Y}(t)|^{2\alpha} e^{\alpha\upsilon|X(t)|^2} \leq |x-y|^{2\alpha} e^{\alpha\upsilon|x|^2} \exp(-\alpha\chi t + \alpha\upsilon\Xi_\gamma).$$

Therefore by the definition of the premetric θ_α (4.30)

$$\mathbb{E}\,\theta_\alpha(X(t), \tilde{Y}(t)) \leq C_\Xi \theta_\alpha(x,y)\exp(-\chi\alpha t) \tag{4.47}$$

where

$$C_\Xi := \mathbb{E}\, \exp(\upsilon\alpha\Xi_\gamma) < \infty \quad \text{if } \upsilon\alpha < 2\gamma \tag{4.48}$$

Putting together these results in (4.43) we have that (4.40) holds for all $\alpha \in (0, \alpha_0)$ setting $\alpha_0 := 2\gamma/\upsilon$.

□

We are now ready to show Theorem 4.1.10:

Proof of Theorem 4.1.10. First we show that V is a Lyapunov function as in Definition 4.1.4 thanks to a classic comparison theorem Lemma A.3. By **A4** we know that there exists γ_1, K strictly positive such that

$$\mathcal{P}_t V(x) - \mathcal{P}_s V(x) \leq \int_s^t (-\gamma_1 \mathcal{P}_\tau V(x) + K)\, d\tau, \quad t \geq s \geq 0. \tag{4.49}$$

Therefore the hypothesis of Lemma A.3 are satisfied by the function $f(t) := \mathcal{P}_t V(x)$, which has non-negative values and it is continuous in time, and we have the desired result

$$\mathcal{P}_t V(x) \leq e^{-\gamma_1 t} V(x) + K_V. \tag{4.50}$$

with $K_V = K/\gamma_1$

(i) the semigroup $\mathcal{P}_t$ is d_N-contracting. We have to show that there exists $\rho < 1$ such that

$$W_{d_N}(P_t(x,\cdot), P_t(y,\cdot)) \le \rho\, d_N(x,y)$$

for all x, y such that $d_N(x,y) < 1$. Thanks to Lemma 4.1.13 we have that for all $\alpha < 2\gamma/\upsilon$

$$W_{d_N}(P_t(x,.), P_t(y,.)) \le \tfrac{1}{2} d_{TV}(\text{Law}\, W_n(t), \text{Law}\, \tilde{W}_n(t)) \\ + N C_\Xi \theta_\alpha(x,y) e^{-\chi\alpha t} \quad (4.51)$$

We want to apply Lemma 4.1.12 to bound the first term on the right hand side so we have to ensure that (4.37) holds. As seen in Proposition 4.1.9, in our context

$$M_\delta = \mathbb{E}\left(\int_0^t |Q_n^{-1/2} G(X,\tilde{Y})|^2\, ds\right)^\delta. \quad (4.52)$$

and we have the bound (4.35), i.e.

$$\int_0^t |Q_n^{-1/2} G(X,\tilde{Y})|^2\, ds \le \tfrac{c\|Q_n^{-1/2}\|^2}{\chi} |x-y|^2 \exp\big(\upsilon\left(|x|^2 + \Xi_\gamma\right)\big)\left(1 - e^{-\chi t}\right)$$

where

$$\upsilon := \frac{\kappa_1}{\kappa_2} \quad \text{and} \quad \chi := \kappa_0 - \upsilon\kappa_3 > 0.$$

By the properties of Ξ_γ (4.27) we have that

$$\mathbb{E}\, \exp(\upsilon\delta\Xi_\gamma) < \infty \quad \text{if } \upsilon\delta < 2\gamma$$

and so for all $0 < \delta < (2\gamma/\upsilon) \wedge 1$

$$M_\delta \le C_\delta |x-y|^{2\delta} \exp\big(\upsilon\delta|x|^2\big) \quad \text{with } C_\delta = \left(\tfrac{c\|Q_n^{-1/2}\|^2}{\chi}\right)^\delta \mathbb{E}\, e^{\upsilon\delta\Xi_\gamma}. \quad (4.53)$$

Therefore, since condition (4.37) holds, the bound (4.38) in Lemma 4.1.12 and (4.53) give

$$d_{TV}(\text{Law}\, W_n(t)), \text{Law}\, \tilde{W}_n(t)) \le 2^{(1-\delta)/(1+\delta)} M_\delta^{\frac{1}{1+\delta}} \\ \le C_\delta \left(|x-y|^2 \exp\big(\upsilon|x|^2\big)\right)^{\frac{\delta}{1+\delta}}$$

where we relabelled the constant C_δ appropriately. Since $\delta \in (0, (2\gamma/\upsilon) \wedge 1)$, the exponent $\delta/(1+\delta) =: \alpha$ is in the interval

$$0 < \alpha < \alpha_0 = \frac{1}{2} \wedge \frac{2\gamma}{\upsilon + 2\gamma} < \frac{2\gamma}{\upsilon}.$$

which is consistent with (4.47). Given the definition of the premetric θ_α (4.30) we have then shown that for all $\alpha \in (0, \alpha_0)$ there exists $C_\alpha > 0$ such that

$$d_{TV}(\mathrm{Law}\, W_n(t)), \mathrm{Law}\, \tilde{W}_n(t)) \leq C_\alpha \theta_\alpha(x, y) \quad \text{for all } \; x, y \in \mathcal{H} \; x \neq y. \tag{4.54}$$

Going back to (4.51), we have proved that

$$\begin{aligned} W_{d_N}(P_t(x, \cdot\,), P_t(y, \cdot\,)) &\leq C_\alpha \theta_\alpha(x, y) + N C_\Xi e^{-\chi\alpha t} \theta_\alpha(x, y) \\ &= N\theta_\alpha(x, y) \left(C_\alpha N^{-1} + C_\Xi e^{-\chi\alpha t} \right) \end{aligned}$$

and, inverting the roles of x and y, we get in the same way

$$W_{d_N}(P_t(x, \cdot\,), P_t(y, \cdot\,)) \leq N\theta_\alpha(y, x) \left(C_\alpha N^{-1} + C_\Xi e^{-\chi\alpha t} \right).$$

Therefore

$$W_{d_N}(P_t(x, \cdot\,), P_t(y, \cdot\,)) \leq (N\theta_\alpha(x, y) \wedge N\theta_\alpha(y, x)) \left(C_\alpha N^{-1} + C_\Xi e^{-\chi\alpha t} \right) \tag{4.55}$$

and since x, y are assumed to be such that $d_N(x, y) < 1$ we have that

$$W_{d_N}(P_t(x, \cdot\,), P_t(y, \cdot\,)) \leq d_N(x, y) \left(C_\alpha N^{-1} + C_\Xi e^{-\chi\alpha t} \right).$$

Then for all $N \in \mathbb{N}$ and $t \in \mathbb{R}_+$ such that

$$\rho := C_\alpha N^{-1} + C_\Xi e^{-\chi\alpha t} < 1 \tag{4.56}$$

we showed that $\mathcal{P}_t$ is d_N-contracting.

(ii) The sublevel set of V is d_N-small. We have now to prove that there exists $\varepsilon > 0$ such that

$$W_{d_N}(P_t(x, \cdot), P_t(y, \cdot)) \leq 1 - \varepsilon \tag{4.57}$$

for all $x, y \in \{z \in \mathcal{H} \,:\, V(z) \leq 4K_V\}$.

In the previous step we have shown that (4.55) holds for all x, y. Since, by Assumption **A4**, $z \mapsto |z|^2$ is bounded over $\{V \leq 4K_V\}$ so it is $\theta_\alpha(x, y)$ i.e. for any $K \geq 0$ there exists $C_K > 0$ such that

$$\theta_\alpha(x, y) = |x - y|^{2\alpha} e^{\alpha v |x|^2} < C_K. \tag{4.58}$$

for all $x, y \in \{V \leq 4K_V\}$. It is clear though that we cannot make $C_\alpha + NC_\Xi e^{-\chi\alpha t}$ arbitrarily small to have (4.57). However we can do it estimating differently $d_{TV}(\mathrm{Law}\, W_n(t), \mathrm{Law}\, \tilde{W}_n(t))$.

Thanks to (4.39) in Lemma 4.1.12, we have that

$$d_{TV}(\mathrm{Law}\, W_n(t), \mathrm{Law}\, \tilde{W}_n(t)) \le 1 - \frac{1}{6}\min\left(\frac{1}{8}, \exp\bigl(-(2^{2-\delta}M_\delta)^{1/\delta}\bigr)\right),$$

where M_δ is as in (4.52). Thanks to (4.53) and (4.58) it follows that

$$M_\delta \le C_\delta |x-y|^{2\delta}\exp\bigl(\upsilon\delta|x|^2\bigr) \le C_\delta C_K$$

for all $0 < \delta < (2\gamma/\upsilon) \wedge 1$. Setting $\varepsilon_1(\delta)$ to be

$$\varepsilon_1 = \frac{1}{6}\min\left(\frac{1}{8}, \exp\bigl(-(2^{2-\delta}C_\delta C_K)^{1/\delta}\bigr)\right), \tag{4.59}$$

we can have that

$$d_{TV}(\mathrm{Law}\, W_n(t), \mathrm{Law}\, \tilde{W}_n(t)) \le 1 - \varepsilon_1. \tag{4.60}$$

Then from (4.51) i.e.

$$W_{\tilde{d}}\left(P_t(x,\cdot), P_t(y,\cdot)\right) \le \tfrac{1}{2} d_{TV}(\mathrm{Law}\, W_n(t), \mathrm{Law}\, \tilde{W}_n(t)) + NC_\Xi e^{-\alpha\chi t}\theta_\alpha(x,y)$$

by (4.58) and (4.60) we have

$$W_{\tilde{d}}\left(P_t(x,\cdot), P_t(y,\cdot)\right) \le 1 - \frac{1+\varepsilon_1}{2} + NC_K C_\Xi e^{-\chi\alpha t}.$$

Picking $N \in \mathbb{N}$ and $t > 0$ such that

$$\frac{1+\varepsilon_1}{2} - NC_K C_\Xi e^{-\chi\alpha t} =: \varepsilon > 0$$

then the level set $\{x \in \mathcal{H} : V(x) \le 4K_V\}$ is d_N-small.

We have shown that (i), the contraction property for d_N, and (ii), the d_N–smallness of V level sets, hold for any N and t such that

$$C_\alpha N^{-1} + C_\Xi e^{-\chi\alpha t} < 1 \quad \text{and} \quad NC_K C_\Xi e^{-\chi\alpha t} < \frac{1+\varepsilon_1}{2}. \tag{4.61}$$

Then setting

$$N_* = \frac{2C_K C_\Xi}{2C_K C_\Xi C_\alpha + C_\Xi(1+\varepsilon_1)} \quad \text{and} \quad t_* = \frac{1}{\chi\alpha}\ln\left(\frac{2N_* C_K C_\Xi}{1+\varepsilon_1}\right) \tag{4.62}$$

concludes the proof of the theorem.

□

Having shown that Assumption A is a sufficient condition for the hypothesis of Theorem 4.1.6 to hold we deduce that there exists $t > 0$ and $\rho \in (0,1)$ such that

$$W_{\tilde{d}}(P_t(x,\cdot), P_t(y,\cdot)) \leq \rho\tilde{d}(x,y).$$

Before applying this methodology to show the same result for our two nonlinear model of interest, the stochastic Navier–Stokes equation and the stochastic two–layer quasi–geostrophic model, we look at a simple example with the following linear stochastic differential equation in $L^2(\mathcal{D})$.

Example 4.1.14 (Ornstein-Uhlenbeck process). As first elementary example let us consider $\mathcal{H} = L^2(\mathcal{D})$ and $\mathcal{V} = H^1(\mathcal{D})$ and the linear equation

$$dX + \alpha AX\, dt = dW, \quad X(0) = x \in \mathcal{H} \tag{4.63}$$

where A is defined as $Au = -\Delta u$ and W is a Q-Wiener process with Q trace class in $\mathcal{H}$. Then we can see that Assumption A is immediately fulfilled with a control $G = 0$. In fact consider

$$dY + \alpha AY\, dt = dW, \quad Y(0) = y \in \mathcal{H}$$

with $x \neq y$ and same realisation of the noise. Then

$$\frac{d}{dt}(X(t) - Y(t)) + \alpha A(X(s) - Y(s))\, ds = 0$$

with initial condition $x - y$ and by now well known manipulations we have

$$\frac{1}{2}\frac{d}{dt}|X(t) - Y(t)|^2 \leq -\alpha\|X(s) - Y(s)\|^2 \leq -\frac{\alpha}{\lambda_1}|X(s) - Y(s)|^2$$

and by Gronwall's lemma

$$|X(t) - Y(t)|^2 \leq |x - y|^2 \exp\left(-\frac{2\alpha}{\lambda_1}t\right).$$

Therefore $X(t)$ and $Y(t)$ converge almost surely to each other exponentially fast for large time despite starting at different initial conditions and Theorem 4.1.10 holds.

It is given by Corollary 4.11 in [44] and, more specifically for this context, by Theorem 4.2 in [12], that from Theorem 4.1.10 we also have *existence* of the invariant measure if $\mathcal{P}_t$ is Feller on $\mathcal{H}$. As a consequence $\mathcal{P}_t$ Markov semigroup associated to $X(t)$ solution of (4.63) has a unique invariant measure and $X(t)$ is ergodic.

4.2 Stochastic Navier–Stokes equations

Consider the stochastic Navier–Stokes equations introduced in section 1.2, namely

$$du + (Au + B(u,u))\,dt = f\,dt + dW, \quad u(0) = u_0 \tag{4.64}$$

with $\mathcal{H}$ and $\mathcal{V}$ as in (1.12). Then for $f \in \mathcal{V}^*$ there exists a unique solution in $L^2(\Omega; C([0,\infty);\mathcal{H}) \cap L^2_{loc}([0,\infty),\mathcal{V}))$.

From the literature we know that this model exhibits convergence of transition probabilities and in particular [50, 12] both showed it with the methodology presented above with respect the Lipschitz seminorm $\|\cdot\|_{\tilde{d}}$ where

$$\begin{aligned} \tilde{d}(x,y)^2 &= d_N(x,y)(1 + V(x) + V(y)) \quad \text{with} \\ d_N(x,y) &= N\theta_\alpha(x,y) \wedge N\theta_\alpha(y,x) \wedge 1 \quad \text{and} \\ \theta_\alpha(x,y) &= e^{\alpha\upsilon|x|^2}|x-y|^{2\alpha}, \quad \alpha \in (0,\alpha_0) \end{aligned} \tag{4.65}$$

for an appropriate choice of the parameters N, υ and α_0 given by Theorem 4.1.10. In the literature the proof of this result is for a finite dimensional noise so, to enhance clarity of exposition, we reformulate it here in a streamlined presentation for our trace class noise showing that Assumption A holds. This result will prove crucial in the next chapter to study the dependence on the parameters of the invariant measure for the stochastic Navier–Stokes equation.

Theorem 4.2.1. *Assume that there exists $M \in \mathbb{N}$ such that $\Pi_M\mathcal{H} \subset \operatorname{range} Q$ and*

$$\lambda_M > \frac{k(\operatorname{Tr} Q + \frac{|f|^2_{-1}}{\nu})}{\nu^2(\nu - \gamma \operatorname{Tr} Q\lambda_1^{-1})} > 0.$$

Then there exist $t > 0$ and $\rho < 1$ such that

$$W_{\tilde{d}}\,(P_t(u_0,\cdot), P_t(\tilde{u}_0,\cdot)) \leq \rho\,\tilde{d}(u_0,\tilde{u}_0) \tag{4.66}$$

for all $u_0, \tilde{u}_0 \in \mathcal{H}$. Here $\tilde{d}$ is as in (4.65) with Lyapunov function $V(x) = |x|^2$ and parameters

$$\upsilon = \frac{k_B}{\nu(\nu - \gamma \operatorname{Tr} Q\lambda_1^{-1})} \quad \text{and} \quad \alpha_0 = \frac{1}{2} \wedge \frac{2\gamma}{\upsilon + 2\gamma}$$

where γ is arbitrary so that $\upsilon > 0$.

Proof. Consider the controlled system

$$d\tilde{u} + (\nu A\tilde{u} + B(\tilde{u},\tilde{u}))\,dt = (f + G(u,\tilde{u}))\,dt + dW, \quad \tilde{u}(0) = \tilde{u}_0 \tag{4.67}$$

with $\tilde{u}_0 \neq u_0$ and G a finite dimensional control which we will specify later. We ensure that Assumption A holds.

A1. Set $v = u - \tilde{u}$, then it satisfies the differential equation

$$\frac{\mathrm{d}v}{\mathrm{d}t} + \nu A v + B(v,u) + B(\tilde{u},v) = -G(u,\tilde{u}) \quad v(0) = u_0 - \tilde{u}_0.$$

Taking the scalar product with v itself and using the fact that $(B(\tilde{u},v),v) = 0$ we have

$$\frac{1}{2}\frac{\mathrm{d}|v|^2}{\mathrm{d}t} + \nu\|v\|^2 + (B(v,u),v) + (G(u,\tilde{u}),v) = 0. \tag{4.68}$$

Given $M \in \mathbb{N}$, set

$$G(u,\tilde{u}) = \alpha_1 \Pi_M (u - \tilde{u}) \tag{4.69}$$

with arbitrary $\alpha_1 > 0$ to be fixed later. By the properties of the bilinearity B, Lemma 1.2.1, we have

$$\frac{\mathrm{d}|v|^2}{\mathrm{d}t} + 2\nu\|v\|^2 + 2\alpha_1|\Pi_M v|^2 \leq 2k_B|v|\|v\|\|u\|.$$

Then taking $\alpha_1 = \nu\lambda_M/2$, the generalised Poincaré inequalities (3.66) give

$$\nu\|v\|^2 + 2\alpha_1|\Pi_M v|^2 \geq \nu\|\Pi_M^{\perp} v\|^2 + \nu\lambda_M\|\Pi_M v\|^2 \geq \nu\lambda_M|v|^2,$$

so that

$$\frac{\mathrm{d}|v|^2}{\mathrm{d}t} + \nu\|v\|^2 + \nu\lambda_M|v|^2 \leq \frac{k_B^2\|u\|^2|v|^2}{\nu} + \nu\|v\|^2.$$

By Gronwall's lemma we have

$$|v(t)|^2 \leq |v(0)|^2 \exp\left(-\nu\lambda_M t + \frac{k_B^2}{\nu}\int_0^t \|u\|^2\, ds\right) \tag{4.70}$$

So that Assumption **A1** holds with

$$\kappa_0 = \nu\lambda_M \quad \text{and} \quad \kappa_1 = \frac{k_B^2}{\nu}. \tag{4.71}$$

A2. This is given by Theorem 2.5.1 part (i) i.e.

$$|u(t)|^2 + (\nu - \gamma \operatorname{Tr} Q \lambda_1^{-1}) \int_0^t \|u(s)\|^2\, ds \leq |u_0|^2 + t\left(\operatorname{Tr} Q + \frac{|f|_{-1}^2}{\nu}\right) + 2\Xi_\gamma. \tag{4.72}$$

where

$$\Xi_\gamma = \sup_{t\geq 0}\left(X_t - \gamma\langle X\rangle_t\right), \quad X_t = \int_0^t \langle u(s),\, dW(s)\rangle$$

for $\gamma < \nu\lambda_1/\operatorname{Tr} Q$. Then

$$\kappa_2 = \nu - \gamma \operatorname{Tr} Q \lambda_1^{-1} > 0 \quad \text{and} \quad \kappa_3 = \operatorname{Tr} Q + \frac{|f|_{-1}^2}{\nu} \tag{4.73}$$

and the condition $\kappa_0 > \kappa_1\kappa_3/\kappa_2$ translate in a lower bound for λ_M

$$\lambda_M > \frac{k_B^2(\operatorname{Tr} Q + \frac{|f|_{-1}^2}{\nu})}{\nu^2(\nu - \gamma \operatorname{Tr} Q \lambda_1^{-1})}.$$

A3. Given the control (4.69) we have immediately

$$|G(u,\tilde{u})|^2 = \frac{\nu\lambda_M}{2}|\Pi_M(u-\tilde{u})|^2 \le \frac{\nu\lambda_M}{2}|u-\tilde{u}|^2$$

with $c = \nu\lambda_M/2$.

