

# On the Nikolaevskiy Equation and the Fractal Dimension of its Attractor

by

**Hans Oeri**

B.Sc., Simon Fraser University, 2015

Thesis Submitted in Partial Fulfillment of the  
Requirements for the Degree of  
Master of Science

in the  
Department of Mathematics  
Faculty of Science

© Hans Oeri 2018  
SIMON FRASER UNIVERSITY  
Spring 2018

Copyright in this work rests with the author. Please ensure that any reproduction or re-use is done in accordance with the relevant national copyright legislation.

# Approval

**Name:** Hans Oeri

**Degree:** Master of Science

**Title:** On the Nikolaevskiy Equation and the Fractal Dimension of its Attractor

**Examining Committee:** **Chair:** JF Williams  
Associate Professor

**Ralf Wittenberg**  
Senior Supervisor  
Associate Professor

**Weiran Sun**  
Co-Supervisor  
Assistant Professor

**Razvan Fetecau**  
Internal Examiner  
Associate Professor

**Date Defended:** April 23, 2018

# Abstract

We investigate the attractor of the Nikolaevskiy equation, a sixth-order partial differential equation (PDE) containing a small parameter whose solutions exhibit spatiotemporal chaos with strong scale separation.

We first prove well-posedness and regularity of the solutions, and derive asymptotic bounds on their derivatives, to put the subsequent results on a firm footing. The rest of the work focuses on showing that the dynamical system associated with the Nikolaevskiy equation possesses an attractor with a finite fractal dimension. Bounds on this dimension are both derived analytically and computed numerically, paying particular attention to their scaling with the parameters. We describe the numerical methods, and present computational results that include the scaling of various norms of the solutions, as well as of the power spectrum and the spectrum of Lyapunov exponents of the PDE.

**Keywords:** Nikolaevskiy Equation; Parameter Scaling; Attractor; Fractal Dimension

# Dedication

I dedicate this thesis to my parents; to my mother who, through her kindness and support, enables me to study, and to the memory of my father whose great appreciation of nature has led me into science.

# Acknowledgements

First and foremost I thank my two wonderful supervisors, Dr. Ralf Wittenberg and Dr. Weiran Sun whose patience and training has brought me to this point and has empowered me to finish this thesis. I thank Weiran for her tireless and kind support and for always being willing to help me. I thank her for sharing her wide knowledge of analysis with me and for providing me with the fundamental skills to begin working with PDEs. I thank Ralf for his immense support and for opening up an incredibly fascinating field of research to me, by introducing me to the material in this work and by having the patience and dedication to see me through this long process. His gentle yet persistent mentoring has taught me things well beyond the scope of academia, and without him this work would not have been possible.

I also thank the chair Dr. JF Williams and the examiner Dr. Razvan Fetecau for the many hours of work involved in reading the thesis and evaluating its contents.

# Table of Contents

Approval	ii
Abstract	iii
Dedication	iv
Acknowledgements	v
Table of Contents	vi
List of Figures	viii
<b>1 Introduction</b>	<b>1</b>
1.1 The Kuramoto-Sivashinsky Equation . . . . .	1
1.2 The Nikolaevskiy Equation . . . . .	4
1.3 Spatiotemporal Chaos . . . . .	6
1.4 Basic Properties of the Nikolaevskiy Equation . . . . .	7
<b>2 Well-Posedness</b>	<b>10</b>
2.1 Spaces of Solutions, and Basic Framework . . . . .	10
2.2 The Galerkin Approximations . . . . .	13
2.3 Bounds on Galerkin Approximations to the Nikolaevskiy Equation . . . . .	15
2.4 Convergence of the Galerkin Approximations . . . . .	21
2.5 Uniqueness . . . . .	26
<b>3 Asymptotic Bounds and Regularity</b>	<b>29</b>
3.1 An Absorbing Set in $\dot{L}^2(0, \ell)$ . . . . .	29
3.2 An Absorbing Set in $\dot{H}^1(0, \ell)$ . . . . .	30
3.3 An Absorbing Set in $\dot{H}^2(0, \ell)$ . . . . .	32
3.4 Regularity . . . . .	33
<b>4 Attractor Dimension</b>	<b>42</b>
4.1 Existence of an Attractor, and Fractal Dimensions . . . . .	42
4.2 Tools and Lemmas . . . . .	44

4.3	Bounding the Fractal Dimension of a Set . . . . .	50
4.4	Growth of Volumes . . . . .	53
4.5	Uniform Differentiability . . . . .	59
4.6	Compactness . . . . .	63
4.7	The Dimension of the Attractor . . . . .	66
<b>5</b>	<b>Numerical Methods</b>	<b>73</b>
5.1	Lyapunov Exponents . . . . .	73
5.2	Computing Lyapunov Exponents . . . . .	75
5.3	Kaplan-Yorke Dimension . . . . .	77
5.4	Numerical Methods . . . . .	81
<b>6</b>	<b>Numerical Results</b>	<b>84</b>
6.1	Solutions of the Nikolaevskiy Equation . . . . .	84
6.2	The Power Spectrum . . . . .	89
6.3	Lyapunov Spectrum . . . . .	90
6.4	Dimension . . . . .	92
<b>7</b>	<b>Conclusion</b>	<b>97</b>
	<b>Bibliography</b>	<b>100</b>
	<b>Appendix A Table of Notation</b>	<b>105</b>
	<b>Appendix B Lemmas and Tools</b>	<b>107</b>
B.1	ODE Theory . . . . .	107
B.2	Functional Analysis and Linear Algebra . . . . .	107
B.3	Young's-type Inequalities . . . . .	111
B.4	Other Inequalities . . . . .	114
B.5	Regularity . . . . .	116
B.6	The Nonlinear Term . . . . .	119

# List of Figures

Figure 1.1	Linear dispersion relation for the KS equation for various values of $\alpha$	3
Figure 1.2	Linear dispersion relation $\omega(k)$ (1.5), for the Nikolaevskiy equation for various values of $r$ . . . . .	5
Figure 6.1	Plots of solutions to equation (1.4) for $r \ll 1$ . Solutions are within the Nikolaevskiy chaos regime. The top panels display the solutions over a range of $t$ , while the bottom panels display a snapshot of each solution at an arbitrarily chosen time. The reader is encouraged to zoom in on the plots to see the short wavelength rolls. Solutions are shown at parameter values (a) $r = 0.0001$ and $\ell = 5\pi \times 10^4$ and at (b) $r = 0.04$ and $\ell = 1600$ . . . . .	85
Figure 6.2	Plot of a typical solution of the KS equation (1.1). The top panel displays the solution over a range of $t$ , while the bottom panel displays a snapshot of the solution at an arbitrary time for $\ell = 200$ . . . . .	86
Figure 6.3	Plots of solutions to the Nikolaevskiy equation (1.4) well into the KS-type chaos regime. The top panels display the solutions over a range of $t$ , while the bottom panels display a snapshot of each solution at an arbitrary time. Solutions are shown at parameter values (a) $r = 0.5$ and $\ell = 100$ and at (b) $r = 1$ and $\ell = 30$ . . . . .	87
Figure 6.4	Plot of solutions to the Nikolaevskiy equation (1.4) in the transition region between Nikolaevskiy-type and KS-type chaos at parameter values $r = 0.1$ and $\ell = 1000$ . The top panel displays the solution over a range of $t$ , while the bottom panel displays a snapshot of the solution at an arbitrary time. . . . .	88
Figure 6.5	Scaling of (a) the $\dot{L}^2$ norm and (b) the $\dot{L}^\infty$ norm of solutions to (1.4) with $\ell$ . Plots show the data and a best fit line on a log-log scale, for fixed $r = 1$ . For the $\dot{L}^2$ norm (a) the scaling of the exponent on $\ell$ is $0.492 \pm 0.032$ . For the $\dot{L}^\infty$ norm (b) the scaling of the exponent on $\ell$ is $-0.024 \pm 0.068$ . . . . .	89
Figure 6.6	Plots of the normalized time-averaged power spectrum $S(k)/\ell$ , for $r = 10^{-n}$ for $n \in \{0, 1, 2, 3, 4\}$ for the Nikolaevskiy equation (blue), as well as for the KS equation (red). . . . .	90



Figure 6.7	Lyapunov spectra, showing the LEs $\mu$ vs. the wavenumber $k$ . The parameter $r$ is fixed at (a) $r = 1$ , (b) $r = 0.9$ , (c) $r = 0.5$ , (d) $r = 0.1$ , (e) $r = 0.03$ and (f) $r = 0.01$ . For each $r$ value, the spectra are shown for system size $\ell \in \{100, 200, 300, 400\}$ . . . . .	91
Figure 6.8	Kaplan-Yorke dimension $d_f$ as a function of $\ell$ for: $\circ r = 1$ with $D = 0.351$ , $\times r = 0.5$ with $D = 0.328$ , $\times r = 0.1$ with $D = 0.244$ and $\circ r = 0.03$ with $D = 0.110$ . . . . .	93
Figure 6.9	(a) Log-log plot of the dimension density $D$ versus $r$ , over a range $0.2 \leq r \leq 1$ . The scaling exponent on $r$ is approximately 0.093. (b) A log-log plot of $D$ vs. $r$ , over a range $0.03 \leq r \leq 1$ . The scaling exponent on $r$ for small $r$ is approximately 0.665. . . . .	94
Figure 6.10	(a) Kaplan-Yorke dimension $d_f$ versus $r$ for fixed $\Delta n$ . (b) Log-log plot of Kaplan-Yorke dimension $d_f$ versus $r$ for fixed $\Delta n$ . In both plots, $\Delta n \in \{24, 36, 50\}$ . . . . .	95

# Chapter 1

## Introduction

The study of chaotic dynamical systems has long been a major focus of mathematics and physics. Defining characteristics of chaotic systems are that nearby solutions separate exponentially fast in phase space, and that their long-term dynamics is confined to a compact set, known as an attractor [Str15]. Much work has gone into characterizing the dynamics on the attractor of these systems, and its dependence on the system's parameters. Although chaos was originally discovered in finite dimensions [Lor63], since the 1980s, partial differential equations (PDEs) whose solutions are chaotic have also been investigated extensively [HN86, Li04]. Although many PDEs can exhibit chaotic behaviour, some PDEs have emerged as canonical models that capture the essential features of various types of chaos, including spatiotemporal chaos in PDEs. Among these, the Kuramoto-Sivashinsky (KS) equation [KT76, Siv77, HN86, WH99], discussed in Section 1.1, has been of particular interest, and has been studied in great detail as a canonical model of chaos in PDEs. More recently, KS-like equations, such as the (de)stabilized KS equations and the Nikolaevskiy equation have also been investigated for their parameter-dependent, chaotic and pattern-forming properties [Wit14].

In this text we will focus on the one-dimensional Nikolaevskiy equation [Nik89, Poo09], a sixth-order analogue of the KS equation, whose solutions possess interesting scale separation properties in the presence of a small bifurcation parameter  $r$ . We perform analytical and numerical studies of solutions of the equation to estimate the dependence of the long-term dynamics of the Nikolaevskiy equation on its parameters.

### 1.1 The Kuramoto-Sivashinsky Equation

Before we talk about the Nikolaevskiy equation, we will discuss the Kuramoto-Sivashinsky (KS) equation, a more widely studied PDE that shares many properties with the Nikolaevskiy equation. Many of the techniques developed to study the KS equation carry over to the Nikolaevskiy equation, and it is therefore helpful to first understand some aspects of the KS equation. The KS equation has been proposed as a model for several physi-

cal phenomena, such as the evolution of disturbed flame fronts [Siv80]. Other applications of this model include plasma physics [LMRT75] and as a simplified model for describing some of the phenomena found in the Navier-Stokes equations, such as wave propagation [KT76]. It combines a fourth-order diffusive term with a second-order destabilizing term and a Burgers-like, nonlinear reaction term. Although other boundary conditions have been studied [Man85], periodic boundary conditions are a natural choice for the KS equation due to certain symmetries that its solutions enjoy. The equation is given by

$$u_t = -\partial_x^4 u - 2\partial_x^2 u - u \partial_x u, \quad x \in [0, \ell]. \quad (1.1)$$

Since the time derivative only appears in the equation to first order, we will just use a subscript  $t$  to denote it, whereas we will use  $\partial_x$  to denote space derivatives. The scaling which leads to the coefficient two is non-standard for the KS equation, but convenient for generalization, and for factoring the linear operator [Wit14, Poo09]. The domain size  $\ell$  plays the role of a bifurcation parameter, and governs the onset of chaos.

A related equation is the (de)stabilized KS equation, which contains a term proportional to  $u$ , which has the effect of stabilizing or destabilizing long wavelength modes, depending on the sign of the coefficient  $\alpha$ . It is given by

$$u_t = -\partial_x^4 u - 2\partial_x^2 u + \alpha u - u \partial_x u, \quad x \in [0, \ell]. \quad (1.2)$$

Its **linear dispersion relation** is

$$\omega_{dKS}(k) = -k^4 + 2k^2 + \alpha = -(1 - k^2)^2 + 1 + \alpha. \quad (1.3)$$

Here,  $k = 2\pi n/\ell$ , with  $n \in \mathbb{Z}$ , is a Fourier wavenumber of the solution, and the sign of  $\omega_{dKS}(k)$  gives the stability of the  $n$ 'th Fourier mode and of  $u$  about the zero solution (the stability of the zero solution, which is a fixed point of equation (1.2)). For  $\alpha = 0$ , this is just the dispersion relation for (1.1). The dispersion relation is shown in Figure 1.1.

It can be shown that the KS equation possesses an absorbing ball in its phase space  $\dot{L}^2$  (as defined in Appendix A), containing the attractor of the system [NST85]. The radius of this absorbing ball depends on the domain size  $\ell$ . Numerical evidence indicates, that the radius of the absorbing ball is bounded by some constant times  $\ell^{1/2}$  [WH99]; that is, for large  $t$ , one has  $\|u(t)\|_{\dot{L}^2} = \mathcal{O}(\ell^{1/2})$ . In numerical experiments, the maximum amplitude of  $|u(x, t)|$  is bounded independent of the domain size  $\ell$ , which implies the bound on the  $\dot{L}^2$  norm of  $u$ .

Since a bound  $\mathcal{O}(\ell^{5/2})$  was initially derived for odd solutions by Nicolaenko *et al.* [NST85], the exponent on  $\ell$  has been subsequently refined to  $8/5$  in [CEES93b] (for any initial condition) and to  $3/2$  in [BG06]. A further improvement to  $o(\ell^{3/2})$  has been proved

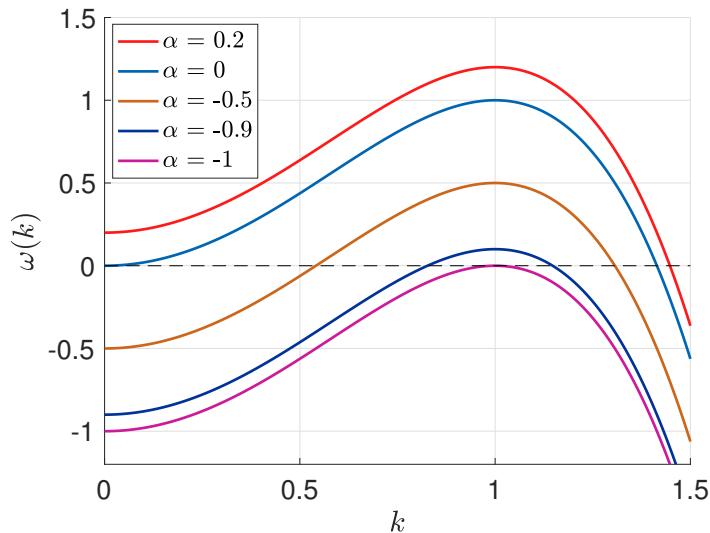


Figure 1.1: Linear dispersion relation for the KS equation for various values of  $\alpha$

in [GO04], and a very recent result using semidefinite programming suggests a method that can give arbitrarily tight bounds, given enough computing power [GF18].

Part of the interest in refining this bound stems from the fact that many other extensive quantities, such as the radius of analyticity of solutions [CEES93a], bounds in other spaces, and the fractal dimension of the attractor [NST85], depend on the radius of the absorbing ball. Since the number of Fourier modes in the unstable band of (1.3) is proportional to  $\ell$  for any fixed  $\alpha$ , the scaling of the dimension is expected to be  $\mathcal{O}(\ell)$ . This proposition is also supported by numerical experiments [Man85].

The KS equation shares similar diffusive linear terms as well as an identical nonlinear term with the Nikolaevskiy equation, introduced below, which means that the methods for the KS equation are generalizable to the Nikolaevskiy equation. In much of the analysis, the main difficulties arise from finding suitable ways to treat the nonlinear term in the equation. In many applications, such as proving analyticity, bounding the long-term degrees of freedom and energy estimates, one can directly apply results obtained for the KS equation to the Nikolaevskiy equation. Furthermore, the KS equation seems to share the same scaling properties with the system size  $\ell$  as the Nikolaevskiy equation for all the quantities that we study here. These properties, together with the considerably larger body of knowledge that exists for the KS equation makes the KS a useful model equation when studying the Nikolaevskiy equation.

## 1.2 The Nikolaevskiy Equation

The Nikolaevskiy equation is sixth-order in space and nonlinear. In this text, we only study the Nikolaevskiy equation in one space dimension. For  $r \in \mathbb{R}$ , it is given by

$$u_t = \partial_x^6 u + 2\partial_x^4 u + (1-r)\partial_x^2 u - u \partial_x u. \quad (1.4)$$

For  $r \leq 0$ , all solutions to equation (1.4) decay to zero [WP09]. Since we are interested in solutions that contain energy for all time (their  $\dot{L}^2$  norm is nonzero), we will only consider  $r > 0$ . We sometimes write  $r$  as  $r := \epsilon^2$  for  $\epsilon \in \mathbb{R}$  to emphasize that  $r$  is positive, and because some quantities, such as the energy of the solutions, scale with  $\sqrt{r}$  rather than with  $r$ . These coefficients allow us to factorize the linear operator in equation (1.4) and write it as

$$u_t + u \partial_x u = -\partial_x^2 (\epsilon^2 - (1 + \partial_x^2)^2) u, \quad x \in [0, \ell],$$

following the notation in [Wit14] and [WP09].

In the following chapters, we will often make use of the linear dispersion relation of equation (1.4):

$$\omega(k) := -k^6 + 2k^4 - (1-r)k^2, \quad (1.5)$$

where  $k = 2\pi j/\ell$ , with  $j \in \mathbb{Z}$ , denotes a wavenumber of the solution  $u$ . The dispersion relation  $\omega(k)$  has the same interpretation for equation (1.4) as for equation (1.2). Figure 1.2 shows some plots of the linear dispersion relation of (1.4).

The Nikolaevskiy equation is physically relevant, in that it has been derived as a model for nonlinear seismic waves in viscoelastic materials, such as marine sands [Nik89]. Later the Nikolaevskiy equation was proposed by Tribelski and Tsuboi [TT96] as a model for describing the transition to turbulence, where it was numerically shown that there exists a supercritical bifurcation in the parameter  $r$  that demarcates the boundary between the chaotic and the spatially uniform regime of solutions. In addition, much work has also gone into investigating its pattern-forming properties (see for instance [MC00, WP09, Poo09]). Most investigations of equation (1.4) have employed numerical and asymptotic methods, but a few exact solutions have also been found [RSK11], albeit on an unbounded domain.

We can view equation (1.4) as a dynamical system mapping time into a function space, which serves as the phase space of the system. We will define this function space in the next chapter. One quantity associated with this dynamical system is the radius  $R$  of an absorbing ball in the phase space  $\dot{L}^2$  of solutions. The estimation of this radius, in particular its scaling with the two parameters  $r$  and  $\ell$ , lies at the heart of much of the analysis that is done to derive extensive properties, such as the number of degrees of freedom of solutions of the

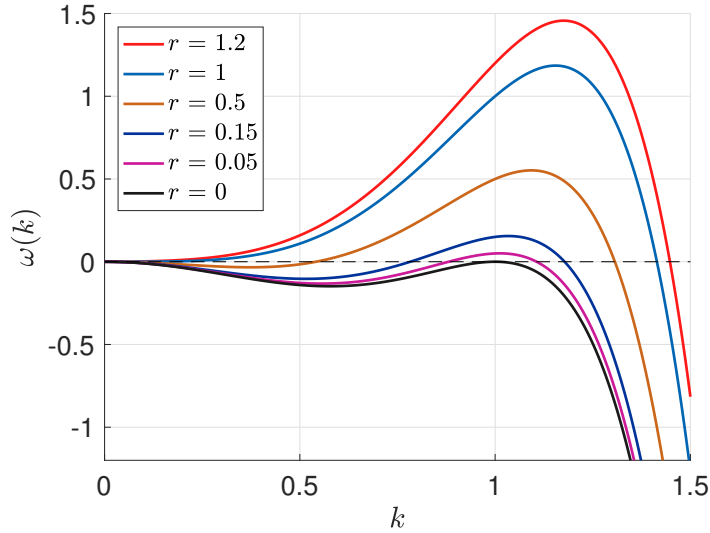


Figure 1.2: Linear dispersion relation  $\omega(k)$  (1.5), for the Nikolaevskiy equation for various values of  $r$ .

equation. The optimal scaling in  $r$  has been derived rigorously for the radius of the absorbing ball [Wit14], but the scaling in  $\ell$  is subject to the same difficulties that are encountered when deriving parameter scalings for the absorbing ball of the better-known KS equation, since the same methods are used on both equations, and thus still remains an open problem. The importance to the analysis of the radius  $R$  for the absorbing ball of equation (1.4) lies in the fact that many other quantities, such as the dimension of the attractor, the radius of analyticity, and the radii of absorbing balls in other Sobolev spaces, depend upon it. A primary focus of this thesis will be to attempt to derive bounds on these quantities, paying particular attention to their dependence on the parameter  $r$ .

For small  $r$ , the solutions to (1.4) display strong scale separation, with the unstable modes being in two distinct bands, one containing long wavelengths and one containing short wavelengths. This gives rise to high frequency oscillations, modulated by a low frequency wave envelope. The strong separation between the short wavelength and long wavelength unstable modes gives rise to the so-called ‘Nikolaevskiy Chaos’, which was first found using asymptotic analysis in [TT96] (where it was called ‘slow turbulence’) and later shown to exist numerically in [Tan05b]. Bifurcation routes to Nikolaevskiy chaos were explored in [Tan05a], and extensive numerical and asymptotic work was later done in [Poo09] to further characterize Nikolaevskiy chaos. It should be noted that this type of chaos is not found in the KS equation (1.1) or its generalization, the (de)stabilized KS equation (1.2), and is one of the main reasons why the Nikolaevskiy equation (1.4) deserves to be studied in its own right.

### 1.3 Spatiotemporal Chaos

In this section we give a brief overview of the type of dynamics that solutions to both the Nikolaevskiy equation (1.4) and the KS equation (1.1) display for the choices of system size  $\ell$  and the parameter  $r$  in this thesis. As stated earlier, we will be studying solutions in the chaotic regime of these equations, but characterizing and quantifying the type of chaos they exhibit is not always straightforward, and we will attempt to provide some definitions to help make these concepts more clear.

One can generally distinguish between two types of chaos: Temporal chaos [Str15, p. 331], which is more often found in low-dimensional systems such as the famous Lorenz equations [Lor63], and spatiotemporal chaos (STC), which is more often found in systems with a high number of degrees of freedom [CH93, p. 941], and which requires solutions to depend on both time and space. Several different definitions have been proposed to quantify STC. Most rely on the asymptotic scaling of the number of degrees of freedom of the system, or on measuring the rate of decay of spatial and temporal correlations of values of the solutions.

As we will see in Section 4.1, asymptotically in time all the dynamics of the system occurs on a bounded set in the phase space of solutions, called an attractor. Furthermore, in Section 5.1, we will discuss how one can use the exponential rate of separation of nearby solutions (the Lyapunov exponents) to estimate the number of degrees of freedom of solutions on the attractor. Later we will make this more precise by defining the concept of the fractal dimension  $d_f$  of the attractor. A straightforward definition of STC is to say that a system exhibits STC if its fractal dimension diverges with the system size  $\ell$  [CH93, p. 945], while keeping any other parameters fixed. Numerical studies by Xi *et al.* [XTGT00] have shown this to be true for equation (1.4), at least for fairly large  $r$ . In Section 6.4 we confirm this result and extend it to a wider range of  $r$  values. This definition is useful for checking for the existence of STC, but it does not directly separate the chaotic behaviour in time from that in space.

To clearly separate the ‘spatial’ from the temporal chaos of solutions, it is often useful to define a correlation length  $\xi$ . Spatiotemporally chaotic systems are characterized by a finite correlation length [CH93, CH94]. Although one can use the  $\ell$  scaling of the fractal dimension of the attractor to define a correlation length [Gre96], a more common way to define it is via a correlation function. A correlation function essentially measures the degree to which some variables depend on each other. If the correlation is high, then one can generally find some relationship between the variables, while if it is low, then the variables are essentially independent from each other, and knowledge about one variable does not provide knowledge about the other variable. In chaotic dynamical systems, two nearby initial conditions (which are thus initially highly correlated) diverge exponentially fast, and after only a short while, their trajectories will be unrelated and it becomes very difficult to distinguish solutions that started out close to each other from ones that started out far away from each other.

The trajectories have then become decorrelated, which is measured by a low value of their correlation function.

Temporal chaos is characterized by a rapidly decaying temporal correlation function, whereas STC is characterized by both a rapidly decaying temporal correlation function *and* a rapidly decaying spatial correlation function. The correlation length quantifies how fast two variables become decorrelated. There are many different ways to define correlation functions [CH93, p. 945]. For instance, one way to define a correlation function is via the so-called autocorrelation functions as follows:

$$C(\Delta x, x') := \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T u(x' - \Delta x, t) u(x', t) dt$$

for the spatial autocorrelation function (for fixed  $x'$ ) as in [WH99], and as

$$C(\Delta \tau, t) := \frac{1}{\ell} \int_0^\ell u(x, t - \Delta \tau) u(x, t) dx,$$

for the temporal autocorrelation function (for fixed  $t$ ). If the system is chaotic, we expect

$$C(\Delta x) = \mathcal{O}(e^{-\Delta x/\xi}),$$

in which case  $\xi$  is the correlation length (or correlation time, if  $x$  is replaced with  $\tau$ ). For equation (1.4), it has been shown that the system has finite correlation lengths and times [Poo09].

## 1.4 Basic Properties of the Nikolaevskiy Equation

### Nondimensionalization:

Since the Nikolaevskiy equation is proposed in order to model physical systems [Nik89], its terms will generally have coefficients that come from the physical derivation. It is usually easiest to work with the equation when it has the form (1.4), and so we will first begin by showing how to non-dimensionalize the equation to obtain this form. Consider therefore the PDE in the following form:

$$\tilde{u}_{\tilde{t}} = A \partial_{\tilde{x}}^6 \tilde{u} + B \partial_{\tilde{x}}^4 \tilde{u} + C \partial_{\tilde{x}}^2 \tilde{u} + D \tilde{u} \partial_{\tilde{x}} \tilde{u}, \quad \tilde{x} \in [0, E], \quad (1.6)$$

where  $A > 0, B > 0, C, D$  and  $E > 0$  are some constants. Without loss of generality, we assume that the coefficient in front of  $\tilde{u}_{\tilde{t}}$  is 1. Let

$$\tilde{u} = \alpha u, \quad x = \beta \tilde{x}, \quad t = \gamma \tilde{t},$$



where  $\alpha, \beta$  and  $\gamma$  are to be determined. With this change of variables the equation (1.6) becomes:

$$u_t = \underbrace{\frac{A\beta^6}{\gamma}}_a \partial_x^6 u + \underbrace{\frac{B\beta^4}{\gamma}}_b \partial_x^4 u + \underbrace{\frac{C\beta^2}{\gamma}}_c \partial_x^2 u + \underbrace{\frac{D\alpha\beta}{\gamma}}_d u \partial_x u, \quad x \in [0, \underbrace{E\beta}_\ell].$$

We now choose  $\alpha, \beta$  and  $\gamma$  to make  $a = 1$ ,  $b = 2$ , and  $d = 1$ , giving

$$\begin{aligned} \alpha &= \frac{A}{D} \left( \frac{B}{2A} \right)^{5/2}, \\ \beta &= \left( \frac{B}{2A} \right)^{1/2}, \\ \gamma &= \frac{B^3}{8A^2}, \end{aligned}$$

and define  $r$  by  $r := 1 - c$  and  $\ell$  by  $\ell := E\beta$ ; then the equation takes on the desired form (1.4).

### Symmetries and Conservation of Mean:

Another property of (1.4) is that its flow preserves the spatial mean of solutions. In everything that follows, we will be imposing periodic boundary conditions on the solutions  $u$  and on their derivatives. Thus the spatial mean evolves as follows:

$$\begin{aligned} \frac{d}{dt} \frac{1}{\ell} \int_0^\ell u \, dx &= \frac{1}{\ell} \int_0^\ell u_t \, dx = \frac{1}{\ell} \int_0^\ell \left( \partial_x^6 u + 2\partial_x^4 u + (1-r)\partial_x^2 u - \frac{1}{2}\partial_x(u^2) \right) dx \\ &= \frac{1}{\ell} \left( \partial_x^5 u + 2\partial_x^3 u + (1-r)\partial_x u - \frac{1}{2}u^2 \right) \Big|_0^\ell = 0. \end{aligned}$$

This shows that the spatial mean is conserved, and for this reason we will work in spaces of functions with zero mean.

Besides conserving the mean, the Nikolaevskiy equation also enjoys several symmetries. The equation is symmetric under space and time shifts,  $x \mapsto x + c_1$  and  $t \mapsto t + c_2$ , for some constants  $c_1$  and  $c_2$ . The equation is also invariant under Galilean transformations of the form  $x \mapsto x - c_3 t$  and  $u(x, t) \mapsto \tilde{u} := u(x - c_3 t, t) + c_3$ , since if  $u$  solves equation (1.4), then

$$\begin{aligned} \tilde{u}_t &= \partial_x^6 \tilde{u} + 2\partial_x^4 \tilde{u} + (1-r)\partial_x^2 \tilde{u} - \tilde{u} \partial_x \tilde{u} \\ \iff -c_3 \partial_x u + u_t &= \partial_x^6 u + 2\partial_x^4 u + (1-r)\partial_x^2 u - (u + c_3) \partial_x u. \end{aligned}$$

Thus both  $u$  and  $\tilde{u}$  solve the PDE, meaning that after applying the above transformations to  $u$ , the result is still a solution to the Nikolaevskiy equation. Another important symmetry which we will make use of is the symmetry under odd reflections. Let  $\tilde{x} := -x$  and  $\tilde{u}(\tilde{x}, t) :=$

$-u(x, t)$ . Then since  $\partial_x = -\partial_{\tilde{x}}$ , if  $u(x, t)$  solves equation (1.4), one has

$$\begin{aligned} u_t &= \partial_x^6 u + 2\partial_x^4 u + (1-r)\partial_x^2 u - u \partial_x u \\ \iff u_t &= \partial_{\tilde{x}}^6 u + 2\partial_{\tilde{x}}^4 u + (1-r)\partial_{\tilde{x}}^2 u + u \partial_{\tilde{x}} u \\ \iff \tilde{u}_t &= \partial_{\tilde{x}}^6 \tilde{u} + 2\partial_{\tilde{x}}^4 \tilde{u} + (1-r)\partial_{\tilde{x}}^2 \tilde{u} - \tilde{u} \partial_{\tilde{x}} \tilde{u}, \end{aligned}$$

and  $\tilde{u}(\tilde{x}, t)$  also solves (1.4). Thus the odd reflection of  $u$  also solves the PDE. This means that solutions that are initially odd, remain so for all time. Thus the subspace of odd solutions is invariant under the flow of (1.4)

# Chapter 2

## Well-Posedness

In this chapter we will set up the framework in which we will conduct the rest of the work in the thesis. We will rigorously prove existence and uniqueness of solutions of the Nikolaevskiy equation (1.4). We begin by motivating the approach that we are taking and defining the function spaces that we will be working in. We then prove existence and uniqueness of solutions. We also state some known results about long-term bounds on solutions in  $\dot{L}^2$ , and then extend these bounds to more function spaces. Finally, it is shown that solutions are analytic in space. Most of the work in this chapter is based on the procedure in [Rob01] for proving existence and uniqueness for the KS equation (1.1).

### 2.1 Spaces of Solutions, and Basic Framework

In this section we list some basic properties that functions need to have in order to be eligible as solutions to equation (1.4):

$$u_t = \partial_x^6 u + 2\partial_x^4 u + (1-r)\partial_x^2 u - u \partial_x u, \quad u(0, x) = u_0 \in \dot{L}^2(0, \ell).$$

These properties are then used to motivate the choice of function spaces in which we seek solutions. The solutions to (1.4) require a degree of temporal as well as spatial differentiability in order to be able to satisfy the equation. At first we will not be looking for classical (the derivatives exist classically and the right hand side of the PDE equals the left hand side pointwise) or even strong solutions (the right hand side of the PDE equals the left hand side in the  $L^2(0, T; \dot{L}^2)$  norm), to the PDE, but rather, to solutions in the sense of distributions. That is, if we look at (1.4) in the sense of distributions, and if  $\phi \in \mathcal{D} := C_0^\infty[0, \ell]$  is a test function on the space of distributions, then we need

$$\langle u_t, \phi \rangle - \langle \partial_x^6 u, \phi \rangle - \langle 2\partial_x^4 u, \phi \rangle - \langle (1-r)\partial_x^2 u, \phi \rangle + \langle u \partial_x u, \phi \rangle = 0, \quad (2.1)$$

for almost every  $t$ , where  $\langle \cdot, \cdot \rangle$  denotes the action of a functional on a function. We think of  $u \in \mathcal{D}'$  (the dual space of  $\mathcal{D}$ ) in the sense that for an integrable function  $u$  on  $[0, \ell]$ , its

action as a distribution on  $\phi$  is given by

$$\langle u, \phi \rangle := \int_0^\ell u \phi \, dx.$$

The  $x$ -derivatives of  $u$  should at least exist in some weak sense.

By the definition of distributional derivatives, (2.1) is equivalent to

$$\langle u_t, \phi \rangle + \langle \partial_x^3 u, \partial_x^3 \phi \rangle - 2 \langle \partial_x^2 u, \partial_x^2 \phi \rangle + (1-r) \langle \partial_x u, \partial_x \phi \rangle + \langle u \partial_x u, \phi \rangle = 0.$$

We will be working in Hilbert spaces, and so we will be replacing the action by an inner product, using the Riesz representation theorem [Kre89, p. 188]. Of course we could move even more (or fewer) derivatives to the right hand side in the expression of the above action; however this would have two undesirable consequences:

First (at least for the second term), it would mean that the function  $\phi$  or the function  $u$  would have at least four derivatives on it instead of just three. Second, it would also break a symmetry between the derivatives on both terms which we might want to exploit later on. Thus the minimum number of spatial derivatives that we will need are three. Furthermore, we will also require the solutions to be integrable in time, and their time derivative  $u_t$  should also exist in some distributional sense.

A further property that we want for solutions is that they satisfy suitable boundary conditions; in particular, we want boundary conditions that make the cross terms in the integration by parts vanish, which will allow us to keep only three derivatives when taking the action of (2.1) on a solution itself. One possibility for such boundary conditions are periodic boundary conditions, that is  $u(0, t) = u(\ell, t)$ .

Given periodic boundary conditions, and that  $u$  is square integrable, it is natural to represent  $u$  by its Fourier series:

$$u(x, t) = \sum_k \hat{u}_k(t) e^{ikx}, \tag{2.2}$$

where the  $\hat{u}_k$  are the Fourier coefficients

$$u_k(t) := \frac{1}{\ell} \int_0^\ell u(x, t) e^{ikx} \, dx,$$

and  $k = 2\pi j/\ell$ , with  $j \in \mathbb{Z}$ , is the wave number. In general, we will let the index of sums range over a countable index set (such as  $\mathbb{Z}$ ) unless otherwise specified.

Putting (2.2) into (1.4), we obtain

$$\sum_k \frac{d}{dt} \hat{u}_k e^{ikx} - \sum_k \omega(k) \hat{u}_k e^{ikx} + \sum_l \hat{u}_l e^{ilx} \sum_m im \hat{u}_k e^{imx} = 0,$$

where  $l$  and  $m$  are defined analogously to  $k$ . In the above,  $\omega(k)$  is the linear dispersion relation given by (1.5) of equation (1.4).

For a fixed  $k$  we can take the inner product of the above with  $e^{ikx}$ , which gives

$$\frac{d}{dt}\hat{u}_k - \omega(k)\hat{u}_k + \sum_l i(k-l)\hat{u}_l\hat{u}_{k-l} = 0. \quad (2.3)$$

This gives us an ordinary differential equation (ODE) for each Fourier coefficient  $\hat{u}_k$  of  $u$ . Using equations (2.3), we can thus represent the original PDE (1.4) as an infinite-dimensional dynamical system. We will make use of this idea in the next section when we introduce Galerkin approximations.

Finally, putting together all these observations, namely three  $x$ -derivatives, a single  $t$ -derivative, integrability, periodicity and zero mean, we are ready to define the spaces in which we will look for solutions. We will first state what the spaces are, and then give a precise definition afterwards. We will look for  $u$  in the space

$$L^2(0, T; \dot{H}^3), \quad (2.4)$$

and for the weak time derivative (which we will define at the end of this section)  $u_t$  of  $u$  in the space

$$L^2(0, T; \dot{H}^{-3}), \quad (2.5)$$

where  $T > 0$  is some maximum time,  $\dot{H}^3$  is the Sobolev space whose elements have three (weak) spatial derivatives, zero mean and periodic boundary conditions on  $[0, \ell]$ . The space  $\dot{H}^{-3}$  is the dual space of  $\dot{H}^3$ .

We will now give a more precise definition of these spaces. A function  $u$  is in (2.4) if its weak space derivatives up to order 3 exist, it is periodic and has zero mean, and it is finite under the following norm:

$$\|u\|_{L^2(0, T; \dot{H}^3)} := \left( \int_0^T \int_0^\ell \sum_{n \leq 3} |\partial_x^n u|^2 dx dt \right)^{1/2}.$$

Similarly, a functional  $u$  is an element of the space (2.5), if it is finite under the norm

$$\|u\|_{L^2(0, T; \dot{H}^{-3})} := \left( \int_0^T \|u\|_{\dot{H}^{-3}}^2 dt \right)^{1/2}.$$

Note that  $L^2(0, T; \dot{H}^3) \subset L^2(0, T; \dot{H}^{-3})$ .

Since we are looking for the weak time derivative in the space (2.5), we will now define precisely what this means.

**Definition 2.1.1** (Weak Time Derivative). We say that  $u_t \in L^2(0, T; \dot{H}^{-3})$  is the **weak time derivative** of  $u \in L^2(0, T; \dot{H}^{-3})$  if for every test function  $\phi \in C_c^\infty[0, T]$  and every  $v \in \dot{H}^3$ , we have

$$\left\langle \int_0^T u_t \phi \, dt + \int_0^T u \phi_t \, dt, v \right\rangle = 0. \quad (2.6)$$

Note that this implies that

$$\int_0^T u_t \phi \, dt = - \int_0^T u \phi_t \, dt$$

in the space  $\dot{H}^{-3}$ . Note also that since  $L^2(0, T; \dot{H}^3) \subset L^2(0, T; \dot{H}^{-3})$ , the above definition is also the definition for the weak time derivative of  $u$  if  $u \in L^2(0, T; \dot{H}^3)$  and  $u_t \in L^2(0, T; \dot{H}^{-3})$ .

## 2.2 The Galerkin Approximations

As we saw in the previous section, the Nikolaevskiy equation (1.4) can be represented by an infinite-dimensional dynamical system. Here we will be approximating (1.4) by a sequence of PDEs that can be represented by finite-dimensional dynamical systems. In other words we will first obtain solutions to a sequence of approximate ‘Nikolaevskiy’ equations, and then show that this sequence of solutions converges to a solution of the full PDE (1.4). That is, we approximate our original problem with problems that can be solved more easily and then show that the solutions to those problems converge to a solution of (1.4). The solutions to the approximate equations are known as the **Galerkin approximations** of the solution to equation (1.4).

To begin, note that the eigenfunctions of the Laplacian ( $\partial_x^2$ ) are also eigenfunctions of all the other linear spatial derivative operators in the Nikolaevskiy equation. In particular, they are the Fourier modes of the series (2.2). Moreover, since the partial Fourier sum

$$\sum_{k=-2\pi n/\ell}^{2\pi n/\ell} \hat{u}_k e^{ikx},$$

converges to  $u$  in  $\dot{L}^2$  as  $n \rightarrow \infty$ , we also know that the eigenfunctions of the Laplacian span  $\dot{L}^2$ . In order to get approximate solutions, we will thus project (1.4) onto the subspace spanned by the first  $2n + 1$  basis functions. By a projection we mean

$$P_n u := \sum_{k=-2\pi n/\ell}^{2\pi n/\ell} \hat{u}_k e^{ikx}. \quad (2.7)$$

Note that  $\partial_x(P_n u) = P_n(\partial_x u)$  and that  $u \in \dot{L}^2$  implies that  $\hat{u}_0 = 0$ . The approximate equations are then given by

$$u_t^n - \partial_x^6 u^n - 2\partial_x^4 u^n - (1-r)\partial_x^2 u^n + P_n(u^n \partial_x u^n) = 0 \quad , \quad u^n(0) = P_n u_0, \quad (2.8)$$

where  $u_0$  is the initial condition of  $u$  of the original problem (1.4). Note that in general,  $P_n u \neq u^n$ . The reason why we have to apply an extra projection to the nonlinear term is because multiplying two series like (2.7) will produce terms that lie outside of the projection space due to cross terms in the multiplication. We assume that the initial condition of  $u$  is in  $\dot{L}^2$ . We will then look for solutions given by

$$u^n(x, t) = \sum_{k=-2\pi n/\ell}^{2\pi n/\ell} \hat{u}_k^n(t) e^{ikx}, \quad (2.9)$$

in the subspace defined by  $P_n$ . We hope that  $u^n$  converges to  $u$ , and a major focus of this chapter is to show that this is indeed the case.

With (2.9), equation (2.8) becomes

$$\sum_{k=-2\pi n/\ell}^{2\pi n/\ell} \left( \frac{d}{dt} \hat{u}_k^n - \omega(k) \hat{u}_k^n + \sum_{\substack{j+m=k \\ |j|, |m| \leq 2\pi n/\ell}} (im) \hat{u}_m^n \hat{u}_j^n \right) e^{ikx} = 0, \quad (2.10)$$

where we have collected the  $\hat{u}_m^n$  and the  $\hat{u}_j^n$  in the last term (the nonlinear term), such that  $j$  and  $m$  add up to  $k$ , so that each term in the sum is multiplied by the  $k$ 'th basis function.

Now recall that the basis functions  $e^{ikx}$  are orthogonal, and so if we take the inner product of the whole equation (2.10) with  $e^{ikx}$ , for each  $k$  we get an equation of the form

$$\frac{d}{dt} \hat{u}_k^n - \omega(k) \hat{u}_k^n + \sum_{\substack{j+m=k \\ |j|, |m| \leq 2\pi n/\ell}} (im) \hat{u}_m^n \hat{u}_j^n = 0, \quad (2.11)$$

for a total of  $2n + 1$  equations. This is now a system of coupled ODEs for the Fourier coefficients of  $u^n$ . In fact, due to the nonlinear term, each coefficient depends on all the other coefficients! But the saving grace is that the system is still finite-dimensional: It is really a system of ODEs in  $\mathbb{C}^{2n+1}$ .

In order to show existence of solutions to this system, we turn to some ODE theory. In what follows, we will be working with vectors in  $\mathbb{C}^{2n+1}$  whose components are associated with particular Fourier coefficients  $\hat{u}_k$ . We will thus formally index the components of a vector by the wavenumber  $k$ , even though  $k$  is not an integer. If we now think of (2.11) as

a system of ODEs then we can let the right hand side for the  $\hat{u}_k$  equation be

$$f_k(u^n) := \omega(k)\hat{u}_k^n - \sum_{\substack{j+m=k \\ |j|,|m|\leq 2\pi n/\ell}} (im) \hat{u}_m^n \hat{u}_j^n.$$

Then the actual vector that we are solving for is the vector  $\mathbf{y}(t) \in \mathbb{C}^{2n+1}$ , whose components are the Fourier coefficients  $\hat{u}_k^n(t)$  at some time  $t$ . That is, if  $(\mathbf{y}(t))_k = \hat{u}_k^n(t)$  and  $(\mathbf{F}(\mathbf{y}))_k = f_k(u^n)$ , then we will show that the solution to the system

$$\frac{d}{dt}\mathbf{y}(t) = \mathbf{F}(\mathbf{y})$$

exists.

In order to conclude that a solution must exist, we use Theorem B.1.1 from the Appendix. We simply need to verify that  $\mathbf{F}$  is locally Lipschitz. But  $\mathbf{F}$  is just the sum of a linear term plus many quadratic terms; thus it can easily be verified that  $\mathbf{F}$  is indeed locally Lipschitz. So we may use Theorem B.1.1 to conclude that there exists some time  $\tilde{T}$  such that the system of ODEs has a unique solution on  $[0, \tilde{T}]$ . Since the Fourier modes are smooth, the approximate Galerkin solutions also possess infinitely many spatial derivatives; that is  $u^n \in C^1([0, \tilde{T}]; C^\infty(0, \ell))$ .

So far we only have some undetermined time  $\tilde{T} > 0$  up to which, solutions are guaranteed to exist. However, according to Lemma B.1.2 in the Appendix, the maximal value of  $\tilde{T}$  is only finite if the solutions to (2.11) become unbounded at some finite time. As we will show in the next section, the solutions remain bounded in any finite time interval  $[0, T]$ . Once we have shown those bounds, then we will be able to conclude that these solutions exist and are unique up to any time  $T$ .

## 2.3 Bounds on Galerkin Approximations to the Nikolaevskiy Equation

We will now derive some bounds on the Galerkin solutions (2.9) of the approximate equations (2.8). Since it will turn out that these bounds are uniform (they do not depend on  $n$ ), we will drop the superscript  $n$ , and work with the full Nikolaevskiy equation (1.4) in the following calculations to simplify notation.

**Bounds in  $L^\infty(0, T; \dot{L}^2)$ .**

First note that there are some distinct cases to consider. The analysis is slightly different for  $(1 - r) < 0$  and for  $(1 - r) \geq 0$ , we will therefore treat each case separately. Before we



begin, we will introduce some simplified notation for the nonlinear term in (1.4). Let

$$\mathcal{N}(u) := u \partial_x u. \quad (2.12)$$

Note that by the periodicity of  $u$ , the  $\dot{L}^2$  inner product of the nonlinear term with  $u$  satisfies

$$(\mathcal{N}(u), u) := (u \partial_x u, u) = \int_0^\ell \frac{1}{3} \partial_x (u^3) dx = \frac{1}{3} u^3 \Big|_0^\ell = 0, \quad (2.13)$$

where  $(\cdot, \cdot)$  denotes the inner product on  $\dot{L}^2$ , with induced norm  $\|\cdot\|_{\dot{L}^2}$ . Furthermore, note that since the projection  $P_n$  (2.7) only removes Fourier modes that are orthogonal to the Galerkin solution  $u^n$ , one has that  $(P_n(u^n \partial_x u^n), u^n) = (u^n \partial_x u^n, u^n) = 0$  by the above identity (2.13). Essentially, since  $u^n$  is an element of the subspace defined by  $P_n$  already, taking an inner product of any vector with  $u^n$  acts as a (weighted) projection of that vector into the subspace defined by  $P_n$ .

In what follows we will use the following trilinear form to make treating the nonlinear term in (1.4) more straightforward. Let

$$b(u, v, w) := \int_0^\ell u \partial_x v w dx. \quad (2.14)$$

We start with equation (2.8), and take the  $\dot{L}^2$  inner product of the equation with  $u$ . Then we get

$$\begin{aligned} (u_t, u) - (\partial_x^6 u, u) - 2(\partial_x^4 u, u) - (1-r)(\partial_x^2 u, u) + (u \partial_x u, u) &= 0 \\ \implies \frac{1}{2} \frac{d}{dt} \|u\|_{\dot{L}^2}^2 + \|\partial_x^3 u\|_{\dot{L}^2}^2 - 2\|\partial_x^2 u\|_{\dot{L}^2}^2 + (1-r)\|\partial_x u\|_{\dot{L}^2}^2 &= 0, \end{aligned} \quad (2.15)$$

where we have used (2.13), to remove the nonlinear term. In what follows we will repeatedly use the generalized Young's inequalities from Appendix B.3 and Poincaré's inequality (B.12). It should be noted that in this section, when we use Poincaré's inequality, we introduce a bounding constant  $C = (\ell/(2\pi))^2$ . Later on, when we derive bounds for which we are interested in the scaling with  $\ell$ , we will try to avoid using Poincaré's inequality whenever possible, as this introduces an extra factor of  $\ell$ . In this section though, we are only interested in the finiteness of the bound, since the bound we will derive is still  $t$ -dependent, and so the  $\ell$  scaling is unimportant. We will also frequently introduce arbitrary, positive constants that come from the Young's-type inequalities from Appendix B.3. We will usually call these constants  $\alpha, \beta, \gamma, \delta, \epsilon, \eta$  and  $\sigma$ .

**Case 1:**  $(1-r) \geq 0$

From (2.15) we begin with

$$\begin{aligned}
\frac{1}{2} \frac{d}{dt} \|u\|_{\dot{L}^2}^2 + \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 + (1-r) \left\| \partial_x u \right\|_{\dot{L}^2}^2 &= 2 \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 \\
&\leq 2 \left\| \partial_x u \right\|_{\dot{L}^2} \left\| \partial_x^3 u \right\|_{\dot{L}^2} \\
&\leq \frac{\left\| \partial_x u \right\|_{\dot{L}^2}^2}{\beta} + \beta \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 \\
&\leq \frac{\|u\|_{\dot{L}^2}^2}{2\beta\gamma} + \frac{\gamma}{2\beta} \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 + \beta \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2,
\end{aligned}$$

for any  $\beta, \gamma > 0$ . Now since  $(1-r) \geq 0$  we can drop the  $(1-r)$  term, and we can bound the  $\partial_x^2 u$  term from above by the third derivative term using Poincaré's inequality (B.12). So we get

$$\frac{1}{2} \frac{d}{dt} \|u\|_{\dot{L}^2}^2 + \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 \leq \frac{\|u\|_{\dot{L}^2}^2}{2\beta\gamma} + \left( \frac{C\gamma}{2\beta} + \beta \right) \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2.$$

Choosing  $\beta = 1/2$  and  $\gamma = 1/(2C)$ , we get

$$\frac{C\gamma}{2\beta} + \beta = 1,$$

so

$$\frac{d}{dt} \|u(t)\|_{\dot{L}^2}^2 \leq 4C \|u(t)\|_{\dot{L}^2}^2.$$

Hence by Gronwall's inequality

$$\|u(t)\|_{\dot{L}^2}^2 \leq \|u_0\|_{\dot{L}^2}^2 e^{4Ct},$$

where  $u(0) = u_0$ . This shows that  $\|u\|_{\dot{L}^2}$  remains bounded in  $t$  for any finite  $T$ .

**Case 2:**  $(1-r) < 0$ .

We begin with the same Ansatz (2.15) as in the first case, but since  $(1-r) < 0$ , we can move the second-order term over to the right hand side. The idea for deriving the bound is the same; however, since there are fewer damping terms when  $r > 1$ , we will use inequality (B.7) with  $n = 1$  and  $\epsilon = \beta$  and inequality (B.8) with  $n = 0$  and  $\epsilon = \alpha$  to get

$$\begin{aligned}
\frac{1}{2} \frac{d}{dt} \|u\|_{\dot{L}^2}^2 + \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 &= 2 \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 - (1-r) \left\| \partial_x u \right\|_{\dot{L}^2}^2 \\
&\leq \left( \frac{2}{\alpha} - \frac{(1-r)\beta^2}{4} \right) \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 + \left( \frac{\alpha^2}{2} - \frac{(1-r)}{\beta} \right) \|u\|_{\dot{L}^2}^2.
\end{aligned}$$

If we now pick  $\alpha = 4$  and  $\beta = \sqrt{-2/(1-r)}$ , then we can drop the third derivative term and get

$$\frac{1}{2} \frac{d}{dt} \|u\|_{\dot{L}^2}^2 \leq \left( \frac{\alpha^2}{2} - \frac{(1-r)}{\beta} \right) \|u\|_{\dot{L}^2}^2 = \left( 8 + \frac{(r-1)^{3/2}}{\sqrt{2}} \right) \|u\|_{\dot{L}^2}^2.$$

If we now once again apply Gronwall's inequality, we get

$$\|u(t)\|_{\dot{L}^2}^2 \leq \|u_0\|_{\dot{L}^2}^2 e^{\left( 8 + \frac{(r-1)^{3/2}}{\sqrt{2}} \right) t}.$$

So  $\|u(t)\|_{\dot{L}^2}$  is also bounded for any finite  $T$  for  $(1-r) < 0$ . This gives us the bound in  $L^\infty(0, T; \dot{L}^2)$  in the second case, and thus we now have uniform boundedness in time for all cases that we are considering.

### **Bounds in $L^2(0, T; \dot{H}^3)$ .**

We will now turn to the problem of getting some bounds on the spatial derivatives of  $u$  in finite time.

As we have just seen,  $u$  is in fact in  $L^\infty(0, T; \dot{L}^2)$ , so now we also want to get a bound in  $L^2(0, T; \dot{H}^3)$ .

#### **Case 1: $(1-r) \geq 0$**

We begin with (2.15) and integrate over  $t$  from 0 to  $T$ . After immediately dropping the  $\|u(T)\|_{\dot{L}^2}^2$  term and the  $(1-r) \|\partial_x u\|_{\dot{L}^2}^2$  term, we get

$$\int_0^T \left( \|\partial_x^3 u\|_{\dot{L}^2}^2 - 2 \|\partial_x^2 u\|_{\dot{L}^2}^2 \right) dt \leq \frac{1}{2} \|u_0\|_{\dot{L}^2}^2.$$

Now if we move the second term to the right hand side and use the inequalities (B.6) and (B.4) twice we have

$$\begin{aligned} \int_0^T \|\partial_x^3 u\|_{\dot{L}^2}^2 dt &\leq \int_0^T \left( \frac{1}{\beta} \|\partial_x u\|_{\dot{L}^2}^2 + \beta \|\partial_x^3 u\|_{\dot{L}^2}^2 \right) dt + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2 \\ &\leq \int_0^T \left( \frac{1}{2\beta\gamma} \|u\|_{\dot{L}^2}^2 + \frac{\gamma}{2\beta} \|\partial_x^2 u\|_{\dot{L}^2}^2 + \beta \|\partial_x^3 u\|_{\dot{L}^2}^2 \right) dt + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2, \end{aligned}$$

which after applying inequality (B.12) becomes

$$\int_0^T \|\partial_x^3 u\|_{\dot{L}^2}^2 dt \leq \int_0^T \left( \frac{1}{2\beta\gamma} \|u\|_{\dot{L}^2}^2 + \frac{C\gamma}{2\beta} \|\partial_x^3 u\|_{\dot{L}^2}^2 + \beta \|\partial_x^3 u\|_{\dot{L}^2}^2 \right) dt + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2.$$

Now if we pick  $\beta = \frac{1}{4}$  and  $\gamma = \frac{1}{8C}$  to get

$$\frac{C\gamma}{2\beta} = \frac{1}{4},$$

then

$$\begin{aligned} \int_0^T \|\partial_x^3 u\|_{\dot{L}^2}^2 dt &\leq \int_0^T \left( 16C \|u\|_{L^2}^2 + \frac{1}{2} \|\partial_x^3 u\|_{\dot{L}^2}^2 \right) dt + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2 \\ \implies \frac{1}{2} \int_0^T \|\partial_x^3 u\|_{\dot{L}^2}^2 dt &\leq \int_0^T (16C \|u\|_{L^2}^2) dt + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2. \end{aligned}$$

Now, repeatedly using Poincaré's inequality (B.12), one can bound the  $L^2(0, T; \dot{H}^3)$  norm of  $u$  by the  $L^2(0, T; \dot{L}^2)$  norm of  $\partial_x^3 u$  as follows:

$$\|u\|_{L^2(0, T; \dot{H}^3)}^2 \leq (C^3 + C^2 + C + 1) \|\partial_x^3 u\|_{L^2(0, T; \dot{L}^2)}^2.$$

Thus, letting  $\tilde{C} := \frac{1}{C^3 + C^2 + C + 1}$ , we get

$$\frac{1}{2} \tilde{C} \|u\|_{L^2(0, T; \dot{H}^3)}^2 \leq \int_0^T \frac{1}{2} \|\partial_x^3 u\|_{\dot{L}^2}^2 dt \leq 16CT \|u\|_{L^\infty(0, T; \dot{L}^2)}^2 + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2 < \infty.$$

This shows that for  $(1 - r) \geq 0$ , we have  $u \in L^2(0, T; \dot{H}^3)$ .

**Case 2:**  $(1 - r) < 0$

We start again with equation (2.15), and integrate it over time from 0 to  $T$ . This gives

$$\frac{1}{2} \|u(T)\|_{\dot{L}^2}^2 + \int_0^T \left( \|\partial_x^3 u\|_{\dot{L}^2}^2 - 2 \|\partial_x^2 u\|_{\dot{L}^2}^2 + (1 - r) \|\partial_x u\|_{\dot{L}^2}^2 \right) dt = \frac{1}{2} \|u_0\|_{\dot{L}^2}^2. \quad (2.16)$$

Dropping the first term, moving the lower derivatives to the other side and repeatedly using inequalities (B.6) and (B.4) we get

$$\begin{aligned}
& \int_0^T \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 dt \int_0^T \left( 2 \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 - (1-r) \left\| \partial_x u \right\|_{\dot{L}^2}^2 \right) dt + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2 \\
& \leq \int_0^T \left( 2 \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 - \frac{(1-r)}{2\beta} \|u\|_{\dot{L}^2}^2 - \frac{(1-r)\beta}{2} \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 \right) dt + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2 \\
& \leq \int_0^T \left( \frac{1}{\gamma} \left\| \partial_x u \right\|_{\dot{L}^2}^2 + \gamma \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 - \frac{(1-r)}{2\beta} \|u\|_{\dot{L}^2}^2 \right. \\
& \quad \left. - \frac{(1-r)\beta}{2} \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 \right) dt + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2 \\
& \leq \int_0^T \left( \frac{\delta}{2\gamma} \|u\|_{\dot{L}^2}^2 + \frac{1}{2\gamma\delta} \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 + \gamma \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 \right. \\
& \quad \left. - \frac{(1-r)}{2\beta} \|u\|_{\dot{L}^2}^2 - \frac{(1-r)\beta}{2} \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 \right) dt + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2,
\end{aligned}$$

then using Poincaré's inequality (B.12)

$$\begin{aligned}
\int_0^T \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 dt &= \int_0^T \left( \left( \frac{\delta}{2\gamma} - \frac{(1-r)}{2\beta} \right) \|u\|_{\dot{L}^2}^2 + \left( \frac{C}{2\gamma\delta} - \frac{C(1-r)\beta}{2} \right) \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 \right. \\
& \quad \left. + \gamma \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 \right) dt + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2.
\end{aligned}$$

Now if we pick

$$\begin{aligned}
\gamma &= \frac{1}{4}, \\
\delta &= 16C, \\
\beta &= \frac{-1}{4C(1-r)},
\end{aligned}$$

then we have

$$\int_0^T \left( \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 dt \leq \int_0^T (32C + 2(1-r)^2C) \|u\|_{\dot{L}^2}^2 + \frac{1}{2} \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 \right) dt + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2.$$

So we get

$$\frac{1}{2} \tilde{C} \|u\|_{L^2(0,T;\dot{H}^3)}^2 \leq \int_0^T \frac{1}{2} \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 dt \leq T(32C + 2(1-r)^2C) \|u\|_{L^\infty(0,T;\dot{L}^2)}^2 + \frac{1}{2} \|u_0\|_{\dot{L}^2}^2 < \infty,$$

where we have bounded the first term below using Poincaré's inequality. Thus for  $(1-r) < 0$ , we also have  $u \in L^2(0, T; \dot{H}^3)$ .

We now know that the  $\dot{L}^2$  norm of  $u$  and of its first three spatial derivatives is bounded in finite time in the mean square sense. Since we also have a time derivative, we will now turn to bounding  $u_t$  in the  $L^2(0, T; \dot{H}^{-1})$  for finite time.

**Bound on  $u_t$  in  $L^2(0, T; \dot{H}^{-3})$ .**

In order to get a bound on  $u_t$  in  $L^2(0, T; \dot{H}^{-3})$ , we simply take the  $L^2(0, T; \dot{H}^{-3})$  norm of (1.4), apply the triangle inequality and use the fact that  $u \in L^2(0, T; \dot{H}^3)$  implies that the first six spatial derivatives are in  $L^2(0, T; \dot{H}^{-3})$ . The nonlinear term (2.12) is also in this space, since for any  $v \in L^2(0, T; \dot{H}^3)$ , we have

$$\begin{aligned} \int_0^T \langle \mathcal{N}(u), v \rangle dt &= \int_0^T b(u, u, v) dt \\ &= \int_0^T \int_0^\ell uu_x v dx dt, \end{aligned}$$

and then using estimate (B.6.3) gives

$$\begin{aligned} \int_0^T b(u, u, v) dt &\leq \ell^{1/2} \int_0^T \|u\|_{\dot{L}^2} \|\partial_x^2 u\|_{\dot{L}^2} \|v\|_{\dot{L}^2} dt \\ &\leq \ell^{1/2} \|u\|_{L^\infty(0, T; \dot{L}^2)} \left( \|u\|_{L^2(0, T; \dot{H}^2)} \|v\|_{L^2(0, T; \dot{L}^2)} \right) < \infty, \end{aligned}$$

by the bounds we derived on  $u$  in these spaces. Now let

$$\mathcal{L} := -\partial_x^6 - 2\partial_x^4 - (1-r)\partial_x^2 \tag{2.17}$$

be the linear operator in equation (1.4). Then since all terms except the  $u_t$  term can be shown to be in  $L^2(0, T; \dot{H}^3)$ , the triangle inequality implies that

$$\|u_t\|_{L^2(0, T; \dot{H}^{-3})} = \|\mathcal{L}u + \mathcal{N}(u)\|_{L^2(0, T; \dot{H}^{-3})} \leq \|\mathcal{L}u\|_{L^2(0, T; \dot{H}^{-3})} + \|\mathcal{N}(u)\|_{L^2(0, T; \dot{H}^{-3})} < \infty.$$

Thus  $u_t \in L^2(0, T; \dot{H}^{-3})$  also.

As stated before, even though we did not explicitly use the Galerkin solutions  $u^n$  for the bounds, we originally required that  $u_0^n := P_n u_0$ , so we can bound all of their bounds from above uniformly using the norm of the initial condition to the full problem. None of the bounds we have derived above depend on  $n$ .

## 2.4 Convergence of the Galerkin Approximations

In the previous section we saw that the sequence of Galerkin approximations  $u^n$  is uniformly bounded in  $L^\infty(0, T; \dot{L}^2)$ , as well as in  $L^2(0, T; \dot{H}^3)$ . We now want to show that these func-

tions converge to a solution to the Nikolaevskiy equation, and that their time derivatives converge to the time derivative of the solution.

We begin by recalling the Alaoglu Compactness theorem B.2.1. Since  $L^2(0, T; \dot{H}^3)$  is a Hilbert space, it is reflexive, and since the sequence  $u^n$  is uniformly bounded in it, there exists some element  $u \in L^2(0, T; \dot{H}^3)$ , such that a subsequence  $u^{n^j}$  converges weakly to  $u$  in  $L^2(0, T; \dot{H}^3)$ . In what follows we will relabel the subsequence  $u^{n^j}$  as  $u^n$ .

We begin by showing that the Galerkin approximations  $u^n$  under the linear operator  $\mathcal{L}$  converge to  $u$  under the linear operator, where  $u$  is the weak limit of a subsequence of  $u^n$ . We will then show that the time derivatives of  $u^n$  converge to the weak time derivative of  $u$ . In order to show that the  $u^n$  converge under the nonlinear operator (2.12), we will first have to show that  $u^n$  converges to  $u$  strongly. We can then prove that  $u^n$  under the nonlinear term converges to  $u$  under the nonlinear term.

### Convergence of the linear spatial terms:

Recall the form of the linear spatial operator in (1.4)

$$\mathcal{L} := -\partial_x^6 - 2\partial_x^4 - (1-r)\partial_x^2. \quad (2.18)$$

We want to show that  $\mathcal{L}u^n \rightharpoonup \mathcal{L}u$  as  $n \rightarrow \infty$  in  $L^2(0, T; \dot{H}^{-3})$  in the weak sense; that is, we want to show that  $\mathcal{L}u^n \rightharpoonup \mathcal{L}u$ . We first note that (2.17) is self-adjoint because of the boundary conditions and from integrating by parts. Thus for  $v \in L^2(0, T; \dot{H}^3)$ ,

$$\begin{aligned} \int_0^T \langle \mathcal{L}u^n, v \rangle dt &= \int_0^T \int_0^\ell (\mathcal{L}u^n)v dx dt \\ &= \int_0^T \int_0^\ell (\mathcal{L}v)u^n dx dt \\ &= \int_0^T \langle \mathcal{L}v, u^n \rangle dt \rightarrow \int_0^T \langle \mathcal{L}v, u \rangle dt \\ &= \int_0^T \langle \mathcal{L}u, v \rangle dt. \end{aligned}$$

Thus  $\mathcal{L}u^n \rightharpoonup \mathcal{L}u$  in  $L^2(0, T; \dot{H}^{-3})$ . Here we have used the fact that for any  $v \in L^2(0, T; \dot{H}^3)$ , we have that  $\mathcal{L}v \in L^2(0, T; \dot{H}^{-3})$  to get the above convergence.

### Convergence of the time derivative terms:

We will now show that the partial time derivatives  $\partial_t u^n$  converge to the weak time derivative  $u_t$  of  $u$  in the sense of (2.6), in the space  $L^2(0, T; \dot{H}^{-3})$ . Since the sequence  $\partial_t u^n$  is uniformly bounded in  $L^2(0, T; \dot{H}^{-3})$ , and  $L^2(0, T; \dot{H}^{-3})$  is reflexive, we can take as corollary

to the Alaoglu compactness theorem B.2.1 that there exist a subsequence  $\partial_t u^{n^j}$  of  $\partial_t u^n$  and an element  $\dot{u} \in L^2(0, T; \dot{H}^{-3})$ , such that  $\partial_t u^{n^j}$  converges weakly to  $\dot{u}$  in  $L^2(0, T; \dot{H}^{-3})$ .

We now have to show that  $\dot{u}$  is indeed the weak time derivative of  $u$ . Thus let  $\phi \in C_c^\infty(0, T)$ , and  $v \in \dot{H}^3$ . Then since the  $\partial_t u^n$  are the classical time derivatives of  $u^n$ , we have by integration by parts that

$$LHS := \int_0^T (\partial_t u^n) \phi \, dt = - \int_0^T u^n \partial_t \phi \, dt := -RHS. \quad (2.19)$$

Now if we take the action of each term on the element  $v$ , then beginning with the left hand side we get

$$\begin{aligned} \langle LHS, v \rangle &= \left\langle \int_0^T \partial_t u^n \phi \, dt, v \right\rangle = \int_0^T \langle \partial_t u^n, v \phi \rangle \, dt \\ &\rightarrow \int_0^T \langle \dot{u}, v \phi \rangle \, dt = \left\langle \int_0^T \dot{u} \phi \, dt, v \right\rangle, \end{aligned}$$

where we have treated  $v\phi$  as an element of  $L^2(0, T; \dot{H}^3)$  since  $\phi$  only depends on  $t$  and  $v$  only depends on  $x$ . Next, the action on the right hand side of (2.19) becomes

$$\begin{aligned} \langle RHS, v \rangle &= \left\langle \int_0^T u^n \partial_t \phi \, dt, v \right\rangle = \int_0^T \langle u^n, v \partial_t \phi \rangle \, dt \\ &\rightarrow \int_0^T \langle u, v \partial_t \phi \rangle \, dt = \left\langle \int_0^T u \partial_t \phi \, dt, v \right\rangle. \end{aligned}$$

Therefore

$$\left\langle \int_0^T \dot{u} \phi \, dt + \int_0^T u \phi_t \, dt, v \right\rangle = \langle LHS + RHS, v \rangle = 0,$$

which implies that  $\dot{u}$ , is the weak time derivative of  $u$ , that is,  $\dot{u} = u_t$ .

### Strong convergence of the sequence $u^n$ .

Recall that by the Rellich-Kondrachov Compactness Theorem [Rob01, p. 143], we have that  $\dot{H}^3$  is compactly embedded in  $\dot{H}^2$  (we write  $\dot{H}^3 \subset\subset \dot{H}^2$ ). That is, there exists a constant  $C$  such that

$$\|u\|_{\dot{H}^2} \leq C \|u\|_{\dot{H}^3},$$

for every  $u \in \dot{H}^3$ . Furthermore, every bounded sequence in  $\dot{H}^3$  contain a subsequence that converges in  $\dot{H}^2$ .