A4. Consider (4.64) over the time interval $[s,t]$ and take the scalar product with u itself to get

$$|u(t)|^2 - |u(s)|^2 + 2\nu \int_s^t \|u(\tau)\|^2\, d\tau = \int_s^t 2\langle f, u\rangle\, d\tau + (t-s)\operatorname{Tr} Q + 2\int_s^t \langle u, \cdot\rangle\, dW_\tau.$$

We estimate the forcing term as

$$\langle f, u\rangle \le \frac{|f|_{-1}^2}{2\nu} + \frac{\nu\|u\|^2}{2}$$

to get

$$|u(t)|^2 - |u(s)|^2 + \nu \int_s^t \|u(\tau)\|^2\, d\tau \le (t-s)\left(\frac{|f|_{-1}^2}{\nu} + \operatorname{Tr} Q\right) + 2\int_s^t \langle u, \cdot\rangle\, dW_\tau.$$

Then simply taking the expectation we have

$$\mathbb{E}\,|u(t)|^2 \le \mathbb{E}\,|u(s)|^2 - \nu\mathbb{E}\int_s^t \|u(\tau)\|^2\, d\tau + (t-s)\left(\frac{|f|_{-1}^2}{\nu} + \operatorname{Tr} Q\right)$$

and by Poincaré inequality

$$\mathbb{E}\,|u(t)|^2 \le \mathbb{E}\,|u(s)|^2 - \nu\lambda_1 \int_s^t \mathbb{E}|u(\tau)|^2\, d\tau + (t-s)\left(\frac{|f|_{-1}^2}{\nu} + \operatorname{Tr} Q\right). \tag{4.74}$$

So **A4** holds with

$$\gamma_1 = \nu\lambda_1 \quad \text{and} \quad K = \frac{|f|_{-1}^2}{\nu} + \operatorname{Tr} Q. \tag{4.75}$$

□

Remark 4.2.2. Given how the distance $\tilde{d}$ is defined i.e. (4.65) and the fact that the Lyapunov function is $V(x) = |x|^2$ we can see that the functions φ with $\|\varphi\|_{\tilde{d}} < \infty$ are α-Hölder functions with respect to the usual metric on $\mathcal{H}$ on the level sets of the Lyapunov function. In fact, given $x, y \in \{V \leq K\}$ for $K \geq 0$ we have that

$$
\begin{aligned}
&|\varphi(x) - \varphi(y)| \leq \|\varphi\|_{\tilde{d}}\, \tilde{d}(x,y) \\
&= \|\varphi\|_{\tilde{d}} \left(N|x-y|^{2\alpha} e^{\alpha\upsilon|x|^2} \wedge N|x-y|^{2\alpha} e^{\alpha\upsilon|y|^2} \wedge 1 \right)^{1/2} (1 + V(x) + V(y))^{1/2} \\
&\leq \|\varphi\|_{\tilde{d}} N^{1/2} e^{\frac{\alpha\upsilon K}{2}} (1+2K)^{1/2} |x-y|^{\alpha}.
\end{aligned}
$$

4.3 Stochastic two–layer quasi–geostrophic model

Let $\mathbf{q}(t, \omega; \mathbf{q}_0)$ be the solution of the stochastic two–layer quasi–geostrophic model (1.66) on the Hilbert space $\mathcal{H} = (\mathbb{H}^{-1}, |||\cdot|||_{-1})$. Then we want to ensure that this system has exponential convergence of transition probabilities and spectral gap condition as described in section 4.1, i.e. with respect to the Lipschitz seminorm $\|\cdot\|_{\tilde{d}}$ where

$$
\begin{aligned}
\tilde{d}(\mathbf{u}, \mathbf{v})^2 &= d_N(\mathbf{u}, \mathbf{v})(1 + V(\mathbf{u}) + V(\mathbf{v})) \quad \text{with} \\
d_N(\mathbf{u}, \mathbf{v}) &= N\theta_\alpha(\mathbf{u}, \mathbf{v}) \wedge N\theta_\alpha(\mathbf{u}, \mathbf{v}) \wedge 1 \quad \text{and} \\
\theta_\alpha(\mathbf{u}, \mathbf{v}) &= e^{\alpha\upsilon |||\mathbf{u}|||_{-1}^2} |||\mathbf{u} - \mathbf{v}|||_{-1}^{2\alpha}, \quad \alpha \in (0, \alpha_0)
\end{aligned}
\tag{4.76}
$$

for an appropriate choice of the parameters N, υ and α_0 given by Theorem 4.1.10. By the results in section 4.1.1 it will be enough to construct a controlled system that makes Assumption A hold, as we will show in the following theorem:

Theorem 4.3.1. *Assume that there exists $N \in \mathbb{N}$ such that $\Pi_N \mathcal{H} \subset \operatorname{range} Q$ and $\nu - 2rh_1\lambda_N^{-1} > 0$. Then there exists r_0 such that for all $r > r_0$ there exist $t > 0$ and $\rho < 1$ such that*

$$
W_{\tilde{d}}(P_t(\mathbf{q}_0, \cdot), P_t(\tilde{\mathbf{q}}_0, \cdot)) \leq \rho\, \tilde{d}(\mathbf{q}_0, \tilde{\mathbf{q}}_0) \tag{4.77}
$$

for all $\mathbf{q}_0, \tilde{\mathbf{q}}_0 \in \mathcal{H}$. Here $\tilde{d}$ is as in (4.76) with Lyapunov function $V(\mathbf{u}) = |||\mathbf{u}|||_{-1}^2$ and parameters

$$
\upsilon = \frac{k_B}{\nu - \frac{2\gamma \operatorname{Tr} Q}{\lambda_1^2}} \quad \text{and} \quad \alpha_0 = \frac{1}{2} \wedge \frac{2\gamma}{\upsilon + 2\gamma}
$$

where γ is arbitrary so that υ is well defined.

Proof. We only have to ensure that the Assumption A is satisfied, as then Theorem 4.1.10 and the general Harris' theorem, Theorem 4.1.6, give the desired result. Consider the controlled system

$$\begin{aligned} &d\tilde{\mathbf{q}} + \Big(B(\tilde{\boldsymbol{\psi}},\tilde{\boldsymbol{\psi}}) + \beta\partial_1\tilde{\boldsymbol{\psi}}\Big)\,dt = \nu\Delta^2\tilde{\boldsymbol{\psi}}\,dt + \begin{pmatrix} f + G(\mathbf{q},\tilde{\mathbf{q}}) \\ -r\Delta\tilde{\psi}_1 \end{pmatrix} dt + d\boldsymbol{W} \\ &\tilde{\mathbf{q}} = (\Delta + M)\tilde{\boldsymbol{\psi}} \end{aligned} \tag{4.78}$$

with initial condition $\tilde{\mathbf{q}}(0) = \tilde{\mathbf{q}}_0 \neq \mathbf{q}_0$, and $\boldsymbol{W} = (W, 0)^t$. Let the feedback control G be as in section 3.2.2 i.e. $G(\mathbf{q},\tilde{\mathbf{q}}) = \alpha_1\Pi_N\Delta(\boldsymbol{\psi} - \tilde{\boldsymbol{\psi}})$ with $\alpha_1 > 0$ arbitrary to be fixed later. Notice that by the same arguments in seen in the proof of Theorem 3.2.9, the controlled system with this choice of G is well defined.

A1. We have already showed the estimate (3.71) for the difference $\mathbf{u} = \mathbf{q} - \tilde{\mathbf{q}}$, i.e.

$$|||\mathbf{u}(t)|||_{-1}^2 \leq |||\mathbf{u}(0)|||_{-1}^2 \exp\bigg(-2rt + 2k_B\int_0^t |\Delta\boldsymbol{\psi}|^2\,ds\bigg)$$

for all $N > 0$ such that $\nu - 2\alpha_1\lambda_N^{-1} > 0$ and $\alpha_1 \geq r$. Then Assumption **A1** follows immediately setting the weak norm $|\mathbf{q}| = |||\mathbf{q}|||_{-1}$, the strong norm $\|\mathbf{q}\| = |\Delta\boldsymbol{\psi}|$ and the constants to be

$$\kappa_0 = 2r \quad \text{and} \quad \kappa_1 = 2k_B. \tag{4.79}$$

A2. In the proof of Theorem 2.4.2 we showed that, setting

$$X_t = \int_0^t (\psi_1, dW),$$

the estimate (2.58) for the quadratic variation $\langle X\rangle_t$ holds and we have that

$$\begin{aligned} |||\mathbf{q}(t)|||_{-1}^2 - |||\mathbf{q}_0|||_{-1}^2 + \bigg(\nu - \frac{2\gamma\operatorname{Tr}Q}{\lambda_1^2}\bigg)\int_0^t |\Delta\boldsymbol{\psi}|^2\,ds \\ - t\left(T_Q + \tfrac{h_1}{2\nu}\|f\|_{-2}^2\right) \leq 2h_1\Xi_\gamma \end{aligned} \tag{4.80}$$

with $\Xi_\gamma = \sup_{t\geq 0}(X_t - \gamma\langle X\rangle_t)$. Setting

$$\kappa_2 = \nu - \frac{2\gamma\operatorname{Tr}Q}{\lambda_1^2} \quad \text{and} \quad \kappa_3 = \tfrac{h_1}{2\nu}\|f\|_{-2}^2 + T_Q, \tag{4.81}$$

Assumption **A2** is satisfied for all arbitrary parameter $\gamma > 0$ such that $\kappa_2 > 0$ and choices of parameters of the system such that $\kappa_0 > \kappa_1\kappa_3/\kappa_2$, i.e.

$$r > \frac{k_B}{\nu - \frac{2\gamma \operatorname{Tr} Q}{\lambda_1^2}} \left(\tfrac{h_1}{2\nu}\|f\|_{-2}^2 + T_Q\right) =: r_0$$

which gives an explicit expression for the minimum value of the bottom friction r_0. Finally Theorem 4.1.10 gives that

$$\upsilon = \frac{\kappa_1}{\kappa_2} = \frac{k_B}{\nu - \frac{2\gamma \operatorname{Tr} Q}{\lambda_1^2}} > 0.$$

A3. Recall the generalized Poincaré inequality $|\Pi_N\varphi|^2 \le \lambda_N\|\Pi_N\varphi\|_{-1}^2$. Then

$$\begin{aligned}|G(\mathbf{q},\tilde{\mathbf{q}})|^2 &= |\alpha_1\Pi_N\Delta(\psi_1 - \tilde{\psi}_1)|^2 \le \lambda_N\alpha_1^2\|\Pi_N\Delta(\psi_1 - \tilde{\psi}_1)\|_{-1}^2 \\ &\le \lambda_N\alpha_1^2\|\psi_1 - \tilde{\psi}_1\|^2 \le \lambda_N\alpha_1^2|||\mathbf{q} - \tilde{\mathbf{q}}|||_{-1}^2\end{aligned}$$

giving the desired inequality with $c = \alpha_1^2\lambda_N$.

A4. Finally we show that $V(\mathbf{u}) := |||\mathbf{u}|||_{-1}^2$ satisfies Assumption **A4**. We take the $\mathbf{L}^2$ product with $\boldsymbol{\psi}$ of the stochastic model (1.66) for times in $[s,t]$ as done in Theorem 2.4.2. In particular using Itô's formula we had (2.55) from which we get

$$|||\mathbf{q}(t)|||_{-1}^2 - |||\mathbf{q}(s)|||_{-1}^2 + \nu\int_s^t |\Delta\boldsymbol{\psi}|^2\,d\tau - (t-s)T_Q \le \frac{h_1\|f\|_{-2}^2}{\nu}(t-s) - 2h_1\int_s^t(\psi_1, dW).$$

Rearranging and taking the expectation we have

$$\mathbb{E}|||\mathbf{q}(t)|||_{-1}^2 \le \mathbb{E}|||\mathbf{q}(s)|||_{-1}^2 + \mathbb{E}\int_s^t\left(-\nu|\Delta\boldsymbol{\psi}|^2 + K\right)ds$$

where $K = \frac{h_1|f|_{-2}^2}{\nu} + T_Q$. By definition of $|||\cdot|||_{-1}$ and Poincaré inequality

$$|||\mathbf{q}|||_{-1}^2 \le \frac{\lambda_1 + F_1}{\lambda_1^2}|\Delta\boldsymbol{\psi}|^2.$$

Then

$$\mathbb{E}|||\mathbf{q}(t)|||_{-1}^2 \le \mathbb{E}|||\mathbf{q}(s)|||_{-1}^2 + \mathbb{E}\int_s^t\left(-\tfrac{\nu\lambda_1^2}{\lambda_1+F_1}|||\mathbf{q}(\tau)|||_{-1}^2 + K\right)d\tau. \tag{4.82}$$

Therefore $V(\mathbf{q}) = \|\!|\mathbf{q}|\!\|_{-1}^2$ satisfies the estimate (4.29) with

$$\gamma_1 := \frac{\nu\lambda_1^2}{\lambda_1 + F_1} \quad \text{and} \quad K = \frac{h_1\|f\|_{-2}^2}{\nu} + T_Q. \tag{4.83}$$

□

Remark 4.3.2. As observed for the stochastic Navier–Stokes equations, given how the distance $\tilde{d}$ is defined i.e. (4.76) and the fact that the Lyapunov function is $V(\mathbf{u}) = \|\!|\mathbf{u}|\!\|_{-1}^2$, the functions φ with $\|\varphi\|_{\tilde{d}} < \infty$ are α-Hölder functions with respect to the usual metric on $\mathcal{H}$ on the level sets of the Lyapunov function.

Summary and remarks

In this chapter we presented a streamlined version of the results in [12] and applied it to our model of interest, the stochastic two–layer quasi–geostrophic model, to ensure the exponential convergence of transition probabilities. This provides information on the stability of the system for large times, in the sense that not only there is a unique invariant measure but the system approaches it exponentially fast in time, whenever the bottom friction r is large enough.

This result is most interesting as it will be one crucial condition to ensure the stability with respect to the parameters, more precisely in the next chapter we will be able to show weak differentiability and Hölder continuity of the invariant measure with respect to the intensity of the forcing f.

Last, let us comment on a possible quantitative interpretation of the results presented in this chapter. Given the significance of the two–layer quasi–geostrophic model in the applications, it would be interesting to quantify the size of the spectral gap, or, in other words, measure the rate of convergence of the transition probabilities. This in fact could give an estimate of the time after which the model reaches a stationary regime, one under which certain configurations of the state space would be forbidden.

However, the approach used in the present work is not the most suitable to answer this type of quantitative questions. From the results presented here, it is possible to provide an estimate for ρ in the general Harris theorem (Theorem 4.1.6) by carefully combing through the original result in [44], and expanding some results there only hinted. This approach would only give a very crude estimate for ρ, though, as the methodology is most appropriate to ensure $\rho < 1$, but not for quantifying it. The scientific insight would then be quite limited despite the complexity of the arguments needed to achieve the result. Estimating the spectral gap, in fact, usually constitutes a research question on its own, and therefore goes beyond the scope of this thesis.

Chapter 5

Linear and fractional response

In this final chapter we focus on the regularity with respect to the parameters of the invariant measure of stochastic Navier–Stokes equations and stochastic two–layer quasi–geostrophic model. Proving the continuity or, even better, differentiability of the invariant measure would ensure a controlled response of the long time average behaviour of the models to small perturbations of the parameters, in the sense that no abrupt change is associated with them. We talk about linear response when the measure is (in an appropriate sense) differentiable with respect to the parameter of choice while we can refer to stability with respect to the parameters in the case of continuity and fractional response when the invariant measure is Hölder continuous with respect to the parameter.

The literature on linear response is vast both in the applications and in more theoretical works for finite dimensional systems, but to the best of this author's knowledge the only theoretical framework for forced dissipative stochastic dynamical systems is the work of Hairer and Majda [40]. The authors proved the weak differentiability of invariant measures μ_a of families of Markov semigroup $\{\mathcal{P}_t^a : t \geq 0, a \in \mathbb{R}\}$ on a Hilbert space $\mathcal{H}$. This means that, given any $a_0 \in \mathbb{R}$, the map

$$a \mapsto \langle \psi, \mu_a \rangle = \int_{\mathcal{H}} \psi(x)\, \mu_a(dx).$$

is differentiable for an appropriate choice of test functions φ, i.e. the limit

$$\lim_{a \to a_0} \frac{\langle \varphi, \mu_a - \mu_{a_0} \rangle}{a - a_0}$$

exists. Moreover an explicit expression for its derivative is proven i.e.

$$\langle \partial_a \mathcal{P}_t^{a_0} (1 - \mathcal{P}_t^{a_0})^{-1} (\varphi - \langle \varphi, \mu_{a_0} \rangle), \mu_{a_0} \rangle.$$

However, as we will better explain later, the techniques used in the previous chapter to ensure the exponential stability of the invariant measure do not gives the necessary conditions for observables which are suitable for [40]. Hence we will here modify the approach in [40] to apply to our family of observables.

In the first section we give a didactic presentation to the proof of the linear response developed by Hairer and Majda in [40], highlighting how their assumptions arise, how they depend on the family of observables chosen, and how they can be modified for our framework.

Then in section 5.1.1 we will develop the methodology for the case where the Markov semigroup in given by the solution of a nonlinear stochastic differential equation in $\mathcal{H}$ and the parameter of interest is the intensity of an external deterministic finite dimensional forcing. This general equation is used as a mean to summarize the main features of both Navier–Stokes equation and the two–layer quasi–geostrophic model studied here and as a consequence the linear response for the two models of interest will follow straightforwardly in section 5.3 and section 5.4.

In section 5.2 we study the continuous dependence on the parameters of the invariant measure focusing directly on the framework for SPDEs laid out for the response. In particular we will show that the invariant measure is weak Hölder continuous with respect to the intensity of an external deterministic forcing, even for forcings not necessarily finite dimensional. Consequently in section 5.3 and section 5.4 this result is derived for the stochastic Navier–Stokes equations and the stochastic two–layer quasi–geostrophic model respectively.

5.1 Linear response

Let $(\mathcal{H}, |\cdot|)$ be the Hilbert space of interest and let $(\mathcal{O}, \|\cdot\|_{\mathcal{O}})$ be a Banach space of measurable functions $\psi : \mathcal{H} \to \mathbb{R}$ that we call observables. Let $\mathcal{P}_t^a$ be a Markov semigroup on $\mathcal{O}$ depending on the parameter $a \in \mathbb{R}$ with corresponding invariant measure μ_a. Fix a time t and drop the dependence on time of the semigroup. Let us start by showing a simple yet crucial identity:

$$\frac{\langle (1 - \mathcal{P}^{a_0})\psi, (\mu_a - \mu_{a_0}) \rangle}{a - a_0} = \frac{\langle (\mathcal{P}^a - \mathcal{P}^{a_0})\psi, \mu_a \rangle}{a - a_0}. \tag{5.1}$$

By the invariance of the measures and simple manipulations we have in fact

$$\begin{aligned}\frac{\langle \psi, \mu_a - \mu_{a_0}\rangle}{a-a_0} &= \frac{\langle \psi, (\mathcal{P}^a)^*\mu_a - (\mathcal{P}^{a_0})^*\mu_{a_0}\rangle}{a-a_0}\\ &= \frac{\langle \psi, (\mathcal{P}^a)^*\mu_a - (\mathcal{P}^{a_0})^*\mu_{a_0}\rangle}{a-a_0} - \frac{\langle \psi, (\mathcal{P}^{a_0})^*\mu_a\rangle}{a-a_0} + \frac{\langle \psi, (\mathcal{P}^{a_0})^*\mu_a\rangle}{a-a_0}\\ &= \frac{\langle \psi, ((\mathcal{P}^a)^* - (\mathcal{P}^{a_0})^*)\mu_a\rangle}{a-a_0} + \frac{\langle \psi, (\mathcal{P}^{a_0})^*(\mu_a - \mu_{a_0})\rangle}{a-a_0}.\end{aligned}$$

Therefore it follows

$$\frac{\langle \psi, (1-(\mathcal{P}^{a_0})^*)(\mu_a - \mu_{a_0})\rangle}{a-a_0} = \frac{\langle \psi, ((\mathcal{P}^a)^* - (\mathcal{P}^{a_0})^*)\mu_a\rangle}{a-a_0},$$

relation equivalent to (5.1).

Suppose the space of observables $\mathcal{O}$ is such that $1-\mathcal{P}^{a_0}$ is invertible over it, or more precisely for every $\varphi \in \mathcal{O}$ there is a unique $\psi \in \mathcal{O}$ such that $\varphi = (1-\mathcal{P}^{a_0})\psi$ and we denote ψ as $(1-\mathcal{P}^{a_0})^{-1}\varphi$ and

$$\tilde{\mathcal{O}} := \operatorname{Im}(1-\mathcal{P}^{a_0})\big|_{\mathcal{O}}. \tag{5.2}$$

Then from (5.1) we have

$$\frac{\langle \varphi, (\mu_a - \mu_{a_0})\rangle}{a-a_0} = \frac{\langle (\mathcal{P}^a - \mathcal{P}^{a_0})(1-\mathcal{P}^{a_0})^{-1}\varphi, \mu_a\rangle}{a-a_0}. \tag{5.3}$$

To compute the derivative of $\langle \varphi, \mu_a\rangle$ at a_0 it is then enough to show that the right hand side of (5.3) converges to the desired quantity as a goes to a_0. Formally if $\mathcal{P}^a\psi$ is differentiable in a_0 and μ_a is weakly continuous in a_0, we can expect the desired convergence to follow.

Let us present how these conditions can be formulated. As in [40], let $C_0^\infty(\mathcal{H})$ be the set of all functions $\varphi : \mathcal{H} \to \mathbb{R}$ such that there exists $N > 0$, a linear map $T : \mathcal{H} \to \mathbb{R}^N$ and a smooth, compactly supported map $\hat{\varphi} : \mathbb{R}^N \to \mathbb{R}$ such that $\varphi = \hat{\varphi} \circ T$. Then let $\mathcal{C}_U$ be the completion of $C_0^\infty(\mathcal{H})$ under the norm

$$\|\varphi\|_U := \sup_{x\in\mathcal{H}} \frac{|\varphi(x)|}{U(x)}, \tag{5.4}$$

with $U : \mathcal{H} \to [1,\infty)$ such that $\mathcal{O} \subset \mathcal{C}_U$. For every $\psi \in \mathcal{O}$ consider the map $a \mapsto \mathcal{P}^a\psi$. If it is differentiable as a function from a neighborhood of a_0 in $\mathcal{C}_U$, then

$$\frac{\langle (\mathcal{P}^a - \mathcal{P}^{a_0})\psi, \mu_a\rangle}{a-a_0} = \langle D_a(\mathcal{P}^a\psi)\big|_{a=a_0}, \mu_a\rangle + \langle r_a, \mu_a\rangle \tag{5.5}$$

with $\|r_a\|_U \to 0$ for $a \to a_0$.

Consequently, if U is such that for all $\xi \in \mathcal{C}_U$ we have

$$\langle \xi, \mu_a \rangle \to \langle \xi, \mu_{a_0} \rangle \quad \text{and} \quad \langle r_a, \mu_a \rangle \to 0 \tag{5.6}$$

for $a \to a_0$, then

$$\frac{\langle (\mathcal{P}^a - \mathcal{P}^{a_0})\psi, \mu_a \rangle}{a - a_0} \to \langle\, D_a(\mathcal{P}^a\psi)\big|_{a=a_0}, \mu_{a_0} \rangle$$

as desired.

For the convergence of the term involving the remainder r_a, note that

$$|\langle r_a, \mu_a \rangle| \le \langle \left| \tfrac{r_a}{U} U \right|, \mu_a \rangle \le \|r_a\|_U \langle U, \mu_a \rangle. \tag{5.7}$$

so, if there is ε such that

$$\sup_{|a-a_0|<\varepsilon} \langle U, \mu_a \rangle < \infty, \tag{5.8}$$

then $|\langle r_a, \mu_a \rangle|$ vanishes on the limit $a \to a_0$ as desired for all a in an ε-neighbourhood of a_0.