Also, since we can use any element in  $\dot{H}^2$  to define a bounded linear functional on  $\dot{H}^3$ , we have that

$$\dot{H}^2 \subset\subset \dot{H}^1 \subset \dot{H}^{-2}.$$

Thus since  $u^n$  is uniformly bounded in  $L^2(0, T; \dot{H}^2)$ , and  $\partial_t u^n$  is uniformly bounded in  $L^2(0, T; \dot{H}^{-3})$ , we get by Theorem B.2.4 in the Appendix that the weakly convergent subsequence we obtained for  $u^n$  also converges strongly in  $L^2(0, T; \dot{H}^2)$ . The proof of Theorem B.2.4 is somewhat lengthy, and relies on Ehrling's lemma B.3 and on its corollary B.2.3. To make this section more clear, we did not include it here, but rather proved Theorem B.2.4 in the Appendix.

Using the result from Theorem B.2.4 we get

$$\|u^n - u\|_{L^2(0, T; \dot{H}^2)} \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (2.20)$$

It follows from  $\dot{H}^2 \subset \dot{H}^1 \subset \dot{L}^2$  that we also have  $u^n \rightarrow u$  (strongly) in  $L^2(0, T; \dot{H}^1)$  and  $L^2(0, T; \dot{L}^2)$ .

### Convergence of the nonlinear term:

Before we begin, we will introduce some new notation for the nonlinear term in the Nikolaevskiy equation. We will use this notation again, throughout this text, whenever the two functions in the non-linear term may not be equal. Let

$$B(u, v) := u \partial_x v. \quad (2.21)$$

We want to show that  $P_n B(u^n, u^n)$  in (2.8) converges to  $B(u, u)$  in the sense of distributions in  $L^2(0, T; \dot{H}^{-3})$ . We will apply this to a test function  $\phi \in \mathcal{D}([0, T] \times [0, \ell])$ , and use the definition of  $\langle B(u^n, u^n), \phi \rangle$  to show the result, but first we will get some more bounds on  $u^n$ . We want to use the uniform bounds on  $u^n$  in  $L^\infty(0, T; \dot{L}^2)$  and in  $L^2(0, T; \dot{H}^3)$ , to derive uniform bounds on  $u^n$  in  $L^2(0, T; \dot{L}^\infty)$ , and hence in  $L^4(0, T; \dot{L}^4)$ .

Recall that the norm of  $u^n$  in  $L^2(0, T; \dot{L}^\infty)$  is just

$$\|u^n\|_{L^2(0, T; \dot{L}^\infty)}^2 := \int_0^T \|u^n(t)\|_{\dot{L}^\infty}^2 dt.$$

Furthermore, recall the bound

$$\|u^n(t)\|_{\dot{L}^\infty} \leq \ell^{1/2} \|\partial_x u^n(t)\|_{\dot{L}^2}.$$

Using this we get

$$\begin{aligned}
\|u^n\|_{L^2(0,T;\dot{L}^\infty)}^2 &= \int_0^T \|u^n(t)\|_{\dot{L}^\infty}^2 dt \\
&\leq \ell \int_0^T \|\partial_x u^n(t)\|_{\dot{L}^2}^2 dt \leq \ell \|u^n\|_{L^2(0,T;\dot{H}^1)}^2 \\
&\leq \ell \|u^n\|_{L^2(0,T;\dot{H}^3)}^2.
\end{aligned}$$

By the uniform boundedness of  $u^n$  in  $L^2(0, T; \dot{H}^3)$  this shows that  $u^n$  is uniformly bounded in  $L^2(0, T; \dot{L}^\infty)$  and we will now use this to get the bound in  $L^4(0, T; \dot{L}^4)$ : We calculate

$$\|u^n\|_{L^4(0,T;\dot{L}^4)}^4 = \int_0^T \int_0^\ell (u^n)^4 dx dt \leq \int_0^T \|(u^n)^2\|_{\dot{L}^\infty} \int_0^\ell (u^n)^2 dx dt;$$

but in the previous section we showed that  $u^n \in L^\infty(0, T; \dot{L}^2)$ , so we can take the  $\int_0^\ell (u^n)^2 dx$  term out of the time integral and bound it by the uniform bound of  $u^n$  in  $L^\infty(0, T; \dot{L}^2)$  to get

$$\begin{aligned}
\|u^n\|_{L^4(0,T;\dot{L}^4)}^4 &\leq \sup_{0 \leq t \leq T} \left( \int_0^\ell (u^n)^2 dx \right) \int_0^T \|(u^n)^2\|_{\dot{L}^\infty} dt \\
&= \|u^n\|_{L^\infty(0,T;\dot{L}^2)} \|u^n\|_{L^2(0,T;\dot{L}^\infty)} \leq M,
\end{aligned}$$

for some new  $M$ .

Using all these bounds, we will now show the convergence of the nonlinear term in (1.4): Let  $\phi \in \mathcal{D}([0, T] \times [0, \ell])$  (whose elements are also assumed to be periodic), then

$$\begin{aligned}
\int_0^T \langle P_n B(u^n, u^n), \phi \rangle - \langle B(u, u), \phi \rangle dt &= \int_0^T \langle B(u^n, u^n), P_n \phi \rangle - \langle B(u, u), \phi \rangle dt \\
&= \int_0^T \int_0^\ell u^n \partial_x u^n P_n \phi - u \partial_x u \phi dx dt.
\end{aligned}$$

Now if we use the reverse chain rule to get  $v \partial_x v = \frac{1}{2} \partial_x (v^2)$  and integrate by parts we have

$$\begin{aligned}
2 \int_0^T \int_0^\ell u^n \partial_x u^n P_n \phi - u \partial_x u \phi dx dt &= \int_0^T \int_0^\ell u^2 \partial_x \phi - (u^n)^2 P_n \partial_x \phi dx dt \\
&= \int_0^T \int_0^\ell \underbrace{u^2 \partial_x \phi - u^2 P_n \partial_x \phi}_{\textcircled{1}} + \underbrace{u^2 P_n \partial_x \phi - (u^n)^2 P_n \partial_x \phi}_{\textcircled{2}} dx dt.
\end{aligned}$$

We will look at each of the terms ① and ② separately. First,

$$\begin{aligned}
\textcircled{1} &= \int_0^T \int_0^\ell u^2 (\partial_x \phi - P_n \partial_x \phi) \, dx \, dt \\
&\leq \left( \int_0^T \int_0^\ell u^4 \, dx \, dt \right)^{1/2} \left( \int_0^T \int_0^\ell (\partial_x \phi - P_n \partial_x \phi)^2 \, dx \, dt \right)^{1/2} \\
&\leq M^{1/2} \|\partial_x \phi - P_n \partial_x \phi\|_{L^2(0,T;\dot{L}^2)} \rightarrow 0
\end{aligned}$$

as  $n \rightarrow \infty$ , since projections converge strongly for any test function  $\phi$ . Next we look at the second term and define  $\|\partial_x \phi\|_\infty := \|\partial_x \phi\|_{L^\infty(0,T;\dot{L}^\infty)}$ . Note that  $\|P_n \partial_x \phi\|_\infty = \|\partial_x P_n \phi\|_\infty \leq \ell^{1/2} \|\partial_x^2 P_n \phi\|_{\dot{L}^2} \leq \ell^{1/2} \|\partial_x^2 \phi\|_{\dot{L}^2}$ . Thus we have

$$\begin{aligned}
\textcircled{2} &= \int_0^T \int_0^\ell (u^2 - (u^n)^2) P_n \partial_x \phi \, dx \, dt \leq \ell^{1/2} \|\partial_x^2 \phi\|_{\dot{L}^2} \int_0^T \int_0^\ell (u - u^n)(u + u^n) \, dx \, dt \\
&\leq \ell^{1/2} \|\partial_x^2 \phi\|_{\dot{L}^2} \left( \int_0^T \int_0^\ell (u - u^n)^2 \, dx \, dt \right)^{1/2} \left( \int_0^T \int_0^\ell (u + u^n)^2 \, dx \, dt \right)^{1/2} \\
&\leq \ell^{1/2} \|\partial_x^2 \phi\|_{\dot{L}^2} \tilde{M} \int_0^T \int_0^\ell (u - u^n)^2 \, dx \, dt = \ell^{1/2} \|\partial_x^2 \phi\|_{\dot{L}^2} \tilde{M} \|u - u^n\|_{L^2(0,T;\dot{L}^2)}^2 \rightarrow 0,
\end{aligned}$$

as  $n \rightarrow \infty$ , by the strong convergence of  $u^n$  in  $L^2(0,T;\dot{L}^2)$ , where  $\tilde{M}$  is twice the uniform bound on  $u$  in  $L^2(0,T;\dot{L}^2)$ . We thus have that

$$\int_0^T \langle P_n B(u^n, u^n), \phi \rangle - \langle B(u, u), \phi \rangle \, dt = \textcircled{1} + \textcircled{2} \rightarrow 0$$

as  $n \rightarrow \infty$ , giving the convergence of the nonlinear term in equation (1.4) in the sense of distributions.

We now know that  $\partial_t u^n$ ,  $\mathcal{L}u^n$  and  $\mathcal{N}(u^n)$  converge, at least in the sense of distributions. Putting all of this together, we get that the Galerkin approximations do converge to a distributional solution of the Nikolaevskiy equation (1.4). In later sections, we will see that these solutions do exist classically as well, and are continuous in time, and even analytic in space! In fact, Theorem B.5.4 shows that  $u \in C([0,T];\dot{L}^2)$ , and therefore we can also conclude that  $\lim_{t \rightarrow 0} u(t) = u_0$  with convergence in  $\dot{L}^2$ .

## 2.5 Uniqueness

Now that we know that weak solutions to the Nikolaevskiy equation (1.4) exist, we will show that they are unique. We begin with the case  $(1-r) \geq 0$ . Thus, suppose that we have two solutions  $u$  and  $v$ , such that  $u(0) = u_0$  and  $v(0) = v_0$ , and they both satisfy the equation (1.4). Define  $w := u - v$ . Then subtracting the equations that  $u$  and  $v$  satisfy, we

get

$$\begin{aligned}
\frac{dw}{dt} - \partial_x^6 w - (1-r)\partial_x^2 w &= 2\partial_x^4 w - B(u, u) + B(v, v) \\
&= 2\partial_x^4 w - B(u, u) + B(u, v) - B(u, v) + B(v, v) \\
&= 2\partial_x^4 w - B(u, w) - B(w, v).
\end{aligned}$$

We now take the inner product of the above with  $w$  and integrate by parts to get

$$\begin{aligned}
\frac{1}{2} \frac{d}{dt} \|w\|_{\mathbb{L}^2}^2 + \|\partial_x^3 w\|_{\mathbb{L}^2}^2 + (1-r)\|\partial_x w\|_{\mathbb{L}^2}^2 &= 2 \|\partial_x^2 w\|_{\mathbb{L}^2}^2 - b(u, w, w) - b(w, v, w) \\
&\leq 2 \|\partial_x^2 w\|_{\mathbb{L}^2}^2 + C \|u\|_{\mathbb{L}^2} \|\partial_x^2 w\|_{\mathbb{L}^2} \|w\|_{\mathbb{L}^2} + \frac{C}{2} \|v\|_{\mathbb{L}^2} \|\partial_x^2 w\|_{\mathbb{L}^2} \|w\|_{\mathbb{L}^2},
\end{aligned}$$

where  $C := \ell^{1/2}$  and we have used Corollary (B.6.2) on the nonlinear term. Using Young's inequality and inequality (B.6) we have

$$\begin{aligned}
\frac{1}{2} \frac{d}{dt} \|w\|_{\mathbb{L}^2}^2 + \|\partial_x^3 w\|_{\mathbb{L}^2}^2 + (1-r)\|\partial_x w\|_{\mathbb{L}^2}^2 &\leq 2 \|\partial_x^2 w\|_{\mathbb{L}^2}^2 + \left( \frac{C}{2} \|u\|_{\mathbb{L}^2}^2 + \frac{C}{4} \|v\|_{\mathbb{L}^2}^2 \right) \|w\|_{\mathbb{L}^2}^2 \\
&\quad + \left( \frac{C}{2} + \frac{C}{4} \right) \|\partial_x^2 w\|_{\mathbb{L}^2}^2.
\end{aligned}$$

Redefining  $C := \ell^{1/2}/2$  and letting  $C_0 = 2 + \frac{3\ell^{1/2}}{4}$ , and repeatedly using Young's inequality and equality (B.6) gives

$$\begin{aligned}
\frac{1}{2} \frac{d}{dt} \|w\|_{\mathbb{L}^2}^2 + \|\partial_x^3 w\|_{\mathbb{L}^2}^2 + (1-r)\|\partial_x w\|_{\mathbb{L}^2}^2 &\leq C_0 \|\partial_x^2 w\|_{\mathbb{L}^2}^2 + \frac{C}{2} (2\|u\|_{\mathbb{L}^2}^2 + \|v\|_{\mathbb{L}^2}^2) \|w\|_{\mathbb{L}^2}^2 \\
&\leq \frac{\beta C_0}{2} \|\partial_x w\|_{\mathbb{L}^2}^2 + \frac{C_0}{2\beta} \|\partial_x^3 w\|_{\mathbb{L}^2}^2 + \frac{C}{4} (2\|u\|_{\mathbb{L}^2}^2 + \|v\|_{\mathbb{L}^2}^2) \|w\|_{\mathbb{L}^2}^2 \\
&\leq \frac{\beta\gamma C_0}{4} \|w\|_{\mathbb{L}^2}^2 + \frac{\beta C_0}{4\gamma} \|\partial_x^2 w\|_{\mathbb{L}^2}^2 + \frac{C_0}{2\beta} \|\partial_x^3 w\|_{\mathbb{L}^2}^2 + \frac{C}{2} (2\|u\|_{\mathbb{L}^2}^2 + \|v\|_{\mathbb{L}^2}^2) \|w\|_{\mathbb{L}^2}^2 \\
&\leq \frac{\beta\gamma C_0}{4} \|w\|_{\mathbb{L}^2}^2 + \frac{\beta C_0}{\gamma} \|w\|_{\mathbb{L}^2}^2 + \frac{\beta C_0}{16\gamma} \|\partial_x^3 w\|_{\mathbb{L}^2}^2 \\
&\quad + \frac{C_0}{2\beta} \|\partial_x^3 w\|_{\mathbb{L}^2}^2 + \frac{C}{2} (2\|u\|_{\mathbb{L}^2}^2 + \|v\|_{\mathbb{L}^2}^2) \|w\|_{\mathbb{L}^2}^2 \\
&\leq \left( \frac{\beta C_0}{16\gamma} + \frac{C_0}{2\beta} \right) \|\partial_x^3 w\|_{\mathbb{L}^2}^2 + \left( \frac{\beta\gamma C_0}{4} + \frac{\beta C_0}{\gamma} + \frac{C}{2} (2\|u\|_{\mathbb{L}^2}^2 + \|v\|_{\mathbb{L}^2}^2) \right) \|w\|_{\mathbb{L}^2}^2.
\end{aligned} \tag{2.22}$$

Since we assumed that  $(1-r) \geq 0$  we can drop it on the left. Then if for instance, we make the choices

$$\begin{aligned}
\beta &= 2C_0 \\
\gamma &= C_0^2,
\end{aligned}$$

then the coefficient in front of  $\|\partial_x^3 w\|_{\dot{L}^2}^2$  on the right hand side is less than 1, so for two universal constants  $C_1 > 0$  and  $C_2 > 0$ , we have

$$\frac{1}{2} \frac{d}{dt} \|w\|_{\dot{L}^2}^2 + C_1 \|\partial_x^3 w\|_{\dot{L}^2}^2 \leq \frac{C}{2} (C_2 + 2\|u\|_{\dot{L}^2}^2 + \|v\|_{\dot{L}^2}^2) \|w\|_{\dot{L}^2}^2. \quad (2.23)$$

This will be useful later on when we bound the attractor dimension, but for now let us drop the spatial derivative term on the left to get

$$\frac{d}{dt} \|w\|_{\dot{L}^2}^2 \leq C(C_2 + 2\|u\|_{\dot{L}^2}^2 + \|v\|_{\dot{L}^2}^2) \|w\|_{\dot{L}^2}^2.$$

Therefore, by Gronwall's inequality, we have

$$\|w(t)\|_{\dot{L}^2}^2 \leq \|w_0\|_{\dot{L}^2}^2 e^{C \int_0^t (C_2 + 2\|u\|_{\dot{L}^2}^2 + \|v\|_{\dot{L}^2}^2) ds}, \quad (2.24)$$

with  $w_0 = u_0 - v_0$ , which is finite for all  $t \leq T$ , since  $u, v \in L^2(0, T; \dot{H}^3)$ . Thus, if  $u_0 = v_0$ , this implies that  $w(t) = 0$  for all  $t \geq 0$ , so  $u(t) = v(t)$ .

This gives us uniqueness when  $(1 - r) \geq 0$ .

Now for the case when  $(1 - r) < 0$ , we begin with (2.22) and interpolate the  $(1 - r)$  term with (B.7) with  $n = 1$  to get

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|w\|_{\dot{L}^2}^2 + \|\partial_x^3 w\|_{\dot{L}^2}^2 &\leq \left( \frac{\beta C_0}{16\gamma} + \frac{C_0}{2\beta} + \frac{(r-1)\eta^2}{4} \right) \|\partial_x^3 w\|_{\dot{L}^2}^2 \\ &+ \underbrace{\left( \frac{(r-1)}{\eta} + \frac{\beta\gamma C_0}{4} + \frac{\beta C_0}{\gamma} \right)}_{C_{1/2}} + \frac{C}{2} (2\|u\|_{\dot{L}^2}^2 + \|v\|_{\dot{L}^2}^2) \|w\|_{\dot{L}^2}^2. \end{aligned}$$

Letting

$$\gamma = C_0^2, \quad \beta = 2C_0, \quad \eta^2 = \frac{1}{(r-1)},$$

we can drop the third derivative term and then, using Gronwall's inequality, we again have that

$$\|w(t)\|_{\dot{L}^2}^2 \leq \|w(0)\|_{\dot{L}^2}^2 e^{\int_0^t (C_1 + C(2\|u\|_{\dot{L}^2}^2 + \|v\|_{\dot{L}^2}^2)) dt}, \quad (2.25)$$

which gives uniqueness for  $(1 - r) < 0$ .

## Chapter 3

# Asymptotic Bounds and Regularity

### 3.1 An Absorbing Set in $\dot{L}^2(0, \ell)$

We will now state an important result that has been developed in the literature of the Nikolaevskiy equation, and that gives one an absorbing set in  $\dot{L}^2$ .

**Definition 3.1.1** (Absorbing Set). *Let  $H$  be a space with a dynamical system*

$$\frac{dx}{dt} = f(x), \quad x(0) = x_0,$$

*defined on it. We say that  $\mathcal{B} \subset H$  is an absorbing set in  $H$  if for all  $x_0 \in H$ , there exists a  $t_0 \geq 0$ , such that*

$$x(t) \in \mathcal{B}$$

*for all  $t \geq t_0$ .*

Technically the above definition does not require  $\mathcal{B}$  to be bounded; however, we will always be interested in the case when  $\mathcal{B}$  is bounded.

Finding optimal bounds on the size of an absorbing ball for equations similar to (1.4), such as the Kuramoto-Sivashinsky (KS) equation (1.1) is an area of active research, and one into which much work has gone in the past [NST85, WH99, Wit14, BG06, GO04] as discussed in Section 1.1. In particular, the scaling with the parameter  $\ell$ , the domain size, has been investigated, since the KS equation can be rescaled so that all possible parameters are absorbed into the parameter  $\ell$ .

For the case of the Nikolaevskiy equation, it has been shown that there exists an absorbing ball in  $\dot{L}^2(0, \ell)$ , and that the bound on its radius depends on  $\ell$  [FR08, Wit14], although [FR08] derived a bound of  $\mathcal{O}(\ell^{9/2})$  for large  $\ell$  and did not derive an  $r$  dependence of the absorbing ball. As we will see in later sections, the radius of analyticity of solution to (1.4), the dimension of the attractor, as well as the absorbing balls of the derivatives of solutions

to (1.4), depend on the size of the  $\dot{L}^2$  absorbing set. It is thus important to obtain the optimal scaling of this ball with the parameters.

The strictest bound on the absorbing set in  $\dot{L}^2(0, \ell)$  for the Nikolaevskiy equation derived to date, taking both the  $\ell$  and the  $r$  scaling into account, is given in [Wit14], and we state it without proof:

**Theorem 3.1.2** (Absorbing Set in  $\dot{L}^2(0, \ell)$ ). *Let  $u(t)$  be a solution to (1.4); then there exists a constant  $K$ , independent of  $\ell$  and  $r$ , such that*

$$\limsup_{t \rightarrow \infty} \|u(t)\|_{\dot{L}^2} \leq K \sqrt{r} \ell^{3/2}, \quad r \leq 1.$$

For a proof see [Wit14].

In particular, define

$$R := K \sqrt{r} \ell^{3/2}, \tag{3.1}$$

then the theorem says that the ball of radius  $R$  constitutes an absorbing set  $\mathcal{B}$  of (1.4) in  $\dot{L}^2(0, \ell)$ .

We should note here that the  $\ell$  scaling in (3.1) is generally not believed to be optimal. However, the scaling with  $r$  is, since it can be shown that (1.4) has unstable stationary solutions that scale like  $\sqrt{r}$  to leading order [TV96]. Such solutions will generally not be seen in numerical experiments, due to their instability, but their existence implies that the  $r$ -dependence of the bound (3.1) cannot be improved.

### 3.2 An Absorbing Set in $\dot{H}^1(0, \ell)$

One ingredient that we shall need in order to prove the existence of an attractor in  $\dot{L}^2$  for the Nikolaevskiy equation (1.4), is a uniform bound on the derivative of solutions to (1.4). Specifically we need an absorbing ball in  $\dot{H}^1$ . In what follows, we require that  $0 < r \leq 1$ , so that the bound (3.1) holds.

We start by taking the  $\dot{L}^2$  inner product of equation (1.4) with  $-\partial_x^2 u$  to get

$$\frac{1}{2} \frac{d}{dt} \|\partial_x u\|_{\dot{L}^2}^2 = - \|\partial_x^4 u\|_{\dot{L}^2}^2 + 2 \|\partial_x^3 u\|_{\dot{L}^2}^2 - (1-r) \|\partial_x^2 u\|_{\dot{L}^2}^2 + b(u, u, \partial_x^2 u).$$

We bound the  $b(u, u, \partial_x^2 u)$  term by  $\ell^{1/2} \|u\|_{\dot{L}^2} \|\partial_x^2 u\|_{\dot{L}^2}^2$  by Proposition B.6.3, and then use the  $\dot{L}^2$  absorbing ball  $\|u\|_{\dot{L}^2} \leq R$  (for large  $t$ ) to get

$$\frac{1}{2} \frac{d}{dt} \|\partial_x u\|_{\dot{L}^2}^2 \leq - \|\partial_x^4 u\|_{\dot{L}^2}^2 + 2 \|\partial_x^3 u\|_{\dot{L}^2}^2 - \|\partial_x^2 u\|_{\dot{L}^2}^2 + M_1 \|\partial_x^2 u\|_{\dot{L}^2}^2, \tag{3.2}$$

where  $M_1 := r + \ell^{1/2}R$ , with  $R$  defined as in (3.1). Applying (B.11) with  $\mu = 2$  and  $n = 0$  to the third order term we get

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\partial_x u\|_{\dot{L}^2}^2 &\leq -\frac{1}{2} \|\partial_x^4 u\|_{\dot{L}^2}^2 + \frac{\alpha}{\eta} \|\partial_x^2 u\|_{\dot{L}^2}^2 - 2\alpha \|\partial_x u\|_{\dot{L}^2}^2 + \alpha\eta \|u\|_{\dot{L}^2}^2 + (M_1 + 2) \|\partial_x^2 u\|_{\dot{L}^2}^2 \\ &= -\frac{1}{2} \|\partial_x^4 u\|_{\dot{L}^2}^2 + \left(\frac{\alpha}{\eta} + M_1 + 2\right) \|\partial_x^2 u\|_{\dot{L}^2}^2 - 2\alpha \|\partial_x u\|_{\dot{L}^2}^2 + \alpha\eta \|u\|_{\dot{L}^2}^2. \end{aligned}$$

If we now define  $M := M_1 + 2$  and apply (B.9) to the second-order term we have

$$\frac{1}{2} \frac{d}{dt} \|\partial_x u\|_{\dot{L}^2}^2 \leq \left[-\frac{1}{2} + \sigma \left(\frac{\alpha}{\eta} + M\right)\right] \|\partial_x^4 u\|_{\dot{L}^2}^2 - 2\alpha \|\partial_x u\|_{\dot{L}^2}^2 + \left[\alpha\eta + \frac{1}{2\sigma} \left(\frac{\alpha}{\eta} + M\right)\right] \|u\|_{\dot{L}^2}^2.$$

Now, in order to apply Gronwall's lemma (B.14) we want the highest order term to be zero, so we pick  $\sigma = (2(\alpha/\eta + M))^{-1}$ , giving

$$\begin{aligned} \frac{d}{dt} \|\partial_x u\|_{\dot{L}^2}^2 &\leq -4\alpha \|\partial_x u\|_{\dot{L}^2}^2 + 2 \left[\alpha\eta + \left(\frac{\alpha}{\eta} + M\right)^2\right] \|u\|_{\dot{L}^2}^2 \\ &\leq -4\alpha \|\partial_x u\|_{\dot{L}^2}^2 + 2 \left[\alpha\eta + \left(\frac{\alpha}{\eta} + M\right)^2\right] R^2. \end{aligned}$$

where we have used the bound (3.1) on  $\|u\|_{\dot{L}^2}^2$ . Then using Gronwall's inequality (B.14), we have that

$$\limsup_{t \rightarrow \infty} \|\partial_x u\|_{\dot{L}^2}^2 \leq \frac{2[\alpha\eta + (\alpha/\eta + M)^2]}{4\alpha} R^2. \quad (3.3)$$

We thus have to choose  $\alpha$  and  $\eta$  to optimize this constant. We want  $\alpha$  to be as large as possible while keeping the product  $\alpha\eta$  small. At the same time we want  $\alpha/\eta = \mathcal{O}(M)$  and  $\alpha\eta = \mathcal{O}(M^2)$ . Therefore, one possible choice is  $\alpha = M^{3/2}/2$  and  $\eta = M^{1/2}/2$ . With these choices of the constants, the above expression becomes:

$$\limsup_{t \rightarrow \infty} \|\partial_x u\|_{\dot{L}^2}^2 \leq \frac{17}{4} M^{1/2} R^2 := R_1^2 = \mathcal{O}(\ell^{1/4} R^{1/2} R^2) = \mathcal{O}(r^{5/4} \ell^4). \quad (3.4)$$

The above bound also automatically gives us an upper bound for the  $\dot{L}^\infty$  norm scaling for large  $t$ . Specifically, since

$$\|u\|_{\dot{L}^\infty} \leq \ell^{1/2} \|\partial_x u\|_{\dot{L}^2},$$

an upper bound on the  $\dot{L}^\infty$  norm of  $u$  is  $\ell^{1/2}R$ . Thus the scaling of the  $\dot{L}^\infty$  norm can be at most  $r^{5/8}$ . Computationally, we expect the  $\dot{L}^\infty$  norm to be  $\ell$  independent, but there does seem to be an  $r$  dependence.



### 3.3 An Absorbing Set in $\dot{H}^2(0, \ell)$

In this section we will derive an absorbing set in  $\dot{H}^2$ . We once more only assume that  $0 < r \leq 1$ , so that the bound (3.1) holds. Such a set is not needed to conclude that an attractor exists; however, we will use it, and specifically its scaling with  $r$ , to derive tighter bounds on the dimension of the attractor in the following chapter.

**Theorem 3.3.1** (An Absorbing Set in  $\dot{H}^2(0, \ell)$ ). *Suppose  $u$  is a solution to (1.4); then one has*

$$\limsup_{t \rightarrow \infty} \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 \leq R_2, \quad (3.5)$$

where  $R_2 = \mathcal{O}(r^{7/8}\ell^3)$  is the radius of the absorbing ball in  $\dot{H}^2$  (3.1).

*Proof.* We start by taking the inner product of the equation (1.4) with  $\partial_x^4 u$  to get

$$\frac{1}{2} \frac{d}{dt} \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 = - \left\| \partial_x^5 u \right\|_{\dot{L}^2}^2 + 2 \left\| \partial_x^4 u \right\|_{\dot{L}^2}^2 - (1-r) \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 - b(u, u, \partial_x^4 u).$$

We now bound the  $b(u, u, \partial_x^4 u)$  term. We can bound it as follows:

$$\begin{aligned} |b(u, u, \partial_x^4 u)| &\leq \left| \int_0^\ell u \partial_x u \partial_x^4 u \, dx \right| \\ &\leq \ell^{1/2} \left\| \partial_x^2 u \right\|_{\dot{L}^2} \int_0^\ell |u \partial_x^4 u| \, dx \\ &\leq \ell^{1/2} R \left\| \partial_x^2 u \right\|_{\dot{L}^2} \left\| \partial_x^4 u \right\|_{\dot{L}^2}, \end{aligned}$$

where  $R$  is the radius of the  $\dot{L}^2$  absorbing ball (3.1). Now let  $M := \ell R^2/4$ , then using Young's inequality and dropping the  $(1-r)$  term we get

$$\frac{1}{2} \frac{d}{dt} \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 \leq - \left\| \partial_x^5 u \right\|_{\dot{L}^2}^2 + 3 \left\| \partial_x^4 u \right\|_{\dot{L}^2}^2 + M \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2.$$

We will now use inequality (B.11) with  $n = 1$  on the fourth-order term, and then follow this up by using inequality (B.9) with  $n = 1$  on the resulting third-order term. This gives

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 &\leq \left( -1 + \frac{3}{2\mu} \right) \left\| \partial_x^5 u \right\|_{\dot{L}^2}^2 + \left( \frac{3\alpha}{2\eta} + \frac{3\mu}{2} \right) \left\| \partial_x^3 u \right\|_{\dot{L}^2}^2 \\ &\quad + (M - 3\alpha) \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 + \frac{3\alpha\eta}{2} \left\| \partial_x u \right\|_{\dot{L}^2}^2 \\ &\leq \left[ -1 + \frac{3}{2\mu} + \frac{1}{\epsilon} \left( \frac{3\alpha}{2\eta} + \frac{3\mu}{2} \right) \right] \left\| \partial_x^5 u \right\|_{\dot{L}^2}^2 + \left[ \frac{\epsilon}{2} \left( \frac{3\alpha}{2\eta} + \frac{3\mu}{2} \right) + \frac{3\alpha\eta}{2} \right] \left\| \partial_x u \right\|_{\dot{L}^2}^2 \\ &\quad + (M - 3\alpha) \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2. \end{aligned}$$

We immediately read off that  $\mu = \mathcal{O}(1)$ ,  $\alpha/(\epsilon\eta) \leq \mathcal{O}(1)$  and  $\alpha = \mathcal{O}(M)$ . We also choose  $\eta = \mathcal{O}(\alpha^{1/2})$  so we must have  $\epsilon = \mathcal{O}(\eta) = \mathcal{O}(\alpha^{1/2})$ . Choosing  $\alpha = 2M/3$ ,  $\mu = 6/2$ ,  $\eta = M^{1/2}$  and  $\epsilon = 4M^{1/2}$  we get

$$\frac{1}{2} \frac{d}{dt} \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 \leq -M \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 + \left[ 2M^{1/2} \left( M^{1/2} + \frac{9}{2} \right) + M^{3/2} \right] \left\| \partial_x u \right\|_{\dot{L}^2}^2,$$

for large enough  $M$ . We can now use Gronwall's Lemma B.14 to get

$$\limsup_{t \rightarrow \infty} \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 \leq \left[ 2 \left( 1 + \frac{9}{2M^{1/2}} \right) + M^{1/2} \right] \left\| \partial_x u \right\|_{\dot{L}^2}^2 \leq \left[ 2 \left( 1 + \frac{9}{2M^{1/2}} \right) + M^{1/2} \right] R_1^2,$$

where  $R_1$  is the bound from the previous section (3.4). Thus

$$\limsup_{t \rightarrow \infty} \left\| \partial_x^2 u \right\|_{\dot{L}^2}^2 := R_2^2 = \mathcal{O}(M^{1/2} R_1^2) = \mathcal{O}(\ell^{1/2} R R_1^2) = \mathcal{O}(r^{7/4} \ell^6).$$

□

We note here that the results on analyticity of solutions of the Nikolaevskiy equation in the next section can be used to derive bounds on the radii of the absorbing balls in all  $\dot{H}^n$  spaces, since we derive uniform bounds on a norm that contains all the derivatives, and hence also individually bounds each derivative of  $u$ . However, the results obtained in that way are not as tight as the ones we obtained in this section.

### 3.4 Regularity

In this section we will show that the Nikolaevskiy equation (1.4) has the effect of ‘smoothing’ out initial data, much like the well-known heat equation. Specifically, we will show that all solutions become analytic once they have entered the absorbing ball in  $\dot{L}^2$ . We will be following the methods developed in [CEES93a], and derive bounds on the radius of analyticity. A bound on the radius of analyticity has been derived for the KS equation [CEES93a] and the (de)stabilized KS equation [Wit02], as well as on its generalizations [IS14]. As in [Wit02], we are also interested in the scaling of the radius of analyticity with the parameter  $r$ . Specifically we will show that for any initial condition in  $\dot{L}^2$  the solution starting at this initial condition is analytic for all  $t > 0$ .

Before we begin, we note that Theorem B.5.4 already gives us some regularity in time. Since we showed that  $u \in L^2(0, T; \dot{H}^3)$  and  $u_t \in L^2(0, T; \dot{H}^{-3})$ , we have that

$$u \in C([0, T]; \dot{L}^2), \tag{3.6}$$

so  $u$  is continuous in time. We will now proceed to show the regularity in space.

In what follows we will be working with the absolute value of the derivative operator, that is, we will work with  $A := |\partial_x|$ , where

$$Au = |\partial_x|u := \sum_k |k| \hat{u}_k e^{ikx}. \quad (3.7)$$

We will show that solutions to equation (1.4) and their derivatives decay in a certain way, and then use the Paley-Wiener Theorem from B.5.3 to conclude that the solutions must be analytic. The following approach comes directly from [CEES93a], but we present the proof with added details, and with a slight modification to the linear term  $\mathcal{L}$  to adapt the results of [CEES93a] to the Nikolaevskiy equation (1.4). We begin by rewriting (1.4) in Fourier space in a slightly different way than before. If we take the nonlinear term to be  $\frac{1}{2} \partial_x(u)^2$  instead of  $u \partial_x u$ , then we get

$$\frac{d}{dt} \hat{u}_k = \omega(k) \hat{u}_k + \frac{ik}{2} \sum_{k'+k''=k} \hat{u}_{k'} \hat{u}_{k''}, \quad (3.8)$$

for each Fourier coefficient.