In summary, if $\varphi \in \tilde{\mathcal{O}}$, i.e. there exists $\psi \in \mathcal{O}$ such that $\varphi = (1 - \mathcal{P}^{a_0})\psi$, thanks to (5.3), (5.5) and (5.6) we get

$$\begin{aligned}
\frac{\langle \varphi, (\mu_a - \mu_{a_0}) \rangle}{a - a_0} &= \langle\, D_a(\mathcal{P}^a\psi)\big|_{a=a_0}, \mu_a \rangle + \langle r_a, \mu_a \rangle \\
&\le \langle\, D_a(\mathcal{P}^a\psi)\big|_{a=a_0}, \mu_a \rangle + \|r_a\|_U \langle U, \mu_a \rangle \\
&\to \langle\, D_a(\mathcal{P}^a\psi)\big|_{a=a_0}, \mu_{a_0} \rangle \quad \text{for } a \to a_0,
\end{aligned}$$

and we have shown the following result, a more general version of [40, Theorem 2.3].

Theorem 5.1.1. *Let $\{\mathcal{P}^a_t : a \in \mathbb{R}\}$ be a family of Markov semigroups acting on $\mathcal{O}$. Suppose there exists a function $U : \mathcal{H} \to [1, \infty)$ such that $\mathcal{O} \subset \mathcal{C}_U$ and such that for some $a_0 \in \mathbb{R}$ and $t > 0$ the following conditions hold:*

(i) for every $\psi \in \mathcal{O}$ and some fixed $t > 0$ the map $a \mapsto \mathcal{P}^a_t\psi$ from $\mathbb{R}$ to $\mathcal{C}_U$ is differentiable in a neighbourhood of a_0;

(ii) given $\xi \in \mathcal{C}_U$, $\langle \xi, \mu_a \rangle \to \langle \xi, \mu_{a_0} \rangle$ for $a \to a_0$;

(iii) there is $\varepsilon > 0$ such that $\sup_{|a-a_0|<\varepsilon} \langle U, \mu_a \rangle < \infty$.

Then the map $a \mapsto \langle \varphi, \mu_a \rangle$ is differentiable at a_0 for every $\varphi \in \tilde{\mathcal{O}}$ and in particular given ψ such that $\varphi = (1 - \mathcal{P}_t^{a_0})\psi$

$$\frac{\mathrm{d}}{\mathrm{d}a}\langle \varphi, \mu_a \rangle \bigg|_{a=a_0} = \langle\, D_a(\mathcal{P}_t^a \psi)|_{a=a_0}\,, \mu_{a_0} \rangle. \tag{5.9}$$

To make this theorem more applicable, we show sufficient conditions to ensure hypothesis (ii) holds, and that the space $\tilde{\mathcal{O}}$ is nothing but $\mathcal{O} \cap \ker \mu_{a_0}$ where $\ker \mu_{a_0}$ is the space of measurable functions on $\mathcal{H}$ which are centred with respect to μ_{a_0}, i.e. $\langle \varphi, \mu_{a_0} \rangle = 0$. More precisely it can be proven that the spectral gap property on the space of observables is a sufficient condition to show that $(1 - \mathcal{P}^{a_0})$ is an invertible operator over $\mathcal{O} \cap \ker \mu_{a_0}$.

Given a function $\varphi \in \mathcal{O}$ we denote by $\bar{\varphi}$ its centred version with respect to μ_{a_0} i.e.

$$\bar{\varphi} = \varphi - \langle \varphi, \mu_{a_0} \rangle.$$

Proposition 5.1.2. *Suppose there exists $\rho < 1$ and $t > 0$ such that*

$$\|\mathcal{P}_t^{a_0}\varphi - \langle \varphi, \mu_{a_0} \rangle\|_{\mathcal{O}} \leq \rho\, \|\varphi - \langle \varphi, \mu_{a_0} \rangle\|_{\mathcal{O}} \quad \textit{for all } \varphi \in \mathcal{O} \tag{5.10}$$

Then for all $\varphi \in \mathcal{O}$ there exists a unique $\psi \in \mathcal{O} \cap \ker \mu_{a_0}$ such that $\bar{\varphi} = (1 - \mathcal{P}_t^{a_0})^{-1}\psi$, namely $(1 - \mathcal{P}_t^{a_0})$ is invertible over $\mathcal{O} \cap \ker \mu_{a_0}$. Furthermore we have an explicit representation for $(1 - \mathcal{P}_t^{a_0})^{-1}$ by the Neumann series

$$(1 - \mathcal{P}_t^{a_0})^{-1} = \sum_{k=0}^{\infty} (\mathcal{P}_t^{a_0})^k\,.$$

Proof. Condition (5.10) ensures that $\mathcal{P}_t^{a_0}$ is a bounded operator on the space of centred functions in $\mathcal{O}$, i.e. given $\bar{\varphi} \in \mathcal{O} \cap \ker \mu_{a_0}$

$$\|\mathcal{P}_t^{a_0}\|_{\mathcal{O}\cap\ker\mu_{\dashv,}} = \sup_{\|\bar{\varphi}\|_{\mathcal{O}}>0} \frac{\|\mathcal{P}_t^{a_0}\bar{\varphi}\|_{\mathcal{O}}}{\|\bar{\varphi}\|_{\mathcal{O}}} \leq \rho.$$

Then, since $\rho < 1$ the Neumann series

$$\sum_{k=0}^{\infty} (\mathcal{P}_t^{a_0})^k \tag{5.11}$$

converges and as a consequence $(1 - \mathcal{P}_t^{a_0})$ is invertible with inverse (5.11). □

Next we can substitute hypothesis (ii) with the requirement that the derivative operator $D_a\mathcal{P}^a : \mathcal{O} \to \mathcal{C}_U$ is bounded locally in a and that the space $\mathcal{O}$ is a dense subset of $\mathcal{C}_U$.

Proposition 5.1.3. *Assume $\mathcal{O}$ is dense in $\mathcal{C}_U$. Suppose for a $t > 0$ condition* (5.10) *and (iii) in Theorem 5.1.1 hold and that there exists $C(a_0) > 0$ such that for all $a \in B_\varepsilon(a_0)$, ε-neighbourhood of a_0, $\varepsilon > 0$,*

$$\|D_a \mathcal{P}_t^a \psi\|_U \leq C\|\psi\|_{\mathcal{O}} \quad \textit{for all } \psi \in \mathcal{O}. \tag{5.12}$$

Then for any $\xi \in \mathcal{C}_U$, $\langle \xi, \mu_a \rangle \to \langle \xi, \mu_{a_0} \rangle$ when $a \to a_0$.

Proof. First we show that it is enough to prove the result for observables in $\mathcal{O}$. Let $\xi \in \mathcal{C}_U$, then by (iii)

$$|\langle \xi, \mu_a \rangle| \leq \langle \left| \frac{\xi}{U} U \right|, \mu_a \rangle \leq \|\xi\|_U \langle U, \mu_a \rangle \leq \|\xi\|_U \sup_{|a-a_0|<\varepsilon} \langle U, \mu_a \rangle < \infty \tag{5.13}$$

i.e. the functional $\xi \mapsto \langle \xi, \mu_a \rangle$ on $\mathcal{C}_U$ is bounded uniformly in a over $B_\varepsilon(a_0)$. By definition, since $\mathcal{O}$ is dense in $\mathcal{C}_U$, for each $\xi \in \mathcal{C}_U$ and for each $\delta > 0$ there exists $\varphi \in \mathcal{O}$ such that

$$\|\xi - \varphi\|_U < \delta.$$

Then we have

$$|\langle \xi, \mu_a \rangle - \langle \xi, \mu_{a_0} \rangle| \leq |\langle \xi - \varphi, \mu_a \rangle| + |\langle \varphi, \mu_a \rangle - \langle \varphi, \mu_{a_0} \rangle| + |\langle \xi - \varphi, \mu_{a_0} \rangle|$$

and by (5.13), setting $C_\varepsilon = \sup_{|a-a_0|<\varepsilon} \langle U, \mu_a \rangle$, we get

$$\begin{aligned} |\langle \xi, \mu_a \rangle - \langle \xi, \mu_{a_0} \rangle| &\leq 2C_\varepsilon \|\xi - \varphi\|_U + |\langle \varphi, \mu_a \rangle - \langle \varphi, \mu_{a_0} \rangle| \\ &\leq 2\delta C_\varepsilon + |\langle \varphi, \mu_a \rangle - \langle \varphi, \mu_{a_0} \rangle|. \end{aligned}$$

Therefore if we have weak convergence for observable in $\mathcal{O}$, it holds also for observables in $\mathcal{C}_U$ as desired.

Now let $\varphi \in \mathcal{O}$, and note that by Proposition 5.1.2 and (5.3) we have

$$\langle \varphi, \mu_a - \mu_{a_0} \rangle = \langle \bar{\varphi}, \mu_a - \mu_{a_0} \rangle = \langle (\mathcal{P}_t^a - \mathcal{P}_t^{a_0})(1 - \mathcal{P}_t^{a_0})^{-1} \bar{\varphi}, \mu_a \rangle.$$

By the mean value theorem

$$\left| \langle (\mathcal{P}_t^a - \mathcal{P}_t^{a_0})(1 - \mathcal{P}_t^{a_0})^{-1} \bar{\varphi}, \mu_a \rangle \right| \leq |a - a_0| \langle D\mathcal{P}_t^{\bar{a}} (1 - \mathcal{P}_t^{a_0})^{-1} \bar{\varphi}, \mu_a \rangle$$

where $\bar{a} \in B_\varepsilon(a_0)$. Thanks to (5.13) and (5.12), the right hand side is bounded as follows

$$\leq C_\varepsilon |a - a_0| \|D\mathcal{P}^{\bar{a}} (1 - \mathcal{P}_t^{a_0})^{-1} \bar{\varphi}\|_U \leq C_1 |a - a_0| \|(1 - \mathcal{P}_t^{a_0})^{-1} \bar{\varphi}\|_{\mathcal{O}}$$

for all $a \in B_\varepsilon(a_0)$ and $C_1 = C \cdot C_\varepsilon$. Therefore μ_a seen as operator on $\mathcal{O}$ is locally Lipschitz with respect to the parameter a, and in particular continuous. $\square$

Note that in the proof we only used that $\bar{\varphi} \in \mathrm{Im}(1 - \mathcal{P}_t^{a_0})$ so the result ensures also that for all $\varphi \in \mathrm{Im}(1 - \mathcal{P}_t^{a_0})$ the function $a \mapsto \langle \varphi, \mu_a \rangle$ is continuous in a_0. However such a result is not useful to conclude $a \mapsto \langle \xi, \mu_a \rangle$ is continuous in a_0 for $\xi \in \mathcal{C}_U$ since $\mathrm{Im}(1 - \mathcal{P}_t^{a_0})$ does not contain all possible centred functions of $\mathcal{O}$ and it cannot be dense in $\mathcal{C}_U$.

Thanks to Proposition 5.1.2 and Proposition 5.1.3 we can reformulate Theorem 5.1.1 with the following assumptions, which are the same assumptions as set out in [40] but for a more general family of observables.

Assumption R There exists $(\mathcal{O}, \|\cdot\|_{\mathcal{O}})$ Banach space and U such that $\mathcal{O}$ is a dense subset of $\mathcal{C}_U$ and the following conditions hold:

R1 There exists $\rho < 1$ and $t > 0$ such that

$$\|\mathcal{P}_t^{a_0} \varphi - \langle \varphi, \mu_{a_0} \rangle\|_{\mathcal{O}} \leq \rho \|\varphi - \langle \varphi, \mu_{a_0} \rangle\|_{\mathcal{O}} \quad \text{for all } \varphi \in \mathcal{O};$$

R2 For every $\psi \in \mathcal{O}$ and some fixed $t > 0$ the map $a \mapsto \mathcal{P}_t^a \psi$ from $\mathbb{R}$ to $\mathcal{C}_U$ is differentiable in a neighbourhood of a_0 and $D_a \mathcal{P}_t^a$ is a bounded operator from $\mathcal{O}$ into $\mathcal{C}_U$ uniformly in $a \in B_\varepsilon(a_0)$;

R3 There is $\varepsilon > 0$ such that $\sup_{|a - a_0| < \varepsilon} \langle U, \mu_a \rangle < \infty$.

Hence the following version of [40, Theorem 2.3] holds.

Theorem 5.1.4. *Let $\{\mathcal{P}_t^a : a \in \mathbb{R}\}$ be a family of Markov semigroups over $\mathcal{O}$ and assume Assumption R holds. Then the map $a \mapsto \langle \varphi, \mu_a \rangle$ is differentiable at a_0 for every $\varphi \in \mathcal{O}$ and in particular*

$$\frac{\mathrm{d}}{\mathrm{d}a} \langle \varphi, \mu_a \rangle \bigg|_{a = a_0} = \langle D_a \mathcal{P}_t^{a_0} (1 - \mathcal{P}_t^{a_0})^{-1} (\varphi - \langle \varphi, \mu_{a_0} \rangle), \mu_{a_0} \rangle. \tag{5.14}$$

Proof. Let $\varphi \in \mathcal{O}$, then by Proposition 5.1.2

$$\frac{\langle \varphi, \mu_a - \mu_{a_0} \rangle}{|a - a_0|} = \frac{\langle \bar{\varphi}, \mu_a - \mu_{a_0} \rangle}{|a - a_0|} = \frac{\langle (\mathcal{P}_t^a - \mathcal{P}_t^{a_0})(1 - \mathcal{P}_t^{a_0})^{-1} \bar{\varphi}, \mu_a \rangle}{|a - a_0|}.$$

By Assumption **R2** we have that

$$\frac{\langle (\mathcal{P}_t^a - \mathcal{P}_t^{a_0})(1 - \mathcal{P}_t^{a_0})^{-1} \bar{\varphi}, \mu_a \rangle}{|a - a_0|} = \langle D_a (\mathcal{P}_t^a \psi) \big|_{a = a_0}, \mu_a \rangle + \langle r_a, \mu_a \rangle$$

with $\|r_a\|_U \to 0$ for $a \to a_0$. Thanks to Proposition 5.1.3, for a converging to a_0 we have

$$\langle D_a (\mathcal{P}_t^a (1 - \mathcal{P}_t^{a_0})^{-1} \bar{\varphi}) \big|_{a = a_0}, \mu_a \rangle \to \langle D_a (\mathcal{P}_t^a (1 - \mathcal{P}_t^{a_0})^{-1} \bar{\varphi}) \big|_{a = a_0}, \mu_{a_0} \rangle.$$

Finally as seen in (5.7)-(5.8), Assumption **R3** ensures that

$$\langle r_a, \mu_a \rangle \to 0 \quad \text{for } a \to a_0$$

and therefore $\langle \varphi, \mu_a \rangle$ is locally differentiable for all $\varphi \in \mathcal{O}$ and (5.14) holds. □

We have shown that, given a family of Markov semigroups $\mathcal{P}_t^a$ with relative invariant measure μ_a, acting on a space of observables $\mathcal{O}$ such that Assumption R holds, we have weak differentiability of the invariant measures with respect to the parameter $a \in \mathbb{R}$. We are now interested in giving possible examples of such space $\mathcal{O}$. In the original work of Hairer and Majda [40] the set $\mathcal{O}$ is the closure of $C_0^\infty(\mathcal{H})$ under the following norm

$$\|\varphi\|_{1;V,W} = \sup_{x \in \mathcal{H}} \left(\frac{|\varphi(x)|}{V(x)} + \frac{\|D\varphi(x)\|}{W(x)} \right) \tag{5.15}$$

where $V, W : \mathcal{H} \to [1, \infty)$ are two continuous functions. As the authors observe, it can be proven that if we quotient this space by the space of all constant function, then there is a distance function $d_{V,W}$ such that this norm is equivalent to the Lipschitz norm corresponding to $d_{V,W}$ i.e.

$$\|\varphi\|_{1;V,W} = \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{d_{V,W}(x, y)}.$$

This choice is probably guided by the results in [42]; indeed this work ensures the spectral gap condition in the norm (5.15) holds for a large class of hypoelliptic diffusions and in particular for the 2D stochastic Navier–Stokes equations.

In our analysis we will not use such a space of observables as by the asymptotic coupling approach we are able to establish the spectral gap condition with respect to a different norm. Indeed, in Corollary 4.1.8 we showed that the general form of Harris' theorem ensures $\mathcal{P}_t$ exhibits the spectral gap condition in the Lipschitz seminorm corresponding to the distance-like function $\tilde{d}$

$$\|\varphi\|_{\tilde{d}} = \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{\tilde{d}(x, y)}$$

with

$$\tilde{d}(x, y) = \sqrt{d_N(x, y)(1 + V(x) + V(y))}. \tag{5.16}$$

Here $V : \mathcal{H} \to [0, \infty]$ is the Lyapunov function as defined in Definition 4.1.4 and d_N is the distance-like function as given by Theorem 4.1.10 i.e.

$$\begin{aligned} d_N(x, y) &= N\theta_\alpha(x, y) \wedge N\theta_\alpha(y, x) \wedge 1, \quad \alpha \in [0, \alpha_0] \\ &= Ne^{\alpha v |x|^2} |x - y|^{2\alpha} \wedge Ne^{\alpha v |y|^2} |x - y|^{2\alpha} \wedge 1 \end{aligned} \tag{5.17}$$

with α_0 and υ as in (4.36).

As observed in section 4.1 (see (4.4) and subsequent calculations) $\|\cdot\|_{\tilde{d}}$ is only a seminorm as $\|\varphi\|_{\tilde{d}} = 0$ only implies that φ is constant. However for the framework presented it is not enough for $\|\cdot\|_{\tilde{d}}$ to be a seminorm and we have to get rid of the constant functions to ensure it is a norm. Then, instead of the space C_0^∞, consider its subset

$$C_c^\infty := \{\phi - \langle \phi, \mu_{a_0}\rangle \ : \ \phi \in C_0^\infty\} \tag{5.18}$$

and define

$$\mathcal{C}_{\tilde{d}} := \overline{C_c^\infty}^{\|\cdot\|_{\tilde{d}}} \quad \text{and} \quad \mathcal{C}_U := \overline{C_c^\infty}^{\|\cdot\|_U}. \tag{5.19}$$

where $\|\cdot\|_U$ is as in (5.4). Then $(\mathcal{C}_{\tilde{d}}, \|\cdot\|_{\tilde{d}})$ is the candidate observable space $(\mathcal{O}, \|\cdot\|_{\mathcal{O}})$.

Remark 5.1.5. Note that we could also defined $\mathcal{C}_{\tilde{d}}$ as the union of the space of constant functions and the completion of C_c^∞. As we have seen in the general framework though we would consider again the subspace of centred functions to invert the operator $(1 - \mathcal{P}_t^{a_0})$ so we prefer to define directly $\mathcal{C}_{\tilde{d}}$ not to include the constant functions. In this case then $\mathcal{O}$ will correspond to $\tilde{\mathcal{O}} = \mathrm{Im}(1 - \mathcal{P}_t^{a_0})$.

First of all to satisfy Assumption R we have to ensure that $(\mathcal{C}_{\tilde{d}}, \|\cdot\|_{\tilde{d}})$ is complete and find U such that it is a dense subset of $\mathcal{C}_U$.

Proposition 5.1.6. *Let $U : \mathcal{H} \to [1, \infty)$ be such that $U \geq \sqrt{1+V}$ and consider $\mathcal{C}_U$ and $\mathcal{C}_{\tilde{d}}$ as in* (5.19)*, where*

$$\|\varphi\|_U = \sup_{x\in\mathcal{H}} \frac{|\varphi(x)|}{U(x)} \quad \textit{and} \quad \|\varphi\|_{\tilde{d}} = \sup_{x\in\mathcal{H}} \frac{|\varphi(x) - \varphi(y)|}{\tilde{d}(x,y)}.$$

Then we have that:

(i) $\mathcal{C}_{\tilde{d}}$ *is a dense subset of* $\mathcal{C}_U$*;*

(ii) $(\mathcal{C}_U, \|\cdot\|_U)$ *is a complete metric space;*

(iii) $(\mathcal{C}_{\tilde{d}}, \|\cdot\|_{\tilde{d}})$ *is a complete metric space.*

Proof. Proof of (i) Since $\mathcal{C}_{\tilde{d}}$ and $\mathcal{C}_U$ are both defined as closure of the same set with respect to different norms, to show that $\mathcal{C}_{\tilde{d}}$ is dense in $\mathcal{C}_U$ it is sufficient to prove that $\mathcal{C}_{\tilde{d}}$ is continuously embedded in $\mathcal{C}_U$ i.e. that there exists a constant $k > 0$ such that

$$\|\varphi\|_U \leq k\|\varphi\|_{\tilde{d}} \quad \text{for all} \ \ \varphi \in \mathcal{C}_{\tilde{d}}. \tag{5.20}$$

Given $\varphi \in \mathcal{C}_{\tilde{d}}$, since it is by definition centred, we have

$$\begin{aligned}\frac{|\varphi(x)|}{U(x)} &= \frac{|\varphi(x) - \langle \varphi, \mu_{a_0}\rangle|}{U(x)} \\ &\leq \frac{1}{U(x)} \int |\varphi(x) - \varphi(y)| \, \mu_{a_0}(dy) \\ &\leq \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{\tilde{d}(x,y)} \int \frac{\tilde{d}(x,y)}{U(x)} \, \mu_{a_0}(dy).\end{aligned}$$

Now by definition of $\tilde{d}$, the fact that $d_N \leq 1$ and $U \geq 1$ we have

$$\begin{aligned}\int \frac{\tilde{d}(x,y)}{U(x)} \, \mu_{a_0}(dy) &= \int \frac{\sqrt{d_N(x,y)(1 + V(x) + V(y))}}{U(x)} \, \mu_{a_0}(dy) \\ &\leq \frac{\sqrt{1 + V(x)}}{U(x)} + \int \sqrt{V(y)} \, \mu_{a_0}(dy)\end{aligned}$$

and so

$$\sup_x \frac{|\varphi(x)|}{U(x)} \leq \|\varphi\|_{\tilde{d}} \left(\sup_x \frac{\sqrt{(1 + V(x)}}{U(x)} + \int \sqrt{V(y)} \, \mu_{a_0}(dy) \right).$$

It follows by the definition of the Lyapunov function V (Definition 4.1.4 and subsequent discussion) that it is integrable against μ_a for all a, and $U \geq \sqrt{1 + V}$ so there exists $k > 0$ such that $\|\varphi\|_U \leq k \|\varphi\|_{\tilde{d}}$ as desired.