In the following we define

$$v := e^{\alpha t A} u \iff \hat{v}_k = e^{\alpha t |k|} \hat{u}_k. \quad (3.9)$$

We will show that the  $\dot{L}^2$  norm of  $v$  defined in (3.9) is bounded. We first have to prove a Lemma concerning the nonlinear term in (1.4). For the trilinear form  $b(u, v, w)$  defined in (2.14) we have

$$\begin{aligned} b(u, v, w) &= \int_0^\ell u \partial_x v w \, dx = \int_0^\ell \left( \sum_k \hat{u}_k e^{ikx} \right) \left( \sum_{k'} ik' \hat{v}_{k'} e^{ik'x} \right) \left( \sum_{k''} \hat{w}_{k''} e^{ik''x} \right) \, dx \\ &= \int_0^\ell \left( \sum_k \hat{u}_k e^{ikx} \right) \left( \sum_{-k'} -ik' \hat{v}_{-k'} e^{-ik'x} \right) \left( \sum_{k''} \hat{w}_{k''} e^{ik''x} \right) \, dx \\ &= \int_0^\ell \left( \sum_k \sum_{k'} \sum_{k''} (-ik') \hat{u}_k \hat{v}_{-k'} \hat{w}_{k''} e^{i(k-k'+k'')x} \right) \, dx \end{aligned}$$

where we have flipped the middle sum, which can be done by simply relabelling. But the only nonzero terms in this triple sum occur when the complex exponents add to zero. Thus we need  $k'' = k' - k$ , so we have

$$b(u, v, w) = -i\ell \sum_k \sum_{k'} k' \hat{u}_k \hat{v}_{-k'} \hat{w}_{k'-k}. \quad (3.10)$$

We will now find a bound on (2.14) in terms of the norm of  $v$  and its derivative.

**Lemma 3.4.1** ([CEES93a]). *Let  $A = |\partial_x|$  be defined as in (3.7) and let  $\alpha > 0$ ; then*

$$|b(u, u, e^{\alpha t A} v)| \leq C \sqrt{\alpha t} \|v\|_{L^2} \|Av\|_{L^2}^2, \quad (3.11)$$

where we define

$$C = \sqrt{\frac{\log(256)}{\pi}}.$$

*Proof.* Recall that we have the identities  $b(u, u, w) = -\frac{1}{2}b(u, w, u)$ , and  $b(u, u, u) = 0$  by Corollary B.6.2. Thus using

$$v = e^{\alpha t A} u,$$

we can write

$$\begin{aligned} b(u, u, e^{\alpha t A} v) &= -\frac{1}{2}b(u, e^{\alpha t A} v, u) \\ &= -\frac{1}{2} \left[ b(e^{-\alpha t A} v, e^{\alpha t A} v, e^{-\alpha t A} v) - b(v, v, v) \right] \\ &= \frac{i\ell}{2} \sum_k \sum_{k'} k' \hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k} \left( e^{\alpha t(|k'| - |k| - |k' - k|)} - 1 \right), \end{aligned}$$

where we have used (3.10) in the last line. It is straightforward to verify considering the different cases, that

$$|k'| - |k| - |k' - k| = \begin{cases} 0, & \text{if } k' \geq k > 0 \text{ or } k' \leq k < 0 \\ -2|k|, & \text{if } k' > 0 > k \text{ or } k' < 0 < k \\ -2|k' - k|, & \text{if } k > k' > 0 \text{ or } k < k' < 0 \end{cases}.$$

We omit the cases with  $k = 0$  or  $k' = 0$ , since  $\hat{v}_0 = 0$ . Thus the above sum can be split into four sums of the form:

$$\begin{aligned}
b(u, u, e^{\alpha t A} v) &= \frac{i\ell}{2} \underbrace{\sum_{k' > 0 > k} k' \hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k} \left( e^{-2\alpha t |k|} - 1 \right)}_{\textcircled{1}} \\
&+ \frac{i\ell}{2} \underbrace{\sum_{k' < 0 < k} k' \hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k} \left( e^{-2\alpha t |k|} - 1 \right)}_{\textcircled{2}} \\
&+ \frac{i\ell}{2} \underbrace{\sum_{k > k' > 0} k' \hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k} \left( e^{-2\alpha t |k'-k|} - 1 \right)}_{\textcircled{3}} \\
&+ \frac{i\ell}{2} \underbrace{\sum_{k < k' < 0} k' \hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k} \left( e^{-2\alpha t |k'-k|} - 1 \right)}_{\textcircled{4}}.
\end{aligned}$$

We will combine some of these terms and then treat the resultant terms individually.

Let us look at terms  $\textcircled{3}$  and  $\textcircled{4}$ . In term  $\textcircled{4}$ , if we replace  $k$  with  $-k$  and  $k'$  with  $-k'$ , we get

$$\begin{aligned}
\textcircled{4} &= \sum_{k < k' < 0} k' \hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k} \left( e^{-2\alpha t |k'-k|} - 1 \right) = \sum_{k > k' > 0} (-k') \hat{v}_{-k} \hat{v}_{k'} \hat{v}_{-k'+k} \left( e^{-2\alpha t |k'-k|} - 1 \right) \\
&= \sum_{k > k' > 0} -k' \overline{\hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k}} \left( e^{-2\alpha t |k'-k|} - 1 \right) \\
&= -\overline{\textcircled{3}},
\end{aligned}$$

where the bar denotes complex conjugate. Thus  $\textcircled{4} + \textcircled{3} = 2i \text{Im} \textcircled{3}$ . Similarly we get that  $\textcircled{1} + \textcircled{2} = 2i \text{Im} \textcircled{1}$ . So we get

$$\begin{aligned}
b(u, u, e^{\alpha t A} v) &= -\ell \text{Im} \left[ \sum_{k' > 0 > k} k' \hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k} \left( e^{-2\alpha t |k|} - 1 \right) \right. \\
&\quad \left. + \sum_{k > k' > 0} k' \hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k} \left( e^{-2\alpha t |k'-k|} - 1 \right) \right],
\end{aligned}$$

so then

$$\begin{aligned}
|b(u, u, e^{\alpha t A} v)| &\leq \underbrace{\ell \sum_{k' > 0 > k} k' |\hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k}| \left| e^{-2\alpha t |k|} - 1 \right|}_{\textcircled{\text{I}}} \\
&\quad + \underbrace{\ell \sum_{k > k' > 0} k' |\hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k}| \left| e^{-2\alpha t |k'-k|} - 1 \right|}_{\textcircled{\text{II}}}.
\end{aligned}$$

We will first estimate  $\textcircled{\text{I}}$ . Multiplying by  $k/k$  and applying Cauchy-Schwarz gives

$$\begin{aligned}
\textcircled{\text{I}} &= \ell \sum_{k' > 0 > k} k' |\hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k}| \left| e^{-2\alpha t |k|} - 1 \right| \left( \frac{k}{k} \right) \\
&\leq \ell \left( \sum_{k' > 0 > k} k^2 |\hat{v}_k|^2 |\hat{v}_{k'-k}|^2 \right)^{1/2} \left( \sum_{k' > 0 > k} k'^2 |\hat{v}_{-k'}|^2 \left( \frac{e^{-2\alpha t |k|} - 1}{k} \right)^2 \right)^{1/2} \\
&= \ell \left( \sum_{k < 0} k^2 |\hat{v}_k|^2 \sum_{k' > 0} |\hat{v}_{k'-k}|^2 \right)^{1/2} \left( \sum_{k' > 0} \sum_{k < 0} k'^2 |\hat{v}_{k'}|^2 \left( \frac{e^{-2\alpha t |k|} - 1}{k} \right)^2 \right)^{1/2},
\end{aligned}$$

so summing first over  $k'$  and then over  $k$  we have, using Parseval's identity

$$\textcircled{\text{I}} \leq \|v\|_{\mathbb{L}^2} \|Av\|_{\mathbb{L}^2} \left( \sum_{k'} |k'|^2 |\hat{v}_{k'}|^2 \right)^{1/2} \left( \sum_{k > 0} \left( \frac{e^{-2\alpha t |k|} - 1}{k} \right)^2 \right)^{1/2}.$$

Now

$$\begin{aligned}
\left( \sum_{k > 0} \left( \frac{e^{-2\alpha t |k|} - 1}{k} \right)^2 \right)^{1/2} &= \left( \sum_{j > 0} \left( \frac{e^{-2\alpha t (2\pi j/\ell)} - 1}{(2\pi j/\ell)} \right)^2 \right)^{1/2} \\
&\leq \left( \int_0^\infty \left( \frac{e^{-2\alpha t (2\pi x/\ell)} - 1}{(2\pi x/\ell)} \right)^2 dx \right)^{1/2} \\
&= \left( \frac{2\alpha t \ell}{2\pi} \int_0^\infty \left( \frac{e^{-y} - 1}{y} \right)^2 dy \right)^{1/2} \\
&= \left( \frac{\alpha t \log(4)\ell}{\pi} \right)^{1/2}.
\end{aligned}$$

We have used the software Maple to compute the integral  $\int_0^\infty \left(\frac{e^{-y}-1}{y}\right)^2 dy = \log(4)$ . Thus by Parseval's identity we have

$$\textcircled{\text{I}} \leq \sqrt{\frac{\alpha t \log(4)}{\pi}} \|v\|_{\dot{L}^2} \|Av\|_{\dot{L}^2}^2.$$

In order to estimate  $\textcircled{\text{II}}$ , we will first make the change of variables  $p = k' - k$ , so then  $k = k' - p$ . Then we have

$$\begin{aligned} \textcircled{\text{II}} &= \ell \sum_{k > k' > 0} k' |\hat{v}_k \hat{v}_{-k'} \hat{v}_{k'-k}| \left( e^{-2\alpha t |k'-k|} - 1 \right) \\ &= \ell \sum_{k'-p > k' > 0} k' |\hat{v}_{k'-p} \hat{v}_{-k'} \hat{v}_p| \left( e^{-2\alpha t |p|} - 1 \right). \end{aligned}$$

But the condition  $k' - p > k' > 0$  implies that  $k' > 0 > p$  and so we take the sum as

$$\textcircled{\text{II}} = \ell \sum_{k' > 0 > p} k' |\hat{v}_{k'-p} \hat{v}_{-k'} \hat{v}_p| \left( e^{-2\alpha t |p|} - 1 \right).$$

Now we have a sum analogous to  $\textcircled{\text{I}}$ , with an index  $p$  instead of  $k$ . Thus the same argument as before shows that

$$\textcircled{\text{II}} \leq \sqrt{\frac{\alpha t \log(4)}{\pi}} \|v\|_{\dot{L}^2} \|Av\|_{\dot{L}^2}^2,$$

and so we finally have

$$|b(u, u, e^{\alpha t A} v)| \leq \textcircled{\text{I}} + \textcircled{\text{II}} \leq \sqrt{\frac{\alpha t \log(256)}{\pi}} \|v\|_{\dot{L}^2} \|Av\|_{\dot{L}^2}^2,$$

giving the result.  $\square$

We are now ready to prove a theorem which, together with the Paley-Wiener Theorem A.9, will give us analyticity of solutions of the Nikolaevskiy equation.

**Theorem 3.4.2** (Analyticity of Solutions of the Nikolaevskiy Equation). *Let  $u(x, t)$  be a solution to (1.4). Then if  $\|u(0, \cdot)\|_{\dot{L}^2} \leq R$ , where  $R$  given in (3.1) is the radius of the absorbing ball of solutions in  $\dot{L}^2$ , we have that*

$$\left\| e^{\alpha \min(t, t_c) A} u(t, \cdot) \right\|_{\dot{L}^2} \leq 2R, \quad (3.12)$$

for any  $t > 0$ . Here  $A = |\partial_x|$ ,  $\alpha = \tilde{\alpha} R^{2/3}$  and  $t_c = \tilde{t} R^{-4/3}$ , where,  $\tilde{\alpha}$  and  $\tilde{t}$  are constants independent of  $\ell$ .

Specifically, solutions are analytic for all  $t > 0$ , and they are analytic in a strip of width at least  $\alpha t_c = \tilde{\alpha} \tilde{t} R^{-2/3}$  about  $[0, \ell]$  in the complex plane, once they have entered the absorbing ball and  $t$  is greater than  $t_c$  plus the time at which they have entered the absorbing ball.

The proof of the above theorem parallels the proof of Theorem 3.1 in [CEES93a].

*Proof.* If we multiply each of the equations (3.8) by  $\hat{v}_{-k} e^{\alpha|k|t}$  and sum over  $k$ , we get

$$\frac{1}{2} \frac{d}{dt} \sum_k |\hat{v}_k|^2 = \alpha \sum_k |k| |\hat{v}_k|^2 + \sum_k \omega(k) |\hat{v}_k|^2 - \frac{1}{\ell} b(u, u, e^{\alpha t A} v),$$

where we have used a product rule on the time derivative term. And so using Lemma 3.4.1 and re-expressing everything in terms of real-space norms using Parseval's inequality, we get

$$\frac{1}{2} \frac{d}{dt} \|v\|_{\dot{L}^2}^2 \leq \alpha \|A^{1/2} v\|_{\dot{L}^2}^2 - \|A^3 v\|_{\dot{L}^2}^2 + 2 \|A^2 v\|_{\dot{L}^2}^2 - (1-r) \|Av\|_{\dot{L}^2}^2 + C\sqrt{\alpha t} \|v\|_{\dot{L}^2} \|Av\|_{\dot{L}^2}^2$$

where  $\|A^n v\|_{\dot{L}^2} = \|\partial_x^n v\|_{\dot{L}^2}$  for  $n = 1, 2, 3$ . We will now repeatedly use Young's inequality as well as inequality (B.6) and the fact that  $r \leq 1$  to get an estimate on the above expression.

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|v\|_{\dot{L}^2}^2 &\leq \alpha \|v\|_{\dot{L}^2} \|Av\|_{\dot{L}^2} - \|A^3 v\|_{\dot{L}^2}^2 + 2 \|Av\|_{\dot{L}^2} \|A^3 v\|_{\dot{L}^2} + C\sqrt{\alpha t} \|v\|_{\dot{L}^2} \|A^2 v\|_{\dot{L}^2} \\ &\leq \frac{\alpha}{2\eta} \|v\|_{\dot{L}^2}^2 + \frac{\alpha\eta}{2} \|Av\|_{\dot{L}^2}^2 - \|A^3 v\|_{\dot{L}^2}^2 + \delta \|Av\|_{\dot{L}^2}^2 + \frac{1}{\delta} \|A^3 v\|_{\dot{L}^2}^2 \\ &\quad + \frac{C\alpha t}{2} \|v\|_{\dot{L}^2}^4 + \frac{C}{2} \|A^2 v\|_{\dot{L}^2}^2 \\ &\leq \frac{\alpha}{2\eta} \|v\|_{\dot{L}^2}^2 + \frac{\alpha\eta}{2} \|Av\|_{\dot{L}^2}^2 - \|A^3 v\|_{\dot{L}^2}^2 + \delta \|Av\|_{\dot{L}^2}^2 + \frac{1}{\delta} \|A^3 v\|_{\dot{L}^2}^2 \\ &\quad + \frac{C\alpha t}{2} \|v\|_{\dot{L}^2}^4 + \frac{C\gamma}{4} \|Av\|_{\dot{L}^2}^2 + \frac{C}{4\gamma} \|A^3 v\|_{\dot{L}^2}^2. \end{aligned}$$

Now if we use the Young's-type inequality (B.7) for the  $\|Av\|_{\dot{L}^2}^2$  terms, with  $\epsilon = 1$  for the first occurrence, then we have

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|v\|_{\dot{L}^2}^2 &\leq \frac{\alpha}{2\eta} \|v\|_{\dot{L}^2}^2 + \frac{\alpha\eta}{2} \|v\|_{\dot{L}^2}^2 + \frac{\alpha\eta}{8} \|A^3 v\|_{\dot{L}^2}^2 - \|A^3 v\|_{\dot{L}^2}^2 \\ &\quad + \frac{\delta}{\rho} \|v\|_{\dot{L}^2}^2 + \frac{\delta\rho^2}{4} \|A^3 v\|_{\dot{L}^2}^2 + \frac{1}{\delta} \|A^3 v\|_{\dot{L}^2}^2 \\ &\quad + \frac{C\alpha t}{2} \|v\|_{\dot{L}^2}^4 + \frac{C\gamma}{4\epsilon} \|v\|_{\dot{L}^2}^2 + \frac{C\gamma\epsilon^2}{16} \|A^3 v\|_{\dot{L}^2}^2 + \frac{C}{4\gamma} \|A^3 v\|_{\dot{L}^2}^2 \\ &= \left( \frac{\alpha}{2\eta} + \frac{\alpha\eta}{8} + \frac{\delta}{\rho} + \frac{C\gamma}{4\epsilon} \right) \|v\|_{\dot{L}^2}^2 \\ &\quad + \left( \frac{\alpha\eta}{8} - 1 + \frac{\delta\rho^2}{4} + \frac{1}{\delta} + \frac{C\gamma\epsilon^2}{16} + \frac{C}{4\gamma} \right) \|A^3 v\|_{\dot{L}^2}^2 + \frac{C\alpha t}{2} \|v\|_{\dot{L}^2}^4. \end{aligned}$$



Thus we need to pick the constants in the Young's inequalities such that

$$\frac{\alpha\eta}{8} + \frac{\delta\rho^2}{4} + \frac{1}{\delta} + \frac{C\gamma\epsilon^2}{16} + \frac{C}{4\gamma} - 1 = 0.$$

Thus if we pick the constants to be equal, for instance to

$$\eta = \frac{8}{5\alpha}, \delta = 5, \rho = \frac{2}{5}, \gamma = \frac{5C}{4}, \epsilon = \frac{8}{5C},$$

then we have

$$\frac{\alpha\eta}{8} + \frac{\delta\rho^2}{4} + \frac{1}{\delta} + \frac{C\gamma\epsilon^2}{16} + \frac{C}{4\gamma} - 1 = \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{1}{5} + \frac{1}{5} - 1 = 0.$$

Thus if we define

$$C_1 := \left( \frac{8}{5} + 25 + \frac{25C^3}{64} \right),$$

then we have

$$\frac{d}{dt} \|v\|_{\dot{L}^2}^2 \leq \left( \frac{5\alpha^2}{8} + C_1 \right) \|v\|_{\dot{L}^2}^2 + C\alpha t \|v\|_{\dot{L}^2}^4.$$

Now, if  $\|v(0)\|_{\dot{L}^2} = \|u(0)\|_{\dot{L}^2} \leq R$ , where  $R$  is any bound on  $\|u(0)\|_{\dot{L}^2}$  if the initial condition is outside of the absorbing ball, but if  $\|u(0)\|_{\dot{L}^2}$  is inside of it, then  $R$  is the radius of the absorbing ball of the Nikolaevskiy equation in  $\dot{L}^2$ , then since it is known that solutions grow at most exponentially in time (see Section 2.3), we know that there exists some time  $t_c$ , such that for  $t \leq t_c$  we must have  $\|v(t)\|_{\dot{L}^2}^2 \leq 4R^2$ . Thus for all  $t \leq t_c$ , we have

$$\frac{d}{dt} \|v\|_{\dot{L}^2}^2 \leq \left( \frac{5\alpha^2}{8} + C_1 \right) \|v\|_{\dot{L}^2}^2 + 4C\alpha t R^2 \|v\|_{\dot{L}^2}^2.$$

Then by Gronwall's inequality we have

$$\begin{aligned} \|v(t)\|_{\dot{L}^2}^2 &\leq \|v(0)\|_{\dot{L}^2}^2 \exp\left(\frac{5\alpha^2}{8}t + C_1t + 2CR^2\alpha t^2\right) \\ &\leq R^2 \exp\left(\frac{5\alpha^2}{8}t + C_1t + 2CR^2\alpha t^2\right). \end{aligned}$$

The above bound is less than  $4R^2$  as long as the exponent is less than  $\log(4)$ , which is true when  $t$  is sufficiently small that

$$\left( \frac{5\alpha^2}{8}t + C_1t + 2CR^2\alpha t^2 \right) \leq \log(4).$$

We need to choose the constant  $\alpha$  and the maximum time in such a way that the exponent scales as  $R^{-a}$  for some  $a \geq 0$ . Thus we will let  $\alpha = \tilde{\alpha}R^p$  and  $t_c = \tilde{t}R^q$ , where  $\tilde{\alpha}$  and  $\tilde{t}$  are two suitably chosen constants independent of  $\ell$ . If we let  $p = 2/3$  and  $q = -4/3$ , then for  $t \leq t_c$  we have

$$\left( \frac{5\alpha^2}{8}t + C_1t + 2CR^2\alpha t^2 \right) \leq \left( \frac{5\tilde{\alpha}^2\tilde{t}}{8} + C_1\tilde{t}R^{-4/3} + 2C\tilde{\alpha}\tilde{t}^2 \right) \leq \log(4),$$

for small enough  $\tilde{\alpha}$  and  $\tilde{t}$ , and for  $R \geq 1$  (This holds for any  $R \geq \|u(0)\|_{\dot{L}^2}$ ).

Thus the width of the strip of analyticity about  $[0, \ell]$  is at least  $\mathcal{O}(tR^{2/3})$  for  $0 < t < t_c$  and  $\alpha t_c = \mathcal{O}(R^{-2/3})$  when  $t \geq t_c$ .

Thus the above theorem and the Paley-Wiener Theorem B.5.3 in the Appendix show that any initial condition  $u_0$  immediately gets smoothed by the flow of (1.4). In fact, this shows that  $u$  is analytic for all  $t > 0$ , since one can use  $t \mapsto t - t_0$  and do the whole argument of before using

$$v = e^{\alpha(t-t_0)A},$$

for an arbitrary  $t_0$ , since then  $\|v(t_0)\|_{\dot{L}^2}^2 \leq R^2$  until  $t_0 + t_c$ . Thus  $u$  is analytic for all time, and in fact any initial condition in  $\dot{L}^2$  immediately gets smoothed out, even if it doesn't start in the absorbing ball, since one can choose  $t_0$  so that the solutions starts outside of the absorbing ball.  $\square$

## Chapter 4

# Attractor Dimension

In this chapter we will show that the Nikolaevskiy equation (1.4) has an attractor in  $\dot{L}^2$ , and we will show that this attractor is finite-dimensional and derive an upper bound on its fractal dimension. The first four sections in this chapter give some necessary background on dynamical systems, which we include here for completeness. We also state and give the proofs of some technical theorems about the dimension of dynamical systems. The first four sections are mainly based on [Rob01], and the proofs of the main theorems can be found in the references provided in the statement of the theorems. For this reason the reader may wish to skip the first four sections of this chapter, and begin reading at Section 4.5, where the main results for the Nikolaevskiy equation begin.

### 4.1 Existence of an Attractor, and Fractal Dimensions

We will first give some general results about dynamical systems and then apply these to equation (1.4).

**Definition 4.1.1** (Strongly Continuous Semigroup (Flow) [Bre10, p. 121]). *Let  $\mathcal{B}$  be a Banach space. A **strongly continuous semigroup** of linear operators on  $\mathcal{B}$  is a family of linear maps  $\{S(t)|t \geq 0\}$  with the following properties:*

- (i) *Each  $S(t) : \mathcal{B} \rightarrow \mathcal{B}$  is a bounded linear operator.*
- (ii) *For every  $s, t \geq 0$ , the composition satisfies  $S(t)S(s) = S(t+s)$  (semigroup property). Moreover,  $S(0) = I$ , where  $I$  is the identity operator.*
- (iii) *For every  $u \in \mathcal{B}$ , the map  $t \mapsto S(t)u$  is continuous from  $[0, \infty)$  into  $\mathcal{B}$ .*

The continuity property will be important for us later on. For a definition of a weakly continuous semigroup, see [Cer94].

We will be using semigroups to define solutions to dynamical systems. We will often call the semigroup of some dynamical system the **flow** of the system. Consider the dynamical

system

$$\frac{dx}{dt} = f(x), \quad x(0) = x_0 \in H, \quad (4.1)$$

whose solution is evolved by the flow  $S(t)x_0$ . Here  $x_0$  and hence  $x(t)$  are in some Hilbert space  $H$ , and  $f$  is some operator on  $H$ . Here we have used semigroups to define the solution. The solution  $x(t)$  of (4.1) is given as  $x(t) := S(t)x_0$ . The linearisation of (4.1) about the solution  $x(t)$  is given by

$$\frac{dy}{dt} = Df(x(t))y, \quad y(0) = y_0 \in H, \quad (4.2)$$

where  $Df$  is the Fréchet derivative if  $H$  is infinite-dimensional, or just the Jacobian if  $H$  is finite-dimensional. For the linearised system (4.2), we will denote its flow by  $y(t) := \Lambda(t)y_0$ .

Having defined dynamical systems and their flows, we will now state some results concerning attractors of systems such as (4.1).

**Definition 4.1.2** (Global Attractor [Rob01, p. 268]). *The **global attractor**  $\mathcal{A} \subset H$  is the maximal compact, invariant set*

$$S(t)\mathcal{A} = \mathcal{A} \quad \text{for all } t \geq 0,$$

*and the minimal set that attracts all bounded sets:*

$$\text{dist}(S(t)X, \mathcal{A}) \rightarrow 0 \quad \text{as } t \rightarrow \infty,$$

*for any bounded set  $X \subset H$ .*

For two sets  $X$  and  $Y$ , by ‘**dist**’ we mean

$$\text{dist}(X, Y) := \sup_{x \in X} \inf_{y \in Y} |x - y|.$$

Note that the above expression is not symmetric in  $X$  and  $Y$ , and is thus not a metric.

We will now state the general theorem that gives the existence of an attractor.

**Theorem 4.1.3** (Existence of an attractor). *Suppose that  $H$  contains a compact, absorbing set under the flow  $S(t)$  of a dissipative dynamical system (4.1). Then there exists a connected global attractor  $\mathcal{A}$  in  $H$ .*

For a proof of this theorem see [Rob01, p. 269]. We will later use this theorem to conclude that the Nikolaevskiy equation (1.4) possesses an attractor in  $\dot{L}^2$ .

Having defined attractors, we will look at the concept of dimension more closely, and give some definitions and results about the fractal dimension of a set, specifically of an attractor. The reason we are interested in the dimension of the attractor of the Nikolaevskiy equation

(1.4) is because, as we saw, equation (1.4) can be thought of as a system of coupled ODEs. The dimension of the attractor then tells us how large this system has to be in order to represent the long-term dynamics of solutions of equation (1.4). If the dimension is finite, then the dynamical system associated with equation (1.4) can essentially be thought of as a finite-dimensional system of ODEs for large  $t$ .

**Definition 4.1.4** (Fractal dimension [YHK93]). *Let  $\mathcal{A}$  be a bounded set. Let  $N(\epsilon)$  be the minimum number of balls of radius  $\epsilon > 0$  needed to cover  $\mathcal{A}$ . Then the **fractal dimension** of  $\mathcal{A}$  is defined as*

$$d_f(\mathcal{A}) := -\lim_{\epsilon \rightarrow 0} \frac{\log(N(\epsilon))}{\log(\epsilon)}. \quad (4.3)$$

We now have all the fundamental definitions that we need. In the following section we will develop some tools which we will later use to prove a theorem about bounding the fractal dimension of an attractor.

## 4.2 Tools and Lemmas

Before we can prove that the dynamical system given by the Nikolaevskiy equation (1.4) is finite-dimensional, we have to prove several important theorems in a more abstract setting first. Our treatment is mainly based on the approach in the book by Robinson [Rob01], in which many of the results in this section can be found. We include these results here to make our discussion self contained, and we also add some additional results for completeness.

The main purpose of the next few sections is to study how a volume in phase space evolves under some map, such as the flow of a dynamical system. We will study the evolution of the volume along trajectories; in general, if the volume decays to zero along all trajectories on the attractor, then we can conclude that the attractor cannot contain volumes of the same dimension as the one whose evolution we are studying.

Let us first prove some general lemmas about the image of volumes under compact maps in a Hilbert space. The results that we present here can be found in [Hun96]. We present the proofs in more detail, and give the proofs of propositions that were assumed to be true in [Hun96], but whose proofs are nevertheless not trivial.

**Lemma 4.2.1** (Lemma [Rob01, p. 439]). *Let  $\Lambda$  be a compact, linear operator from a separable Hilbert space  $H$  to itself. Then  $\Lambda(B)$ , the image of the unit ball  $B$  in  $H$  under  $\Lambda$ , is an ellipse, with the lengths of its semi-axes given by the eigenvalues of  $(\Lambda^*\Lambda)^{1/2}$  (the singular values of  $\Lambda$ ), where  $\Lambda^*$  is the adjoint of  $\Lambda$ .*

*Proof.* The self-adjoint operator  $(\Lambda^*\Lambda)^{1/2}$  has non-negative eigenvalues  $\lambda_n$ , with corresponding orthonormal eigenvectors  $e_n$ . Then the vectors  $\Lambda e_n$  are orthogonal with lengths

$\|\Lambda e_n\| = \lambda_n$ , since in general we have

$$(\Lambda e_n, \Lambda e_j) = (\Lambda^* \Lambda e_n, e_j) = \lambda_n^2 (e_n, e_j) = \lambda_n^2 \delta_{nj}.$$

This shows that the vectors  $\Lambda e_n$  are orthogonal and form a basis for  $\Lambda(H)$ , since  $e_n$  is an orthonormal basis for  $H$  by the Hilbert-Schmidt theorem [Rob01, p. 75]. Furthermore, for  $\lambda_n \neq 0$  the  $\Lambda e_n/\lambda_n$  are normalized, while if  $\lambda_n = 0$ , we exclude this term from the sum.

Now let  $u \in B$ , where  $B$  is the unit ball in  $H$ , by which we mean  $u \in B$  if and only if

$$u = \sum_{n=1}^{\infty} c_n e_n.$$

and

$$\sum_{n=1}^{\infty} |c_n|^2 \leq 1.$$

Then we conclude

$$\Lambda u = \sum_{n=1}^{\infty} c_n \Lambda e_n = \sum_{n=1}^{\infty} c_n \lambda_n \frac{\Lambda e_n}{\lambda_n}.$$

It follows that if  $k_n := c_n \lambda_n$ , we have that

$$\Lambda u \in \left\{ y = \sum_{n=1}^{\infty} k_n \frac{\Lambda e_n}{\lambda_n} \mid \sum_{n=1}^{\infty} \frac{|k_n|^2}{\lambda_n^2} \leq 1 \right\},$$

which is an ellipse with semi-axes of lengths  $\lambda_n$ . □

**Definition 4.2.2.** We define the *maximum factor of expansion* of a  $j$ -dimensional volume in  $H$  by

$$\omega_j := \lambda_j \cdots \lambda_i \cdots \lambda_1 = \prod_{i=1}^j \lambda_i,$$

where the  $\lambda_i$  are the lengths of the semi-axes of the ellipse in the previous lemma. Alternatively, we will sometimes write

$$\omega_j(F, x) = \omega_j(DF(x)) := \lambda_j \cdots \lambda_i \cdots \lambda_1,$$

where the  $\lambda_i$  are now understood to be the singular values of the linearisation of  $F$  about  $x$ . That is, the  $\lambda_i$  are the eigenvalues of the operator  $(DF(x)^* DF(x))^{-1/2}$ . In both cases they are ordered  $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ .

We now use the result from Lemma 4.2.1 to get a bound on the number of  $\epsilon$ -balls needed to cover an ellipse.

**Lemma 4.2.3** ([Hun96]). *Let  $G$  be a  $C^1$  map defined on a neighbourhood of a compact set  $X$ . Then the linearisation of  $G$ ,  $DG(x)$ , exists for all  $x \in X$ . Assume that  $DG(x)$  is compact and let  $\lambda_i$  be the eigenvalues of  $(DG(x)^*DG(x))^{1/2}$ . Now assume that there exists an integer  $d$  and constants  $\rho \geq \sigma > 0$ , such that for all  $x \in X$ ,*

$$(i) \quad \lambda_{d+1}(x) \leq \rho/2, \text{ and}$$

$$(ii) \quad \omega_j(x) \leq (\sigma/2)^{j-d} \text{ for } 1 \leq j \leq d.$$

Let  $\beta(x) := \max(2\lambda_{d+1}(x), \sigma) \in [\sigma, \rho]$ . Then there exist constants  $c > 1$  and  $\epsilon_0 > 0$ , such that for all  $x \in X$  and  $0 < \epsilon < \epsilon_0$ , the set  $G(B(x, \epsilon))$  can be covered by at most

$$c\beta(x)^{-d}$$

balls of radius  $\beta(x)\epsilon$ .

*Proof.* Let  $\beta := \beta(x) = \max(2\lambda_{d+1}(x), \sigma)$ . Then since  $2\lambda_{d+1} \leq \rho$  by (i), we must have that  $\beta \in [\sigma, \rho]$ . If  $\beta > \sigma$  take  $j = d$ , but if  $\beta = \sigma$  take  $j$  to be the smallest natural number such that  $2\lambda_{j+1} < \beta$ . This is a strict inequality, because if  $2\lambda_{d+1} = \beta$  were true, then  $2\lambda_{d+1} = \sigma$  since by assumption  $\beta = \sigma$ , and then we would take  $j = d$ . Thus if  $\beta = \sigma$  we can assume that at least  $2\lambda_{d+1} < \beta$ . The smallest natural number  $j$  such that  $2\lambda_{j+1} < \beta$  must thus satisfy  $j \leq d$ . Furthermore, by (ii) we have

$$\omega_j(x) \leq \left(\frac{\sigma}{2}\right)^{j-d} = \left(\frac{\beta}{2}\right)^{j-d},$$

since we either have  $j = d$  or  $\beta = \sigma$ .

Now let us look at the image of the unit ball under  $DG(x)$ . Let  $E = DG(x)[B(0, 1)]$ , then by Lemma 4.2.1, the lengths of the semi-axes of  $E$  are the  $\lambda_i(x)$ . Set  $\delta = \sigma/4$ , and let  $E'$  be the  $\delta$ -neighbourhood of  $E$ . Since  $2\lambda_{i+1} \leq \beta$  for every  $i \geq j$ , we have that every point of  $E'$  lies within  $(\beta/2 + \delta)$  of  $E \cap \Pi$ , where  $\Pi$  is the hyperplane spanned by the first  $j$  semi-axes of  $E$ . This is true since all the semi-axes not in the span of  $\Pi$  have lengths less than  $\beta/2$ , ( $\lambda_{i+1} \leq \beta/2$ ).

So if  $E \cap \Pi$  is covered by a collection of balls of radius  $(\beta/2 - \delta)$  with their centres in  $\Pi$ , then  $E'$  is covered by balls with the same centres but with radius  $\beta$ .

To estimate the number of balls to cover  $E \cap \Pi$ , we first cover it by a grid of  $j$ -dimensional boxes with sides of length  $(\beta - 2\delta)/\sqrt{j}$ . Then the number of balls is bounded by

$$N_b = \left( \frac{(2\lambda_1 + 2\delta)}{(\beta - 2\delta)/\sqrt{j}} + 1 \right) \cdot \left( \frac{(2\lambda_2 + 2\delta)}{(\beta - 2\delta)/\sqrt{j}} + 1 \right) \cdot \dots \cdot \left( \frac{(2\lambda_j + 2\delta)}{(\beta - 2\delta)/\sqrt{j}} + 1 \right). \quad (4.4)$$

The reason is that if  $r = (\beta/2 - \delta)$  (half of the length of the diagonal of the box) and  $s = (\beta - 2\delta)/\sqrt{j}$  (the side length of the box), then we need at most  $N_b$  boxes to cover the ellipse. We need the same number of balls of radius  $r$  to cover the ellipse, since

$$(2r)^2 = js^2 \implies s = (\beta - 2\delta)/\sqrt{j},$$

is the largest box that can be inscribed in a ball of radius  $r$ .

Now continuing with the estimate, since  $\beta \geq \sigma = 4\delta$ , we have  $\beta - 2\delta \geq (\beta + 2\delta)/3$ , since

$$\begin{aligned} 3\beta - 6\delta &= \beta + 2\beta - 6\delta \geq \beta + 2\delta \geq \beta + 2\delta \\ \implies \beta - 2\delta &\geq (\beta + 2\delta)/3. \end{aligned}$$

Thus

$$\left( \frac{(2\lambda_i + 2\delta)}{(\beta - 2\delta)/\sqrt{j}} + 1 \right) \leq \left( \frac{3\sqrt{j}(2\lambda_i + 2\delta)}{(\beta + 2\delta)} + 1 \right).$$

By the ordering of the  $\lambda_i$ , we have for  $1 \leq i \leq j$

$$(2\lambda_i + 2\delta) \geq (\beta + 2\delta),$$

so

$$\left( \frac{3\sqrt{j}(2\lambda_i + 2\delta)}{(\beta + 2\delta)} + \frac{(\beta + 2\delta)}{(\beta + 2\delta)} \right) \leq \frac{(3\sqrt{j} + 1)(2\lambda_i + 2\delta)}{\beta + 2\delta}.$$

But in general if  $a > b > 0$  and  $c > 0$ , then

$$\frac{(a + c)}{b + c} < \frac{a}{b},$$

so we can write

$$\frac{(3\sqrt{j} + 1)(2\lambda_i + 2\delta)}{\beta + 2\delta} \leq \frac{(3\sqrt{j} + 1)2\lambda_i}{\beta}.$$

Thus the number of balls  $N_b$  is bounded by

$$N_b \leq \frac{(3\sqrt{j} + 1)2\lambda_1}{\beta} \cdot \frac{(3\sqrt{j} + 1)2\lambda_2}{\beta} \cdot \dots \cdot \frac{(3\sqrt{j} + 1)2\lambda_j}{\beta} \quad (4.5)$$

$$= (6\sqrt{j} + 2)^j \frac{\omega_j}{\beta^j} \leq c \frac{\beta^{j-d}}{\beta^j} = c\beta^{-d}, \quad (4.6)$$

with  $c = (6\sqrt{j} + 2)^d$ .



Now let  $\{B_i(y)\}_{i=1}^M$  be a covering of balls of radius  $(\beta/2 - \delta)$  of  $E \cap \Pi$ , with  $y \in \Pi$ . The above estimate showed that  $N_b \leq c\beta^{-d}$ .

Now

$$DG(x)[B(x, \epsilon)] = DG(x)[x] + \epsilon DG(x)[B(0, 1)],$$

where  $+$  is taken in the sense of element-wise addition of vectors in a set. Thus if  $\tilde{B}_i(y) := DG(x)[x] + \epsilon DG(x)[B_i(y)]$ , then  $\{\tilde{B}_i(y)\}_{i=1}^M$  is a covering of at most  $c\beta^{-d}$  balls of radius  $\epsilon(\beta - \delta)/2$  of the set  $DG(x)[B(x, \epsilon)]$ .

Now since  $G$  is  $C^1$  on the compact set  $X$ , we can pick a small enough  $\epsilon_0$ , so that we have

$$|(G(x) + DG(x)u) - G(u)| \leq \delta\epsilon,$$

for all  $0 < \epsilon < \epsilon_0$  and  $u \in B(x, \epsilon)$ . Thus for this choice of  $\epsilon$  we have

$$\begin{aligned} G(B(x, \epsilon)) &\subset G(x) + \epsilon DG(x)[B(0, 1)] + \epsilon\delta B(0, 1) \\ &= G(x) + \epsilon E + \epsilon\delta B(0, 1) = G(x) + \epsilon E', \end{aligned}$$

where  $E$  and  $E'$  are defined as before. Now, since  $E'$  can be covered by at most  $c\beta^{-d}$  balls of radius  $\beta$ , we have that  $G(x) + \epsilon E'$  can be covered by at most  $c\beta^{-d}$  of radius  $\epsilon\beta$ . Thus we have the result that  $G(B(x, \epsilon))$  can be covered by at most  $c\beta^{-d}$  balls of radius  $\epsilon\beta$  too, for any  $0 < \epsilon < \epsilon_0$ .  $\square$

The above is the main Lemma that we will need in the next section. However, we still need some smaller results whose proofs can nevertheless be rather involved, which we will also use in the proof of the main theorem.

**Lemma 4.2.4** (Compact Gradients of Composite Operators). *Suppose  $F \in C^1(X)$  on some set  $X \subset H$ , and suppose that the gradient (linearisation)  $DF(x)$  of  $F$  is compact on  $H$  for every  $x \in X$ , where  $H$  is some Hilbert space. Then for any  $r \in \mathbb{N}$ ,  $D(F^r)(x)$  is also compact.*

*Proof.* By the chain rule we have

$$D(F^r)(x) = [DF(F^{r-1}(x))][DF(F^{r-2}(x))] \dots [DF(F(x))][DF(x)].$$

For any given  $x \in X$ , by the compactness of  $DF(x)$  we know that for any bounded sequence  $v_n$ , there exists a subsequence  $v_{n_j}$  and some element  $v$  in the space such that

$$\|DF(x)v_{n_j} - v\| \rightarrow 0$$

as  $j \rightarrow \infty$ . But then if we let  $w := [DF(F^{r-1}(x))][DF(F^{r-2}(x))] \dots [DF(F(x))]v$ , we have

$$\begin{aligned} & \|D(F^r(x))v_{n_j} - w\| \\ & \leq \left( \|[DF(F^{r-1}(x))][DF(F^{r-2}(x))] \dots [DF(F(x))]\| \right) \left( \|DF(x)v_{n_j} - v\| \right) \rightarrow 0 \end{aligned}$$

as  $j \rightarrow \infty$ . Thus  $D(F^r)(x)$  is compact. □

We now prove another lemma which will allow us to deduce some stricter bounds on the product of numbers.

**Lemma 4.2.5.** *Let  $\beta_0, \beta_1, \dots, \beta_j, \dots$  be a sequence of real numbers such that  $\beta_j \in [\sigma, \rho]$  for all  $j \geq 0$ , with  $0 < \sigma < \rho < 1$ . Then for every integer  $m \geq 1$ , there exists an integer  $0 \leq n \leq m - 1$ , such that*

$$\beta_n \cdot \beta_{n+1} \cdots \beta_{m-1} \in [\sigma \rho^m, \rho^m]. \quad (4.7)$$

*Proof.* Clearly  $\beta_0 \cdots \beta_{m-1} \leq \rho^m$ . Let  $n \leq m-1$  be the largest integer such that  $\beta_n \cdots \beta_{m-1} \leq \rho^m$ . We want to show that  $\sigma \rho^m \leq \beta_n \cdots \beta_{m-1}$ . Thus suppose for a contradiction that  $\sigma \rho^m > \beta_n \cdots \beta_{m-1}$  then we have

$$\left( \frac{\beta_n}{\sigma} \right) \beta_{n+1} \cdots \beta_{m-1} < \rho^m.$$

But  $\frac{\beta_n}{\sigma} \geq 1$ , and so

$$\beta_{n+1} \cdots \beta_{m-1} < \rho^m,$$

but this contradicts the assumption that  $n$  is the largest integer such that  $\beta_n \cdots \beta_{m-1} \leq \rho^m$ . Thus we must have  $\sigma \rho^m \leq \beta_n \cdots \beta_{m-1}$ , which gives the result. □

**Theorem 4.2.6** (Bound on Expansion Factors [Tem98, p. 345]). *Let  $Q, \Lambda : E \rightarrow E$  be compact, linear operators; then we have*

$$\omega_d(Q\Lambda) \leq \omega_d(Q)\omega_d(\Lambda). \quad (4.8)$$

*Proof.* The proof of this theorem requires tools and concepts from the theory of exterior algebras on Hilbert spaces. It is somewhat involved and requires a lot of machinery that, for the sake of brevity, we will not give in this thesis. For a proof and for all of the relevant background, see [Tem98, p. 345]. □

### 4.3 Bounding the Fractal Dimension of a Set

We are now ready to prove a very general theorem that tells us about the dimension of a set, if we know something about the long-time behaviour of expansion factors in Definition 4.2.2. From there, it is only a small step to get a result about the attractor dimension of the Nikolaevskiy equation. The proof of the theorem below was first given by Hunt in 1996 [Hun96]. The proof was modified from [Hun96] and reproduced the Appendix of [Rob01]. The proof below is reproduced from the Appendix of [Rob01] with some details added.

**Theorem 4.3.1** (Bound on the Fractal Dimension of a Set [Hun96]). *Let  $X$  be a compact set, and let  $F$  be a  $C^1$  map defined on a neighbourhood of  $X$  whose derivative  $DF$  is compact. Let  $\lambda_1 \geq \dots \geq \lambda_j \geq \dots \geq 0$  be the singular values of  $DF$ . Suppose also that  $X$  is invariant under  $F$ , that is  $F(X) = X$ . If for some integer  $d$ ,*

$$\omega_d(F, x) \leq \gamma \quad \text{for all } x \in X$$

for some  $0 < \gamma < 1$ , then we have

$$d_f(X) \leq d. \tag{4.9}$$

Here  $d_f(X)$  is the fractal dimension (4.3) of the set  $X$ .

*Proof.* Let  $\omega_d$  be defined as in the previous section in Definition 4.2.2. By assumption we have

$$\omega_{d+1}(F, x) \leq \gamma < 1$$

for all  $x \in X$ . Since  $X$  is invariant under  $F$ , by inequality (4.8) and Lemma 4.2.4 we get that for any  $r \geq 1$ ,

$$\omega_{d+1}(F^r, x) \leq \omega_{d+1}(F, x)\omega_{d+1}(F, F(x)) \dots \omega_{d+1}(F, F^{r-1}(x)) \leq \gamma^r.$$

Now since  $\lambda_{d+1}(x) \leq \lambda_j(x)$  for all  $j \leq d$ , we also have

$$\omega_{d+1}(F^r, x) \geq [\lambda_{d+1}(F^r, x)]^{d+1},$$

where  $\lambda_{d+1}(F^r, x)$  is the  $d + 1$ 'th singular value of  $D(F^r)$ . This implies that

$$\lambda_{d+1}(F^r, x) \leq \gamma^{r/(d+1)}.$$

Since  $\gamma < 1$ , we can pick  $r$  large enough, so that

$$\rho := 2\gamma^{r/(d+1)} < \frac{1}{2}.$$

This means that if we define  $G := F^r$ , then  $G$  satisfies condition (i) of Lemma 4.2.3.

To show condition (ii) of Lemma 4.2.3, note that by Lemma 4.2.4 and since  $F$  is  $C^1$ ,  $DG(x)$  is compact on  $X$ . We must therefore have that  $\omega_{d+1}(F, x)$  is a continuous function of  $x$  in  $X$ , and so it takes on its maximum on the compact set  $X$ , and hence we can choose  $\sigma > 0$  small enough, such that  $\omega_j(DG, x) \leq (\sigma/2)^{j-d}$  for  $j \leq d$ . Thus condition (ii) of Lemma 4.2.3 is also satisfied.

Define  $\beta(x) := \max\{2\lambda_{d+1}(x), \sigma\}$  as in Lemma 4.2.3. We will now iteratively cover  $X$  and obtain a bound on its dimension. Observe that on  $X$  we have  $\beta(x) > 0$  and in fact  $\beta(x) \geq h > 0$  for some  $h$ , since  $X$  is compact. This means that for every  $x \in X$ , there exists an  $\epsilon > 0$ , such that for all  $y$  with  $|y - x| < \epsilon$  we have

$$\frac{1}{2}\beta(y) \leq \beta(x) \leq 2\beta(y). \quad (4.10)$$

By the uniform continuity of  $\beta(x)$  we can choose an  $\epsilon$  such that the above holds for all  $x \in X$ . Thus let  $C_0$  be a covering of  $N_0$ , consisting of balls of radius  $\epsilon$  of  $X$ .

What we will do now is estimate the number of balls of radius  $(2\rho)^m \epsilon$ , that are needed to cover  $X$ . We will do this by covering  $X$  with some balls, then applying  $G$  to those balls, and then covering the resulting ellipses with smaller balls, to which we then apply Lemma 4.2.3, to get a covering of each ellipse made up of smaller balls. We repeat this process until we have balls that are small enough, and then we will estimate their number.

We begin with our cover  $C_0$  of balls of radius  $\epsilon$ , and apply Lemma 4.2.3 to the balls. We then know that the ellipse  $G(B(x_0, \epsilon))$  can be covered by at most  $c\beta(x_0)^{-d}$  balls of radius  $\beta(x_0)\epsilon$ , where  $x_0$  denotes a center of a ball in  $C_0$ . Let  $C_1$  be union of  $C_0$  and the cover of balls of radius  $\beta(x_0)\epsilon$  of all the ellipses  $G(B(x_0, \epsilon))$ . So the number of balls in  $C_1$  is  $N_1 \leq N_0 + \sum_{x_0} c\beta(x_0)^{-d}$ . For some  $x_0$ , we then apply Lemma 4.2.3 to the balls of radius  $\beta(x_0)\epsilon$ , and conclude that each set  $G(B(x_1, \beta(x_0)\epsilon))$  can be covered by at most  $c\beta(x_1)^{-d}$  balls of radius  $\beta(x_1)\beta(x_0)\epsilon$ .

We continue with these covers, until we arrive at a cover  $C_m$ , which can be covered by at most  $N_m \leq N_{m-1} + \sum_{x_m} c\beta(x_m)^{-d}$  balls of radius  $\beta(x_0) \cdot \dots \cdot \beta(x_j) \cdot \dots \cdot \beta(x_{m-1})\epsilon$ , where the  $x_j$  are the centres of the balls used in each stage of the construction. Note that  $C_i \subset C_j$  for  $i \leq j$ , since we take the union of  $C_{j-1}$  and the balls centred at  $x_j$  to construct  $C_j$ . We will now show that the sub-collection  $C'_m$  of balls in  $C_m$  whose radii lie in the range  $[(\rho/2)^m \sigma \epsilon, (2\rho)^m \epsilon]$ , still covers  $X$ . To see this, let  $y$  be a point in  $X$  and consider the points

$$y_0, y_1, \dots, y_m \quad \text{with} \quad y_m = y, \quad y_0 \in X \quad \text{and} \quad y_{j+1} = G(y_j).$$

The above is well defined, since we assumed that  $X$  is invariant under  $G$ . Since  $\beta(x) \in [\sigma, \rho]$ , for all  $x \in X$ , we can apply Lemma 4.2.5, and conclude that there exists an  $n$  with  $0 \leq n \leq$

$m - 1$  such that

$$\beta(y_n)\beta(y_{n+1}) \cdots \beta(y_{m-1}) \in [\sigma\rho^m, \rho^m]. \quad (4.11)$$

Now if we follow the trajectory of  $y_n$  under iterations of  $G$ , we get a sequence of balls as follows: Take  $B(x_0, \epsilon)$  to be a ball in  $C_0$  that contains  $y_n$ . Then choose  $x_1$  to be the center of a ball in  $C_1 \setminus C_0$  that contains  $y_{n+1}$ , and in general, let  $x_j$  be the center of a ball in  $C_j \setminus C_{j-1}$  that contains  $y_{n+j}$ . Then eventually we get a point  $x_{m-n}$ , which is the center of a ball of radius  $\beta(x_{m-n-1}) \cdots \beta(x_0)\epsilon$  in  $C_{m-n} \setminus C_{m-n-1} \subset C_m$  that contains  $y = y_m$ . Furthermore, we also have that  $|x_j - y_{n+j}| \leq \rho^j \epsilon \leq \epsilon$  for any  $0 \leq j \leq m - n$ . Thus by the continuity result (4.10), we have that for each  $0 \leq j \leq m - n$ ,

$$\frac{1}{2}\beta(y_{n+j}) \leq \beta(x_j) \leq 2\beta(y_{n+j}),$$

which implies that we have the bound

$$\beta(x_{m-n-1}) \cdots \beta(x_0)\epsilon \in [2^{n-m}\beta(y_n)\beta(y_{n+1}) \cdots \beta(y_{m-1})\epsilon, 2^{m-n}\beta(y_n)\beta(y_{n+1}) \cdots \beta(y_{m-1})\epsilon].$$

But then because of (4.11), we have

$$\beta(x_{m-n-1}) \cdots \beta(x_0)\epsilon \in [(\rho/2)^m \sigma \epsilon, (2\rho)^m \epsilon].$$

But this means that there is a ball whose radius is in  $[(\rho/2)^m \sigma \epsilon, (2\rho)^m \epsilon]$ , that contains  $y$ , and so we have found a ball in  $C'_m$  that contains  $y$  for an arbitrary point  $y \in X$ , so  $C'_m$  is a cover of  $X$ .

We will now look at how many balls of  $C'_m$  are needed to cover  $X$ . The following uses the weighted system from [Rob01, p. 443]. We assign a weight of one to the balls with centres  $x_0$  of the  $N_0$  balls in the original cover. For every  $x_0$ , we divide the weight evenly among all the balls centred at  $x_1$  that are needed to cover the ellipse formed from the ball centred at  $x_0$  under  $G$ . We continue in this way for each stage of the construction. Since we are dividing the weights equally, the sum of the weights for each collection of balls centred at  $x_n$  ( $1 \leq n \leq m$ ) is at most  $N_0$ , and the sum of the weights of all the balls in  $C_m$  is at most  $mN_0$ .

Since the weights are divided equally, an upper bound on the number of balls centred at  $x_{n+1}$  coming from a ball centred at  $x_n$  implies a lower bound on the weights of the balls at the  $n + 1$  level. The weight assigned to a ball at the  $n$ th stage is therefore at least

$$\frac{1}{c\beta(x_0)^d \cdots \beta(x_n - 1)^d} = \frac{(\beta(x_0)^d \cdots \beta(x_{n-1}))^d}{c^n}.$$

Now, if a ball centred at  $x_n$  lies in  $C'_m$  then its weight must be at least

$$\frac{((\rho/2)^m \sigma)^d}{c^n} \geq \frac{((\rho/2)^m \sigma)^d}{c^m},$$

and since the total number of weights of all the balls in  $C'_m$  cannot exceed  $mN_0$ , the number of balls in  $C'_m$  cannot exceed

$$\tilde{N}'_m = \frac{mN_0 c^m}{((\rho/2)^m \sigma)^d}.$$

We are now finally ready to estimate the dimension of  $X$ . We have that

$$\begin{aligned} d_f(X) &\leq \limsup_{m \rightarrow \infty} \frac{\log(\tilde{N}'_m)}{-\log((2\rho)^m \sigma \epsilon)} = \limsup_{m \rightarrow \infty} \frac{\log(mN_0 \sigma^{-d}) + m \log(c(\rho/2)^{-d})}{-\log(\sigma \epsilon) - m \log(2\rho)} \\ &= \frac{\log(c(\rho/2)^{-d})}{-\log(\rho/2)} = d \frac{\log(\rho/2) + \log(4)}{\log(2\rho)} - \frac{\log(c)}{\log(2\rho)} = d - \frac{\log(c) - \log(4)}{\log(2\rho)}. \end{aligned}$$

The last term can be made arbitrarily small by choosing  $\rho$  small enough, so we finally have the result,

$$d_f(X) \leq d. \tag{4.12}$$

□

## 4.4 Growth of Volumes

We will now return to the more concrete setting of a dynamical system, and derive an ODE that governs the evolution of volumes in phase space. This will then give us a quantity which is the analogue of the expansion factor in Definition 4.2.2, and which we can use in Theorem 4.3.1. The derivation we give in this section comes from Section 13.2 in [Rob01], but we have added some additional details, and we prove some lemmas that are not given in [Rob01].

Consider the dynamical system (4.1) and its linearisation (4.2). Then any displacement  $\delta x_0^{(j)}$  away from some initial point  $x_0$  will evolve according to

$$\frac{d\delta x^{(j)}}{dt} = Df(x) \delta x^{(j)}, \quad \delta x^{(j)}(0) = \delta x_0^{(j)}, \tag{4.13}$$

along the trajectory  $x(t)$ , where  $x(t)$  is the solution to the full system (4.1), with initial condition  $x(0) = x_0$ .

If we have  $n$  linearly independent initial displacements  $\delta x_0^{(j)}$ ,  $j = 1, \dots, n$ , then the volume of the parallelepiped spanned by these displacements is given by

$$V_n(0) = \|\delta x_0^{(1)} \wedge \dots \wedge \delta x_0^{(n)}\|, \quad (4.14)$$

where

$$\|\delta x_0^{(1)} \wedge \dots \wedge \delta x_0^{(n)}\|^2 := \det \left\{ (\delta x_0^{(i)}, \delta x_0^{(j)})_H \right\}, \quad (4.15)$$

and  $H$  is the Hilbert space in (4.1) [Tem98, p. 338]. Letting

$$M(t)_{ij} := (\delta x^{(i)}(t), \delta x^{(j)}(t))_H,$$

the volume  $V_n(t)$  is given by

$$V_n(t)^2 = \det(M(t)). \quad (4.16)$$

We will now try to derive a differential equation that governs the growth of  $V_n(t)$  directly, without having to know the  $\delta x^{(j)}$ 's explicitly. To do this, we will first prove two short lemmas.

**Lemma 4.4.1** ([Rob01, Exercise 13.8]). *Suppose  $M$  is a positive definite, self-adjoint matrix; then*

$$\log(\det(M)) = \text{Tr}(\log(M)), \quad (4.17)$$

where  $\text{Tr}$  denotes the trace of a matrix.

*Proof.* Let  $A := \log(M)$ , in the sense of [Cul66], which is defined because  $M$  is invertible. Note however that in general, the logarithm of a matrix is not unique. Fortunately, the trace and the determinant of the logarithm of a matrix are unique [Cul66], so the result we derive is well-defined. Then  $M = e^A$ . Note that if  $v_i$  is an eigenvector of  $A$ , with eigenvalue  $\lambda_i$ , then

$$Mv_i = e^A v_i = e^{\lambda_i} v_i,$$

so  $A$  and  $M$  have the same eigenvectors, and  $e^{\lambda_i}$  is the corresponding eigenvalue of  $M$ . Then since  $\text{Tr}(A) = \sum_i \lambda_i$ , and  $\det(e^A) = \prod_i e^{\lambda_i}$ , we have

$$\det(M) = \det(e^A) = \prod_i e^{\lambda_i} = e^{\sum_i \lambda_i} = e^{\text{Tr}(A)} = e^{\text{Tr}(\log(M))}.$$

Taking the logarithm of both sides gives the result. □

**Lemma 4.4.2** ([Rob01, Exercise 13.8]). *Suppose  $M(t)$  is a positive definite, self-adjoint matrix; then*

$$\frac{d}{dt} \text{Tr}(\log(M)) = \text{Tr}\left(M^{-1} \frac{dM}{dt}\right). \quad (4.18)$$

*Proof.* The quantity  $\det M(t)$  is just a function of  $t$ , and so we can differentiate its logarithm using the chain rule. Thus, differentiating (4.17) with respect to  $t$  gives

$$\frac{d}{dt} \text{Tr}(\log(M)) = \frac{d}{dt} \log(\det(M)) = \frac{1}{\det M} \frac{d}{dt} \det M.$$

By Jacobi's formula of Theorem B.2.5 we have

$$\frac{d}{dt} \det M = \det M \text{Tr}\left(M^{-1} \frac{dM}{dt}\right).$$

Putting the two together gives

$$\frac{d}{dt} \text{Tr}(\log(M)) = \frac{1}{\det M} \frac{d}{dt} \det M = \text{Tr}\left(M^{-1} \frac{dM}{dt}\right),$$

as desired. □

Using the above results, we can now proceed. Where no confusion can arise, we will usually not write that the quantities in question depend on time explicitly; however, in what follows the vectors and matrices do depend on  $t$ . The following derivation can be found in [Rob01, p. 336]. In order to get an equation that governs the evolution of  $V_n(t)$ , we start by taking the following derivative, using Lemmas 4.4.1 and 4.4.2:

$$\begin{aligned} \frac{d}{dt} \log(V_n(t)) &= \frac{1}{2} \frac{d}{dt} \log(V_n(t)^2) = \frac{1}{2} \frac{d}{dt} \log(\det(M(t))) \\ &= \frac{1}{2} \frac{d}{dt} \text{Tr}(\log(M(t))) = \frac{1}{2} \text{Tr}\left(M^{-1} \frac{dM(t)}{dt}\right). \end{aligned}$$

We will now find a suitable expression for  $\frac{dM}{dt}$  that will allow us to relate it back to the linearised operator in (4.2). Let  $\phi^{(i)}(t)$ ,  $i = 1, \dots, n$ , be an orthonormal basis for the time-dependent subspace spanned by the  $\delta x^{(i)}(t)$ . Then we define

$$m := \left\{ (\phi^{(i)}(t), \delta x^{(j)}(t))_H \right\}, \quad (4.19)$$



so that  $m$  is an  $n \times n$  matrix whose components are  $m_{ij} = (\phi^{(i)}(t), \delta x^{(j)}(t))_H$ . Now since the  $\phi^{(i)}$  are orthonormal and span the same subspace as the  $\delta x^{(i)}(t)$ , we can write

$$\delta x^{(i)} = \sum_{l=1}^n (\delta x^{(i)}, \phi^{(l)})_H \phi^{(l)}.$$

Thus we have

$$\begin{aligned} M_{ij} &= (\delta x^{(i)}, \delta x^{(j)})_H = \sum_{l=1}^n (\delta x^{(i)}, \phi^{(l)})_H (\phi^{(l)}, \delta x^{(j)})_H \\ &= \sum_{l=1}^n (m^T)_{il} m_{lj} = (m^T m)_{ij}, \end{aligned}$$

which shows that

$$M = m^T m, \tag{4.20}$$

and  $M^{-1} = m^{-1}(m^T)^{-1}$ . Furthermore, if we look at the derivative of  $M$ , by the product rule we have

$$\begin{aligned} \frac{dM_{ij}}{dt} &= \left( \frac{d}{dt} \delta x^{(i)}(t), \delta x^{(j)}(t) \right)_H + \left( \delta x^{(i)}(t), \frac{d}{dt} \delta x^{(j)}(t) \right)_H \\ &= \left( Df(x) \delta x^{(i)}(t), \delta x^{(j)}(t) \right)_H + \left( \delta x^{(i)}(t), Df(x) \delta x^{(j)}(t) \right)_H, \end{aligned}$$

by (4.13). We can express the displacements in terms of the orthonormal basis by projecting  $\delta x^{(i)}$  onto the vectors  $\phi^{(j)}$ :

$$\delta x^{(i)} = \sum_{j=1}^n (\delta x^{(i)}, \phi^{(j)})_H \phi^{(j)}.$$

We now define

$$a_{ij} := (\phi^{(i)}, Df(x) \phi^{(j)})_H, \tag{4.21}$$

with which we can get

$$\begin{aligned}
\frac{dM_{ij}}{dt} &= \left( \sum_{l=1}^n Df(x)\phi^{(l)}(\phi^{(l)}, \delta x^{(i)})_H, \sum_{k=1}^n \phi^{(k)}(\phi^{(k)}, \delta x^{(j)})_H \right)_H \\
&\quad + \left( \sum_{l=1}^n \phi^{(l)}(\phi^{(l)}, \delta x^{(i)})_H, \sum_{k=1}^n Df(x)\phi^{(k)}(\phi^{(k)}, \delta x^{(j)})_H \right)_H \\
&= \sum_{l=1}^n \sum_{k=1}^n (\phi^{(l)}, \delta x^{(i)})_H \left( Df(x)\phi^{(l)}, \phi^{(k)} \right) (\phi^{(k)}, \delta x^{(j)})_H \\
&\quad + \sum_{l=1}^n \sum_{k=1}^n (\phi^{(l)}, \delta x^{(i)})_H \left( \phi^{(l)}, Df(x)\phi^{(k)} \right) (\phi^{(k)}, \delta x^{(j)})_H \\
&= \sum_{l=1}^n \sum_{k=1}^n (\phi^{(l)}, \delta x^{(i)})_H [a_{kl} + a_{lk}] (\phi^{(k)}, \delta x^{(j)})_H.
\end{aligned}$$

Thus

$$\frac{dM}{dt} = m^T (a^T + a) m. \quad (4.22)$$

Thus using (4.20) and (4.22), we have that

$$\begin{aligned}
2 \frac{d}{dt} \log(V_n(t)) &= \text{Tr} \left( M^{-1} \frac{dM}{dt} \right) = \text{Tr} \left( m^{-1} (m^T)^{-1} m^T (a^T + a) m \right) \\
&= \text{Tr} \left( m^{-1} (a^T + a) m \right) = \text{Tr} \left( m m^{-1} (a^T + a) \right) = \text{Tr} (a^T + a) = 2 \text{Tr} (a),
\end{aligned}$$

since for any  $n \times n$  matrices  $A$  and  $B$  one has  $\text{Tr}(AB) = \text{Tr}(BA)$ . So

$$\frac{d}{dt} \log(V_n(t)) = \text{Tr}(a). \quad (4.23)$$

Now consider the projection

$$P_n := \sum_{i=1}^n \phi^{(i)}(\phi^{(i)}, \cdot)_H,$$

that is  $P_n v$  is the projection of any  $v \in H$  into the subspace spanned by the  $\phi^{(j)}$  (and the  $\delta x^{(j)}$ , since they span the same space).

We define the **trace** of a linear operator  $A$  on  $H$  as

$$\text{Tr}(A) := \sum_{i=1}^{\infty} (\phi^{(i)}, A\phi^{(i)})_H, \quad (4.24)$$

where  $\phi^{(i)}$  is an orthonormal basis of  $H$ . The above is well-defined since it is independent of our choice of orthonormal basis. To see this let  $\psi^{(i)}$  be another orthonormal basis for  $H$ ; then we have  $\psi^{(i)} = \sum_{j=1}^{\infty} (\phi^{(j)}, \psi^{(i)})_H \phi^{(j)}$ , so then since we also have  $\phi^{(i)} =$

$\sum_{j=1}^{\infty} (\psi^{(j)}, \phi^{(i)})_H \psi^{(j)}$  we get

$$\begin{aligned}
Tr(A) &:= \sum_{i=1}^{\infty} (\psi^{(i)}, A\psi^{(i)})_H = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sum_{l=1}^{\infty} (\phi^{(j)}, \psi^{(i)})_H (\phi^{(l)}, \psi^{(i)})_H (\phi^{(l)}, A\phi^{(j)})_H \\
&= \sum_{l=1}^{\infty} \sum_{j=1}^{\infty} (\phi^{(l)}, \phi^{(j)})_H (\phi^{(l)}, A\phi^{(j)})_H \\
&= \sum_{j=1}^{\infty} (\phi^{(j)}, A\phi^{(j)})_H.
\end{aligned}$$

Therefore the value of  $Tr(A)$  is the same no matter which orthonormal basis we choose. This means that we have

$$\begin{aligned}
Tr(Df(x)P_n) &= \sum_{i=1}^{\infty} (\phi^{(i)}, Df(x)P_n\phi^{(i)})_H \\
&= \sum_{i=1}^{\infty} \sum_{j=1}^n (\phi^{(i)}, Df(x)(\phi^{(j)}, \phi^{(i)})_H \phi^{(j)})_H \\
&= \sum_{i=1}^{\infty} \sum_{j=1}^n (\phi^{(i)}, Df(x)\phi^{(j)})_H (\phi^{(j)}, \phi^{(i)})_H \\
&= \sum_{j=1}^n (\phi^{(j)}, Df(x)\phi^{(j)})_H.
\end{aligned}$$

Now, for the matrix  $a$  we have

$$\begin{aligned}
Tr(a) &= \sum_{i=1}^n (\phi^{(i)}, Df(x)\phi^{(i)})_H = \sum_{i=1}^n \sum_{j=1}^n (\phi^{(i)}, Df(x)\phi^{(j)})_H (\phi^{(i)}, \phi^{(j)})_H \\
&= \sum_{j=1}^n \sum_{i=1}^n (\phi^{(i)}(\phi^{(i)}, Df(x)\phi^{(j)})_H, \phi^{(j)})_H \\
&= \sum_{j=1}^n (\phi^{(j)}, Df(x)\phi^{(j)})_H = Tr(Df(x)P_n).
\end{aligned}$$

So we get that

$$\frac{d}{dt} \log(V_n(t)) = Tr(Df(x)P_n),$$

which gives us the differential equation

$$\frac{d}{dt} V_n(t) = V_n(t) Tr(Df(x)P_n), \quad (4.25)$$

whose solution is

$$V_n(t) = V_n(0)e^{\int_0^t \text{Tr}(Df(x(s))P_n) ds}. \quad (4.26)$$

Of course the exponent can be negative or positive for different values of  $t$ , so we have to look at the long-time average of the exponent to determine whether how the volume grows as  $t$  gets large. We thus look at

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t \text{Tr}(Df(x(s))P_n) ds. \quad (4.27)$$

If the integrand is positive on average, then the volume will grow, but if the integrand is negative on average, then the volume will decay as  $t \rightarrow \infty$ . To get an upper bound on this growth rate of volumes along trajectories starting in some set  $\mathcal{A}$ , we define

$$\mathcal{TR}_n(\mathcal{A}) := \sup_{x_0 \in \mathcal{A}} \sup_{P_n(0)} \limsup_{t \rightarrow \infty} \frac{1}{t} \int_0^t \text{Tr}(Df(x(s))P_n) ds, \quad (4.28)$$

where  $x_0$  is the initial condition of  $x(t)$ , and the supremum over  $P_n(0)$  is a supremum over  $n$ -dimensional projections into  $H$ . The quantity  $\mathcal{TR}_n(\mathcal{A})$  can be thought of as the maximum growth rate of any  $n$ -dimensional volume in  $H$  which evolves along trajectories of (4.1).