Proof of (ii). Let $(\varphi_n)_n$ be a Cauchy sequence in $\mathcal{C}_U$ i.e. for all $\varepsilon > 0$ there exists $N > 0$ such that for all $n, m > N$

$$\|\varphi_n - \varphi_m\|_U = \sup_{x \in \mathcal{H}} \frac{|\varphi_n(x) - \varphi_m(x)|}{U(x)} < \varepsilon. \tag{5.21}$$

It follows that for all x the sequence $(\varphi_n(x))_n$ is Cauchy in $\mathbb{R}$ and there exists $\varphi(x)$ such that for all $\varepsilon > 0$ there exists $M_x > 0$ such that for all $m > M_x$

$$|\varphi_m(x) - \varphi(x)| < \varepsilon.$$

Then for all $n > N$ and for all x pick $m_x > \max(N, M_x)$

$$\begin{aligned}\frac{|\varphi_n(x) - \varphi(x)|}{U(x)} &\leq \frac{|\varphi_n(x) - \varphi_{m_x}(x)|}{U(x)} + \frac{|\varphi_{m_x}(x) - \varphi(x)|}{U(x)} \\ &\leq \frac{|\varphi_n(x) - \varphi_{m_x}(x)|}{U(x)} + |\varphi_{m_x}(x) - \varphi(x)| < 2\varepsilon\end{aligned}$$

using that $U(x) \geq 1$ and the fact that since (5.21) holds for all $n, m > N$, in particular it holds for any such choice of m_x. We have then shown that for for all $n > N$

$$\sup_{x \in \mathcal{H}} \frac{|\varphi_n(x) - \varphi(x)|}{U(x)} < \varepsilon.$$

Finally it is easy to see that if two Cauchy sequences differ by a sequence converging to zero, then they must have the same limit. Let φ_n and ψ_n be two Cauchy sequences in $\mathcal{C}_U$ such that $\psi_n \to \psi$ and $\varphi_n = \psi_n + a_n$ with $\|a_n\|_U \to 0$ for $n \to \infty$. Then for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$\|\varphi_n - \psi\|_U \leq \|\varphi_n - a_n - \psi\|_U + \|a_n\|_U < 2\varepsilon$$

for all $n > N$.

Proof of (iii). Let $(\varphi_n)_n$ be a Cauchy sequence in $\mathcal{C}_{\tilde{d}}$, i.e. for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n, m > N$

$$\|\varphi_n - \varphi_m\|_{\tilde{d}} = \sup_{x \neq y} \frac{|\varphi_n(x) - \varphi_m(x) - \varphi_n(y) + \varphi_m(y)|}{\tilde{d}(x,y)} < \varepsilon \tag{5.22}$$

In particular then by (i) it is Cauchy also in $\mathcal{C}_U$ and by (ii), it is convergent in $\mathcal{C}_U$. We will show that the convergence holds in $\mathcal{C}_{\tilde{d}}$ as well. As seen in (ii), since $(\varphi)_n$ is Cauchy in $\mathcal{C}_U$, for all x there exists $\varphi(x)$ such that $\varphi_n(x) \to \varphi(x)$ for all $n \to \infty$. For an arbitrary $\varepsilon > 0$ then for all $x \neq y$ there exists $M_{x,y}$ such that for all $m_{x,y} > M_{x,y}$

$$|\varphi_m(x) - \varphi(x)| \leq \varepsilon\tilde{d}(x,y) \quad \text{and} \quad |\varphi_m(y) - \varphi(y)| \leq \varepsilon\tilde{d}(x,y)$$

For all $n > N$ and for all x, y we can pick an $m > \max(N, M_{x,y})$ and note that

$$\frac{|\varphi_n(x) - \varphi(x) - \varphi_n(y) + \varphi(y)|}{\tilde{d}(x,y)} \leq \frac{|\varphi_n(x) - \varphi_{m_{x,y}}(x) - \varphi_n(y) + \varphi_{m_{x,y}}(y)|}{\tilde{d}(x,y)} + \frac{|\varphi_{m_{x,y}}(x) - \varphi(x)|}{\tilde{d}(x,y)} + \frac{|\varphi_{m_{x,y}}(y) - \varphi(y)|}{\tilde{d}(x,y)}$$

hence

$$\frac{|\varphi_n(x) - \varphi(x) - \varphi_n(y) + \varphi(y)|}{\tilde{d}(x,y)} \leq 3\varepsilon \quad \text{for all } \; x \neq y.$$

□

Having shown that $\mathcal{C}_{\tilde{d}}$ is a complete metric space we set the space of observables

$$(\mathcal{O}, \|\cdot\|_{\mathcal{O}}) = (\mathcal{C}_{\tilde{d}}, \|\cdot\|_{\tilde{d}}).$$

Then the results in section 4.1.1, specifically Corollary 4.1.8, will ensure Assumption **R1** holds. In the next subsection we will focus on semigroups $\mathcal{P}_t^a$ associated to a particular type nonlinear stochastic differential equation on the Hilbert space $\mathcal{H}$ and show verifiable conditions to ensure the rest of Assumption R holds.

5.1.1 Application to SPDEs

Let $(\mathcal{H}, |\cdot|)$ and $(\mathcal{V}, \|\cdot\|)$ be Hilbert spaces with $\mathcal{V} \subset\subset \mathcal{H}$ and consider the stochastic equation

$$dX = (AX + F(X))\, dt + af\, dt + dW, \quad X(0) = x \tag{5.23}$$

where A is a nonnegative selfadjoint linear operator, $f \in \mathcal{H}$ is a deterministic forcing modulated in intensity by the constant $a \in \mathbb{R}$, W is a Wiener process on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathcal{H}$ and trace class covariance operator Q, and F a nonlinear function such that there exists a unique solution for any initial condition $X(0) = x$. Assume then that the associated Markov semigroup

$$(\mathcal{P}_t^a \varphi)(x) = \mathbb{E}\, \varphi(X(t, \cdot; a, x)) \quad \text{for all } \varphi \in B_b(\mathcal{H})$$

is well defined and admits an invariant measure μ_a.

Given the space of observables $(\mathcal{C}_{\tilde{d}}, \|\cdot\|_{\tilde{d}})$ we want to show that the model exhibit linear response with respect to the parameter a, so we will show sufficient conditions for Assumption R to hold. In chapter 4 we saw that if Assumption A in section 4.1.1 is satisfied, the spectral gap condition **R1** holds for $\|\cdot\|_{\tilde{d}}$ i.e. there exists $t > 0$ and $\rho < 1$ such that

$$\|\mathcal{P}_t^{a_0} \varphi\|_{\tilde{d}} \le \rho\, \|\varphi\|_{\tilde{d}}$$

for all $\varphi \in \mathcal{C}_{\tilde{d}}$.

We move on looking at the condition **R2**. We have to ensure that for an appropriate choice of the function U, the function $a \mapsto \mathcal{P}_t^a \varphi$ with values in $\mathcal{C}_U$ is differentiable at a_0, for any $\varphi \in \mathcal{C}_{\tilde{d}}$, namely

$$\lim_{a \to a_0} \left\| \frac{\mathcal{P}_t^a \varphi - \mathcal{P}_t^{a_0} \varphi}{a - a_0} - D_a \mathcal{P}_t^a \varphi|_{a=a_0} \right\|_U = 0. \tag{5.24}$$

In order to do so we restrict our analysis to *finite dimensional* forcings f and transforming (5.24) to a simpler formulation using Girsanov theorem. Recall that we can see the infinite dimensional Wiener process as sum of two Wiener processes

$$W(t) = W_n(t) + W^n(t) = \sum_{k=1}^{n} \sqrt{\sigma_k} \beta_k(t) e_k + \sum_{k=n+1}^{\infty} \sqrt{\sigma_k} \beta_k(t) e_k,$$

respectively with covariance matrix Q_n and Q^n as in (3.54).

Theorem 5.1.7. *Consider the system* (5.23) *with* $f \in \mathrm{Range}(Q_n)$ *for some* $n \in \mathbb{N}$. *Suppose* $\mathcal{P}_t^{a_0}$ *admits a Lyapunov function* V_{a_0} *and consider a function* $U : \mathcal{H} \to [1, \infty)$ *such that* $U \geq \sqrt{1 + V_{a_0}}$. *Then for all* $\varphi \in \mathcal{C}_{\tilde{d}}$, *the function* $a \mapsto \mathcal{P}_t^a \varphi \in \mathcal{C}_U$ *is differentiable in* a_0. *Furthermore the derivative* $(D_a \mathcal{P}_t^a)|_{a=a_0}$ *is a bounded operator from* $\mathcal{C}_{\tilde{d}}$ *into* $\mathcal{C}_U$, *i.e. there exists* $C = C(a_0) > 0$ *such that*

$$\| \, D_a \mathcal{P}_t^a \varphi|_{a=a_0} \, \|_U \leq C \|\varphi\|_{\tilde{d}} \quad \textit{for all } \varphi \in \mathcal{C}_{\tilde{d}} \tag{5.25}$$

Proof. Let $\{e_k\}$ be an orthonormal basis of $\mathcal{H}$ made of eigenfunctions of Q. As $f \in \mathrm{Range}(Q_n)$ there exists $f_k \in \mathbb{R}$, $k = 1, \ldots, k$ such that

$$f = \sum_{k=1}^{n} f_k e_k. \tag{5.26}$$

Set $h = a - a_0$. Note that on a finite time domain $[0, T]$ we immediately have

$$\int_0^T |h Q_n^{-1/2} f|^2 \, ds < \infty$$

so that, by Girsanov theorem (Theorem 3.2.5), the process

$$\tilde{W}_n(t) := hft + W_n(t) \tag{5.27}$$

is a Q_n-Wiener process on $(\Omega, \tilde{\mathbb{P}})$ where $\tilde{\mathbb{P}}$ is equivalent to the original probability measure $\mathbb{P}$ with density

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}(t) = \exp\left(hM(t) - \frac{h^2}{2} \langle M \rangle_t \right) \tag{5.28}$$

where

$$M(t) = Q_n^{-1} f W_n(t) \quad \text{and} \quad \langle M \rangle_t = |Q_n^{-1/2} f|^2 t.$$

As this transformation depends only on W_n, and since W_n and W^n are independent, we get that $\tilde{W}_n$ and $\tilde{W} := \tilde{W}_n + W^n$ are Wiener processes on $(\Omega, \tilde{\mathbb{P}})$. Therefore $\tilde{Y}$, solution of

$$d\tilde{Y} = (A\tilde{Y} + F(\tilde{Y}))\, dt + a_0 f \, dt + d\tilde{W} \quad \tilde{Y}(0) = x \tag{5.29}$$

is equivalent to X, solution of (5.23). Denoting as Y the solution of

$$dY = (AY + F(Y))\, dt + a_0 f \, dt + dW \quad Y(0) = x \tag{5.30}$$

it follows from (5.28) that

$$\mathbb{E}\,\varphi(X(t,x)) = \tilde{\mathbb{E}}\,\varphi(\tilde{Y}(t,x)) = \mathbb{E}\left[\varphi(Y(t,x))\exp\Big(hM(t) - \frac{h^2}{2}\langle M\rangle_t\Big)\right]. \tag{5.31}$$

We have to show that $\mathbb{E}\varphi(X(t,x))$, for which we just derived an equivalent formulation, is differentiable in a_0. Now, taking *formally* the derivative of (5.31) with respect to h in zero we have

$$D_h\mathbb{E}\left[\varphi(Y(t,x))\exp\Big(hM(t) - \frac{h^2}{2}\langle M\rangle_t\Big)\right]\Bigg|_{h=0} = \mathbb{E}\left[\varphi(Y(t,x))M(t)\right]. \tag{5.32}$$

We have to make sure this candidate is indeed the derivative of $\mathcal{P}^a\varphi$ at a_0, namely ensure that

$$\lim_{h\to 0}\sup_x \frac{1}{U(x)}\left|\mathbb{E}\left[\frac{\varphi(X(t,x)) - \varphi(Y(t,x))}{h} - \varphi(Y(t,x))M(t)\right]\right| = 0. \tag{5.33}$$

It is evident that (5.33) does not change if we consider $\varphi + k$ with k being a constant, which can depend on the initial condition x. Therefore using (5.31) we have

$$\mathbb{E}\left[\frac{\varphi(X(t,x)) - \varphi(Y(t,x))}{h} - \varphi(Y(t,x))M(t)\right] =$$

$$\mathbb{E}\left[(\varphi(Y(t,x)) - k)\frac{\exp\Big(hM(t) - \frac{h^2}{2}\langle M\rangle_t\Big) - 1 - hM(t)}{h}\right].$$

Hence setting

$$F(h) = \tfrac{1}{h}\exp\Big(hM(t) - \tfrac{h^2}{2}\langle M\rangle_t\Big) - \tfrac{1}{h} - M(t) \tag{5.34}$$

it suffices to show

$$\lim_{h\to 0}\sup_x \frac{1}{U(x)}\mathbb{E}\left[|\varphi(Y(t,x)) - k|\,|F(h)|\right] = 0,$$

for some k which is free to chose.

Since $\|\varphi\|_{\tilde{d}} < \infty$, for any fixed $t > 0$, we have, by definition of $\tilde{d}$ (5.16)

$$|\varphi(Y(t)) - \varphi(x)| \le \|\varphi\|_{\tilde{d}}\,\tilde{d}(Y(t),x) \le \|\varphi\|_{\tilde{d}}(1 + V_{a_0}(Y(t)) + V_{a_0}(x))^{1/2} \tag{5.35}$$

where V_{a_0} is the Lyapunov function. Therefore

$$\begin{aligned}\mathbb{E}\left[|\varphi(Y(t,x)) - \varphi(x)|\,|F(h)|\right] &\le \|\varphi\|_{\tilde{d}}\,\mathbb{E}\left[(1 + V_{a_0}(Y(t)) + V_{a_0}(x))^{1/2}\,|F(h)|\right]\\ &\le \|\varphi\|_{\tilde{d}}\sqrt{\mathbb{E}(1 + V_{a_0}(Y(t)) + V_{a_0}(x))}\sqrt{\mathbb{E}\,|F(h)|^2},\end{aligned}$$

by Cauchy-Schwartz inequality, so that

$$\lim_{h\to 0}\sup_{x\in\mathcal{H}}\frac{1}{U(x)}\mathbb{E}|\varphi(Y(t,x))-\varphi(x)|\,|F(h)|\leq$$
$$\|\varphi\|_{\tilde{d}}\underbrace{\sup_{x\in\mathcal{H}}\frac{\sqrt{1+V_{a_0}(x)+\mathbb{E}V_{a_0}(Y(t))}}{U(x)}}_{(I)}\underbrace{\lim_{h\to 0}\sqrt{\mathbb{E}\,|F(h)|^2}}_{(II)}.$$

Let us examine the terms on the right hand side of this expression.

(I) By the definition of the Lyapunov function V_{a_0}, and in particular as shown in the proof of Theorem 4.1.10, there exist constants $\gamma_1, K(a_0) > 0$ such that
$$\mathcal{P}_t^{a_0}V_{a_0}(x)\leq e^{-\gamma_1 t}V_{a_0}(x)+\tfrac{K(a_0)}{\gamma_1}.$$
It follows
$$\sup_x\frac{\sqrt{1+V_{a_0}(x)+\mathbb{E}V_{a_0}(Y(t,x))}}{U(x)}\leq\sup_{x\in\mathcal{H}}\frac{\sqrt{1+\frac{K(a_0)}{\gamma_1}+V_{a_0}(x)(1+e^{-\gamma_1 t})}}{U(x)}.$$
Then, since $U(x)\geq\sqrt{1+V_{a_0}(x)}$ the right hand side stays bounded.

(II) By the differentiability of the exponential function $\lim_{h\to 0}|F(h)| = 0$ almost surely. We have to ensure that $|F(h)|^2$ is bounded by an integrable function, so that the dominated convergence theorem gives the convergence also in expectation. Rewrite F using the mean value theorem
$$\begin{aligned}&\exp\Big(hM(t)-\tfrac{h^2}{2}\langle M\rangle_t\Big)\\&=\int_0^1\exp\Big(rhM(t)-r\tfrac{h^2}{2}\langle M\rangle_t\Big)\Big(hM(t)-\tfrac{h^2}{2}\langle M\rangle_t\Big)\,dr+1.\end{aligned}$$
So we get that
$$\begin{aligned}|F(h)|&=\left|\frac{1}{h}\exp\Big(hM(t)-\tfrac{h^2}{2}\langle M\rangle_t\Big)-\frac{1}{h}-M(t)\right|\\&=\left|\int_0^1\exp\Big(rhM(t)-r\tfrac{h^2}{2}\langle M\rangle_t\Big)\Big(M(t)-\frac{h}{2}\langle M\rangle_t\Big)\,dr-M(t)\right|\\&\leq\left|\int_0^1\exp\Big(rhM(t)-r\tfrac{h^2}{2}\langle M\rangle_t\Big)\,dr-1\right|\,|M(t)|.\end{aligned}$$

It follows that, take h in a ε-neighbourhood of 0, we have

$$|F(h)| \le \exp(\varepsilon\,|M(t)|)\,|M(t)|$$

and consequently, since $M(t)$ is normally distributed,

$$\mathbb{E}|F(h)|^2 \le \mathbb{E}\exp(2\varepsilon|M(t)|)|M(t)|^2 < \infty.$$

for ε small enough.

We can conclude that $a \mapsto \mathcal{P}_t^a\varphi$ is differentiable and its derivative at a_0 is

$$\left.(D_a\mathcal{P}^a\varphi)\right|_{a=a_0}(x) = \mathbb{E}\left[\,\varphi(Y(t,x))M(t)\,\right]. \tag{5.36}$$

We are left to show that the derivative is bounded as an operator from $\mathcal{C}_{\tilde{d}}$ to $\mathcal{C}_U$, i.e. there exists C such that

$$\|\left.(D_a\mathcal{P}^a\varphi)\right|_{a=a_0}\|_U = \sup_x \frac{|\,\mathbb{E}\left[\varphi(Y(t,x))M(t)\right]|}{U(x)} \le C\|\varphi\|_{\tilde{d}}$$

for all $\varphi \in \mathcal{C}_{\tilde{d}}$. Let us start noticing that

$$\begin{aligned}|\,\mathbb{E}\left[\,\varphi(Y(t,x))M(t)\,\right]| &= |\mathbb{E}\left[((\varphi(Y(t,x)) - \varphi(x))\,M(t)\right]| \\ &\le \mathbb{E}\left[|\varphi(Y(t,x)) - \varphi(x)||M(t)|\right] \\ &\le |Q_n^{-1}f|\mathbb{E}\left[|\varphi(Y(t,x)) - \varphi(x)||W_n(t)|\right].\end{aligned}$$

Then by (5.35) and Cauchy-Schwartz inequality we have that

$$\begin{aligned}|\mathbb{E}\left[\,\varphi(Y(t,x))W_n(t)\,\right]| &\le \|\varphi\|_{\tilde{d}}\,\mathbb{E}\left[(1+V_{a_0}(x)+V_{a_0}(Y(t,x)))^{1/2}\,|W_n(t)|\,\right] \\ &\le \|\varphi\|_{\tilde{d}}\,(1+V_{a_0}(x)+\mathbb{E}\,V_{a_0}(Y(t,x)))^{1/2}\left(\mathbb{E}\,|W_n(t)|^2\right)^{1/2}.\end{aligned}$$

As seen above, U is such that

$$K_U := \sup_{x\in\mathcal{H}} \frac{\sqrt{1+V_{a_0}(x)+\mathbb{E}\,V_{a_0}(Y(t,x))}}{U(x)} < \infty,$$

so that

$$\|\left.(D_a\mathcal{P}_t^a\varphi)\right|_{a=a_0}\|_U \le C\|\varphi\|_{\tilde{d}}$$

with $C = |Q_n^{-1}f|K_U t^{1/2}$. □

We are ready to show the following result for an equation like (5.23).

Theorem 5.1.8. *Consider the system* (5.23) *with* $f \in \mathrm{Range}(Q_n)$ *for some* $n \in \mathbb{N}$ *and suppose Assumption* ***A1****,* ***A2****,* ***A3*** *holds for* $a_0 \in \mathbb{R}$*. Assume there exists a measurable function* $V : \mathcal{H} \to \mathbb{R}_+$ *and positive constants* γ_1 *and* K *satisfying* ***A4*** *for all* a *in a* ε*-neighbourhood of* a_0*. Then, setting the function* $U : \mathcal{H} \to [1, \infty)$ *to be* $U = \sqrt{1+V}$*, the map* $a \mapsto \langle \varphi, \mu_a \rangle$ *is differentiable at* a_0 *for every* $\varphi \in \mathcal{C}_{\tilde{d}}$ *and the identity*

$$\frac{\mathrm{d}}{\mathrm{d}a}\langle \varphi, \mu_a \rangle \bigg|_{a=a_0} = \langle D_a \mathcal{P}_t^{a_0}(1 - \mathcal{P}_t^{a_0})^{-1}\varphi, \mu_{a_0} \rangle$$

holds.

Proof. In order to apply Theorem 5.1.4 we should show that Assumption R holds. Since there exists V such that **A4** holds uniformly in a we have

$$\mathbb{E}V(X(t,a)) \leq \mathbb{E}V(X(s,a)) + \int_s^t \left(-\gamma_1 \mathbb{E}V(X(\tau,a)) + K\right) \, d\tau, \quad t \geq s \geq 0$$

for all $a \in B_\varepsilon(a_0)$, ε-neighbourhood of a_0, hence in particular V is a Lyapunov function for any such a i.e.

$$\mathcal{P}_t^a V(x) \leq e^{-\gamma_1 t} V(x) + \tfrac{K}{\gamma_1} \quad \text{for all } a \in B_\varepsilon(a_0). \tag{5.37}$$

It follows that Theorem 5.1.7 holds with $U = \sqrt{1+V} \geq \sqrt{1+V_a}$ for all $a \in B_\varepsilon(a_0)$. This implies that the map $a \mapsto \mathcal{P}_t^a \varphi$ is differentiable in $B_\varepsilon(a_0)$ and $D_a \mathcal{P}_t^a$ is a bounded operator uniformly in $a \in B_\varepsilon(a_0)$, satisfying Assumption **R2**.

Next, Assumption **R1**, namely the spectral gap for $\mathcal{P}_t^{a_0}$, is given by the fact that Assumption A holds. Therefore we only have to ensure that Assumption **R3** holds for the choice of U we made i.e. there exists $\varepsilon > 0$ such that

$$\sup_{|a-a_0|<\varepsilon} \langle \sqrt{1+V},\, \mu_a \rangle < \infty.$$

Since $\langle \sqrt{1+V},\, \mu_a \rangle \leq \sqrt{\langle 1+V,\, \mu_a \rangle}$ it is enough to show

$$\sup_{|a-a_0|<\varepsilon} \langle 1+V,\, \mu_a \rangle < \infty.$$

As V is a Lyapunov function for any $a \in B_\varepsilon(a_0)$, by definition of Lyapunov function and by (5.37), as we saw in (4.10), we have that

$$\langle V, \mu_a \rangle \leq \frac{K}{\gamma_1(1 - e^{-\gamma_1 t})},$$

giving the desired result.

□

5.2 Fractional response for SPDEs

So far we showed that, as a function of the parameter a, the invariant measure μ_a is weakly differentiable for observables in the space $\mathcal{C}_{\tilde{d}}$ when the forcing f is finite dimensional and Assumption R holds.

On the other hand, weaker regularity is still attainable under less restrictive conditions on the forcing, in fact we will show that for every $\varphi \in \mathcal{C}_{\tilde{d}}$, the map $a \mapsto \langle \varphi, \mu_a \rangle$ is local α-Hölder continuous i.e. there exists $C = C(a_0, \varphi)$ such that

$$|\langle \varphi, \mu_a - \mu_{a_0} \rangle| \leq C(\varphi)|a - a_0|^\alpha \tag{5.38}$$

for an appropriate range of $\alpha \in (0, \alpha_0)$.

In the literature weaker dependence of the invariant measure on model parameter has been studied in [42, Section 5.5], for the 2D stochastic Navier–Stokes equations, and later in [44, Section 4.1] despite it not being labelled as response. As mentioned in the previous chapter, in the first work, Hairer and Mattingly developed a method for the existence of spectral gaps in an appropriate Wasserstein distance making use of Malliavin calculus and previous results on hypoelliptic operators by the same authors [41]. As an example of application it is shown that the invariant measure of 2D Navier–Stokes equation with additive noise is locally continuous with respect to the parameters ν, f and Q, namely the viscosity, the deterministic external forcing and the covariance operator. There the distance on the space of measures is the Wasserstein distance associated to the norm (5.15) with $V(x) = W(x) = \exp(\eta|x|^2)$ with η an appropriate positive parameter.

On the other hand in [44] the authors, given the spectral gap in $\mathcal{C}_{\tilde{d}}$, show that if $\tilde{d}$ satisfies a weak form of the triangular inequality, i.e. there exists a positive constant k such that

$$\tilde{d}(x, y) \leq k(\tilde{d}(x, z) + \tilde{d}(z, y)) \quad \text{for all } x, y, z \in \mathcal{H} \tag{5.39}$$

and the transition probabilities P_t^a are such that

$$W_{\tilde{d}}(P_t^a(x, \cdot), P_t^{a_0}(y, \cdot)) \leq (a - a_0)C(t)\tilde{V}(x) \tag{5.40}$$

for some positive function $\tilde{V}$ and $C(t)$ function bounded on bounded sets of $\mathbb{R}$, then

$$W_{\tilde{d}}(\mu_a, \mu_{a_0}) \leq 2(a - a_0)C(t)\langle \tilde{V}, \mu_a \rangle. \tag{5.41}$$

Lipschitz continuity follows if $\tilde{V}$ is such that $\langle \tilde{V}, \mu_a \rangle$ is bounded uniformly in a.

We will show that indeed a condition similar to (5.40) holds when $\tilde{d}$ is defined as in (5.16), and that, without requiring (5.39) to hold, we obtain

(5.38), which is a statement somewhat weaker than (5.41) as

$$|\langle \varphi, \mu_a - \mu_{a_0} \rangle| \leq \|\varphi\|_{\tilde{d}} W_{\tilde{d}}(\mu_a, \mu_{a_0}). \tag{5.42}$$

In order to prove (5.38) we will do a similar construction to that in Theorem 4.1.10 but now X and Y do not differ for their initial condition but for the value of the parameter a, i.e. consider

$$\begin{aligned} dX &= (AX + F(X))\,dt + a_0 f\,dt + dW \quad X(0) = x \\ dY &= (AY + F(Y))\,dt + af\,dt + dW \quad Y(0) = x \end{aligned} \tag{5.43}$$

and we will impose the following conditions which take both from Assumption A and Assumption R.

Assumption H.

H1 Assumption A holds for all a in a ε-neighbourhood about a_0 and the relative constants κ_i and K are such that

$$K_\varepsilon := \sup_{|a-a_0|<\varepsilon} K(a) < \infty \quad \text{and} \quad \sup_{|a-a_0|<\varepsilon} \kappa_i(a) < \infty, \quad i = 0, \dots 3. \tag{5.44}$$

H2 Given $X(t), Y(t)$ solutions of (5.43), there exists a positive constant C such that

$$|X(t) - Y(t)|^2 \leq C|a - a_0|^2 \exp\left(\kappa_1 \int_0^t \|X(s)\|^2\,ds \right) \quad \text{for all } t \geq 0 \tag{5.45}$$

where κ_1 is as given in Assumption A.

H3 For all a in a ε-neighbourhood of a_0, given $\kappa_1, \kappa_2, \gamma$ as given by Assumption A, set

$$\upsilon = \frac{\kappa_1}{\kappa_2} \quad \text{and} \quad \alpha_0 = \frac{1}{2} \wedge \frac{2\gamma}{\upsilon + 2\gamma}.$$

The Lyapunov function V is such that for all $\alpha \in (0, \alpha_0)$

$$\sup_{|a-a_0|<\varepsilon} \int e^{\alpha\upsilon|x|^2/2} (1 + V(x))^{1/2}\, \mu_a(dx) < \infty.$$

Theorem 5.2.1. *Consider the system* (5.43) *and suppose Assumption H holds. Then for all* $\alpha \in (0, \alpha_0)$ *there exists* $C > 0$ *such that*

$$|\langle \varphi, \mu_a - \mu_{a_0} \rangle| \leq C \|(1 - \mathcal{P}^{a_0})^{-1} \varphi\|_{\tilde{d}} |a - a_0|^\alpha \quad \textit{for all } \varphi \in \mathcal{C}_{\tilde{d}}$$

for all a *in the* ε*-neighbourhood of* a_0*, i.e.* μ_a *is locally* α*-Hölder continuous with respect to the parameter* a*.*

Proof. First of all we show that it is enough to show that there exists $C = C(\alpha, t, a_0, \varepsilon)$ and a positive function $\tilde{V}(x)$ such that

$$W_{\tilde{d}}(P_t^a(x,\cdot), P_t^{a_0}(x,\cdot)) \leq |a-a_0|^\alpha C(t)\tilde{V}(x) \quad \text{with} \tag{5.46}$$

$$\sup_{|a-a_0|<\varepsilon} \langle \tilde{V}(x), \mu_a\rangle < \infty. \tag{5.47}$$

By **H1** Assumption A holds for all $a \in B_\varepsilon(a_0)$ and so, as we have seen in Theorem 5.1.8 and more specifically in Corollary 4.1.8, the spectral gap property holds, namely there exists $\rho < 1$ and $t > 0$ such that

$$\|\mathcal{P}_t^{a_0}\varphi - \langle \varphi, \mu_{a_0}\rangle\|_{\tilde{d}} \leq \rho\|\varphi - \langle \varphi, \mu_{a_0}\rangle\|_{\tilde{d}} \quad \text{for all } \varphi \in \mathcal{C}_{\tilde{d}}.$$

Then, since $\mathcal{C}_{\tilde{d}}$ is by definition made of functions centred with respect to μ_{a_0}, Proposition 5.1.2 ensures that for each $\varphi \in \mathcal{C}_{\tilde{d}}$ there is a unique $\psi \in \mathcal{C}_{\tilde{d}}$ such that $\varphi = (1-\mathcal{P}_t^{a_0})\psi$, and we denote such ψ as $(1-\mathcal{P}_t^{a_0})^{-1}\varphi$. Therefore as seen at the beginning of the chapter from (5.1) we have

$$\langle \varphi, \mu_a - \mu_{a_0}\rangle = \langle (\mathcal{P}_t^a - \mathcal{P}_t^{a_0})\psi, \mu_a\rangle.$$

Now thanks to the definition of the $\tilde{d}$-Wasserstein semidistance and the relation (5.42), it is easy to see that

$$\begin{aligned}
|\langle (\mathcal{P}_t^a - \mathcal{P}_t^{a_0})\psi, \mu_a\rangle| &\leq \langle |(\mathcal{P}_t^a - \mathcal{P}_t^{a_0})\psi|\,, \mu_a\rangle \\
&= \int_H |\langle \psi, P_t^a(x,\cdot)\rangle - \langle \psi, P_t^{a_0}(x,\cdot)\rangle|\; \mu_a(dx) \\
&\leq \|\psi\|_{\tilde{d}} \int_H W_{\tilde{d}}(P_t^a(x,\cdot), P_t^{a_0}(x,\cdot))\, \mu_a(dx).
\end{aligned}$$

Therefore by (5.46) and (5.47)

$$|\langle \varphi, \mu_a - \mu_{a_0}\rangle| \leq |a-a_0|^\alpha \|\psi\|_{\tilde{d}}\, C(t) \sup_{|a-a_0|<\varepsilon} \langle \tilde{V}(x), \mu_a\rangle \tag{5.48}$$

i.e. μ_a is weakly α-Hölder continuous in a_0.

We can then focus on showing (5.46) holds. By definition of the Wasserstein semidistance $W_{\tilde{d}}$ (4.2) we have

$$W_{\tilde{d}}(P_t^{a_0}(x,\cdot), P_t^a(x,\cdot)) \leq \mathbb{E}\, \tilde{d}(X(t), Y(t)) \tag{5.49}$$

and, thanks to Cauchy-Schwartz inequality,

$$\begin{aligned}
\mathbb{E}\, \tilde{d}(X(t), Y(t)) &= \mathbb{E}\, \sqrt{d_N(X(t), Y(t))(1 + V(X(t)) + V(Y(t)))} \\
&\leq \sqrt{\mathbb{E} d_N(X(t), Y(t))}\sqrt{1 + \mathbb{E}\, V(X(t)) + \mathbb{E}\, V(Y(t))}.
\end{aligned} \tag{5.50}$$

Let us first bound $\mathbb{E}\, d_N(X(t), Y(t))$ similarly as done in the proof of Theorem 4.1.10. From Assumption A, **A2** gives that there exists $\kappa_2 > 0$, $\kappa_3 = \kappa_3(a_0) \geq 0$ and a random variable $\Xi_\gamma^{a_0}$ depending on a parameter $\gamma > 0$, such that

$$|X(t)|^2 + \kappa_2 \int_0^t \|X(s)\|^2\, ds \leq |x|^2 + \kappa_3 t + \Xi_\gamma^{a_0} \quad t \geq 0 \tag{5.51}$$

with

$$\mathbb{P}(\Xi_\gamma^{a_0} \geq R) \leq e^{-2\gamma R}, \quad \text{for all } R \geq 0. \tag{5.52}$$

Then, using (5.51) in the estimate **H2**, we have that

$$|X(t) - Y(t)|^2 \leq C|a - a_0|^2 \exp\Big(\tfrac{\kappa_1}{\kappa_2}(|x|^2 - |X(t)|^2 + \kappa_3 t + \Xi_\gamma^{a_0})\Big). \tag{5.53}$$

Recall that by definition of d_N (5.17)

$$d_N(X, Y) \leq N\theta_\alpha(X(t), Y(t)) = N|X(t) - Y(t)|^{2\alpha} e^{\alpha\upsilon|X(t)|^2}$$

where $\upsilon = \kappa_1/\kappa_2$. Then by (5.53)

$$\mathbb{E}\,|X(t) - Y(t)|^{2\alpha} e^{\alpha\upsilon|X(t)|^2} \leq C|a - a_0|^{2\alpha} \exp\big(\alpha\upsilon|x|^2 + \alpha\upsilon\kappa_3 t\big)\mathbb{E}\, \exp\big(\alpha\upsilon\Xi_\gamma^{a_0}\big).$$

Since we consider $\alpha \in (0, \alpha_0)$ where

$$\alpha_0 = \frac{1}{2} \wedge \frac{2\gamma}{\upsilon + 2\gamma} < \frac{2\gamma}{\upsilon}$$

then $\mathbb{E}\exp\big(\alpha\upsilon\Xi_\gamma^{a_0}\big) =: C_\Xi$ is finite for all $\alpha \in (0, \alpha_0)$ and we have

$$\mathbb{E}\, d_N(X(t), Y(t)) \leq N\mathbb{E}\, \theta_\alpha(X(t), Y(t)) \leq NC_\Xi|a - a_0|^{2\alpha} e^{\alpha\upsilon|x|^2 + \alpha\upsilon\kappa_3 t}.$$

Looking back at (5.50) we have found that

$$\mathbb{E}\, \tilde{d}(X(t), Y(t)) \leq |a - a_0|^\alpha N^{1/2} C_\Xi^{1/2} e^{\alpha\upsilon|x|^2 + \alpha\upsilon\kappa_3 t/2} \cdot \sqrt{1 + \mathbb{E}V(X(t)) + \mathbb{E}V(Y(t))}.$$

By definition of Lyapunov function V (Definition 4.1.4), we have that

$$\mathbb{E}V(X(t)) + \mathbb{E}V(Y(t)) \leq 2e^{-\gamma_1 t}V(x) + \frac{2K_\varepsilon}{\gamma_1} \leq 2V(x) + \frac{2K_\varepsilon}{\gamma_1}$$

where $K_\varepsilon = \sup_{|a - a_0| < \varepsilon} K(a)$ stays finite by the assumptions on K in **H1**.

Then we have showed

$$\mathbb{E}\,\tilde{d}(X(t),Y(t)) \le |a-a_0|^\alpha C(t)\tilde{V}(x)$$

with

$$C(t) = N^{1/2}C_\Xi^{1/2}e^{\alpha\upsilon\kappa_3 t/2} \quad \text{and} \quad \tilde{V}(x) = e^{\alpha\upsilon|x|^2/2}\Big(1+\tfrac{2K_\varepsilon}{\gamma_1}+2V(x)\Big)^{1/2} \tag{5.54}$$

and thanks to **H3** it follows that $\sup_{|a-a_0|<\varepsilon}\langle\tilde{V},\mu_a\rangle$ stays finite.

Finally by (5.49), the bound (5.46) holds and consequently (5.48) is fulfilled. □

In the next section we show that Theorem 5.1.8 and Theorem 5.2.1 apply to the stochastic Navier–Stokes equations laying out the proofs in such a way that they generalise easily to the model of interest, the stochastic two–layer quasi–geostrophic model.

5.3 Stochastic Navier–Stokes equations

Consider the Navier–Stokes equations as in (1.14) with deterministic forcing $f = f(x) \in \mathcal{H} = \mathbf{L}^2(\mathcal{D})$ modulated by a multiplicative constant $a > 0$,

$$du + (\nu Au + B(u,u))\,dt = af\,dt + dW \quad u(0) = u_0 \tag{5.55}$$

and the Markov semigroup, acting on bounded measurable observables, associated to its unique solution $u = u(t,\omega;u_0,a)$

$$(\mathcal{P}_t^a\varphi)(u_0) = \mathbb{E}\,\varphi(u(t,\cdot;u_0,a)). \tag{5.56}$$

In section 4.2 we showed that the model exhibits exponential convergence of transition probabilities with respect to the $\tilde{d}$-Wasserstein distance where $\tilde{d}$ is the distance-like function (5.16) with parameters

$$\upsilon = \frac{k_B}{\nu(\nu-\gamma\lambda_1^{-1}\operatorname{Tr}Q)} > 0 \quad \text{and} \quad \alpha_0 = \frac{1}{2}\wedge\frac{2\gamma}{\upsilon+2\gamma}. \tag{5.57}$$

In Table 5.1 we summarize the parameters of Assumption A and their values for the stochastic Navier–Stokes equation which will be used in the results of this section.

Now, for finite dimensional deterministic forcing f we can show that the stochastic Navier–Stokes equations exhibit linear response, namely weak differentiability of the invariant measure with respect to the parameter $a > 0$. We will prove so by means of Theorem 5.1.8 presented in the previous section.

Condition	Reference	Parameter	Value
A1	(4.71)	κ_1	k_B^2/ν
A2	(4.73)	κ_2	$\nu - \gamma\lambda^{-1}\operatorname{Tr} Q$
		κ_3	$\operatorname{Tr} Q + \|f\|_{-1}^2/\nu$
A4	(4.75)	γ_1	$\nu\lambda$
		K	$\operatorname{Tr} Q + \|f\|_{-1}^2/\nu$

Table 5.1: Summary of the parameters in Assumption A for the stochastic Navier–Stokes equations where k_B is given by the estimates of the trilinear form in Lemma 1.2.1, ν the viscosity parameter, γ is an arbitrary positive constant chosen in such a way that κ_2 is positive, λ is the smallest eigenvalue of the Stokes operator, Q the covariance operator of the noise and f is the deterministic external forcing.

Theorem 5.3.1. *Consider the Navier–Stokes equation* (5.55) *with deterministic forcing* $f \in \operatorname{range}(Q_n)$ *for some* $n \in \mathbb{N}$. *Then for all* $\alpha \in (0, \alpha_0)$ *the map* $a \mapsto \langle \varphi, \mu_a \rangle$ *is locally differentiable for every* $\varphi \in \mathcal{C}_{\tilde{d}}$.

Proof. In order to apply Theorem 5.1.8 we have to ensure that Assumption **A1**, **A2**, **A3** for a_0 and Assumption **A4** uniformly in $B_\varepsilon(a_0)$. As already stated in the previous chapters, it is known that this model has a unique invariant measure, see e.g. [37, Section 3.1], and as seen in section 4.2, it satisfies Assumption A with Lyapunov function $V(x) = |x|^2$ and parameters as in Table 5.1. Tracking the dependence on a in the proof of Theorem 4.2.1, we see in (4.75) that **A4** holds with

$$\gamma_1 = \nu\lambda_1, \quad \text{and} \quad K(a) = \operatorname{Tr} Q + \frac{\|af\|_{-1}^2}{\nu},$$

so that, setting $K := \sup_{|a-a_0|<\varepsilon} K(a) < \infty$, we have that

$$\mathcal{P}_t^a V(x) = \mathbb{E}|u(t,x,a)|^2 \le e^{-\gamma_1 t}|x|^2 + \frac{K}{\gamma_1}$$

as desired. □

When the forcing f is not necessarily finite dimensional we can show α-Hölder continuity by ensuring that Assumption H holds. In particular proving **H3** will need new arguments to be developed. First of all we show the following crucial result:

Theorem 5.3.2. *Consider the Navier–Stokes equation* (5.55) *with* $f \in H^{-1}$ *and corresponding unique invariant measure* μ_a*. Suppose Assumption* ***H1*** *holds, then for any choice of the parameters* $a_0 > 0$ *and* $\varepsilon > 0$*, there exists* $\eta_1 > 0$ *such that for all* $\eta \in (0, \eta_1)$

$$\sup_{|a-a_0|<\varepsilon} \int \exp\big(\eta |x|^2\big)\, \mu_a(dx) < \infty.$$

Theorem 5.3.2 appears necessary also in [42, Section 5.5] to study the dependence on the parameters for stochastic Navier–Stokes and there the following lemma was given as a justification for this result.

Lemma 5.3.3 ([42, Lemma 5.1])**.** *Let* M *be a real-valued semimartingale*

$$dM(t,\omega) = F(t,\omega)\, dt + G(t,\omega)\, dW$$

where W *is a standard Brownian motion. Assume there exists a process* Z *and positive constants* b_1, b_2, b_3 *with* $b_2 > b_3$*, such that*

(i) $F \leq b_1 - b_2 Z$ *a.s.,*

(ii) $M \leq Z$ *a.s.,*

(iii) $G^2 \leq b_3 Z$ *a.s.*

Then the bound

$$\mathbb{E} \exp\left(M(t) + \frac{b_2 e^{-b_2 t/4}}{4} \int_0^t Z(s)\, ds \right) \leq \frac{b_2 \exp\left(\frac{2b_1}{b_2}\right)}{b_2 - b_3} \exp\big(M(0) e^{-b_2 t/2}\big)$$

holds for any $t \geq 0$*.*

We will show explicitly how such a statement leads to that of Theorem 5.3.2.

Proof of Theorem 5.3.2. Let $\eta > 0$ and take the $\mathcal{H}$ product of (5.55) with ηu itself to get

$$d(\eta |u(t)|^2) = \eta \left(2a\langle f, u\rangle + q - 2\nu \|u\|^2\right)\, dt + 2\eta \langle u, \cdot \rangle dW(t)$$

where $q = \operatorname{Tr} Q$. We apply Lemma 5.3.3 setting $M(t) = \eta |u_t|^2$ and $Z(t) = \eta \lambda^{-1} \|u_t\|^2$ with λ the smallest eigenvalue of the Stokes operator. Indeed we have:

(i) the first condition of the lemma is satisfied for $b_1 = \eta\left(\frac{a^2}{\nu}\|f\|_{-1}^2 + q\right)$ and $b_2 = \nu\lambda$ as

$$\begin{aligned} F(t) &= \eta\left(2a\langle f, u\rangle + q - 2\nu\|u\|^2\right) \\ &\leq \eta\left(\frac{a^2}{\nu}\|f\|_{-1}^2 + \nu\|u\|^2 + q - 2\nu\|u\|^2\right) = b_1 - b_2 Z(t). \end{aligned}$$

(ii) Simply by Poincaré inequality $M(t) = \eta|u|^2 \leq \eta\lambda^{-1}\|u\|^2 = Z(t)$.