Note that since we want to know how much a volume expands in phase space, the quantity

$$e^{\mathcal{TR}_n(\mathcal{A})}$$

plays the role of the expansion factor  $\omega_n(x)$  of Definition 4.2.2. To be more precise, it is an upper bound on  $\omega_n(x)$ . Thus if  $\mathcal{TR}_n(\mathcal{A}) < 0$ , then we can define  $\gamma$  in Theorem 4.3.1 to be  $\gamma := e^{\mathcal{TR}_n(\mathcal{A})} < 1$ , giving a uniform upper bound on  $\omega_n(x)$  less than one.

## 4.5 Uniform Differentiability

In the following three sections, we will restrict our estimates to  $r \leq 1$ . We made this choice, since all the bounds we obtain will involve the  $\dot{L}^2$  norm of  $u$  on the attractor  $\mathcal{A}$ . The calculations for  $r > 1$  are essentially the same as for  $r \leq 1$ , but since there is no published bound on the  $r$ -dependence of the  $\dot{L}^2$  norm of  $u$  for  $r < 1$ , we will omit this case.

Two more ingredients that we need in order to get a bound on the dimension of the attractor of a dynamical system are uniform differentiability of the flow and compactness of the linearised flow. The reason we need these ingredients is so that we can satisfy the conditions of Theorem 4.3.1, which we want to use to get a bound on the fractal dimension of the attractor.

We say that the flow  $S(t)$  of (4.1) is **uniformly differentiable** on  $\mathcal{A}$ , if for every  $u_0 \in \mathcal{A}$  there exists a linear operator  $\Lambda(t, u_0)$  such that for all  $t \geq 0$ ,

$$\sup_{u_0, v_0 \in \mathcal{A} ; 0 < \|u_0 - v_0\| \leq \epsilon} \frac{\|S(t)v_0 - S(t)u_0 - \Lambda(t, u_0)(v_0 - u_0)\|}{\|v_0 - u_0\|} \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0 \quad (4.29)$$

and

$$\sup_{u_0 \in \mathcal{A}} \|\Lambda(t, u_0)\|_{\text{op}} < \infty \quad \text{for each } t \geq 0.$$

In the above  $\|\cdot\|$  is the norm on the Hilbert space  $H$  on which (4.1) is defined, and  $\|\cdot\|_{\text{op}}$  is the induced operator norm. The operator  $\Lambda(t, u_0)$  should be thought of as the flow of the linearised dynamical system (4.2), or more concretely, the flow of the linearisation of the Nikolaevskiy equation (1.4). Thus proving uniform differentiability is equivalent to proving that an equation can be linearised and that the linearised problem is well-posed.

In what follows, we will be proving uniform differentiability for the Nikolaevskiy equation (1.4). The approach we are taking has been developed for the 2D Navier-Stokes equations and is based on Section 13.4 of [Rob01], following the suggestions of [Rob01, Section 17.4]. We adapted this approach to the Nikolaevskiy equation.

The linearisation of (1.4) about a solution  $u$  is given by

$$U_t = \partial_x^6 U + 2\partial_x^4 U + (1-r)\partial_x^2 U - B(U, u) - B(u, U), \quad (4.30)$$

and its linear operator is then

$$\mathcal{G}_u := \partial_x^6 + 2\partial_x^4 + (1-r)\partial_x^2 - B(\cdot, u) - B(u, \cdot), \quad (4.31)$$

where  $B(u, U) := u \partial_x U$  as in (2.21). We will use the linearised equation to show that the flow of the full equation is uniformly differentiable. Essentially, this is equivalent to rigorously showing that (4.30) is the linearisation of equation (1.4).

We consider  $r \leq 1$ . We begin by defining  $\theta(t) := S(t)u_0 - S(t)v_0 - \Lambda(t, u_0)(u_0 - v_0) = u(t) - v(t) - U(t)$ , where  $u$  and  $v$  satisfy the full equation (1.4) with initial conditions  $u_0$  and  $v_0$ , and  $U$  satisfies the linearised equation (4.30), with  $U_0 = u_0 - v_0$ . Then we derive the equation that  $\theta$  satisfies as follows: We insert each term in  $\theta$  into its respective equation and add them together in the order in which they appear in  $\theta$ , to get

$$\frac{d\theta}{dt} = \mathcal{L}\theta - B(u, u) + B(v, v) + B(u, U) + B(U, u),$$

where  $\mathcal{L}$  is the operator (2.17). Upon adding zero in strategic places we get

$$\begin{aligned}\frac{d\theta}{dt} &= \mathcal{L}\theta - B(u, \theta) - B(u, v) + B(v, v) + B(U, u) \\ &= \mathcal{L}\theta - B(u, \theta) - B(\theta, u) + B(u, u) - B(u, v) - B(v, u) + B(v, v) \\ &= \mathcal{L}\theta - B(u, \theta) - B(\theta, u) + B(u - v, u - v).\end{aligned}$$

As in Section 2.5, we define  $w := u - v$ . Now take the inner product of the above with  $\theta$ , and use the bounds on the trilinear form in Proposition B.6.3 to get

$$\begin{aligned}\frac{1}{2} \frac{d}{dt} \|\theta\|_{\dot{L}^2}^2 + \|\partial_x^3 \theta\|_{\dot{L}^2}^2 + (1-r) \|\partial_x \theta\|_{\dot{L}^2}^2 &\leq 2 \|\partial_x^2 \theta\|_{\dot{L}^2}^2 + \ell^{1/2} \|u_0\|_{\dot{L}^2} e^{C_0 t} \|\theta\|_{\dot{L}^2} \|\partial_x^2 \theta\|_{\dot{L}^2} \\ &\quad + \ell^{1/2} \|w\|_{\dot{L}^2} \|\partial_x^2 w\|_{\dot{L}^2} \|\theta\|_{\dot{L}^2},\end{aligned}\tag{4.32}$$

where we have used the identities on the trilinear term, as well as the fact that  $\|u(t)\|_{\dot{L}^2} \leq \|u_0\|_{\dot{L}^2} e^{C_0 t}$  for some (possibly parameter-dependent)  $C_0$ , which we derived in Section 2.3 where we derived uniform bounds on Galerkin solutions. Since  $r \leq 1$  we can drop the second-order term with the  $(1-r)$  coefficient. We then use our Young's inequality as well as inequality (B.8) with  $\epsilon = 4$  in the following estimate:

$$\begin{aligned}\frac{1}{2} \frac{d}{dt} \|\theta\|_{\dot{L}^2}^2 + \|\partial_x^3 \theta\|_{\dot{L}^2}^2 &\leq 2 \|\partial_x^2 \theta\|_{\dot{L}^2}^2 + \frac{\beta \ell^{1/2} \|u_0\|_{\dot{L}^2} e^{C_0 t}}{2} \|\theta\|_{\dot{L}^2}^2 + \frac{\ell^{1/2} \|u_0\|_{\dot{L}^2} e^{C_0 t}}{2\beta} \|\partial_x^2 \theta\|_{\dot{L}^2}^2 \\ &\quad + \frac{\ell^{1/2}}{2} \|\theta\|_{\dot{L}^2}^2 + \frac{\ell^{1/2}}{2} \|w\|_{\dot{L}^2}^2 \|\partial_x^2 w\|_{\dot{L}^2}^2 \\ &= \left(2 + \frac{\ell^{1/2} \|u_0\|_{\dot{L}^2} e^{C_0 t}}{2\beta}\right) \|\partial_x^2 \theta\|_{\dot{L}^2}^2 + \left(\frac{\ell^{1/2}}{2} + \frac{\beta \ell^{1/2} \|u_0\|_{\dot{L}^2} e^{C_0 t}}{2}\right) \|\theta\|_{\dot{L}^2}^2 \\ &\quad + \frac{\ell^{1/2}}{2} \|w\|_{\dot{L}^2}^2 \|\partial_x^2 w\|_{\dot{L}^2}^2 \\ &\leq \left(\frac{1}{2} + \frac{\ell^{1/2} \|u_0\|_{\dot{L}^2} e^{C_0 t}}{8\beta}\right) \|\partial_x^3 \theta\|_{\dot{L}^2}^2 + \left(C(t) + \frac{\ell^{1/2}}{2} + \frac{\beta \ell^{1/2} \|u_0\|_{\dot{L}^2} e^{C_0 t}}{2}\right) \|\theta\|_{\dot{L}^2}^2 \\ &\quad + \frac{\ell^{1/2}}{2} \|w\|_{\dot{L}^2}^2 \|\partial_x^2 w\|_{\dot{L}^2}^2,\end{aligned}$$

for arbitrary  $\beta > 0$  where

$$C(t) = \left(8 + \frac{2\ell^{1/2} \|u_0\|_{\dot{L}^2} e^{C_0 t}}{\beta}\right).$$

Choosing  $\beta = \ell^{1/2} \|u_0\|_{\dot{L}^2} e^{C_0 t}$  (technically Young's inequality is stated with  $\beta$  as a constant, but it also works for nonconstant  $\beta$ , as long as it remains positive), we can drop all the

$\|\partial_x^3 \theta\|_{\dot{L}^2}$  terms; and if we define:

$$G(t) := 2 \left( C(t) + \frac{\ell^{1/2}}{2} + \frac{\ell \|u_0\|_{\dot{L}^2}^2 e^{2C_0 t}}{2} \right)$$

we have

$$\frac{d}{dt} \|\theta\|_{\dot{L}^2}^2 \leq G \|\theta\|_{\dot{L}^2}^2 + C_2 \|w\|_{\dot{L}^2}^2 \|\partial_x^2 w\|_{\dot{L}^2}^2, \quad (4.33)$$

where  $C_2 = \ell^{1/2}$ . Using Gronwall's inequality we get

$$\|\theta(t)\|_{\dot{L}^2}^2 \leq C_2 e^{\int_0^t G(s) ds} \int_0^t \|w(s)\|_{\dot{L}^2}^2 \|\partial_x^2 w(s)\|_{\dot{L}^2}^2 e^{-\int_0^s G(t') dt'} ds,$$

since  $\theta(0) = 0$ . Thus since  $e^{-\int_0^t G(t') dt'} \leq 1$  we get

$$\begin{aligned} \|\theta(t)\|_{\dot{L}^2}^2 &\leq C_2 e^{\int_0^t G(s) ds} \int_0^t \|w(s)\|_{\dot{L}^2}^2 \|\partial_x^2 w(s)\|_{\dot{L}^2}^2 ds \\ &\leq \tilde{C}_2 e^{\int_0^t G(s) ds} \int_0^t \|w(s)\|_{\dot{L}^2}^2 \|\partial_x^3 w(s)\|_{\dot{L}^2}^2 ds, \end{aligned} \quad (4.34)$$

where we have used a Poincaré inequality (B.12), and  $\tilde{C}_2 := (\ell/2\pi)^2 C_2$ .

Recall that during the proof of uniqueness of solutions of the Nikolaevskiy equation (see equation (2.25) in Section 2.5), we got an estimate similar to

$$\|w\|_{\dot{L}^2}^4 \leq \|w_0\|_{\dot{L}^2}^4 e^{C_3 t},$$

which we got from the inequality (2.23)

$$\frac{d}{dt} \|w\|_{\dot{L}^2}^2 + C_4 \|\partial_x^3 w\|_{\dot{L}^2}^2 \leq C_5 (C_6 + 2 \|u\|_{\dot{L}^2}^2 + \|v\|_{\dot{L}^2}^2) \|w\|_{\dot{L}^2}^2,$$

where  $0 \leq C_5 (C_6 + 2 \|u\|_{\dot{L}^2}^2 + \|v\|_{\dot{L}^2}^2) \leq C_3$  is some constant to bound the time-dependent exponent in (2.25). We multiply the above expression by  $\|w\|_{\dot{L}^2}^2$  to get:

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|w\|_{\dot{L}^2}^4 + C_4 \|\partial_x^3 w\|_{\dot{L}^2}^2 \|w\|_{\dot{L}^2}^2 &\leq C_5 (C_6 + 2 \|u\|_{\dot{L}^2}^2 + \|v\|_{\dot{L}^2}^2) \|w\|_{\dot{L}^2}^4 \\ &\leq C_5 \left( C_6 + 2 \|u\|_{L^\infty(0,T;\dot{L}^2)}^2 + \|v\|_{L^\infty(0,T;\dot{L}^2)}^2 \right) \|w_0\|_{\dot{L}^2}^4 e^{C_3 t}, \end{aligned}$$

which we integrate over time to get

$$\begin{aligned} & \frac{1}{2} \|w\|_{\dot{L}^2}^4 + C_4 \int_0^t \left\| \partial_x^3 w \right\|_{\dot{L}^2}^2 \|w\|_{\dot{L}^2}^2 ds \\ & \leq \underbrace{\left[ \frac{C_5}{C_3} \left( C_6 + 2\|u\|_{L^\infty(0,T;\dot{L}^2)}^2 + \|v\|_{L^\infty(0,T;\dot{L}^2)}^2 \right) (e^{C_3 t} - 1) + \frac{1}{2} \right]}_{=: C_4 G_2(t)} \|w_0\|_{\dot{L}^2}^4. \end{aligned}$$

We can drop the  $\|w\|_{\dot{L}^2}^4$  term to the left, and get

$$\int_0^t \left\| \partial_x^3 w \right\|_{\dot{L}^2}^2 \|w\|_{\dot{L}^2}^2 ds \leq G_2(t) \|w_0\|_{\dot{L}^2}^4,$$

which we insert into (4.34), to get

$$\|\theta(t)\|_{\dot{L}^2}^2 \leq \tilde{C}_2 e^{\int_0^t G(s) ds} G_2(t) \|w_0\|_{\dot{L}^2}^4.$$

Taking the square root of both sides and dividing by  $\|w_0\|_{\dot{L}^2}$  gives

$$\frac{\|u - v - U\|_{\dot{L}^2}}{\|u_0 - v_0\|_{\dot{L}^2}} \leq \sqrt{C_2 e^{\int_0^t G(s) ds} G_2(t)} \|u_0 - v_0\|_{\dot{L}^2}.$$

Now taking the limit as  $v_0 \rightarrow u_0$  shows the uniform differentiability of the flow of the Nikolavskiy equation (1.4), where  $U(t)$  plays the role of  $\Lambda(t, u_0)(u_0 - v_0)$  in (4.29).

## 4.6 Compactness

In this section we will show that the flow of the linearised equation  $\Lambda(t, u_0)$  (4.30) is compact on  $\mathcal{A}$ . We will do this by showing that any uniformly bounded sequence  $\tilde{U}_n \in \dot{L}^2$ , the domain of  $\Lambda(t, u_0)$ , maps to a uniformly bounded sequence in  $\dot{H}^1$  under  $\Lambda(t, u_0)$  for any finite  $t$ . We will then use the Rellich-Kondrachov compactness theorem to conclude that the sequence  $\Lambda(t, u_0)\tilde{U}_n$  must be pre-compact in  $\dot{L}^2$ .



In order to show compactness, we again consider the case  $r \leq 1$ . We then take the inner product of equation (4.30) with  $U$  to get

$$\begin{aligned}
\frac{1}{2} \frac{d}{dt} \|U\|_{\dot{L}^2}^2 &= - \left\| \partial_x^3 U \right\|_{\dot{L}^2}^2 + 2 \left\| \partial_x^2 U \right\|_{\dot{L}^2}^2 - (1-r) \left\| \partial_x u \right\|_{\dot{L}^2}^2 - \int_0^\ell U \partial_x u U + u \partial_x U U \, dx \\
&\leq - \left\| \partial_x^3 U \right\|_{\dot{L}^2}^2 + 2 \left\| \partial_x^2 U \right\|_{\dot{L}^2}^2 + \int_0^\ell U u \partial_x U \, dx \\
&\leq - \left\| \partial_x^3 U \right\|_{\dot{L}^2}^2 + 2 \left\| \partial_x^2 U \right\|_{\dot{L}^2}^2 + \ell^{1/2} \left\| \partial_x u \right\|_{\dot{L}^2} \|U\|_{\dot{L}^2} \left\| \partial_x U \right\|_{\dot{L}^2} \\
&\leq - \left\| \partial_x^3 U \right\|_{\dot{L}^2}^2 + 2 \left\| \partial_x^2 U \right\|_{\dot{L}^2}^2 + \frac{C\alpha}{2} \|U\|_{\dot{L}^2}^2 + \frac{C}{2\alpha} \left\| \partial_x U \right\|_{\dot{L}^2}^2,
\end{aligned}$$

where  $C = \ell^{1/2} R_1$ , where  $R_1$  is the bound on  $\partial_x u$  in  $\dot{L}^2$  from (3.4). Picking  $\alpha = C/2$  and using Young's inequality and inequality (B.8) we get

$$\begin{aligned}
\frac{1}{2} \frac{d}{dt} \|U\|_{\dot{L}^2}^2 &\leq - \left\| \partial_x^3 U \right\|_{\dot{L}^2}^2 + 3 \left\| \partial_x^2 U \right\|_{\dot{L}^2}^2 + \left( \frac{C^2}{4} + \frac{1}{4} \right) \|U\|_{\dot{L}^2}^2 \\
&\leq - \left\| \partial_x^3 U \right\|_{\dot{L}^2}^2 + \frac{3}{\eta} \left\| \partial_x^3 U \right\|_{\dot{L}^2}^2 + \left( \frac{3\eta^2}{4} + \frac{C^2}{4} + \frac{1}{4} \right) \|U\|_{\dot{L}^2}^2. \tag{4.35}
\end{aligned}$$

So choosing  $\eta = 3$  we get

$$\frac{d}{dt} \|U\|_{\dot{L}^2}^2 \leq C_1 \|U\|_{\dot{L}^2}^2,$$

where  $C_1 = \left( 27 + \frac{C^2}{2} + \frac{1}{2} \right)$ , which by Gronwall's inequality implies that

$$\|U(t)\|_{\dot{L}^2}^2 \leq \|U(0)\|_{\dot{L}^2}^2 e^{C_1 t}, \tag{4.36}$$

for any  $0 \leq t$ , so that  $U \in L^\infty(0, T; \dot{L}^2)$ . Returning once more to inequality (4.35), choosing  $\eta = 6$  and  $T > 0$  and integrating over time from 0 to  $T$  we get

$$\|U(T)\|_{\dot{L}^2}^2 + \int_0^T \left\| \partial_x^3 U \right\|_{\dot{L}^2}^2 \, dt \leq \left( 108 + \frac{C^2}{2} + \frac{1}{2} \right) \int_0^T \|U\|_{\dot{L}^2}^2 \, dt + \|U(0)\|_{\dot{L}^2}^2.$$

Thus, using (4.36), we have

$$\int_0^T \left\| \partial_x^3 U \right\|_{\dot{L}^2}^2 \, dt \leq \underbrace{\left[ \frac{1}{C_1} \left( 108 + \frac{C^2}{2} + \frac{1}{2} \right) e^{C_1 T} - \frac{1}{C_1} \left( 108 + \frac{C^2}{2} + \frac{1}{2} \right) + 1 \right]}_{\tilde{C}} \|U(0)\|_{\dot{L}^2}^2. \tag{4.37}$$

This implies that  $U \in L^2(0, T; \dot{H}^3)$ . We now just need one more bound on  $U$  to get the result. We take the inner product of equation (4.30) with  $-\partial_x^2 U$  to get

$$\frac{1}{2} \frac{d}{dt} \|\partial_x U\|_{\dot{L}^2}^2 + \|\partial_x^4 U\|_{\dot{L}^2}^2 - 2 \|\partial_x^3 U\|_{\dot{L}^2}^2 + (1-r) \|\partial_x^2 U\|_{\dot{L}^2}^2 = b(u, U, \partial_x^2 U) + b(U, u, \partial_x^2 U).$$

Dropping the  $(1-r)$  term and applying our inequalities on the trilinear form given in Proposition B.6.3, we get

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|\partial_x U\|_{\dot{L}^2}^2 &\leq - \|\partial_x^4 U\|_{\dot{L}^2}^2 + 2 \|\partial_x^3 U\|_{\dot{L}^2}^2 + \int_0^\ell (u \partial_x U + U \partial_x u) \partial_x^2 U \, dx \\ &= - \|\partial_x^4 U\|_{\dot{L}^2}^2 + 2 \|\partial_x^3 U\|_{\dot{L}^2}^2 - \int_0^\ell u U \partial_x^3 U \, dx \\ &\leq - \|\partial_x^4 U\|_{\dot{L}^2}^2 + 2 \|\partial_x^3 U\|_{\dot{L}^2}^2 + C_2 \|\partial_x U\|_{\dot{L}^2} \|\partial_x^3 U\|_{\dot{L}^2} \\ &\leq - \|\partial_x^4 U\|_{\dot{L}^2}^2 + 2 \|\partial_x^3 U\|_{\dot{L}^2}^2 + \frac{C_2}{2\beta} \|\partial_x U\|_{\dot{L}^2}^2 + \frac{C_2}{2} \beta \|\partial_x^3 U\|_{\dot{L}^2}^2, \end{aligned}$$

where  $C_2 = \ell^{1/2} R$ , and we have used Young's inequality. That is,  $\ell^{1/2} \|u\|_{\dot{L}^2} \leq C_2$ . Then we can use (B.3.5) to get

$$\frac{1}{2} \frac{d}{dt} \|\partial_x U\|_{\dot{L}^2}^2 + \|\partial_x^4 U\|_{\dot{L}^2}^2 \leq \frac{1}{\eta} \left( \frac{\beta C_2}{2} + 2 \right) \|\partial_x^4 U\|_{\dot{L}^2}^2 + \left[ \frac{\eta^2}{4} \left( \frac{\beta C_2}{2} + 2 \right) + \frac{C_2}{2\beta} \right] \|\partial_x U\|_{\dot{L}^2}^2. \quad (4.38)$$

If we let  $\eta = 3$  and choose  $\beta = 2/C_2$ , the  $\|\partial_x^4 U\|_{\dot{L}^2}^2$  terms cancel and we get

$$\frac{d}{dt} \|\partial_x U\|_{\dot{L}^2}^2 \leq \left( \frac{27}{4} + \frac{C_2^2}{4} \right) \|\partial_x U\|_{\dot{L}^2}^2.$$

Letting  $C_3 := \left( \frac{27}{4} + \frac{C_2^2}{4} \right)$ , we integrate the above expression over time from  $t_0$  to  $t$  to get

$$\|\partial_x U(t)\|_{\dot{L}^2}^2 \leq C_3 \int_{t_0}^t \|\partial_x U\|_{\dot{L}^2}^2 \, dt + \|\partial_x U(t_0)\|_{\dot{L}^2}^2.$$

Now using Poincaré's inequality with constant  $C_4 := (\ell/2\pi)^2$  and using inequality (4.37) we get

$$\|\partial_x U(t)\|_{\dot{L}^2}^2 \leq C_3 C_4^2 \tilde{C} \|U(0)\|_{\dot{L}^2}^2 + \|\partial_x U(t_0)\|_{\dot{L}^2}^2,$$

which we can integrate over  $t_0$  from 0 to  $T$  to get

$$\begin{aligned}
T \|\partial_x U(t)\|_{\dot{L}^2}^2 &\leq TC_3 C_4^2 \tilde{C} \|U(0)\|_{\dot{L}^2}^2 + \int_0^T \|\partial_x U(t_0)\|_{\dot{L}^2}^2 dt_0 \\
&\leq TC_3 C_4^2 \tilde{C} \|U(0)\|_{\dot{L}^2}^2 + C_4^2 \int_0^T \|\partial_x^3 U(t_0)\|_{\dot{L}^2}^2 dt_0 \\
&\leq TC_3 C_4^2 \tilde{C} \|U(0)\|_{\dot{L}^2}^2 + C_4^2 \tilde{C} \|U(0)\|_{\dot{L}^2}^2.
\end{aligned}$$

We therefore have

$$\|\partial_x U(t)\|_{\dot{L}^2}^2 \leq C_3 C_4^2 \tilde{C} \|U(0)\|_{\dot{L}^2}^2 + \frac{C_4^2 \tilde{C}}{T} \|U(0)\|_{\dot{L}^2}^2. \quad (4.39)$$

Thus for any finite  $t \leq T$ , if the initial condition  $U(0)$  is bounded in  $\dot{L}^2$ , then  $\partial_x U(t)$  is also bounded in  $\dot{L}^2$ . This is valid for any  $t$ , since we can always choose  $T$  large enough. However, note that  $\tilde{C}$  depends on  $T$ , so we cannot take the limit as  $T \rightarrow \infty$ .

We will now use this to show compactness. Let  $\Lambda(t, u_0)$  be the flow associated with  $U(t)$ . That is,  $\Lambda(t, u_0)U(0) = U(t)$  with initial condition  $U(0) \in \dot{L}^2$ . Then if  $U_n(0)$  is a uniformly bounded sequence in  $\dot{L}^2$  such that  $\|U_n(0)\|_{\dot{L}^2} \leq M$  for some  $M > 0$  and all  $n \in \mathbb{N}$ , we can use inequality (4.39) to conclude that

$$\begin{aligned}
\|\partial_x \Lambda(t, u_0)U_n(0)\|_{\dot{L}^2}^2 &= \|\partial_x U(t)\|_{\dot{L}^2}^2 \leq C_3 C_4^2 \tilde{C} \|U(0)\|_{\dot{L}^2}^2 + \frac{C_4^2 \tilde{C}}{T} \|U(0)\|_{\dot{L}^2}^2 \\
&\leq C_3 C_4^2 \tilde{C} M^2 + \frac{C_4^2 \tilde{C}}{T} M^2.
\end{aligned}$$

Thus the sequence  $\Lambda(t, u_0)U_n(0)$  is also uniformly bounded in  $\dot{H}^1$ , and so by the Rellich-Kondrachov compactness theorem [Eva15, p. 288], the sequence  $\Lambda(t, u_0)U_n(0)$  must be pre-compact in  $\dot{L}^2$ . Therefore  $\Lambda(t, u_0)$  is a compact operator from  $\dot{L}^2$  to  $\dot{L}^2$  for any finite  $t > 0$ .

## 4.7 The Dimension of the Attractor

In this section we will state the main theorem of this chapter, concerning the dimension of the attractor of a dynamical system. We will then prove that the growth rate  $\mathcal{TR}_n(\mathcal{A})$  (4.28) of the Nikolaevskiy equation (1.4) satisfies the necessary bounds. Furthermore we will attempt to get the tightest possible bounds on the scaling of the attractor dimension with  $r$  and  $\ell$ .

**Theorem 4.7.1** (Dimension of the Attractor [Rob01, p. 341]). *Suppose that the flow  $S(t)$  of (4.1) is uniformly differentiable on  $\mathcal{A}$ , and that for any  $u_0 \in \mathcal{A}$  there exists a  $t_0$  such that  $\Lambda(t, u_0)$  is compact for all  $t \geq t_0$ . If  $\mathcal{TR}_n(\mathcal{A}) < 0$ , then the fractal dimension  $d_f$  (4.3)*

satisfies

$$d_f(\mathcal{A}) \leq n.$$

The above theorem is very general, and the results from the previous sections in this chapter constitute the proof of this theorem. Specifically, the map  $F$  in Theorem 4.3.1 is the flow  $S(t)$  of equation (1.4), and  $DF$  is the flow of the linearised system (4.30) given by  $\Lambda(t, u_0)$ . The previous two sections proved that  $\Lambda(t, u_0)$  exists and is compact. In this section we will apply this general theorem to equation (1.4). Since we already have compactness and uniform differentiability, we only need to show the condition  $\mathcal{TR}_n(\mathcal{A}) < 0$ , and find the smallest  $n$  for which this is true. We will define  $\gamma$  from Theorem 4.3.1 to be  $\gamma := e^{\mathcal{TR}_n(\mathcal{A})} < 1$ .

The highest order term in (1.4) will play an important role, as it is the term which we can make ‘negative enough’ to make  $\mathcal{TR}_n(\mathcal{A})$  negative. We will thus derive some bounds that let us deal with it concretely, and relate it back to dimension.

To bound (4.28), we will get some bounds on powers of the projection of the Laplacian. Recall that the projection onto an  $n$ -dimensional subspace of a Hilbert space  $H$  is

$$P_n := \sum_{i=1}^n \phi_i(\phi_i, \cdot), \quad (4.40)$$

where  $\phi_i$  is an orthonormal basis that spans the space of the volume that we are studying, and that for any linear operator  $A : H \rightarrow H$ , the **trace** is defined as

$$\text{Tr}(AP_n) := \sum_{i=1}^n (\phi_i, A\phi_i)_H. \quad (4.41)$$

**Lemma 4.7.2** (Bounds on Traces of an operator). *Let  $P_n$  be a projection onto an  $n$ -dimensional subspace in the Hilbert space  $H$ . Then for any symmetric operator  $A$  whose eigenvalues can be ordered,  $\lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \dots$ , and whose eigenvectors form an orthonormal basis of  $H$ , we have*

$$\text{Tr}(AP_n) \geq \lambda_1 + \dots + \lambda_n, \quad (4.42)$$

where  $\lambda_1, \dots, \lambda_n$  are the  $n$  smallest eigenvalues of  $A$ .

*Proof.* By assumption there exists an orthonormal set of eigenvectors  $w_k$ ,  $k \in \mathbb{N}$  of  $A$ , which form an orthonormal basis of  $H$ . We can thus express each  $\phi_i$  in the definition of  $P_n$  in terms of this basis as

$$\phi_i = \sum_{k=1}^{\infty} (\phi_i, w_k) w_k.$$

Thus

$$\text{Tr}(AP_n) = \sum_{i=1}^n \sum_{k=1}^{\infty} \lambda_k |\langle \phi_i, w_k \rangle|^2 = \sum_{k=1}^{\infty} \lambda_k \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2.$$

We want to show that

$$\text{Tr}(AP_n) \geq \sum_{i=1}^n \lambda_i.$$

Note that  $\lambda_i \leq \lambda_{i+1}$ . Also, note that since  $\|\phi_i\| = 1$  for each  $i$ , we have

$$\sum_{k=1}^{\infty} |\langle \phi_i, w_k \rangle|^2 = 1,$$

which implies that

$$\sum_{k=1}^n |\langle \phi_i, w_k \rangle|^2 \leq 1, \quad i = 1, \dots, n$$

and that,

$$\sum_{k=1}^{\infty} \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 = n.$$

Now the trace satisfies

$$\begin{aligned} \text{Tr}(AP_n) &= \sum_{k=1}^{\infty} \lambda_k \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 \\ &= \sum_{k=1}^n \lambda_k \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 + \sum_{k=n+1}^{\infty} \lambda_k \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 \\ &\geq \sum_{k=1}^n \lambda_k \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 + \lambda_{n+1} \sum_{k=n+1}^{\infty} \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 \\ &= \sum_{k=1}^n \lambda_k \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 + \lambda_{n+1} \left( \sum_{k=1}^{\infty} \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 - \sum_{k=1}^n \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 \right), \end{aligned}$$

by the ordering of the eigenvalues. But then using the fact that  $\sum_{k=1}^{\infty} \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 = \sum_{k=1}^n 1 = n$ , we have

$$\begin{aligned} \text{Tr}(AP_n) &\geq \sum_{k=1}^n \lambda_k \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 + \lambda_{n+1} \left( \sum_{k=1}^n 1 - \sum_{k=1}^n \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 \right) \\ &= \sum_{k=1}^n \lambda_k \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 + \sum_{k=1}^n \lambda_{n+1} \left( 1 - \sum_{i=1}^n |\langle \phi_i, w_k \rangle|^2 \right). \end{aligned}$$

But  $\lambda_{n+1} \geq \lambda_k$  for each  $k \leq n+1$ , and so we have

$$\begin{aligned} \text{Tr}(AP_n) &\geq \sum_{k=1}^n \lambda_k \sum_{i=1}^n |(\phi_i, w_k)|^2 + \sum_{k=1}^n \lambda_k \left(1 - \sum_{i=1}^n |(\phi_i, w_k)|^2\right) \\ &= \sum_{k=1}^n \lambda_k. \end{aligned}$$

□

**Corollary 4.7.3.** *Let  $A := (-1)^m \partial_x^{2m}$ , and recall that the eigenvalues of  $A$  are just  $\lambda_{2j-1} = \lambda_{2j} = (2\pi j/\ell)^{2m}$  (so they have multiplicity two). Then  $A$  satisfies the conditions of the above lemma. For the special cases  $m = 1$ ,  $m = 2$  and  $m = 3$ , we have that*

$$\text{Tr}(-\partial_x^2 P_{2n}) \geq 2 \left(\frac{2\pi}{\ell}\right)^2 \left(\frac{n(n+1)(2n+1)}{6}\right),$$

$$\text{Tr}(\partial_x^4 P_{2n}) \geq 2 \left(\frac{2\pi}{\ell}\right)^4 \left(\frac{6n^5 + 15n^4 + 10n^3 - n}{30}\right),$$

and

$$\text{Tr}(-\partial_x^6 P_{2n}) \geq 2 \left(\frac{2\pi}{\ell}\right)^6 \left(\frac{6n^7 + 21n^6 + 21n^5 - 7n^3 + n}{42}\right),$$

respectively.