(iii) The third condition is fulfilled with $b_3 = 4\eta q$ as by (2.49) and Poincaré inequality we have

$$4\eta^2\|\langle u, \cdot\rangle\|_{L_2^0}^2 \leq 4\eta^2 q|u_t|^2 \leq 4\eta^2 q\lambda^{-1}\|u_t\|^2.$$

To ensure that $b_2 > b_3$ i.e. $4\eta q < \nu\lambda$ we take

$$\eta < \nu\lambda_1/4q. \tag{5.58}$$

Then Lemma 5.3.3 gives

$$\mathbb{E}\exp\left(\eta|u_t|^2 + \frac{\nu\eta e^{-\nu\lambda t/4}}{4}\int_0^t \|u_s\|^2\,ds\right) \leq C(a)\exp\left(\eta|u_0|^2 e^{-\nu\lambda t/2}\right)$$

with

$$C(a) = \frac{\nu\lambda\exp\left(\frac{2\eta(q+\|af\|_{-1}^2/\nu)}{\nu\lambda}\right)}{\nu\lambda - 4\eta q} \tag{5.59}$$

which stays uniformly bounded for all a in a ε-neighbourhood of a_0. Consequently we have

$$\mathbb{E}\exp\left(\eta|u_t|^2\right) \leq C(a)\exp\left(\eta|u_0|^2 e^{-\nu\lambda t/2}\right). \tag{5.60}$$

Next, we observe that as μ_a is the invariant measure for $\mathcal{P}_t^a$

$$\int \exp\left(\eta|x|^2\right)\mu_a(dx) = \int \mathbb{E}\,\exp\left(\eta|u_t(a,x)|^2\right)\mu_a(dx) \quad \text{for all } t > 0,$$

so that, by the bound (5.60) just obtained,

$$\int \exp\left(\eta|x|^2\right)\mu_a(dx) \leq C(a)\int \exp\left(\eta|x|^2 e^{-\nu\lambda t/2}\right)\mu_a(dx) \tag{5.61}$$

for all $0 < \eta < \nu\lambda_1/4q$ and any $t > 0$. In particular then

$$\int \exp\left(\eta|x|^2\right)\mu_a(dx) \leq C(a)\underbrace{\lim_{t\to\infty}\int \exp\left(\eta|x|^2 e^{-\nu\lambda t/2}\right)\mu_a(dx)}_{\text{(I)}}. \tag{5.62}$$

We have the desired result if (I) is finite uniformly in a. In order to prove it we use the following approximation argument.

Define the function

$$\varphi_t(x) := \exp\bigl(\eta|x|^2 e^{-\nu\lambda t/2}\bigr)$$

and introduce an increasing sequence of cut-off functions $\chi_n \in [0,1]$, i.e. smooth functions supported on $[-n,n]$ with $\chi_n = 1$ over $[-n+1, n-1]$ and $\chi_n \to 1$ for $n \to \infty$. Then define the series of functions

$$\varphi_{n,t}(x) := \chi_n(|x|^2)\varphi_t(x), \quad n \in \mathbb{N}.$$

so that $\lim_{n\to\infty} \varphi_{n,t} = \varphi_t$. By the monotone convergence theorem we have

$$\lim_{n\to\infty} \langle \varphi_{n,t}, \mu_a\rangle = \langle \lim_{n\to\infty} \varphi_{n,t}, \mu_a\rangle = \langle \varphi_t, \mu_a\rangle \tag{5.63}$$

therefore we want to show that

$$\lim_{t\to\infty} \lim_{n\to\infty} \langle \varphi_{n,t}, \mu_a\rangle < \infty \tag{5.64}$$

uniformly in a.

With a simple manipulation we see that for any $u_0 \in \mathcal{H}$

$$\langle \varphi_{n,t}, \mu_a\rangle = \langle \varphi_{n,t}, \mu_a\rangle - \langle \varphi_{n,t}, P_n^a(u_0,\cdot)\rangle + \langle \varphi_{n,t}, P_n^a(u_0,\cdot)\rangle.$$

Now, we suppose that $\varphi_{t,n}$ are such that $\|\varphi_{n,t}\|_{\tilde{d}} < \infty$ and in particular that for some $C_1 > 0$

$$\|\varphi_{n,t}\|_{\tilde{d}} \le C_1 n^{1/2} \exp\bigl(\eta n e^{-\nu\lambda t/2}\bigr) \tag{5.65}$$

and we will indeed verify such a statement at the end of the proof. Then from relation (4.6), since $\|\varphi_{n,t}\|_{\tilde{d}} < \infty$ we have the following bound

$$\langle \varphi_{n,t}, \mu_a\rangle \le \|\varphi_{n,t}\|_{\tilde{d}} W_{\tilde{d}}(\mu_a, P_n^a(u_0,\cdot)) + \langle \varphi_{n,t}, P_n^a(u_0,\cdot)\rangle. \tag{5.66}$$

By Corollary 4.1.7 i.e. the exponential convergence of transition probabilities to the invariant measure

$$\le C\|\varphi_{n,t}\|_{\tilde{d}} \exp(-\gamma_1 n)(1 + V(u_0)) + \langle \varphi_{n,t}, P_n^a(u_0,\cdot)\rangle$$

and by (5.65), relabelling appropriately the constant C

$$\langle \varphi_{n,t}, \mu_a\rangle \le C n^{1/2} \exp(\eta n - \gamma_1 n)(1 + V(u_0)) + \langle \varphi_{n,t}, P_n^a(u_0,\cdot)\rangle. \tag{5.67}$$

Then for

$$\eta < \gamma_1 = \nu\lambda_1 \tag{5.68}$$

the first term on the right hand side of (5.67) converges to zero when sending n to infinity.

Next we focus on showing the second term on the right hand side of (5.67) is uniformly bounded in a. From Assumption **A2** we derive the following estimate

$$|u_n|^2 \le |u_0|^2 + \kappa_3(a)n + \Xi^a_\gamma$$

where we know κ_3 is bounded uniformly for all a in a ε-neighbourhood of a_0. Consequently, using the definition of $\varphi_{n,t}$ we get

$$\begin{aligned}\langle \varphi_{n,t}, P^a_n(u_0,\cdot)\rangle &= \mathbb{E}\,\varphi_{n,t}(u_n(a,u_0)) \\ &= \mathbb{E}\,\exp\big(\eta|u_n(a,u_0)|^2 e^{-\nu\lambda t/2}\big) \\ &\le \exp\big(\eta e^{-\nu\lambda t/2}(|u_0|^2+\kappa_3(a)n)\big)\mathbb{E}\exp\big(\eta e^{-\nu\lambda t/2}\Xi^a_\gamma\big).\end{aligned}$$

Recall from Lemma 2.4.1, that if $k < 2\gamma$, then $\mathbb{E}\exp\big(k\Xi^a_\gamma\big) = k/(2\gamma - k) : C_\Xi$ is finite and in particular it is independent of a. Therefore for t large enough (i.e. $t > \frac{2}{\nu}\ln\frac{\eta}{2\gamma}$) and any arbitrary $u_0 \in \mathcal{H}$ from (5.67) we have

$$\begin{aligned}\langle \varphi_{n,t}, \mu_a\rangle \le Cn^{1/2}\exp((\eta-\gamma_1)n)(1+V(u_0)) \\ + C_\Xi \exp\big(\eta e^{-\nu\lambda t/2}(|u_0|^2+\kappa_3(a)n)\big).\end{aligned}$$

for any $n \in \mathbb{N}$, and t large enough.

We have then shown that for all $\eta < \gamma_1$

$$\lim_{t\to\infty}\lim_{n\to\infty}\langle \varphi_{n,t}, \mu_a\rangle \le C_\Xi. \tag{5.69}$$

Recall that to apply Lemma 5.3.3 we required in (5.58) that $\eta < \nu\lambda_1/4q$, then for all $\eta > 0$ such that

$$\eta < \frac{\nu\lambda_1}{4q} \wedge \gamma_1 =: \eta_1, \tag{5.70}$$

given (5.62), (5.63) and (5.69) and the fact that

$$\sup_{|a-a_0|<\varepsilon} \kappa_3(a) < \infty \quad \text{and} \quad \sup_{|a-a_0|<\varepsilon} C(a) < \infty$$

we conclude that

$$\sup_{|a-a_0|<\varepsilon}\int \exp\big(\eta|x|^2\big)\,\mu_a(dx) \le \sup_{|a-a_0|<\varepsilon} C(a)\lim_{t\to\infty}\lim_{n\to\infty}\langle \varphi_{n,t}, \mu_a\rangle < \infty.$$

We conclude the proof by showing that the estimate (5.65) for $\|\varphi_{n,t}\|_{\tilde{d}}$ holds. By the mean value theorem, given $z \in [x,y]$,

$$|\varphi_{n,t}(x) - \varphi_{n,t}(y)| \le \|D_x\varphi_{n,t}(z)\||x-y|$$

and so if $\|D_x\varphi_{n,t}(z)\|$ is bounded uniformly in z, we have the following bound

$$\|\varphi_{n,t}\|_{\tilde{d}} \le \left(\sup_{z\in H}\|D_x\varphi_{n,t}(z)\|\right)\left(\sup_{x\neq y}\frac{|x-y|}{\tilde{d}(x,y)}\right). \tag{5.71}$$

Focusing on the derivative of the functions $\varphi_{n,t}$ with respect to x it is easy to see that

$$\|D_x\varphi_{n,t}(z)\| \le 2|z|\varphi_t(z)\left|\chi_n'(|z|^2)+\chi_n(|z|^2)\eta e^{-\nu\lambda t/2}\varphi_t(z)\right|.$$

The smooth cut-off function $\chi_n(|z|^2)$ has by definition support $[-n,n]$ and it is bounded above by 1, whereas its derivative $\chi_n'(|z|^2)$ is well defined for any choice of n and has support $[-n,-n+1]\cup[n-1,n]$. Therefore

$$\begin{aligned}\sup_{z\in\mathcal{H}}\|D_x\varphi_{n,t}(z)\| &\le \sup_{|z|^2\le n} 2|z|\exp\big(\eta|z|^2e^{-\nu\lambda t/2}\big)|\chi_n'(|z|^2)+\eta e^{-\nu\lambda t/2}\chi_n(|z|^2)| \\ &\le 2n^{1/2}\exp\big(\eta n e^{-\nu\lambda t/2}\big)\left(\sup_{z\in\mathcal{H}}|\chi_n'(|z|^2)|+\eta\right).\end{aligned}$$

Note that, by the construction of χ_n, the length of the interval where the function is not constantly one or zero does not change with n, in fact it is always of length one, and all χ_n for $n\in\mathbb{N}$ behave identically in those intervals. Therefore

$$\sup_{t\in[n-1,n]}|\chi_n'(t)| = \sup_{t\in[0,1]}|\chi_1'(t)| \quad \text{for all } n\in\mathbb{N}$$

and

$$\sup_{z\in\mathcal{H}}|\chi_n'(|z|^2)| = \sup_{t\in[0,1]}|\chi_1'(t)|.$$

Therefore we showed that there exists a positive constant

$$C_1 := 2\left(\sup_{t\in[0,1]}|\chi_1'(t)|+\eta\right)$$

such that the derivative of $\varphi_{n,t}$ satisfies

$$\sup_{z\in\mathcal{H}}\|D_x\varphi_{n,t}(z)\| \le C_1 n^{1/2}\exp(\eta n).$$

Finally from (5.71) we see

$$\|\varphi_{n,t}\|_{\tilde{d}} \le C_1 n^{1/2}\exp\big(\eta n e^{-\nu\lambda t/2}\big)\sup_{x\neq y}\frac{|x-y|}{\tilde{d}(x,y)}. \tag{5.72}$$

By the definition of the distance-like function $\tilde{d}$ in (5.16) we have

$$\sup_{x\neq y} \frac{|x-y|}{\tilde{d}(x,y)} < \infty$$

so that, relabelling C_1 appropriately, the desired result holds. □

Theorem 5.3.4. *Consider the Navier–Stokes equation* (5.55) *with* $f \in H^{-1}$. *Then for all* $\alpha \in (0, \alpha_0)$, *with* α_0 *as in* (5.57), *the map* $a \mapsto \langle \varphi, \mu_a \rangle$ *is locally* α*-Hölder continuous for every* $\varphi \in \mathcal{C}_{\tilde{d}}$.

Proof. Given Theorem 5.2.1 we have to ensure that the solution of (5.55) satisfies Assumption H.

Tracking the dependence on a in Theorem 4.2.1 we see in (4.71) and (4.73) that the constants κ_0, κ_1 and κ_2 are independent of the parameter a and that

$$\kappa_3(a) = K(a) = \operatorname{Tr} Q + \frac{\|af\|_{-1}^2}{\nu}$$

so that Assumption **H1** holds.

Assumption **H2** is verified thanks to the computations carried out in section 2.5. In fact we derived the energy estimate (2.90) for $w_t := u_t(a) - u_t(a_0)$, i.e.

$$\tfrac{d}{dt}|w|^2 + \nu\|w\|^2 \leq \tfrac{2k_B^2}{\nu}\|u(a_0)\|^2|w|^2 + \tfrac{2\|f\|_{-1}^2}{\nu}|a-a_0|^2$$

and by Gronwall's lemma we have

$$\begin{aligned} |w(t)|^2 &\leq \tfrac{2\|f\|_{-1}^2}{\nu}|a-a_0|^2 \int_0^t \exp\Big(\tfrac{2k_B^2}{\nu}\int_s^t \|u(r,a_0)\|^2\,dr\Big)\,ds \\ &\leq \tfrac{2\|f\|_{-1}^2}{\nu}|a-a_0|^2 t \exp\Big(\tfrac{2k_B^2}{\nu}\int_0^t \|u(r,a_0)\|^2\,dr\Big)\,ds \end{aligned}$$

so that **H2** is satisfied with $C = 2\|f\|_{-1}^2/\nu$, $\kappa_1 = 2k_B^2/\nu$.

Recall that for this model the Lyapunov function is $V(x) = |x|^2$. Then we are left to show **H3** i.e.

$$\sup_{|a-a_0|<\varepsilon} \langle \tilde{V}, \mu_a \rangle < \infty$$

where

$$\tilde{V}(x) := e^{\alpha v|x|^2/2}\left(1 + 2K_\varepsilon + 2e^{-\gamma t}|x|^2\right)^{1/2}.$$

Introduce an auxiliary function $U : \mathcal{H} \to (0,\infty)$

$$U(x) = \exp\!\big(\eta|x|^2\big).$$

and we want to show that there exist positive constants η_0, η_1 such that for all $\eta \in [\eta_0, \eta_1]$

$$\sup_{|a-a_0|<\varepsilon} \langle \tilde{V}, \mu_a \rangle \leq \underbrace{\sup_{x\in\mathcal{H}} \frac{e^{\alpha\upsilon|x|^2/2}(1+\frac{K_\varepsilon}{\gamma}+2|x|^2)^{1/2}}{U(x)}}_{(I)} \underbrace{\sup_{|a-a_0|<\varepsilon} \langle U, \mu_a \rangle}_{(II)} < \infty\,.$$

In Theorem 5.3.2 we showed that (II) is finite for all $\eta < \eta_1$ with η_1 as in (5.70). On the other hand part (I) is finite as long as η is strictly larger than $\alpha\upsilon/2$, or since $\alpha \in (0, \alpha_0)$

$$\eta \geq \frac{\alpha_0\upsilon}{2}.$$

Then we only have to make sure that

$$\frac{\alpha_0\upsilon}{2} < \eta_1 = \frac{\nu\lambda_1}{4q} \wedge \gamma_1$$

First note that given the definition of α_0 we have

$$\frac{\alpha_0\upsilon}{2} = \frac{\upsilon}{4} \wedge \frac{\gamma\upsilon}{2\gamma+\upsilon} < \gamma$$

where $\gamma > 0$ is an arbitrary parameter such that $\gamma < \nu\lambda_1/q$. Therefore if we consider

$$0 < \gamma < \eta_1 = \frac{\nu\lambda_1}{4q} \wedge \nu\lambda_1,$$

we have a non empty interval of possible values for the parameter η

$$\frac{\alpha_0\upsilon}{2} < \gamma \leq \eta < \eta_1$$

for which (I) is finite and Theorem 5.3.2 holds.

□

5.4 Stochastic two–layer quasi–geostrophic model

Consider again the system (2.43) studied in section 2.4 with the intensity of the deterministic forcing f modulated by a multiplicative constant $a > 0$,

$$d\mathbf{q} + (B(\boldsymbol{\psi},\boldsymbol{\psi}) + \beta D_x\boldsymbol{\psi})\,dt = \nu\Delta^2\boldsymbol{\psi}\,dt + \begin{pmatrix} af \\ -r\Delta\psi_2 \end{pmatrix} dt + d\mathbf{W} \qquad (5.73)$$

Condition	Reference	Parameter	Value
A1	(4.79)	κ_1	$2k_B$
A2	(4.81)	κ_2	$\nu - 2\gamma\lambda^{-2}\operatorname{Tr} Q$
		κ_3	$T_Q + h_1\|f\|_{-2}^2/2\nu$
A4	(4.83)	γ_1	$\nu\lambda^2/(\lambda_1 + F_1)$
		K	$T_Q + h_1\|f\|_{-2}^2/2\nu$

Table 5.2: Summary of the parameters in Assumption A for the stochastic two–layer quasi–geostrophic equations where k_B is given by the estimates of the trilinear form in Lemma 1.3.4, ν the viscosity parameter, γ is an arbitrary positive constant chosen in such a way that κ_2 is positive, λ is the smallest eigenvalue of the Stokes operator, Q the covariance operator of the noise, T_Q is given by (2.53), f is the deterministic external forcing, h_1 is the height of the top layer and F_1 is a positive parameter defined by (1.23).

and $\mathcal{P}_t^a\varphi(\mathbf{q}_0) = \mathbb{E}\,\varphi(\mathbf{q}(t,\cdot;\mathbf{q}_0,a))$ with unique invariant measure μ_a.

In section 4.3 we showed that the stochastic two–layer quasi–geostrophic model satisfies Assumption A with parameters as in Table 5.2 and consequently the exponential convergence of transition probabilities was established with respect to the $\tilde{d}$-Wasserstein distance with $\tilde{d}$ (5.16) defined with parameters

$$\upsilon = \frac{k_B}{\nu - 2\gamma\lambda_1^{-2}\operatorname{Tr} Q} > 0 \quad \text{and} \quad \alpha_0 = \frac{1}{2} \wedge \frac{2\gamma}{2\gamma + \upsilon}. \tag{5.74}$$

As established in Theorem 5.1.8 and Theorem 5.2.1, this is a crucial ingredient to ensure respectively weak differentiability of the invariant measure with respect to the parameter a when f is finite dimensional, and weak α-Hölder continuity for any $f \in H^{-2}$.

We will start with the first of these two results, the weak differentiability or linear response for the two–layer quasi-geostrophic model. This was our main aim for developing the general methodology presented above and one of the main achievements of this work, despite its concise proof.

Theorem 5.4.1. *Consider the two–layer quasi–geostrophic equation* (5.73) *with* $f \in \operatorname{range}(Q_n)$ *for some* $n \in \mathbb{N}$ *and invariant measure* μ_a*. Then for all* $\alpha \in (0, \alpha_0)$ *the map* $a \mapsto \langle\varphi, \mu_a\rangle$ *is locally differentiable for every* $\varphi \in \mathcal{C}_{\tilde{d}}$*.*

Proof. To apply Theorem 5.1.8 we have to ensure that **A1**, **A2**, **A3** holds for a_0 and **A4** holds uniformly in $B_\varepsilon(a_0)$. We know Theorem 4.3.1 showed

Assumption A to hold with Lyapunov function $V(\mathbf{x}) = |||\mathbf{x}|||_{-1}^2$ and parameters as in Table 5.2. Then, tracking the dependence on the parameter a in the calculations leading to (4.83), we have in fact that **A4** holds with

$$\gamma_1 = \frac{\nu\lambda_1^2}{\lambda_1 + F_1} \quad \text{and} \quad K(a) = \frac{h_1}{2\nu}\|af\|_{-2}^2 + T_Q.$$

Therefore setting $K := \sup_{|a-a_0|<\varepsilon} K(a)$ we have

$$\mathcal{P}_t^a V(\mathbf{q}_0) = \mathbb{E}\,|||\mathbf{q}(t,\mathbf{q}_0;a)|||_{-1}^2 \leq e^{-\gamma_1 t}|||\mathbf{q}_0|||_{-1}^2 + \frac{K}{\gamma_1}$$

for all $a \in B_\varepsilon(a_0)$ as desired. □

When the forcing f is not necessarily finite dimensional we can show weak α-Hölder continuity by ensuring that Assumption H holds. Again we start with the following result on the invariant measure of the model:

Theorem 5.4.2. *Consider the model* (5.73) *with* $f \in H^{-2}$ *and corresponding unique invariant measure* μ_a. *Suppose Assumption* ***H1*** *holds, then for any choice of the parameters* $a_0 > 0$ *and* $\varepsilon > 0$, *there exists* $\eta_1 > 0$ *such that for all* $\eta \in (0, \eta_1)$

$$\sup_{|a-a_0|<\varepsilon} \int \exp\big(\eta|||\mathbf{x}|||_{-1}^2\big)\,\mu_a(d\mathbf{x}) < \infty.$$

Proof. By definition of invariant measure we have for all $t \geq 0$

$$\int \exp\big(\eta|||\mathbf{x}|||_{-1}^2\big)\,\mu_a(d\mathbf{x}) = \int \mathbb{E}\exp\big(\eta|||\mathbf{q}_t(a,\mathbf{x})|||_{-1}^2\big)\,\mu_a(d\mathbf{x}), \tag{5.75}$$

and thanks to Lemma 5.3.3 we will derive an upper bound for the right hand side. In fact taking the $\mathbb{L}^2$ product of (5.73) with $\eta\boldsymbol{\psi}$, as seen in (2.55), we have

$$d(\eta|||\mathbf{q}|||^2) = -2\eta\left(ah_1\,(f,\psi_1) + rh_2\|\psi_2\|^2 + \nu|\Delta\boldsymbol{\psi}|^2\,dt\right)\,dt \\ + \eta T_Q\,dt - 2\eta h_1(\psi_1, dW(t)).$$

We want to ensure that conditions (i)-(iii) in Lemma 5.3.3 are satisfied: first set $M(t) = \eta|||\mathbf{q}(t)|||_{-1}^2$ and by (1.41) i.e. $|||\mathbf{q}(t)|||_{-1}^2 \leq \lambda_1^{-1}|||\mathbf{q}(t)|||_0^2$, we have precisely $M(t) \leq Z(t) := \eta\lambda_1^{-1}|||\mathbf{q}(t)|||_0^2$ so that condition (ii) is fulfilled. Moving on to condition (i) setting the function $F(t)$ to be

$$F(t) := -2\eta\left(ah_1\,(f,\psi_1) + rh_2\|\psi_2\|^2 + \nu|\Delta\boldsymbol{\psi}|^2\right) + \eta T_Q$$

by the usual application of Cauchy-Schwartz, Young and Poincaré inequalities we get

$$
\begin{aligned}
F(t) &\leq \frac{\eta h_1}{\nu}\|af\|_{-2}^2 + \eta\nu h_1|\Delta\psi_1|^2 - 2\eta\nu|\Delta\boldsymbol{\psi}|^2 + \eta T_Q \\
&\leq \eta\left(\frac{h_1}{\nu}\|af\|_{-2}^2 + T_Q\right) - \eta\nu|\Delta\boldsymbol{\psi}|^2.
\end{aligned}
$$