*Proof.* For the case  $\sum_j j^2$ , the formula is well known. For the other cases we have [Bea96]

$$\sum_{j=1}^n j^4 = \frac{6n^5 + 15n^4 + 10n^3 - n}{30},$$

and

$$\sum_{j=1}^n j^6 = \frac{6n^7 + 21n^6 + 21n^5 - 7n^3 + n}{42}.$$

Thus

$$\begin{aligned} \text{Tr}(\partial_x^4 P_{2n}) &\geq \sum_{j=1}^{2n} \lambda_j = 2 \sum_{j=1}^n \left(\frac{2\pi j}{\ell}\right)^4 \\ &= 2 \left(\frac{2\pi}{\ell}\right)^4 \frac{6n^5 + 15n^4 + 10n^3 - n}{30}, \end{aligned}$$

and

$$\begin{aligned} \text{Tr}(-\partial_x^6 P_{2n}) &\geq 2 \sum_{j=1}^{2n} \lambda_j = \sum_{j=1}^n \left( \frac{2\pi j}{\ell} \right)^6 \\ &= 2 \left( \frac{2\pi}{\ell} \right)^6 \frac{6n^7 + 21n^6 + 21n^5 - 7n^3 + n}{42}. \end{aligned}$$

□

We will now show that the trace of the operator of the Nikolaevskiy equation is negative for large enough  $n$ . Using the definition (4.41) of the trace of the linearised operator of (1.4), it is given by

$$\text{Tr}(\mathcal{G}_u P_n) = \sum_{i=1}^n \left[ (\phi_i, \partial_x^6 \phi_i) + 2(\phi_i, \partial_x^4 \phi_i) + (1-r)(\phi_i, \partial_x^2 \phi_i) + b(\phi_i, \phi_i, u) \right], \quad (4.43)$$

since  $b(\phi_i, u, \phi_i) + b(u, \phi_i, \phi_i) = -b(\phi_i, \phi_i, u)$ . Note that if we can bound  $\text{Tr}(\mathcal{G}_u P_n)$  uniformly, such that the bound does not depend on  $t$ ,  $u_0$  or the projection  $P_n$ , then this immediately implies a bound on  $\mathcal{TR}_n(\mathcal{A})$ , for the dynamical system defined by equation (1.4).

**Case 1,  $r < 1$ :**

Sine we are interested in the chaotic regime of equation (1.4), we will therefore assume that  $r\ell^2$  is large, which will allow us to neglect some of the additive constants in our bounds. We will now proceed with the derivation. We start with (4.43). This time, we will bound the nonlinear term as follows: Using Proposition B.6.3 we get

$$\begin{aligned} |b(\phi_i, \phi_i, u)| &= \left| -\frac{1}{2} \int_0^\ell (\phi_i)^2 \partial_x u \, dx \right| \leq \frac{\ell^{1/2}}{2} \left\| \partial_x^2 u \right\|_{\mathbb{L}^2} \|\phi_i\|_{\mathbb{L}^2}^2 \\ &= \frac{\ell^{1/2}}{2} \left\| \partial_x^2 u \right\|_{\mathbb{L}^2} \leq \frac{\ell^{1/2}}{2} R_2, \end{aligned} \quad (4.44)$$

for large enough  $t$ , since  $\phi_i$  is a unit vector. Then we can bound the inner products in the trace as

$$\begin{aligned} (\phi_i, \mathcal{G}_u \phi_i) &\leq - \left\| \partial_x^3 \phi_i \right\|_{\mathbb{L}^2}^2 + 2 \left\| \partial_x^2 \phi_i \right\|_{\mathbb{L}^2}^2 - (1-r) \left\| \partial_x \phi_i \right\|_{\mathbb{L}^2}^2 + \frac{\ell^{1/2}}{2} R_2 \\ &\leq \left( -1 + \frac{1}{\sigma} \right) \left\| \partial_x^3 \phi_i \right\|_{\mathbb{L}^2}^2 + (-1 + \sigma + r) \left\| \partial_x \phi_i \right\|_{\mathbb{L}^2}^2 + \frac{\ell^{1/2}}{2} R_2. \end{aligned}$$

Then with  $\sigma = 1 + r^p$  for  $0 < p \leq 1$  arbitrary, using inequality (B.7) we have

$$\begin{aligned} (\phi_i, \mathcal{G}_u \phi_i) &\leq \left(-1 + \frac{1}{1 + r^p}\right) \left\| \partial_x^3 \phi_i \right\|_{\dot{L}^2}^2 + (r^p + r) \left\| \partial_x \phi_i \right\|_{\dot{L}^2}^2 + \frac{\ell^{1/2}}{2} R_2 \\ &\leq \frac{-r^p}{1 + r^p} \left\| \partial_x^3 \phi_i \right\|_{\dot{L}^2}^2 + 2r^p \left\| \partial_x \phi_i \right\|_{\dot{L}^2}^2 + \frac{\ell^{1/2}}{2} R_2 \\ &\leq \left( \frac{-r^p}{1 + r^p} + \frac{2\eta^2 r^p}{4} \right) \left\| \partial_x^3 \phi_i \right\|_{\dot{L}^2}^2 + \frac{2r^p}{\eta} \left\| \phi_i \right\|_{\dot{L}^2}^2 + \frac{\ell^{1/2}}{2} R_2, \end{aligned}$$

for  $\eta > 0$  arbitrary. Using the fact that  $\phi_i$  is a unit vector and letting  $\eta^2 = 1/(1 + r^p)$  we get

$$(\phi_i, \mathcal{G}_u \phi_i) \leq \left( \frac{-r^p}{2(1 + r^p)} \right) \left\| \partial_x^3 \phi_i \right\|_{\dot{L}^2}^2 + 2r^p \sqrt{1 + r^p} + \frac{\ell^{1/2}}{2} R_2. \quad (4.45)$$

Then if we take the limit  $p \rightarrow 0$ , and sum over  $i$  we get:

$$Tr(\mathcal{G}_u P_{2n}) \leq \sum_{i=1}^{2n} \left[ -\frac{1}{4} \left\| \partial_x^3 \phi_i \right\|_{\dot{L}^2}^2 + 2\sqrt{2} + \frac{\ell^{1/2}}{2} R_2 \right],$$

which by Corollary 4.7.3 is bounded by

$$\begin{aligned} Tr(\mathcal{G}_u P_{2n}) &\leq 2n \left[ 2\sqrt{2} + \frac{\ell^{1/2}}{2} R_2 \right] - \frac{1}{2} \left( \frac{2\pi}{\ell} \right)^6 \frac{6n^7 + 21n^6 + 21n^5 - 7n^3 + n}{42} \\ &\leq 2n \left[ 2\sqrt{2} + \frac{\ell^{1/2}}{2} R_2 \right] - \frac{1}{2} \left( \frac{2\pi}{\ell} \right)^6 \frac{6n^7}{42}. \end{aligned}$$

Then the trace is less than zero if

$$n > \frac{\ell}{2\pi} \left( 14 \left[ \sqrt{32} + \ell^{1/2} R_2 \right] \right)^{1/6}.$$

In the present work, we are generally interested in the large system limit. That is, the limit in which the number of unstable modes is large enough to capture the full PDE behaviour. Recall that the number of unstable modes (the number of modes in the unstable band of the dispersion relation) is  $N_{unstable} \sim (\sqrt{r}\ell)/(2\pi)$ . Thus we will assume that  $\sqrt{r}\ell$  is large and then, since  $R_2 = \mathcal{O}(r^{7/8}\ell^3)$ , we have an upper bound on the dimension of

$$d_f(\mathcal{A}) \leq \frac{\ell}{2\pi} \left( 14 \left[ \sqrt{32} + \ell^{1/2} R_2 \right] \right)^{1/6} = \mathcal{O}(r^{7/48} \ell^{19/12}). \quad (4.46)$$

This gives the bound for small  $r$ . The bound (4.46), gives us a scaling with  $\ell$ , but really, it does not give us a scaling with  $r$ . Deriving such a scaling is difficult, because it is not true, in general, that  $\lim_{r \rightarrow 0} d_f(\mathcal{A}) = 0$ . The reason for this is that we know that for large enough  $\sqrt{r}\ell$ , the system is in the chaotic regime for any  $r > 0$ . Thus the Poincaré-Bendixson



theorem [Per00] tells us that the dimension must be greater than 2 for  $r > 0$ . This lower bound on the dimension is reflected in our derivation by the fact that our bound for the attractor contains an additive constant, independent of  $r$ .

**Case 2,  $r = 1$ :**

Beginning with a term in (4.43) and using the bound on  $b(\phi_i, \phi_i, u)$  (4.44) we get

$$\begin{aligned} (\phi_i, \mathcal{G}_u \phi_i) &= (\phi_i, \partial_x^6 \phi_i) + 2(\phi_i, \partial_x^4 \phi_i) + (1-r)(\phi_i, \partial_x^2 \phi_i) + b(\phi_i, \phi_i, u) \\ &\leq - \left\| \partial_x^3 \phi_i \right\|_{\mathbb{L}^2}^2 + 2 \left\| \partial_x^2 \phi_i \right\|_{\mathbb{L}^2}^2 + \ell^{1/2} \left\| \partial_x^2 u \right\|_{\mathbb{L}^2} \\ &\leq \left( -1 + \frac{2}{\eta} \right) \left\| \partial_x^3 \phi_i \right\|_{\mathbb{L}^2}^2 + \frac{\eta^2}{2} \|\phi_i\|_{\mathbb{L}^2}^2 + \ell^{1/2} R_2, \end{aligned}$$

where we have used integration by parts, Young's inequality, and inequality (B.8). So if we note that  $\phi_i$  is a unit vector and we choose, for instance  $\eta = 4$ , and sum over  $i$  we have

$$\text{Tr}(\mathcal{G}_u P_{2n}) \leq 2 \left( 8 + \ell^{1/2} R_2 \right) n - \left( \frac{6n^7 + 21n^6 + 21n^5 - 7n^3 + n}{42} \right) \left( \frac{2\pi}{\ell} \right)^6.$$

For any given value of  $\ell$ , we can thus find an  $n$  large enough, such that the above expression is less than zero. Specifically, if we assume that  $\ell$  is large, then the above shows that for  $r = 1$  the dimension of the attractor can be bounded by

$$d_f(\mathcal{A}) = \mathcal{O}(\ell^{19/12}),$$

by the same reasoning as in the first case.

The above is of course the same scaling as in (4.46) and so the bound (4.46) holds for all  $0 < r \leq 1$ .

## Chapter 5

# Numerical Methods

In this chapter, we will describe the numerical tools that we used to compute the solutions of the Nikolaevskiy equation. We will also explain some procedures that have been developed to compute the dimension of a dynamical system in general. The concept of Lyapunov exponents will play a vital role in this chapter. Since, as we have seen, the dimension of the attractor of the Nikolaevskiy equation (1.4) is finite (4.46), we can expect the dynamics of (1.4) to resemble that of a finite-dimensional dynamical system on  $\mathbb{R}^n$ . Thus it is reasonable to assume that the same techniques that work for finite-dimensional systems will also work for (1.4), after possibly some minor modifications.

### 5.1 Lyapunov Exponents

One characteristic of chaotic systems is that even though the long-term behaviour of solutions is bounded (the attractor lies in an absorbing ball), the solutions still depend sensitively on initial conditions. This means that on average, solutions starting close to each other locally diverge exponentially fast.

Consider the dynamical system (4.1)

$$\frac{dx(t)}{dt} = f(x(t)).$$

We can ask how rapidly trajectories starting at the points  $x$  and  $\tilde{x}$  separate if  $x$  and  $\tilde{x}$  are close. We will now define Lyapunov exponents, which will allow us to make this more precise.

A Lyapunov exponent (LE) is a scalar  $\mu$  that measures the rate at which nearby trajectories separate. Suppose  $\tilde{x}(t) = x(t) + \epsilon y(t)$ , where  $y(t)$  is some vector in the direction of separation of  $\tilde{x}$  and  $x$ , and  $\epsilon$  is a small scalar. Then we have

$$\frac{d}{dt}(\tilde{x} - x) = \epsilon \frac{dy}{dt} = f(x(t) + \epsilon y(t)) - f(x(t)).$$

Assuming  $f$  is differentiable with respect to  $x(t)$ , this becomes

$$\epsilon \frac{dy}{dt} = f(x(t) + \epsilon y(t)) - f(x(t)) = f(x(t)) - f(x(t)) + \epsilon Df(x(t))y(t) + \mathcal{O}(\epsilon^2),$$

where  $Df$  is the Jacobian or Fréchet derivative of  $f$  as in the linearization of (4.1). Taking the limit as  $\epsilon \rightarrow 0$  gives [Mei17, p. 239]

$$\frac{dy}{dt} = Df(x(t))y(t).$$

In fact, this shows that the separation vector between two infinitesimally close initial conditions evolves exactly according to the linearisation (4.2) of (4.1).

To compute Lyapunov exponents, we want to know how fast  $\|y(t)\|$  grows, where  $\|\cdot\|$  is the norm on the space the system lives in. Specifically, we are interested in cases when  $\|y(t)\|$  grows (or decays) exponentially fast. One thus defines the Lyapunov exponent associated with the vector  $y(t)$  as follows [PP16, p. 12]:

**Definition 5.1.1** (Lyapunov Exponents). *Let  $x(t)$  be a solution to (4.1), and let  $y(t)$  be a solution to (4.2) in the tangent space of  $x(t)$ , then the **Lyapunov exponent** (LE)  $\mu$  of the trajectory  $x(t)$  in the direction of  $y(t)$  is defined by*

$$\mu := \limsup_{t \rightarrow \infty} \frac{\log(\|y(t)\|)}{t}. \quad (5.1)$$

The set of  $\mu$  computed from all  $y(t)$  in the tangent space of  $x(t)$  is called the **Lyapunov spectrum** (LS) of  $x(t)$ .

The reason for the logarithm is that, intuitively,  $\|y(t)\| \sim \|y_0\|e^{\mu t}$  on average, as is certainly the case for linear systems where the coefficient matrix is constant. We are taking the limsup here because the above limit may not be unique [Mei17, p. 241] and we are generally interested in the maximum growth rate in a given direction.

In general, in an  $n$ -dimensional Hilbert space  $H$  (where  $n$  may be infinite), we can have at most  $n$  distinct LEs for some trajectory  $x(t)$ . Since the LEs give the growth rate in a certain direction, we will also need the concept of a **Lyapunov vector** (LV) [Mei17, p. 243]. The LVs, if they exist, are a set of linearly independent vectors  $\{y_1(t), y_2(t), \dots, y_n(t)\}$  which are solutions to the linearised dynamical system (4.2) which are elements of the tangent space of  $x(t)$ , such that

$$\sum_j \mu_j := \sum_j \limsup_{t \rightarrow \infty} \frac{\log(\|y_j(t)\|)}{t}$$

is minimized. We call such a set of vectors a **Lyapunov basis**. In this text, we will always assume that the LEs and LVs exist and are well defined for the system that we are working

with, and we will always order the LEs from largest to smallest

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_n.$$

Another assumption that we will make is that the LEs are the same no matter what trajectory  $x(t)$  we choose. For most systems of interest this is true for almost every trajectory, by Oseledets' theorem [Ose68]. For our system, LEs and LVs are assumed to exist and to be independent of initial condition. It should be noted, however, that these assumptions cannot always be made, and in fact, there are examples of systems for which the LVs do not exist [OY08], because the limit in (5.1) does not exist and we really have to take the lim sup instead.

We will end this section with a useful property of LEs, which we can later exploit in the computation of these numbers.

**Proposition 5.1.2.** *Let  $c_1$  and  $c_2$  be scalars, and let  $y_1(t)$  and  $y_2(t)$  be linearly independent and evolve according to the linearisation (4.2) along some trajectory  $x(t)$ . Let  $\mu_1$  and  $\mu_2$  be the LEs computed from  $y_1(t)$  and  $y_2(t)$  along this trajectory  $x(t)$ . Then*

$$\limsup_{t \rightarrow \infty} \frac{\log(\|c_1 y_1(t) + c_2 y_2(t)\|)}{t} = \max\{\mu_1, \mu_2\}. \quad (5.2)$$

For a proof see [Mei17, p. 242]. Intuitively, the above result is true by basic properties of logarithms, and from the fact that the LV associated with  $\max\{\mu_1, \mu_2\}$  has a component along the vector  $c_1 y_1(t) + c_2 y_2(t)$ , since it must have a component along at least one of  $y_1(t)$  or  $y_2(t)$ .

## 5.2 Computing Lyapunov Exponents

In this section we will give the algorithm that we used to compute the Lyapunov spectrum of (1.4). Several algorithms have been proposed to compute the LEs of a dynamical system [GPT<sup>+</sup>07, EP98, SN79, PC89]. Most of them are based on [SN79]. Even though Shimada and Nagashima [SN79] formulated their algorithm in terms of a finite-dimensional system, it works perfectly well for our infinite-dimensional system too. The reason is that we approximate it by a finite-dimensional system anyway, because we can only compute a finite number of Fourier modes.

The idea for computing the first  $n$  LEs is simple, and we will outline it here: Consider an arbitrary vector  $v_1(t) \in H$  that evolves under the flow of (4.2). With probability one it contains a component along  $y_1(t)$ , the initial condition of the LV corresponding to the largest LE. Thus, (5.2) implies that with probability one,

$$\limsup_{t \rightarrow \infty} \frac{\log(\|v_1(t)\|)}{t} = \limsup_{t \rightarrow \infty} \frac{\log(\|y_1(t)\|)}{t} = \mu_1.$$

Similarly, if we consider a time-dependent plane spanned by two linearly independent vectors  $v_1(t)$  and  $v_2(t)$ , then with probability one the vectors will contain components along the LVs  $y_1(t)$  and  $y_2(t)$ , corresponding to the largest two LEs. Thus the ellipse, defined by  $v_1(t)$  and  $v_2(t)$  will expand at an exponential rate  $\mu_1$  in the direction of  $y_1(t)$  and at a rate  $\mu_2$  in the direction of  $y_2(t)$  under the flow of (4.2). Thus the area of the ellipse will grow at a rate of  $\mu_1 + \mu_2$ , and we find

$$\limsup_{t \rightarrow \infty} \frac{\log(\|v_1 \wedge v_2\|)}{t} = \limsup_{t \rightarrow \infty} \frac{\log(\|y_1 \wedge y_2\|)}{t} = \mu_1 + \mu_2,$$

where  $\wedge$  denotes the outer product [Tem98] and  $\|v_1 \wedge v_2\|$  is the area of the parallelogram defined by  $v_1(t)$  and  $v_2(t)$  (which is proportional to the area of the disc defined by these vectors). Continuing along this line of reasoning, we get that the volume of the parallelepiped defined by  $n$  linearly independent vectors  $v_1(t), \dots, v_n(t)$  grows at a rate proportional to the sum of the largest  $n$  LEs under the flow of (4.2), that is,

$$\limsup_{t \rightarrow \infty} \frac{\log(\|v_1 \wedge \dots \wedge v_n\|)}{t} = \limsup_{t \rightarrow \infty} \frac{\log(\|y_1 \wedge \dots \wedge y_n\|)}{t} = \sum_{j=1}^n \mu_j. \quad (5.3)$$

So in principle, to compute  $n$  LEs we could just follow the evolution of volumes under the flow of the linearised dynamical system. The problem with this, however, is that with probability one, each of the vectors  $v_j(t)$  contains a component along  $y_1(t)$ . This means that the parallelepiped defined by the  $v_j(t)$ s will become very ‘stretched’ and ‘thin’ after a short time. Speaking more precisely,

$$\|v_1(t) \wedge \dots \wedge v_n(t)\| = \sqrt{\det\{M(t)\}}, \quad (5.4)$$

where  $M(t)$  is the matrix with components  $M_{ij}(t) := (v_i(t), v_j(t))$  and  $(\cdot, \cdot)$  is the inner product of  $H$ . Thus computing this volume becomes unstable as the  $v_j(t)$  become close to linearly dependent and their norm grows exponentially. Thus the matrix  $M(t)$  becomes ill-conditioned, and so in practice computing the volume is not feasible for larger times.

The way we can get around this difficulty is by periodically re-orthonormalizing the vectors  $v_j(t)$ , using a QR or Gram-Schmidt algorithm. As long as we keep track of how much each vector grew, we can compute better and better approximations to the LEs. One advantage of this approach is that we don’t have to compute the LEs in sequence and can actually compute all of them simultaneously. Another advantage is that since we are already performing a Gram-Schmidt procedure, which is equivalent to reshaping the parallelepiped into a rectangular box, it becomes very easy to compute the volume, since we just have to take the product of the lengths of the orthogonalized vectors.

We will now give the specific algorithm, based on [PC89], which we used for (1.4), with the specific norm and  $\dot{L}^2$  inner product. The flow of the linearised Nikolaevskiy equation (4.30) is denoted by  $\Lambda(t)$ .

---

**Algorithm 1:** Algorithm for computing the largest  $n$  Lyapunov exponents of (1.4)

---

**Input** : Linearly independent vectors  $v_1, \dots, v_n$ , final time  $T_{\max}$ , time increment  $\Delta t$   
**Output:** Largest  $n$  LEs  $\mu_1, \dots, \mu_n$ .

```

1 while  $T < T_{\max}$  do
2    $v_j = \Lambda(\Delta t)v_j$                                Evolve the vectors for a time  $\Delta t$ .
3    $T = T + \Delta t$                                      Update the current time
4   for  $j=1:n$  do
5      $\tilde{v}_j = v_j - \sum_{i=1}^{j-1} \frac{(v_i, v_j)}{(v_i, v_i)} v_i$            Do ‘Gram-Schmidt’
6   end
7    $S_j = S_j + \log(\|\tilde{v}_j\|_{\dot{L}^2})$                    Add the current length to the previous lengths
8    $\mu_j = S_j/T$                                        Compute the LE
9    $v_j = \tilde{v}_j / \|\tilde{v}_j\|_{\dot{L}^2}$                        Normalize the vectors
10 end

```

---

Some more notes about the above algorithm are in order. We have written ‘Gram-Schmidt’ in the algorithm, because we performed the orthogonalization with a (modified) Gram-Schmidt algorithm [TB97, p. 58]. One could however also do this step using a QR factorization, in which case the LEs are the logarithms of the time-averaged diagonal elements of the ‘R’ matrix of the QR factorization [GPT<sup>+</sup>07]. We found that using QR instead of an explicit Gram-Schmidt algorithm does not speed up the computation.

As we will discuss in the section on the numerical procedures, most of the numerical steps in obtaining a solution to (1.4) are done in Fourier space, which means that one could also do this orthogonalization on the Fourier transform of the solution, using the inner product of the Fourier coefficients of the function. This would have the advantage of reducing the number of fast Fourier transforms one has to perform, but we found that the computation becomes unstable when performed in this way. For this reason we computed all the inner products in the above algorithm using  $\dot{L}^2$  inner products in real space.

### 5.3 Kaplan-Yorke Dimension

We will now describe a procedure for calculating the fractal dimension of an attractor that makes use of LEs. Although there are procedures for calculating the dimension of an attractor using the definition (4.3) directly [Mol93], they tend to scale badly with dimension, as the number  $N$  of boxes required to cover a set in  $\mathbb{R}^m$  scales as a powers of  $m$ . For this reason, it would be desirable to be able to calculate the fractal dimension of the attractor using a less computationally costly method. Some such methods are proposed in [KY79] and [SBH99], which use invariants of the flow of a dynamical system (4.1) to estimate the

fractal dimension of the attractor. We will be using the approach proposed by Kaplan and Yorke [KY79], which uses the Lyapunov exponents of the system to get an upper bound on the attractor dimension. This approach is desirable since it only requires the computation of LEs, which means that the computational cost only scales linearly with the true dimension of the attractor.

Kaplan and Yorke proposed the following formula:

**Definition 5.3.1** (Kaplan-Yorke Formula [KY79]). *Let the Lyapunov exponents  $\mu_1 \geq \dots \geq \mu_j \geq \dots$  be ordered as usual, from largest to smallest. Furthermore, assume that the dynamical system is dissipative and that it is ergodic, so that the LEs are equal for all trajectories, and solutions remain bounded. Since the system is dissipative, the sum of LEs is less than zero. Let  $\tilde{n}$  be the smallest integer such that*

$$\sum_{j=1}^{\tilde{n}} \mu_j < 0.$$

Then the **Kaplan-Yorke formula** for the dimension of the attractor  $\mathcal{A}$  of (4.1) is

$$d_{KY} := \tilde{n} - 1 + \frac{\sum_{j=1}^{\tilde{n}-1} \mu_j}{-\mu_{\tilde{n}}}. \quad (5.5)$$

The **Kaplan-Yorke conjecture** states that the above formula exactly gives the fractal dimension (4.3) of  $\mathcal{A}$  [KY79]. For this reason, the quantity (5.5) is sometimes called the **Kaplan-Yorke dimension** of  $\mathcal{A}$ . We will proceed to give a heuristic argument of why this should be true, but a rigorous proof has not yet been found. It can however be shown rigorously that (5.5) gives an upper bound on the fractal dimension (4.3) [Hun96]. The rigorous proof is essentially exactly what we did in Chapter 4.

We will give a heuristic argument for the Kaplan-Yorke conjecture. The intuition [KY79] is as follows: Suppose our dynamical system has  $M$  LEs that are ordered from most positive to most negative (as usual)  $\mu_1 \geq \dots \geq \mu_j \geq \dots \geq \mu_M$ . For any  $n \in \{1, \dots, M\}$ , if we take a parallelepiped spanned by the corresponding first  $n$  LVs, then its volume  $V(t)$ , in the sense of (5.4), at time  $t$  is given by:

$$V(t) = \prod_{j=1}^n \left( a e^{\mu_j t} \right) = a^n e^{(\mu_1 + \dots + \mu_n)t},$$

where  $a$  is the side length, and we have chosen a cube as the initial volume  $V(0)$ . Clearly, if  $\mu_1 + \dots + \mu_n < 0$ , then the volume contracts, so  $V(t) \rightarrow 0$  as  $t \rightarrow \infty$ . Thus we can estimate the evolution of the volume of an  $n$ -dimensional cube in the basin of attraction of  $\mathcal{A}$ . In particular, if we cover  $\mathcal{A}$  by an  $\epsilon$ -cover formed from a lattice of cubes of side length  $\epsilon$  as in Section 4.2, then we can study its evolution under the flow of the system. To do this, suppose we are given  $\epsilon$ , and that  $N(\epsilon)$  denotes the minimum number of cubes of side-length

$\epsilon$  that are needed to cover  $\mathcal{A}$ . We can assume that  $N(\epsilon)$  is finite, since  $\mathcal{A}$  is compact. We now study the evolution of its volume. First, note that since the system is assumed to be dissipative, the sum of all of its LEs has to be negative. Now let  $\tilde{n}$  denote the smallest integer such that the sum of the first  $\tilde{n}$  elements is negative. That is,

$$\sum_{j=1}^{\tilde{n}} \mu_j < 0 \quad , \quad \sum_{j=1}^{\tilde{n}-1} \mu_j \geq 0,$$

implying that  $\mu_{\tilde{n}} < 0$ . Then the volume of the  $\tilde{n}$ -dimensional  $\epsilon$ -cube is given by:  $V(0) = \epsilon^{\tilde{n}}$ , and  $V(t) = \epsilon^{\tilde{n}} e^{(\mu_1 + \dots + \mu_{\tilde{n}})t}$ . But now there is a problem! Even though the union of all the  $V(t)$  is still a cover of  $\mathcal{A}$ , since the attractor is invariant, it is no longer an  $\epsilon$  cover, since the side-lengths of the boxes are no longer equal. To remedy this, we divide each  $V(t)$  into smaller cubes, each with side-length  $\tilde{\epsilon} = \epsilon e^{\mu_{\tilde{n}} t}$ , then  $\tilde{\epsilon} \leq \epsilon e^{\mu_j t}$  for all  $j \leq \tilde{n}$ , since  $\mu_{\tilde{n}}$  is the most negative exponent we consider. Thus if we let  $N_V$  be the number of  $\tilde{\epsilon}$ -cubes needed to cover  $V(t)$ , then for an  $\tilde{n}$ -dimensional volume, we would have

$$N_V = \prod_{j=1}^{\tilde{n}} \frac{\epsilon e^{\mu_j t}}{\epsilon e^{\mu_{\tilde{n}} t}} + \text{Err} = \prod_{j=1}^{\tilde{n}-1} \frac{\epsilon e^{\mu_j t}}{\epsilon e^{\mu_{\tilde{n}} t}} + \text{Err} = e^{(\mu_1 + \dots + \mu_{\tilde{n}-1})t - (\tilde{n}-1)\mu_{\tilde{n}} t} + \text{Err},$$

where Err accounts for overlap, that is, it accounts for the fact that  $\frac{\epsilon e^{\mu_j t}}{\epsilon e^{\mu_{\tilde{n}} t}}$  may not be an integer. It can be bounded as follows: Clearly the amount of  $\tilde{n}$ -dimensional  $\tilde{\epsilon}$ -cubes needed to encase  $V(t)$  is an upper bound for Err. Therefore, recall that the surface ‘area’ (really a  $(\tilde{n} - 1)$ -dimensional volume) of an  $\tilde{n}$ -dimensional hyperrectangle of side lengths  $\epsilon e^{\mu_j t}$  is given by

$$S' = 2 \sum_{q=1}^{\tilde{n}} \prod_{\substack{j=1 \\ j \neq q}}^{\tilde{n}} \epsilon e^{\mu_j t}.$$

The amount of  $\tilde{n}$ -dimensional  $\tilde{\epsilon}$ -cubes needed to encase  $V(t)$  is  $S' \tilde{\epsilon} / \epsilon^{\tilde{n}}$ . But now we have to remember that by definition,  $\tilde{\epsilon}$  divides the  $\tilde{n}$ 'th side length exactly, and so there is no overlap in that direction and hence no contribution to Err. This means that we actually don't need to include the contribution from the parts of the surface ‘area’ that are orthogonal to the vertex of side length  $\tilde{\epsilon}$ . This means that we only have to consider the volume of the  $(\tilde{n} - 1)$ -dimensional ‘area’

$$S = 2 \sum_{q=1}^{\tilde{n}-1} \prod_{\substack{j=1 \\ j \neq q}}^{\tilde{n}} \epsilon e^{\mu_j t}.$$



Thus we get an upper bound for Err of

$$\text{Err} \leq \frac{2}{\tilde{\epsilon}^{\tilde{n}-1}} \sum_{q=1}^{\tilde{n}-1} \prod_{\substack{j=1 \\ j \neq q}}^{\tilde{n}} \epsilon e^{\mu_j t}.$$

Note that

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{\text{Err}}{e^{(\mu_1 + \dots + \mu_{\tilde{n}-1})t - (\tilde{n}-1)\mu_{\tilde{n}}t}} &= \lim_{t \rightarrow \infty} \frac{\tilde{\epsilon}^{\tilde{n}-1} \text{Err}}{\epsilon^{\tilde{n}-1} e^{(\mu_1 + \dots + \mu_{\tilde{n}-1})t}} \\ &\leq \lim_{t \rightarrow \infty} 2e^{-(\mu_1 + \dots + \mu_{\tilde{n}-1})t} \sum_{q=1}^{\tilde{n}-1} \prod_{\substack{j=1 \\ j \neq q}}^{\tilde{n}} e^{\mu_j t} = \lim_{t \rightarrow \infty} 2 \sum_{q=1}^{\tilde{n}-1} e^{(\mu_{\tilde{n}} - \mu_q)t} = 0, \end{aligned}$$

by the ordering of the  $\mu_j$  and because  $\mu_{\tilde{n}} < 0$ . Clearly, since Err is just a sum of exponentials, we also have that the above limit holds if we replace Err by  $\frac{d}{dt} \text{Err}$ .

Now that we know how Err behaves, we can proceed with the argument of the Kaplan-Yorke conjecture. The number of  $\tilde{\epsilon}$ -cubes needed to cover  $\mathcal{A}$ , is thus

$$N(\tilde{\epsilon}) = N(\epsilon) N_V.$$

Now for any fixed  $\epsilon$ ,  $\frac{\log(N(\epsilon))}{-\log(\epsilon)}$  is an approximation to  $d_f(\mathcal{A})$ , which becomes exact as  $\epsilon \rightarrow 0$ . So the  $\tilde{\epsilon}$  approximation to  $d_f$  is

$$\begin{aligned} d_f(\mathcal{A}) &\approx \frac{\log(N(\tilde{\epsilon}))}{-\log(\tilde{\epsilon})} = \frac{\log(N(\epsilon) N_V)}{-\log(\tilde{\epsilon})} \\ &= \frac{\log\left(N(\epsilon) \left(e^{(\mu_1 + \dots + \mu_{\tilde{n}-1})t - (\tilde{n}-1)\mu_{\tilde{n}}t} + \text{Err}\right)\right)}{-\log(\epsilon e^{\mu_{\tilde{n}}t})} \\ &= \frac{\log(N(\epsilon)) + \log\left(e^{(\mu_1 + \dots + \mu_{\tilde{n}-1})t - (\tilde{n}-1)\mu_{\tilde{n}}t} + \text{Err}\right)}{-\log(\epsilon) - \log(e^{\mu_{\tilde{n}}t})}. \end{aligned}$$

Now, in the limit as  $t \rightarrow \infty$ , we have that  $\tilde{\epsilon} \rightarrow 0$ . Thus

$$\begin{aligned} d_f(\mathcal{A}) &= \lim_{t \rightarrow \infty} \frac{\log(N(\epsilon)) + \log\left(e^{(\mu_1 + \dots + \mu_{\tilde{n}-1})t - (\tilde{n}-1)\mu_{\tilde{n}}t} + \text{Err}\right)}{-\log(\epsilon) - \log(e^{\mu_{\tilde{n}}t})} \\ &= \lim_{t \rightarrow \infty} \frac{\frac{((\mu_1 + \dots + \mu_{\tilde{n}-1}) - (\tilde{n}-1)\mu_{\tilde{n}})e^{(\mu_1 + \dots + \mu_{\tilde{n}-1})t - (\tilde{n}-1)\mu_{\tilde{n}}t}}{e^{(\mu_1 + \dots + \mu_{\tilde{n}-1})t - (\tilde{n}-1)\mu_{\tilde{n}}t} + \text{Err}}}{-\frac{\mu_{\tilde{n}} e^{\mu_{\tilde{n}}t}}{e^{\mu_{\tilde{n}}t}}}} \\ &= \frac{\mu_1 + \dots + \mu_{\tilde{n}-1} - (\tilde{n}-1)\mu_{\tilde{n}}}{-\mu_{\tilde{n}}} \\ &= \tilde{n} - 1 + \frac{\sum_{j=1}^{\tilde{n}-1} \mu_j}{-\mu_{\tilde{n}}}, \end{aligned}$$

where we have used L'Hôpital's rule and  $\text{Err}$  has gone to zero in the above limit by the properties that we proved about it above. This is now exactly formula (5.5). The above argument suggests that the Kaplan-Yorke formula does indeed give the fractal dimension of an attractor. The reason why the above is not a proof, is because volumes don't always grow at a rate proportional to the LEs, but only do so on average.

However, for the purposes of this text, we will assume that the Kaplan-Yorke conjecture is true, and we will use (5.5) to compute the fractal dimension of the attractor of (1.4), as has been done, for instance, in [XTGT00] for the Nikolaevskiy equation, and in [Man85] for the KS equation.

## 5.4 Numerical Methods

In this section we will describe the numerical procedure that we used to solve equation (1.4) and its linearisation (4.30). All computations were performed in MATLAB.