Then estimating $|\Delta\boldsymbol{\psi}|^2$ by (1.45), i.e. $|||\mathbf{q}|||_0^2 \leq c_0|\Delta\boldsymbol{\psi}|^2$,

$$
F(t) \leq b_1 - b_2\eta\lambda_1^{-1}|||\mathbf{q}(t)|||_0^2
$$

with

$$
b_1 = \eta\left(\frac{h_1}{\nu}\|af\|_{-2}^2 + T_Q\right) \quad \text{and} \quad b_2 = \frac{\nu\lambda_1}{c_0}.
$$

Finally for condition (iii) we have to ensure that there exists $b_3 \in (0, b_2)$ such that $G^2 \leq b_3 Z$ almost surely namely

$$
4\eta^2 h_1^2\|(\psi_1,\,\cdot)\|_{L_2^0}^2 \leq b_3\eta\lambda_1^{-1}|||\mathbf{q}(t)|||_0^2. \tag{5.76}
$$

Given the estimate (2.49) for $\|(\psi_1,\,\cdot)\|_{L_2^0}^2$ and Poincaré inequality we have

$$
\|(\psi_1,\cdot)\|_{L_2^0}^2 \leq |\psi_1|^2\operatorname{Tr}Q \leq \frac{\operatorname{Tr}Q}{\lambda_1^2 h_1}|\Delta\boldsymbol{\psi}|^2
$$

and consequently (5.76) holds setting $b_3 = 4\eta h_1\lambda_1^{-1}\operatorname{Tr}Q$. As we require $b_2 > b_3$ i.e.

$$
\frac{\nu\lambda_1}{c_0} > 4\eta h_1\lambda_1^{-1}\operatorname{Tr}Q
$$

we get that for all

$$
0 < \eta < \frac{\nu\lambda_1^2}{4c_0 h_1\operatorname{Tr}Q} =: \eta_0, \tag{5.77}
$$

the hypothesis of Lemma 5.3.3 hold, giving

$$
\mathbb{E}\exp\left(\eta|||\mathbf{q}(t)|||_{-1}^2 + \frac{b_2 e^{-b_2 t/4}}{4}\int_0^t \eta\lambda_1^{-1}|||\mathbf{q}(s)|||_0^2\,ds\right) \leq \frac{b_2\exp\left(\frac{2b_1}{b_2}\right)}{b_2 - b_3}\exp\bigl(\eta|||\mathbf{q}(0)|||_{-1}^2 e^{-b_2 t/2}\bigr).
$$

Using this estimate back in (5.75) we have

$$
\int \exp\bigl(\eta|||\mathbf{x}|||_{-1}^2\bigr)\,\mu_a(d\mathbf{x}) \leq C(a)\int \exp\bigl(\eta|||\mathbf{x}|||_{-1}^2 e^{-b_2 t/2}\bigr)\,\mu_a(d\mathbf{x})
$$

for all $\eta < \eta_0$ and all $t > 0$ with

$$C(a) = \frac{\frac{\nu\lambda_1}{c_0}\exp\Big(\frac{2c_0\eta}{\nu\lambda_1}\left(\frac{h_1}{\nu\lambda_1}\|af\|_{-1}^2 + T_Q\right)\Big)}{\frac{\nu\lambda_1}{c_0} - 4\eta h_1\lambda_1^{-1}\operatorname{Tr} Q}.$$

In particular setting

$$C_\varepsilon = \sup_{|a-a_0|<\varepsilon} C(a) = C(a_0+\varepsilon)$$

we have for all $\eta < \eta_0$

$$\sup_{|a-a_0|<\varepsilon}\int \exp\big(\eta\|\!|\mathbf{x}|\!\|_{-1}^2\big)\,\mu_a(d\mathbf{x}) \le C_\varepsilon \underbrace{\sup_{|a-a_0|<\varepsilon}\lim_{t\to\infty}\int\exp\big(\eta\|\!|\mathbf{x}|\!\|_{-1}^2 e^{-b_2 t/2}\big)\,\mu_a(d\mathbf{x})}_{(A)} \tag{5.78}$$

and we are left to show that part (A) is finite.

Define the function $\varphi_t : \mathcal{H} \to \mathbb{R}_+$ as

$$\varphi_t(\mathbf{x}) = \exp\big(\eta\|\!|\mathbf{x}|\!\|_{-1}^2 e^{-b_2 t/2}\big)$$

and its approximation

$$\varphi_{n,t}(x) = \chi_n(\|\!|\mathbf{x}|\!\|_{-1}^2)\exp\big(\eta\|\!|\mathbf{x}|\!\|_{-1}^2 e^{-b_2 t/2}\big)$$

where $(\chi_n)_n$ is an increasing sequence of cut-off functions as in the proof of Theorem 5.3.4. By the monotone convergence theorem

$$\lim_{n\to\infty}\langle\varphi_{n,t},\mu_a\rangle = \langle\lim_{n\to\infty}\varphi_{n,t},\,\mu_a\rangle = \langle\varphi_t,\,\mu_a\rangle \tag{5.79}$$

and we focus on ensuring

$$\sup_{|a-a_0|<\varepsilon}\lim_{t\to\infty}\lim_{n\to\infty}\langle\varphi_{n,t},\mu_a\rangle < \infty.$$

In a similar way to showing (5.65), for Navier–Stokes, it can be proved that

$$\|\varphi_{n,t}\|_{\tilde d} \le C_1 n^{1/2}\exp(\eta n)\sup_{x\neq y}\frac{\|\!|\mathbf{x}-\mathbf{y}|\!\|_{-1}}{\tilde d(\mathbf{x},\mathbf{y})} < \infty \tag{5.80}$$

so that

$$\langle\varphi_{n,t},\mu_a\rangle \le \|\varphi_{n,t}\|_{\tilde d} W_{\tilde d}(\mu_a, P_n^a(\mathbf{q}_0,\cdot)) + \langle\varphi_{n,t}, P_n^a(\mathbf{q}_0,\cdot)\rangle.$$

Thanks to Assumption A the exponential convergence of transition probabilities holds (Theorem 4.3.1) and Corollary 4.1.7 gives in particular

$$W_{\tilde{d}}(\mu_a, P_n^a(\mathbf{q}_0,\cdot)) \le Ce^{-\gamma_1 n}(1+V(\mathbf{q}_0))$$

where V is the Lyapunov function with parameters γ_1 and K as in Table 5.2. Relabelling appropriately the constant C, by (5.80) we have

$$\langle \varphi_{n,t}, \mu_a\rangle \le Cn^{1/2}\exp(\eta n-\gamma_1 n)(1+V(\mathbf{q}_0)) + \langle \varphi_{n,t}, P_n^a(\mathbf{q}_0,\cdot)\rangle. \tag{5.81}$$

Next thanks to **A2** we know that

$$|||\mathbf{q}(n,\cdot;a,\mathbf{q}_0)|||_{-1}^2 \le |||\mathbf{q}_0|||_{-1}^2 + \kappa_3(a)n + \Xi_\gamma^a$$

so that

$$\begin{aligned}\langle \varphi_{n,t}, P_n^a(\mathbf{q}_0,\cdot)\rangle &= \mathbb{E}\varphi_{n,t}\left(\mathbf{q}(n,\cdot;a,\mathbf{q}_0)\right)\\ &\le \exp\bigl(\eta e^{-b_2 t/2}(|||\mathbf{q}_0|||_{-1}^2+\kappa_3(a)n)\bigr)\,\mathbb{E}\exp\bigl(\eta e^{-b_2 t/2}\Xi_\gamma^a\bigr).\end{aligned} \tag{5.82}$$

Picking t large enough so that

$$C_\Xi := \mathbb{E}\exp\bigl(\eta e^{-b_2 t/2}\Xi_\gamma^a\bigr)$$

is well defined, we get from (5.82)

$$\begin{aligned}\langle \varphi_{n,t}, \mu_a\rangle \le\ & Cn^{1/2}\exp(\eta n-\gamma_1 n)(1+V(\mathbf{q}_0))\\ &+ C_\Xi\exp\bigl(\eta e^{-b_2 t/2}(|||\mathbf{q}_0|||_{-1}^2+\kappa_3(a)n)\bigr)\end{aligned}$$

In particular, by Assumption **H1** the parameter κ_3 is uniformly bounded on a ε-neighbourhood of a_0 so that for all $\eta<\gamma_1$

$$\sup_{|a-a_0|<\varepsilon}\lim_{t\to\infty}\lim_{n\to\infty}\langle \varphi_{n,t},\mu_a\rangle \le C_\Xi \tag{5.83}$$

Finally, (5.78), (5.79) and (5.83) show that

$$\sup_{|a-a_0|<\varepsilon}\int \exp\bigl(\eta|||\mathbf{x}|||_{-1}^2\bigr)\,\mu_a(d\mathbf{x}) < \infty$$

for all parameters $\eta\in(0,\eta_1)$ with

$$\eta_1 := \eta_0\wedge\gamma_1 = \frac{\nu\lambda_1^2}{4c_0h_1\operatorname{Tr}Q}\wedge\gamma_1. \tag{5.84}$$

□

We are now ready to prove the last theorem of this chapter, namely the local Hölder continuity of the invariant measure of the stochastic 2LQG model.

Theorem 5.4.3. *Consider* (5.73) *with* $f \in H^{-2}$. *Then for all* $\alpha \in (0, \alpha_0)$, *with* α_0 *as in* (5.74), *the map* $a \mapsto \langle \varphi, \mu_a \rangle$ *is locally* α*-Hölder continuous for every* $\varphi \in \mathcal{C}_{\tilde{d}}$.

Proof. We want to ensure that Assumption H holds. In Theorem 4.3.1 we showed that Assumption A holds and in particular only the parameters depending on the external forcing f will be dependent on the parameter a. In Table 5.2 we see that κ_0, κ_1 and κ_2 as well as γ_1 are independent of a and

$$\kappa_3(a) = K(a) = T_Q + \frac{h_1 \|af\|_{-2}^2}{\nu}$$

which satisfy immediately the second part of **H1**.

Assumption **H2** is verified thanks to the computations carried out in section 2.4.2 which led to (2.64) namely

$$|||\mathbf{q}_t(a) - \mathbf{q}_t(a_0)|||_{-1}^2 \leq |a - a_0|^2 \frac{h_1 \|f\|_{-2}^2}{\nu} \int_0^t \exp\left(2k_B \int_s^t |\Delta \boldsymbol{\psi}_\tau(a_0)|^2 \, d\tau \right) ds.$$

Therefore **H2** is satisfied with $C = h_1 \|f\|_{-2}^2 / \nu$ and $\kappa_1 = 2k_B$.

For Assumption **H3**, using as Lyapunov function $V(\mathbf{x}) = |||\mathbf{x}|||_{-1}^2$, we have to show that

$$\sup_{|a-a_0|<\varepsilon} \int e^{\alpha \upsilon |||\mathbf{x}|||_{-1}^2} \left(1 + |||\mathbf{x}|||_{-1}^2\right)^{1/2} \mu_a(d\mathbf{x}) < \infty.$$

Given the auxiliary function $U : \mathcal{H} \to (0, \infty)$

$$U(\mathbf{x}) = \exp\big(\eta |||\mathbf{x}|||_{-1}^2\big),$$

it is sufficient to find a suitable value for $\eta > 0$ giving

$$\underbrace{\sup_{\mathbf{x} \in \mathcal{H}} \exp\left(\left(\frac{\alpha \upsilon}{2} - \eta\right) |||\mathbf{x}|||_{-1}^2\right) \left(1 + |||\mathbf{x}|||_{-1}^2\right)^{1/2}}_{(I)} \underbrace{\sup_{|a-a_0|<\infty} \langle U, \mu_a \rangle}_{(II)} < \infty. \quad (5.85)$$

In order that (I) stays finite we have to assume that $\eta > \alpha \upsilon / 2$ for all $\alpha \in (0, \alpha_0)$ i.e.

$$\eta > \alpha_0 \upsilon / 2,$$

while Theorem 5.4.2 ensures (II) stays finite for all $\eta < \eta_1$ with η_1 defined as in (5.84). We have only to show that $\alpha_0 \upsilon/2 < \eta_1$ to have an interval of suitable values for η.

By definition of α_0 (5.74) we have that

$$\frac{\alpha_0 \upsilon}{2} = \frac{\upsilon}{2}\left(\frac{1}{2} \wedge \frac{2\gamma}{2\gamma + \upsilon}\right) < \gamma.$$

Recall that γ is an arbitrary parameter introduced so that υ (5.74) is a well defined positive constant, namely

$$\gamma < \frac{\nu \lambda_1^2}{2 \operatorname{Tr} Q}.$$

From the proof of Lemma 1.3.2 we know that $c_0 = 2F_1/\lambda_1$ so

$$\eta_1 = \frac{\nu \lambda_1^2}{4 c_0 h_1 \operatorname{Tr} Q} = \frac{\nu \lambda_1^2}{2 \operatorname{Tr} Q} \frac{\lambda_1}{4 h_1 F_1}$$

and picking a $\gamma > 0$ such that

$$\gamma < \frac{\nu \lambda_1^2}{2 \operatorname{Tr} Q}\left(1 \wedge \frac{\lambda_1}{4 h_1 F_1}\right) =: \eta_2$$

equation (5.85) holds for all $\eta \in [\gamma, \eta_2)$.

□

Summary and remarks

In this final chapter we analysed the dependence of the invariant measure on parameters of the model, showing new results on linear response and fractional response for both stochastic Navier-Stokes and the stcohastic 2LQG model.

We first presented under a new light the best result available in literature for linear response in the infinite dimensional context, which holds true for a wide range of Markov semigroups and parameters. Then we focused our analysis on Markov semigroups associated to SPDEs like (5.23) where the parameter of interest is the intensity of a deterministic forcing. For this type of systems, on the one hand Theorem 5.1.8 ensured linear response, whenever the deterministic forcing is finite dimensional. On the other hand, in Theorem 5.2.1 we established Hölder continuity of the invariant measure, regardless of the dimension of the forcing.

Having developed two toolboxes to obtain these results for quite general SPDEs, we then applied them to the two model of interest. In the Conclusions we will expand further on the interpretation, significance as well as limitations of these results.

Last, we discuss the possibility of quantifying the Hölder constant. This information would allow to compute how much the average of the observables changes with respect to the intensity of the forcing, as it is clear that the larger is the Hölder constant, the less we can control the changes in these statistics. The proof of Theorem 5.2.1 provides an explicit expression for the Hölder constant namely, for all observables $\varphi \in \mathcal{C}_{\tilde{d}}$

$$\langle \varphi, \mu_a - \mu_{a_0} \rangle \leq |a - a_0|^\alpha \|(1 - \mathcal{P}_t^{a_0})^{-1}\varphi\|_{\tilde{d}}\, C(t) \sup_{|a-a_0|<\varepsilon} \langle \tilde{V}, \mu_a \rangle$$

where $C(t)$ and $\tilde{V}(x)$ are defined in (5.54). In particular then

$$\sup_{\|\varphi\|_{\tilde{d}} \neq 0} \frac{\langle \varphi, \mu_a - \mu_{a_0} \rangle}{\|\varphi\|_{\tilde{d}}} \leq |a - a_0|^\alpha \|(1 - \mathcal{P}_t^{a_0})^{-1}\| C(t) \sup_{|a-a_0|<\varepsilon} \langle \tilde{V}, \mu_a \rangle \tag{5.86}$$

where

$$\|(1 - \mathcal{P}_t^{a_0})^{-1}\| = \sup_{\|\varphi\|_{\tilde{d}} \neq 0} \frac{\|(1 - \mathcal{P}_t^{a_0})^{-1}\varphi\|_{\tilde{d}}}{\|\varphi\|_{\tilde{d}}}.$$

From the results in this thesis it is possible to derive estimates for $C(t)$ and $\sup_{|a-a_0|<\varepsilon} \langle \tilde{V}, \mu_a \rangle$:

(i) The function $C(t)$ is given by

$$C(t) = \left(N C_\Xi e^{\alpha \upsilon \kappa_3 t}\right)^{1/2}$$

where α, υ and κ_3 are as in (5.57) and Table 5.1, for Navier-Stokes, and as in (5.74) and Table 5.2, for the 2LQG model. The constant C_Ξ is defined as

$$C_\Xi = \mathbb{E} \exp\left(\alpha \upsilon \Xi_\gamma^{a_0}\right) = \frac{\alpha \upsilon}{2\gamma - \alpha \upsilon} \quad \text{by (2.52).}$$

Last, N is an arbitrary natural number, which defines the distance–like function d_N and has to be larger that N_* as introduced in Theorem 4.1.10, more precisely in (4.62). By carefully tracking all the parameters appearing in the expression for N_*, it is possible to provide an estimate for it.

(ii) We saw for both the stochastic Navier-Stokes equation and the stochastic 2LQG model that

$$\sup_{|a-a_0|<\varepsilon} \langle \tilde{V}, \mu_a \rangle \leq \sup_{x\in\mathcal{H}} \frac{\tilde{V}(x)}{U(x)} \sup_{|a-a_0|<\varepsilon} \langle U, \mu_a \rangle$$

where U is chosen in such a way that

$$\sup_{x\in\mathcal{H}} \frac{\tilde{V}(x)}{U(x)} = 1 \quad \text{and} \quad \sup_{|a-a_0|<\varepsilon} \langle U, \mu_a \rangle < \infty.$$

Moreover the proof of Theorem 5.3.2 (equivalently for 2LQG Theorem 5.4.2) can provide an upper bound for the value of $\sup_{|a-a_0|<\varepsilon}\langle U, \mu_a\rangle$.

However to obtain an estimate on $\|(1-\mathcal{P}_t^{a_0})^{-1}\|$ we need to compute the size of the spectral gap. Indeed, as seen in Proposition 5.1.2, we know that

$$\|(1-\mathcal{P}_t^{a_0})^{-1}\| \leq \sum_{k=0}^{\infty} \|\mathcal{P}_t^{a_0}\|^k = \sum_{k=0}^{\infty} \left(\sup_{\|\varphi\|_{\tilde{d}}\neq 0} \frac{\|\mathcal{P}_t^{a_0}\varphi\|_{\tilde{d}}}{\|\varphi\|_{\tilde{d}}} \right)^k$$

and the series converges because of the spectral gap i.e. $\|\mathcal{P}_t^{a_0}\varphi\|_{\tilde{d}} \leq \rho\|\varphi\|_{\tilde{d}}$ with $\rho < 1$. As discussed in chapter 4, however, it is beyond the scope of this thesis to provide a good estimate for the size of the spectral gap as from the techniques here used we would expect it to be very crude.

Conclusions

In this work we studied the two–layer quasi–geostrophic (2LQG) model with a forcing acting only on the top layer, composed of a deterministic and an additive stochastic part to represent for example the stochastic wind stress on the ocean. A number of original results regarding this model have been shown, some using classic methodology, others, like exponential stability and especially linear response, requiring proper modifications to recently developed techniques. Moreover we used the the two–dimensional stochastic Navier–Stokes (SNS) equations as test model for the novel techniques.

An explicit proof of the model's well–posedness was provided, in section 2.2 giving existence and uniqueness of weak solutions, and in section 2.3 of strong solutions. Moreover in section 2.4 we studied the regularity of the solutions with respect to small changes in the intensity of the deterministic forcing showing they are locally Lipschitz continuous (Theorem 2.4.3) and locally differentiable (Theorem 2.4.6).

Extending results available for the stochastic Navier–Stokes equations, in Theorem 3.1.9 we have ensured the existence of an invariant measure for the dynamical system associated with the stochastic 2LQG model. Furthermore uniqueness of the invariant measure was established (section 3.2.2) by means of the asymptotic coupling method, under restrictions on some parameter values of the model. In particular we required the bottom friction to be large enough to satisfy Equation 3.73. It is interesting to notice that the asymptotic coupling method used provides also a description of a potential way by which the system stabilizes. Indeed the feedback control we introduced, namely $\Delta(\psi_1 - \tilde{\psi}_1)$, contains information only from the first layer. Then the condition on the bottom friction corresponds to a scenario in which the first layer stabilizes by the influence of the stochastic forcing and the second layer stabilizes mainly thanks to the friction. A more sophisticated choice of the coupling may allow synchronisation of the layers via their interaction instead (see outlook).

Using a similar control we also obtained exponential stability, namely exponential convergence of transition probabilities to the invariant measure, provided the same condition on the bottom friction established in section 3.2.2 holds. To prove exponential stability by means of the asymptotic coupling method, we reformulated the results in [12]. Specifically Theorem 4.1.10 provides sufficient conditions for the general version of Harris theorem in [44]. In doing so we drew a direct link from the conditions for SPDEs in Assumption A to Harris' theorem and gave the explicit formulation of the distance–like function with respect to which we have convergence. In fact by Theorem 4.1.10 and Theorem 4.1.6 we know that such a distance–like function is $\tilde{d}(x, y)^2 = d_N(x, y)(1 + V(x) + V(y))$, with d_N as in (4.31) and V a Lyapunov function of the system. Note that, by definition, the distance-like function d_N is comparable to the α_0–root of the original metric on the space, with α_0 as in Theorem 4.1.10. Therefore, as already observed in [40], the asymptotic coupling approach shows the spectral gap property for the semigroup acting on Hölder continuous observables.

From the point of view of the study of geophysical fluid dynamics, the ergodicity and exponential stability results for the 2LQG model provide a way to describe the long term average behaviour of the ocean dynamics at the mid–latitudes under stochastic wind stress. In particular ergodicity is often a tacitly underlying assumption in the applications when drawing conclusions on the statistics of the system from its time series, which now is shown to hold.

Finally we studied the dependence of the invariant measure on the intensity of the deterministic forcing establishing its weak local differentiability and Hölder continuity. As the observables for which we show the spectral gap property are Hölder continuous, they are less regular than those considered in [40]. Hence to show linear response we had to develop a different framework from the one set out in [40] because in addition to the spectral gap property, we also have to ensure the differentiability of the semigroup when acting on such Hölder continuous observables. The lack of differentiability of the observables effectively requires some regularization property of the semigroup. In order to obtain this, the forcing was assumed to be finite dimensional and in the range of the noise, so for the stochastic 2LQG, it has to act on the same layer as the noise. In particular one potential application of this result accounts for changes in the intensity of the average wind forcing on the upper ocean.

To the best of this author's knowledge, the present work is the first to provide a mathematical framework for linear response theory for dissipative SPDEs with the exception of [40]. Furthermore we expand upon [40] by allowing for observables that are only Hölder continuous. Equally importantly the thesis provides a compact recipe for linear response (Theorem 5.1.8) which can be easily applied to other models with similar forcings.

Similarly, for fractional response (section 5.2), the main arguments have been formulated in a context general enough so that for equations other than Navier-Stokes and 2LQG one only has to check that Assumption H holds. It is clear that the main difference between linear and fractional response lies in the disparity between Assumption **R2** and Assumption **H2**, which comes down to how the Markov semigroup alters the regularity of the observables. In linear response this is related to $\mathcal{P}_t^a\varphi$ being differentiable in the parameter for observables φ only Hölder continuous, effectively improving their regularity. On the other hand, for fractional response it is enough for $\mathcal{P}_t^a\varphi$ to be Hölder continuous. In particular in this case we do not require the forcing to be neither finite dimensional nor in the range of the noise, contrary to the methodology for linear response described above. This would then allow to consider models with deterministic and stochastic forcings acting on different sets of degrees of freedom.

Outlook. There are two main restrictions to the present analysis which would be interesting to lift, especially from the point of view of applications. First one is the condition on the bottom friction made to show ergodicity of the stochastic 2LQG model. Dropping this passivity condition of the second layer would allow investigating mechanisms under which the stochastic wind forcing on the upper ocean stabilizes on the long run the entire dynamics. This could also give an insight on how to treat other stratified models with noise acting only on one component.

In order to remove such restrictions we need to apply a much more sophisticated coupling or control than the one chosen in section 3.2.2 and section 4.3. The desired control needs to ensure that a general solution converges asymptotically in time to a specific subset of trajectories independent of the choice of the parameters. This new control will have to use information from all layers and not only the top layer. One possible approach might be to have the control actively steer the dynamics into regions where the lower layers become passive, before eventually synchronizing the upper layer.

We expect the results presented in this work to be applicable with appropriate tweaks to models with more than two layers. The condition on the passivity of the bottom layer would have to apply to the additional layers without stochastic forcing. Although this condition might still be considered realistic in two layers with the second accounting for the lower ocean, for multiple layers a strong friction is harder to justify from a physical point of view, especially since the density difference between the layers is not so sharp. Therefore removing the passivity condition would also naturally lead to looking at multi–layer models directly, rather than only two–layer models.

The second restriction interesting to be lifted is that applied on the forcing to establish linear response, namely for it to be finite dimensional and in the range of the noise. Regarding forcings acting on infinitely many degrees of freedom, our linear response result should be easily extendable when the high modes are sufficiently weak so to allow application of the infinite dimensional version of Girsanov theorem. More challenging is to consider forcings acting also on lower layers while the noise only acts on the top layer. In this case we would have to use the asymptotic coupling method in a novel way by showing it to be differentiable in the forcing. This approach also impacts on other SPDEs where a deterministic forcing and a stochastic one act on different sets of degrees of freedom. Such a scenario is particularly interesting in applications when the noise represents uncertainty in the small scales and an external forcing acts mainly on larger scales.

Furthermore, one may want to investigate linear response with respect to other parameters of the system. For example it would be interesting to consider the densities of the layers as parameter so to account for changes in composition of the ocean. The densities are embedded in the definition of the QG potential vorticity itself, in particular in the parameters F_1, F_2 (1.23) hence the computations will surely be more challenging in this case. Studying response with respect to density variations would be even more valuable for multi–layer quasi–geostrophic models involving the temperature as a variable. In fact gradients of density and temperature are the major drivers of the thermohaline circulation, a crucial part of the ocean large–scale dynamics.

Another step forward would be to establish response theory for time–dependent perturbations. This is particularly interesting for the applications, for example related to the study of anthropogenic climate change, as external forcings like the radiative forcing associated to greenhouse gas emissions are essentially time–dependent. However it is not clear whether this work extends to time–dependent external forcings, as it is not evident even whether the results we took inspiration from, like [44] on exponential stability and [40] on linear response, can be extended to nonautonomous systems.

Appendix A

Gronwall lemmas

Throughout this work we made extensive use of the following classic versions of Gronwall lemma, for reference see for example [62, Lemma 2.7].

Lemma A.1 (Differential Gronwall Lemma)**.** *Let $f = f(t)$ be an absolutely continuous function on $[t_0, T]$, which satisfies for a.e. t the differential inequality*

$$\frac{\mathrm{d}f}{\mathrm{d}t} \leq g_1(t)f(t) + g_2(t)$$

where g_1 and g_2 are summable functions on $[t_0, T]$. Then

$$f(t) \leq f(t_0)\exp\left(\int_{t_0}^{t} g_1(s)\,ds\right) + \int_{t_0}^{t} g_2(s)\exp\left(\int_{s}^{t} g_1(r)\,dr\right) ds$$

for all $t \in [0, T]$.

A straightforward consequence is the following integral version:

Lemma A.2 (Integral Gronwall Lemma)**.** *Let $f = f(t)$ be a continuous function on $[t_0, T]$, which satisfies for a.e. t the inequality*

$$f(t) \leq g_2(t) + \int_{t_0}^{t} g_1(s)f(s)\,ds \quad \textit{for all } t \in (t_0, T)$$

where g_1 and g_2 are respectively a nonnegative and non–decreasing function on $[t_0, T]$. Then

$$f(t) \leq g_2(t)\exp\left(\int_{t_0}^{t} g_1(s)\,ds\right) \quad \textit{for all } t \in [t_0, T].$$

The last lemma is a modification of Lemma A.2 which will be particularly useful in section 1.4.2 and section 4.1.1. Since it is not quite as classic as the previous lemmas we give a proof for it.

Lemma A.3 (Comparison Theorem)**.** *Let f be a continuous function, for which*

$$f(s) - f(r) \le -\gamma \int_r^s f(\tau)\, d\tau + K(s-r) \quad \textit{for all } r < s \tag{A.1}$$

with $\gamma, K > 0$ then

$$f(t) \le f(t_0) e^{-\gamma(t-t_0)} + K/\gamma \quad \textit{for all } t \ge t_0. \tag{A.2}$$

Proof. Take the partition of $[t_0, t]$ with elements $\tau_k = t_0 + tk/n$, $k = 1, \dots, n$ for $n \in \mathbb{N}$, so that

$$e^{\gamma t} f(t) - e^{\gamma t_0} f(t_0) = \lim_{n\to\infty} \sum_{k=1}^{n} \left(e^{\gamma\tau_k} f(\tau_k) - e^{\gamma\tau_{k-1}} f(\tau_{k-1})\right)$$

$$= \lim_{n\to\infty} \sum_{k=1}^{n} e^{\gamma\tau_k} \left(f(\tau_k) - f(\tau_{k-1})\right) + \left(e^{\gamma\tau_k} - e^{\gamma\tau_{k-1}}\right) f(\tau_{k-1}).$$

Using the estimate (A.1) we have

$$\le \lim_{n\to\infty} \sum_{k=1}^{n} e^{\gamma\tau_k} \left(-\gamma \int_{\tau_{k-1}}^{\tau_k} f(s)\, ds + K(\tau_k - \tau_{k-1})\right) + \left(e^{\gamma\tau_k} - e^{\gamma\tau_{k-1}}\right) f(\tau_{k-1}). \tag{A.3}$$

By definition of the Riemann-Stieltjes integral and the regularity of the exponential function, we have

$$\lim_{n\to\infty} \sum_{k=1}^{n} \left(e^{\gamma\tau_k} - e^{\gamma\tau_{k-1}}\right) f(\tau_{k-1}) = \int_{t_0}^{t} f(s)\, d(e^{\gamma s}) = \int_{t_0}^{t} \gamma f(s) e^{\gamma s}\, ds \tag{A.4}$$

and by the definition of Riemann integral

$$K \lim_{n\to\infty} \sum_{k=1}^{n} e^{\gamma\tau_k} (\tau_k - \tau_{k-1}) = K \int_{t_0}^{t} e^{\gamma s}\, ds = \frac{K}{\gamma}\left(e^{\gamma t} - e^{\gamma t_0}\right). \tag{A.5}$$

On the other hand, set

$$F(t) = \int_{t_0}^{t} f(\tau)\, d\tau \tag{A.6}$$

and note that it has bounded variation since f is continuous. Again given the definition of Riemann-Stieltjes integral and the fundamental theorem of calculus we have

$$\lim_{n\to\infty} \sum_{k=1}^{n} e^{\gamma\tau_k} \int_{\tau_{k-1}}^{\tau_k} f(s)\, ds = \lim_{n\to\infty} \sum_{k=1}^{n} e^{\gamma\tau_k} \left(F(\tau_k) - F(\tau_{k-1})\right)$$

$$=: \int_{t_0}^{t} e^{\gamma s} dF(s) = \int_{t_0}^{t} e^{\gamma s} f(s)\, ds.$$

In summary, using this relation, (A.4) and (A.5) in (A.3) we get

$$f(t)e^{\gamma t} - f(t_0)e^{\gamma t_0} \leq -\int_{t_0}^{t} \gamma f(s)e^{\gamma s}\, ds + \frac{K}{\gamma}\left(e^{\gamma t} - e^{\gamma t_0}\right) + \gamma \int_{t_0}^{t} e^{\gamma s} f(s)\, ds$$

hence the desired bound

$$f(t) \leq f(t_0)e^{-\gamma(t-t_0)} + \frac{K}{\gamma} \quad \text{for all } \; t_0 \leq t.$$

□

Bibliography

[1] R. V. Abramov and A. J. Majda. Low-frequency climate response of quasigeostrophic wind-driven ocean circulation. *Journal of Physical Oceanography*, 42(2):243 – 260, 2012.

[2] W. Bahsoun, M. Ruziboev, and B. Saussol. Linear response for random dynamical systems. *Advances in Mathematics*, 364:107011, 2020.

[3] V. Baladi. Linear response, or else. *Proceedings of the International Congress of Mathematicians (Seoul)*, 3:525–45, 2014.

[4] A. Bensoussan and R. Temam. Equations stochastiques du type Navier-Stokes. *Journal of Functional Analysis*, 13(2):195 – 222, 1973.

[5] P. S. Berloff. Random-forcing model of the mesoscale oceanic eddies. *Journal of Fluid Mechanics*, 529:71–95, 2005.

[6] J. Berner, U. Achatz, L. Batté, L. Bengtsson, A. de la Cámara, H. M. Christensen, M. Colangeli, D. R. B. Coleman, D. Crommelin, S. I. Dolaptchiev, C. L. E. Franzke, P. Friederichs, P. Imkeller, H. Järvinen, S. Juricke, V. Kitsios, F. Lott, V. Lucarini, S. Mahajan, T. N. Palmer, C. Penland, M. Sakradzija, J.-S. von Storch, A. Weisheimer, M. Weniger, P. D. Williams, and J.-I. Yano. Stochastic parameterization: Toward a new view of weather and climate models. *Bulletin of the American Meteorological Society*, 98(3):565 – 588, 2017.

[7] C. Bernier. Existence of attractor for the quasi-geostrophic approximation of the Navier-Stokes equations and estimate of its dimension. Research Report RR-1733, INRIA, 1992.

[8] C. Bernier. Existence of attractor for the quasi-geostrophic approximation of the Navier-Stokes equations and estimate of its dimension. *Adv. Math. Sci. Appl.*, 4(2):465–489, 1994.

[9] J. R. Brannan, J. Duan, and T. Wanner. Dissipative quasi-geostrophic dynamics under random forcing. *Journal of Mathematical Analysis and Applications*, 228(1):221 – 233, 1998.

[10] J. Bricmont, A. Kupiainen, and R. Lefevere. Ergodicity of the 2D Navier–Stokes equations with random forcing. *Communications in Mathematical Physics*, 224(1):65–81, Nov 2001.

[11] O. Butkovsky. Subgeometric rates of convergence of Markov processes in the Wasserstein metric. *Ann. Appl. Probab.*, 24(2):526–552, 04 2014.

[12] O. Butkovsky, A. Kulik, and M. Scheutzow. Generalized couplings and ergodic rates for SPDEs and other Markov models. *Ann. Appl. Probab.*, 30(1):1–39, 02 2020.

[13] J. G. Charney. On the scale of atmospheric motions. *Geofysiske Publikasjoner*, 17(2):1–17, 1948.

[14] K. C. Chhak, A. M. Moore, R. F. Milliff, G. Branstator, W. R. Holland, and M. Fisher. Stochastic forcing of the north atlantic wind-driven ocean circulation. part i: A diagnostic analysis of the ocean response to stochastic forcing. *Journal of Physical Oceanography*, 36(3):300 – 315, 2006.

[15] I. Chueshov, J. Duan, and B. Schmalfuss. Probabilistic dynamics of two-layer geophysical flows. *Stochastics and Dynamics*, 01(04):451–475, 2001.

[16] C. Cotter, D. Crisan, D. Holm, W. Pan, and I. Shevchenko. Modelling uncertainty using stochastic transport noise in a 2-layer quasi-geostrophic model. *Foundations of Data Science*, 2(2):173, 2020.

[17] H. Crauel and F. Flandoli. Attractors for random dynamical systems. *Probability Theory and Related Fields*, 100(3):365–393, 1994.

[18] A. B. Cruzeiro. Solutions et mesures invariantes pour des équations d'évolution stochastiques du type Navier-Stokes. *Exposition. Math.*, 7(1):73–82, 1989.

[19] G. Da Prato. *Nonlinear Stochastic Partial Differential Equations*, pages 1–21. Springer New York, New York, NY, 2009.

[20] G. Da Prato, D. Gatarek, and J. Zabczyk. Invariant measures for semilinear stochastic equations. *Stochastic Analysis and Applications*, 10(4):387–408, 01 1992.

[21] G. Da Prato and J. Zabczyk. *Ergodicity for infinite dimensional systems*, volume 229 0521579007. Cambridge University Press, 1996.

[22] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*. Cambridge university press, 2014.

[23] T. Delsole. Stochastic models of quasigeostrophic turbulence. *Surveys in Geophysics*, 25(2):107–149, 2004.

[24] D. Dolgopyat. On differentiability of SRB states for partially hyperbolic systems. *Inventiones mathematicae*, 155(2):389–449, 2004.

[25] J. Duan, H. Gao, and B. Schmalfuss. Stochastic dynamics of a coupled atmosphere-ocean model. *Stochastics and Dynamics*, 02(03):357–380, 2002.

[26] J. Duan and B. Goldys. Ergodicity of stochastically forced large scale geophysical flows. *International J. Math. Math. Sci*, pages 313–320, 2001.

[27] J. Duan, P. E. Kloeden, and B. Schmalfuss. Exponential stability of the quasigeostrophic equation under random perturbations. *Progress in Probability*, 49:241–256, 2000.

[28] R. M. Dudley. *Real Analysis and Probability*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2002.

[29] W. E, J. C. Mattingly, and Y. Sinai. Gibbsian dynamics and ergodicity for the stochastically forced Navier–Stokes equation. *Communications in Mathematical Physics*, 224(1):83–106, 2001.

[30] B. Ferrario. Ergodic results for stochastic Navier-Stokes equation. *Stochastics and Stochastic Reports*, 60(3-4):271–288, 04 1997.

[31] B. Ferrario. Pathwise regularity or nonlinear Itô equations: application to a stochastic Navier-Stokes equation. *Stochastic Analysis and Applications*, 19(1):135–150, 02 2001.

[32] F. Flandoli. Dissipativity and invariant measures for stochastic Navier-Stokes equations. *Nonlinear Differential Equations and Applications NoDEA*, 1(4):403–423, 1994.

[33] F. Flandoli and B. Maslowski. Ergodicity of the 2-D Navier-Stokes equation under random perturbations. *Communications in Mathematical Physics*, 172(1):119–141, 1995.

[34] C. Frankignoul and P. Müller. Quasi-geostrophic response of an infinite β-plane ocean to stochastic forcing by the atmosphere. *Journal of Physical Oceanography*, 9(1):104 – 127, 1979.

[35] S. Galatolo and P. Giulietti. A linear response for dynamical systems with additive noise. *Nonlinearity*, 32(6):2269–2301, may 2019.

[36] M. Ghil and V. Lucarini. The physics of climate variability and climate change. *Rev. Mod. Phys.*, 92:035002, Jul 2020.

[37] N. Glatt-Holtz, J. C. Mattingly, and G. Richards. On unique ergodicity in nonlinear stochastic partial differential equations. *Journal of Statistical Physics*, 166(3-4):618–649, 2017.

[38] A. Griffa and S. Castellari. Nonlinear general circulation of an ocean model driven by wind with a stochastic component. *Journal of Marine Research*, 49:53–73, 02 1991.

[39] M. Hairer. Ergodic Theory of Stochastic PDEs. Lecture Notes, July 2008.

[40] M. Hairer and A. J. Majda. A simple framework to justify linear response theory. *Nonlinearity*, 23(4):909, 2010.

[41] M. Hairer and J. C. Mattingly. Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing. *Annals of Mathematics*, 164(3):993–1032, 2006.

[42] M. Hairer and J. C. Mattingly. Spectral gaps in Wasserstein distances and the 2D stochastic Navier–Stokes equations. *Ann. Probab.*, 36(6):2050–2091, 11 2008.

[43] M. Hairer and J. C. Mattingly. Yet another look at Harris'ergodic theorem for Markov chains. In R. Dalang, M. Dozzi, and F. Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications VI*, pages 109–117, Basel, 2011. Springer Basel.

[44] M. Hairer, J. C. Mattingly, and M. Scheutzow. Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probability Theory and Related Fields*, 149(1):223–259, Feb 2011.

[45] K. Hasselmann. Stochastic climate models Part I. Theory. *Tellus*, 28(6):473–485, 1976.

[46] J. Holton. *An Introduction to Dynamic Meteorology.* Number v. 1 in An Introduction to Dynamic Meteorology. Elsevier Science, 2004.

[47] I. Karatzas. *Brownian Motion and Stochastic Calculus.* Graduate Texts in Mathematics (113) (Book 113). Springer New York, 1991.

[48] S. Kuksin and A. Shirikyan. A Coupling Approach to Randomly Forced Nonlinear PDE's. I. *Communications in Mathematical Physics*, 221(2):351–366, 2001.

[49] A. Kulik. *Ergodic Behavior of Markov Processes : With Applications to Limit Theorems.* De Gruyter, Inc., Berlin/Boston, Germany, 2017.

[50] A. Kulik and M. Scheutzow. Generalized couplings and convergence of transition probabilities. *Probability Theory and Related Fields*, 171(1):333–376, 2018.

[51] R. Liptser and A. Shiryaev. *Statistics of Random Processes: I. General Theory.* Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2013.

[52] V. Lucarini. Stochastic perturbations to dynamical systems: a response theory approach. *Journal of Statistical Physics*, 146(4):774–786, 2012.

[53] V. Lucarini, F. Ragone, and F. Lunkeit. Predicting climate change using response theory: Global averages and spatial patterns. *Journal of Statistical Physics*, 166(3-4):1036–1064, 2017.

[54] A. Majda. *Introduction to PDEs and Waves for the Atmosphere and Ocean*, volume 9 of *Courant Lecture Notes.* New York University, Courant Institute of Mathematical Sciences and American Mathematical Society, 2003.

[55] A. J. Majda, R. Abramov, and B. Gershgorin. High skill in low-frequency climate response through fluctuation dissipation theorems despite structural instability. *Proceedings of the National Academy of Sciences*, 107(2):581–586, 2010.

[56] J. C. Mattingly. Ergodicity of 2D Navier–Stokes equations with random forcing and large viscosity. *Communications in Mathematical Physics*, 206(2):273–288, Oct 1999.

[57] S. Meyn, R. L. Tweedie, and P. W. Glynn. *Markov Chains and Stochastic Stability.* Cambridge Mathematical Library. Cambridge University Press, 2 edition, 2009.

[58] A. M. Moore. Wind-induced variability of ocean gyres. *Dynamics of Atmospheres and Oceans*, 29(2):335–364, 1999.

[59] P. Müller and C. Frankignoul. Direct atmospheric forcing of geostrophic eddies. *Journal of Physical Oceanography*, 11(3):287 – 308, 1981.

[60] C. H. O'Reilly, A. Czaja, and J. LaCasce. The emergence of zonal ocean jets under large-scale stochastic wind forcing. *Geophysical research letters*, 39(11), 2012.

[61] J. Pedlosky. *Geophysical fluid dynamics*. Springer Science & Business Media, 2013.

[62] J. Robinson. *Infinite-Dimensional Dynamical Systems: An Introduction to Dissipative Parabolic PDEs and the Theory of Global Attractors*. Cambridge University Press, 2001.

[63] M. Röckner, R. Zhu, X. Zhu, et al. Sub and supercritical stochastic quasi-geostrophic equation. *Annals of Probability*, 43(3):1202–1273, 2015.

[64] L. Rogers and D. Williams. *Diffusions, Markov Processes, and Martingales: Volume 1, Foundations*. Cambridge Mathematical Library. Cambridge University Press, 2000.

[65] D. Ruelle. Differentiation of SRB states. *Communications in Mathematical Physics*, 187(1):227–241, 1997.

[66] P. Sura, K. Fraedrich, and F. Lunkeit. Regime transitions in a stochastically forced double-gyre model. *Journal of Physical Oceanography*, 31(2):411 – 426, 2001.

[67] P. Sura and C. Penland. Sensitivity of a double-gyre ocean model to details of stochastic forcing. *Ocean Modelling*, 4(3):327–345, 2002.

[68] R. Temam. *Navier-Stokes equations: theory and numerical analysis*, volume 343. American Mathematical Soc., 2001.

[69] R. Temam. *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*. Applied Mathematical Sciences. Springer New York, 2012.

[70] A. M. Treguier and B. L. Hua. Oceanic quasi-geostrophic turbulence forced by stochastic wind fluctuations. *Journal of Physical Oceanography*, 17(3):397 – 411, 1987.

[71] G. K. Vallis. *Atmospheric and Oceanic Fluid Dynamics.* Cambridge University Press, Cambridge, U.K., 2006.

[72] C. Villani. *Optimal transport: old and new.* Springer, Berlin, 2008.

[73] D. Yang and J. Duan. Large deviations for the stochastic quasi-geostrophic equation with multiplicative noise. *Journal of Mathematical Physics*, 51(5):053301, 2010.