For the numerical solution, we used the exponential time-differencing, fourth-order Runge-Kutta scheme (ETDRK4), which was derived by Cox and Matthews [CM02]. The method ETDRK4 is an example of exponential time-differencing schemes, which themselves are similar to the so-called integrating factor (IF) schemes [KT05]. We will be using the slightly improved version of ETDRK4 presented by Kassam and Trefethen [KT05]. A good introduction and examples of the scheme can also be found in [LeV07]. These types of schemes are particularly well suited for equations such as (1.4), because they eliminate the difficulties that arise from having a stiff linear term in the equation, by solving the linear term exactly.

In theory, both IF schemes and exponential time-differencing schemes involve a change of variables. In order to get some intuition for these types of schemes, we will outline the procedure behind IF schemes as done in [KT05] and [Tre00]. Consider an equation of the form

$$u_t = \mathcal{L}u + \mathcal{N}(u), \tag{5.6}$$

where  $\mathcal{L}$  is a linear and stiff term, and  $\mathcal{N}$  is a nonlinear term. We can introduce the change of variables

$$v = e^{-\mathcal{L}t}u,$$

where  $e^{-\mathcal{L}t}$  is the integrating factor, which is given by the operator exponential of  $\mathcal{L}$  as in [Bre10, p. 118]. In our case, this operator diagonalizes in the form of the linear dispersion relation (1.5) in Fourier space, and so this exponential is easy to compute for us. That is,  $\hat{\mathcal{L}} = \omega(k)$ , and thus, in Fourier space (where we perform our computations),  $e^{-\mathcal{L}t} \rightarrow e^{-\omega(k)t}$

and  $\mathcal{L} \rightarrow \omega(k)$ . Then

$$v_t = -e^{-\mathcal{L}t} \mathcal{L}u + e^{-\mathcal{L}t} u_t,$$

so we get

$$v_t = e^{-\mathcal{L}t} \mathcal{N}(e^{\mathcal{L}t} v).$$

The above can now be solved using a suitable time-stepping method; for instance a Runge-Kutta method.

This is the basic intuition behind IF schemes, and it also applies to ETDRK4, although the details are more involved for ETDRK4, and involve more intermediate steps.

In this section we will give the general scheme for ETDRK4 as applied to (5.6), and then give a brief discussion of how we implemented it in MATLAB.

Consider the solution  $u_n$  to (5.6) at some time  $t_n = t_0 + hn$ , where  $n$  denotes the index of the time step and  $h$  is the time step size. We compute  $u_{n+1}$  using ETDRK4 as follows, using the table in [KT05]:

$$\begin{aligned} a_n &= e^{\mathcal{L}h/2} u_n + \mathcal{L}^{-1} \left( e^{\mathcal{L}h/2} - I \right) \mathcal{N}(u_n, t_n), \\ b_n &= e^{\mathcal{L}h/2} u_n + \mathcal{L}^{-1} \left( e^{\mathcal{L}h/2} - I \right) \mathcal{N}(a_n, t_n + h/2), \\ c_n &= e^{\mathcal{L}h/2} a_n + \mathcal{L}^{-1} \left( e^{\mathcal{L}h/2} - I \right) \left( 2\mathcal{N}(b_n, t_n + h/2) - \mathcal{N}(u_n, t_n) \right), \\ u_{n+1} &= e^{\mathcal{L}h} u_n + h^{-2} \mathcal{L}^{-3} \left\{ \left[ -4 - \mathcal{L}h + e^{\mathcal{L}h} (4 - 3\mathcal{L}h + (\mathcal{L}h)^2) \right] \mathcal{N}(u_n, t_n) \right. \\ &\quad + 2 \left[ 2 + \mathcal{L}h + e^{\mathcal{L}h} (-2 + \mathcal{L}h) \right] \left( \mathcal{N}(a_n, t_n + h/2) + \mathcal{N}(b_n, t_n + h/2) \right) \\ &\quad \left. + \left[ -4 - 3\mathcal{L}h - (\mathcal{L}h)^2 + e^{\mathcal{L}h} (4 - \mathcal{L}h) \right] \mathcal{N}(c_n, t_n + h) \right\}, \end{aligned} \tag{5.7}$$

where  $I$  denotes the identity operator. The above expression may look somewhat daunting to implement, but in Fourier space, it becomes a lot tamer. Specifically, what we do is take the fast Fourier transform (FFT) of  $u$  and then do the time stepping in Fourier space. As mentioned before, this has the effect of turning  $\mathcal{L}$  into the linear dispersion relation (1.5), which is easy to compute and to exponentiate, since it is just a diagonal matrix multiplying the vector of Fourier modes. The nonlinear term is treated as follows: Since  $\mathcal{N}(u) = -\frac{1}{2} \partial_x(u^2)$ , we get that

$$\widehat{\mathcal{N}(\hat{u})} = -\frac{ik}{2} \left( \mathcal{F}^{-1}(\hat{u}) \right)^2,$$

where  $\mathcal{F}^{-1}$  denotes the inverse FFT and  $\hat{\cdot}$  denotes the FFT.

We implemented the scheme in MATLAB, by modifying the code **kursiv.m** from [KT05]. The code **kursiv.m** uses ETDRK4 to solve the KS equation (1.1), and so it can

easily be modified to solve the Nikolaevskiy equation (1.4). Since the equations share the same nonlinear term (2.12), one simply has to modify the linear dispersion relation to be that given by (1.5).

The main modification we introduced to **kursiv.m** was by coupling the linearised equations (4.30) to the full Nikolaevskiy equation (1.4). Since the linear part of the linearised equation is the same as that of the original equation, it did not have to change. The  $B(u, U)$  and  $B(U, u)$  terms (2.21), where  $u$  is the solution to (1.4) and  $U$  is the solution to the linearised equation, had to be implemented in a slightly modified way. In Fourier space,  $B(u, U) = u \partial_x U$  becomes

$$\mathcal{F} \left( B(\mathcal{F}^{-1}\hat{u}, \mathcal{F}^{-1}\hat{U}) \right) = \mathcal{F} \left( (\mathcal{F}^{-1}(\hat{u}))(\mathcal{F}^{-1}(ik\hat{U})) \right),$$

since we are working with the Fourier transforms  $\hat{u}$  and  $\hat{U}$ . A similar expression holds for  $B(U, u)$ . The solutions to the linearised equations are then computed by doing one time step for  $u(t)$ , and then using this value of  $u(t)$  to do one time step of the linearised equations.

## Chapter 6

# Numerical Results

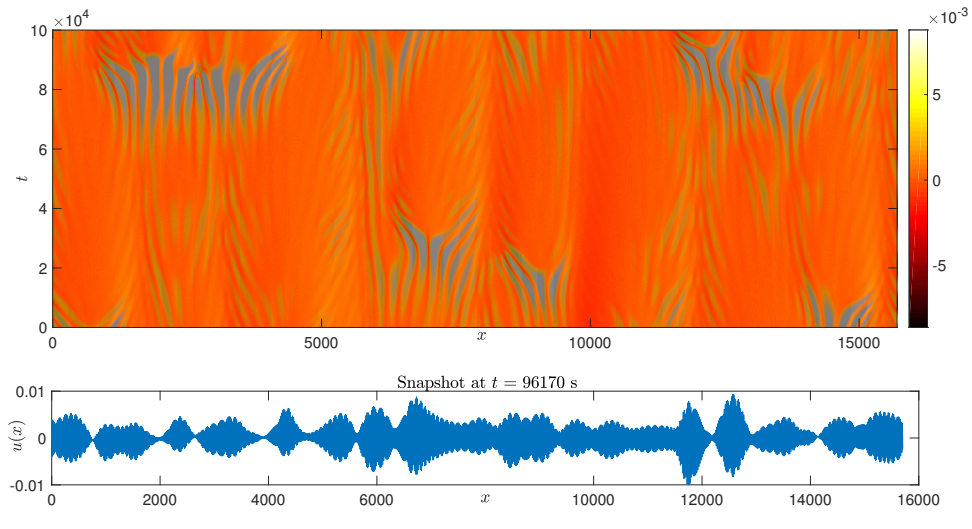
In this chapter, we will be presenting the results that we obtained by the methods described in the previous section. We will first show some characteristic examples of solutions of the Nikolaevskiy equation (1.4), and then proceed to discuss the Lyapunov spectrum and the Kaplan-Yorke dimension, and how these quantities depend on the parameters in the system.

### 6.1 Solutions of the Nikolaevskiy Equation

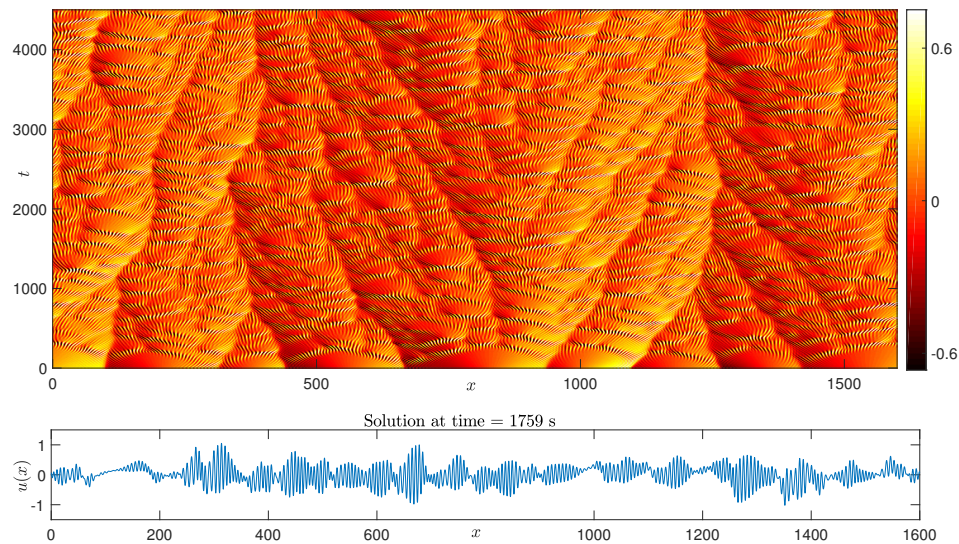
We computed solutions of the Nikolaevskiy equation (1.4) for various values of  $r$  and  $\ell$ . The aim was to determine the dependence of various quantities of the system on these parameters. In the chaotic regime ( $r > 0$ ) equation (1.4) displays two qualitatively distinct types of behaviour, that distinguish what is sometimes known as ‘Nikolaevskiy chaos’ [Tan05b, WP09], from the spatiotemporal chaos that can also be found in other, similar equations such as the Kuramoto-Sivashinsky (KS) equation [WH99]. We shall refer to the latter as KS-type chaos.

KS-type chaos is characterized by instabilities for long wavelengths, all the way down to the  $k = 0$  mode, whereas Nikolaevskiy chaos is characterized by energy in both high and low wavelengths. The most striking difference between the two types of chaos is the separation of scales that occurs in the Nikolaevskiy equation (1.4). For small  $r$  values, solutions possess energy around two well-separated frequencies. This behaviour is clearly visible in Figure 6.1, where the  $r$  values are sufficiently small to capture the separation of scales that is characteristic of Nikolaevskiy chaos. We plotted these solutions on very large spatial domains and over long time intervals to fully capture the small  $k$  structures that emerge at the onset of Nikolaevskiy chaos. On plots with shorter  $t$  and  $\ell$ , only the large  $k$  rolls are visible and the solution looks like a more streamlined KS-type chaos plot. In Figure 6.1b, one can see the short wavelength peaks between the larger, long wavelength peaks clearly. The snapshot of the solution shows how the long wavelength modes modulate the short wavelength modes, by effectively acting as an ‘envelope’ for the short wavelength modes. In Figure 6.1a the scale separation between the wavelengths of the small and large

modes has become so strong that the wavelength of the large  $k$  modes is barely visible any more. The large  $k$  rolls are only visible upon zooming in to the solution.



(a)



(b)

Figure 6.1: Plots of solutions to equation (1.4) for  $r \ll 1$ . Solutions are within the Nikolaevskiy chaos regime. The top panels display the solutions over a range of  $t$ , while the bottom panels display a snapshot of each solution at an arbitrarily chosen time. The reader is encouraged to zoom in on the plots to see the short wavelength rolls. Solutions are shown at parameter values (a)  $r = 0.0001$  and  $\ell = 5\pi \times 10^4$  and at (b)  $r = 0.04$  and  $\ell = 1600$ .

For comparison, in Figure 6.2 we show spatiotemporal chaos in the KS equation (1.1) (using blue tones, to distinguish from the orange tones used for the Nikolaevskiy equation (1.4)).

Figure 6.3 shows two solutions of equation (1.4), for  $r \approx 1$ . This is well into the KS-type chaotic regime, and there is no separation of scales visible any more. The wavelengths of the rolls in the solution are now all of the same order. Note that despite coming from different equations, the solutions in Figure 6.3 and Figure 6.2 appear very similar. We have chosen to display these solutions on relatively small spatial domains, since the solutions do not possess any apparent structure at larger scales and the rolls shown in these solutions just look like ‘noise’ on larger plots.

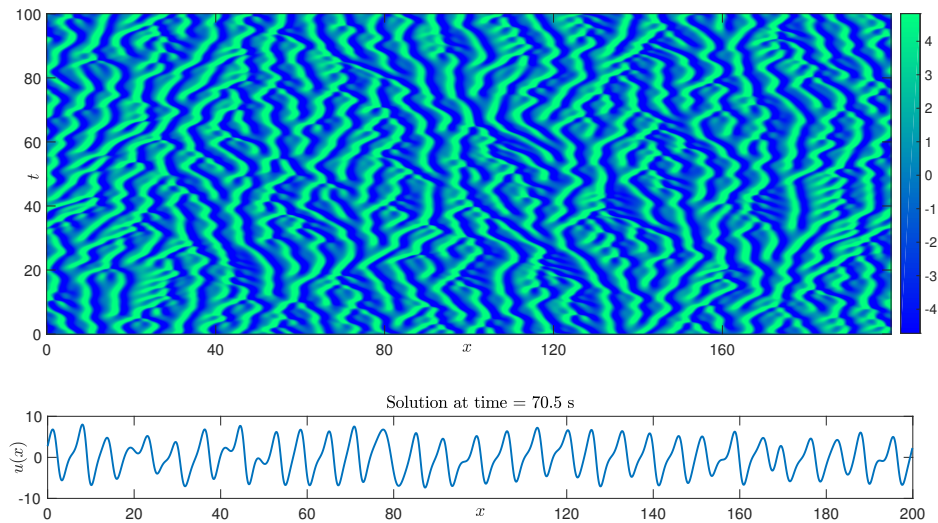


Figure 6.2: Plot of a typical solution of the KS equation (1.1). The top panel displays the solution over a range of  $t$ , while the bottom panel displays a snapshot of the solution at an arbitrary time for  $\ell = 200$ .

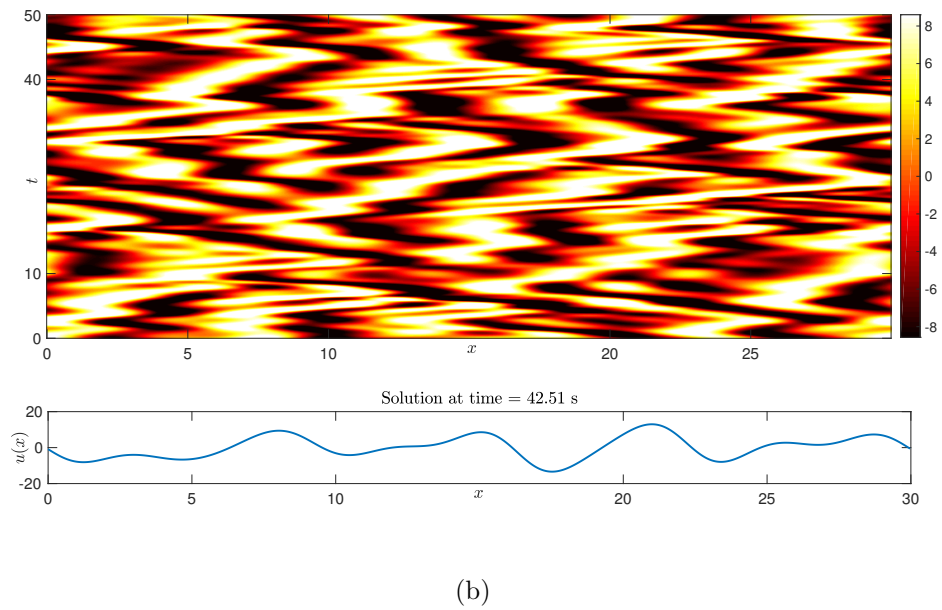
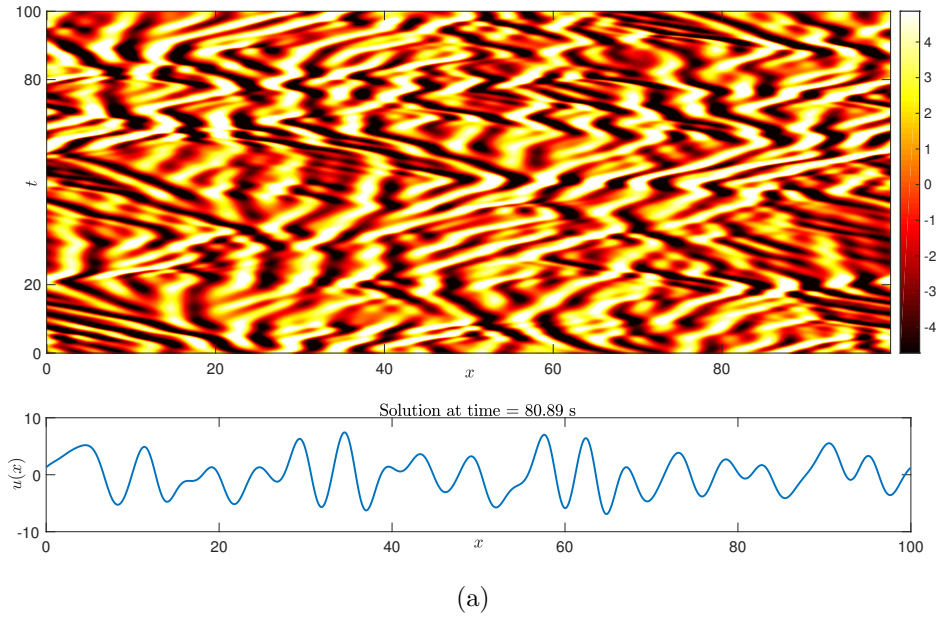


Figure 6.3: Plots of solutions to the Nikolaevskiy equation (1.4) well into the KS-type chaos regime. The top panels display the solutions over a range of  $t$ , while the bottom panels display a snapshot of each solution at an arbitrary time. Solutions are shown at parameter values (a)  $r = 0.5$  and  $\ell = 100$  and at (b)  $r = 1$  and  $\ell = 30$ .

The transition between these two types of spatiotemporal chaos occurs around  $r \approx 0.1$ . The solution shown in Figure 6.4 is in this regime; we see that solutions can readily be distinguished from those of the KS equation (1.1) with the naked eye, but we do not yet see the clear separation of scales visible in solutions for smaller  $r$  values.



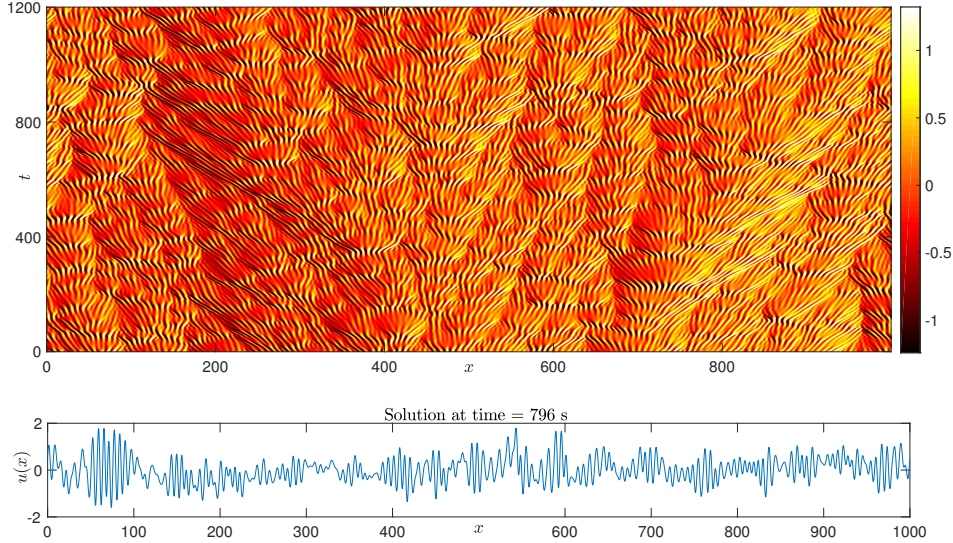


Figure 6.4: Plot of solutions to the Nikolaevskiy equation (1.4) in the transition region between Nikolaevskiy-type and KS-type chaos at parameter values  $r = 0.1$  and  $\ell = 1000$ . The top panel displays the solution over a range of  $t$ , while the bottom panel displays a snapshot of the solution at an arbitrary time.

### Scaling of Norms:

One more result we should mention is the scaling of the norms of the solutions of equation (1.4). As mentioned in Section 3.1, we expect  $\|u\|_{\dot{L}^2} = \mathcal{O}(\ell^p)$  and  $\|u\|_{\dot{L}^\infty} = \mathcal{O}(\ell^{p'})$  for some exponents  $p$  and  $p'$ . Figures 6.5a and 6.5b show the scaling of these exponents for the  $\dot{L}^2$  norm and the  $\dot{L}^\infty$  norm, respectively. As for the KS equation, the  $\dot{L}^\infty$  norm appears to be  $\ell$ -independent [Wit14], since we compute the scaling exponent to be  $p' = -0.024$  with standard deviation 0.068. The scaling of the  $\dot{L}^2$  norm also seems to be the same as the scaling that is believed to be the case for the KS equation, namely 0.5. The computed scaling is  $p = 0.4923$  with standard deviation 0.0319. The results were obtained from five trials for each  $\ell$ , with  $\ell$  ranging from 500 to 1000 in increments of 100 and  $r$  fixed at  $r = 1$ .

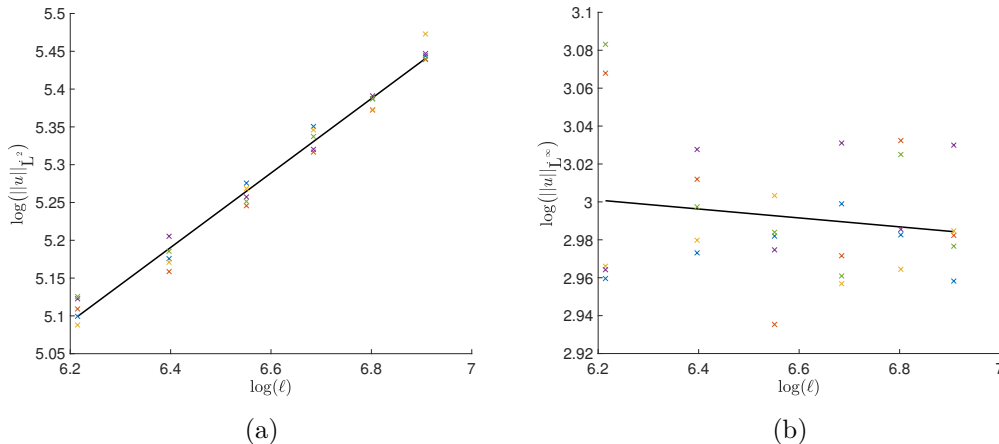


Figure 6.5: Scaling of (a) the  $\dot{L}^2$  norm and (b) the  $\dot{L}^\infty$  norm of solutions to (1.4) with  $\ell$ . Plots show the data and a best fit line on a log-log scale, for fixed  $r = 1$ . For the  $\dot{L}^2$  norm (a) the scaling of the exponent on  $\ell$  is  $0.492 \pm 0.032$ . For the  $\dot{L}^\infty$  norm (b) the scaling of the exponent on  $\ell$  is  $-0.024 \pm 0.068$ .

## 6.2 The Power Spectrum

When trying to understand certain aspects of the behaviour of solutions, such as the appearance of distinct wavelengths of the rolls, it is instructive to measure which Fourier modes carry the most energy. To see which modes carry the most energy on average, and which modes on the other hand, do not contribute significantly to the solution, one can compute the time-averaged power spectrum, defined as follows:

$$S(k) := \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t |\hat{u}_k(s)|^2 ds. \quad (6.1)$$

We will rescale  $S(k)$  with  $\ell$ , and study  $S(k)/\ell$  to minimize the effects of the domain size [WP09]. It has been suggested however, that  $S(k)/\ell$  is independent of  $\ell$  and that the energy distribution is an extensive property of the system [Tan05b]. Our computations are done for  $\ell = 50\pi/\epsilon$ , where  $\epsilon = \sqrt{r}$ , paralleling the approach in [Poo09], in order to have 25 (plus 25 complex conjugate) modes in the unstable band of the dispersion relation (1.5). Taking the time average of the power spectrum tells us which modes consistently appear in the solution. Figure 6.6 shows the time-averaged power spectra for various values of the parameter  $r$ .

Figure 6.6 indicates that for  $r \approx 1$ , the energy is concentrated in a single peak about  $k \approx 1$ . In this range of  $r$ , the power spectrum is qualitatively the same as that of the KS equation, also shown in the figure. In this range, we also generally observe KS-type chaos.

For small  $r$  ( $r \lesssim 0.01$ ) [Tan05a] the peaks are well separated. This scale separation can also readily be seen in the solutions themselves. For  $r \approx 1$  the solutions look qualitatively similar to those of the KS equation, whereas for  $r \ll 1$ , one can clearly see the peaks which

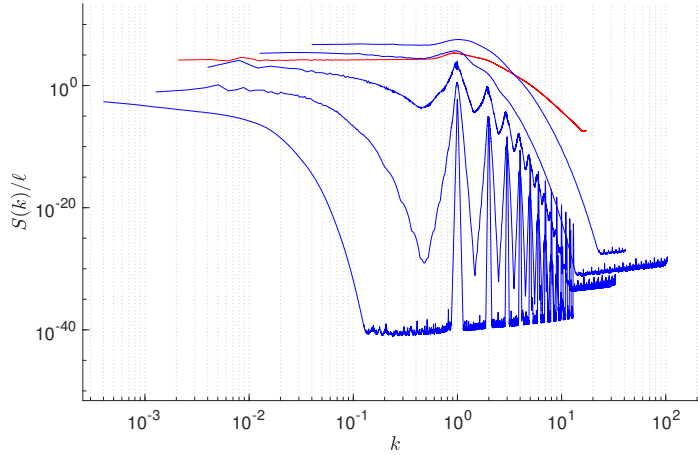


Figure 6.6: Plots of the normalized time-averaged power spectrum  $S(k)/\ell$ , for  $r = 10^{-n}$  for  $n \in \{0, 1, 2, 3, 4\}$  for the Nikolaevskiy equation (blue), as well as for the KS equation (red).

are caused by the long wavelength instabilities, together with the small-scale rolls which are caused by the short wave instabilities.

### 6.3 Lyapunov Spectrum

In this section we will plot the Lyapunov Spectra (LS) for various values of the parameters  $r$  and  $\ell$ . Figure 6.7 shows the LS, the ordered sets of Lyapunov exponents of the system for various values of  $r$ . We plotted the exponents against the wavenumber  $k$ , rather than simply against the index of the exponent. As a function of  $k$ , the shapes of the spectra do not appear to depend on  $\ell$  when  $\ell$  is large enough. For  $\ell = 100$  the values of the negative LEs are slightly bigger than for larger  $\ell$ , but the spectra converge to the same curve as  $\ell$  becomes large. We computed the exponents in the spectra using Definition 5.1.1 and Algorithm 1. We stored the approximations to each exponent periodically during the time integration, and produced a time series of exponents in this way. We computed the standard deviation of the maximum exponents for the last 25% of the integration, and used them later on to test for convergence of the time series. To minimize the effects of random fluctuations, we took as the value of the LEs the average of the last 25% of the time series of approximate LEs.

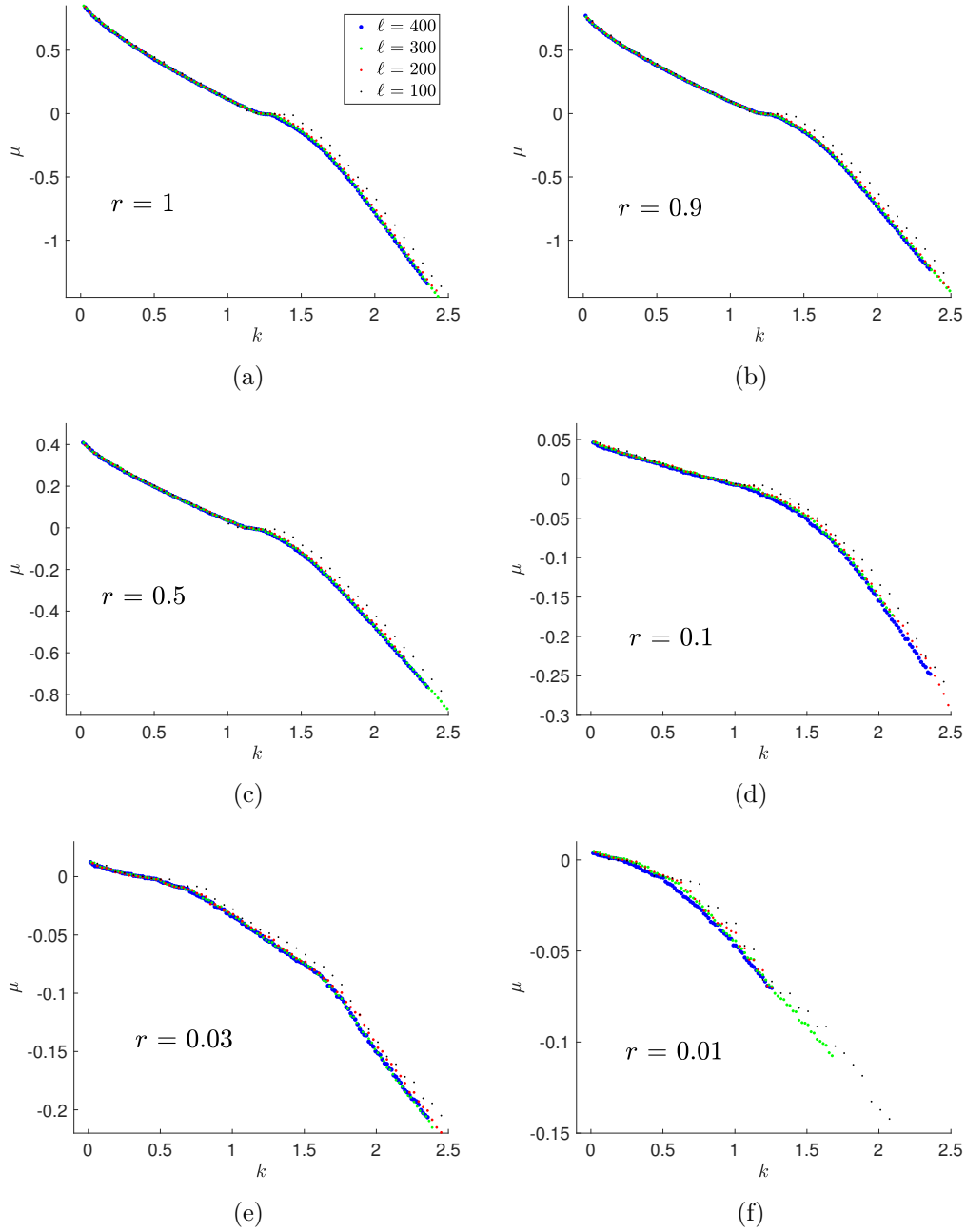


Figure 6.7: Lyapunov spectra, showing the LEs  $\mu$  vs. the wavenumber  $k$ . The parameter  $r$  is fixed at (a)  $r = 1$ , (b)  $r = 0.9$ , (c)  $r = 0.5$ , (d)  $r = 0.1$ , (e)  $r = 0.03$  and (f)  $r = 0.01$ . For each  $r$  value, the spectra are shown for system size  $\ell \in \{100, 200, 300, 400\}$ .

Again we can see the difference between the two types of chaos that can be found in solutions of equation (1.4) in the spectra in Figure 6.7. For  $r \approx 1$  the spectra have a very similar shape, with only one ‘ledge’ at around  $k \approx 1.2$  beyond which the slope of the spectrum decreases. Xi *et al.* [XTGT00] computed the LS for  $r = 0.5$  for various  $\ell$  values

$77.5 \leq \ell \leq 186$ . Unlike in Figure 6.7c the spectra in [XTGT00] lie on top of each other. It is possible that our plots for  $\ell = 100$  are not completely converged yet, but the spectra in Figure 6.7c are also computed for larger  $\ell$  values than those in [XTGT00], and thereby give a better approximation to the large- $\ell$  limit.

The shape of the LS does not seem to change appreciably as  $r$  gets large, although their height certainly changes. The maximum Lyapunov exponent and the number of positive Lyapunov exponents increases as  $r$  increases. This is to be expected, since  $r$  and  $\ell$  determine the number of modes in the unstable band of the dispersion relation (Figure 1.2). Thus larger  $r$  implies more instabilities in the system, which means that the system becomes more chaotic, which is characterized by larger and more positive LEs.

When  $r$  is decreased so that it is close to the Nikolaevskiy chaos range, the shape of the LS undergoes a clear change. Instead of just one ledge at  $k \approx 1.2$ , two ledges appear, one at  $k \approx 0.8$  and the second one at  $k \approx 1.7$ , as in Figure 6.7e. For even smaller  $r$  (Figure 6.7f) the spectrum loses much of its structure. For  $r = 0.01$ , the shape of the spectrum seems to become discontinuous for large  $k$ , in the sense that the exponents appear in pairs that are separated by a clear gap. This behaviour of the exponents has been observed in the KS equation [TYG<sup>+</sup>11], and we also observed it for the Nikolaevskiy equation. However, it is not clear if the behaviour seen in Figure 6.7f is the same, since it is only visible in the  $\ell = 100$  spectrum. It is possible that these spectra aren't fully resolved yet, because for smaller  $r$ , significantly longer time series have to be computed to obtain well-converged averages. In the limit as  $r \rightarrow 0$  the system's largest LE seems to converge to zero. This is to be expected, since  $r = 0$  marks the transition from the chaotic to the non-chaotic regime, which is characterized by an absence of positive LEs.

## 6.4 Dimension

Using the LS and the Kaplan-Yorke formula (5.5), we have computed the fractal dimension of the attractor. We will attempt to determine the dependence of the attractor dimension on both the system size  $\ell$  and the bifurcation parameter  $r$ . In order to reduce the effects that varying system sizes can have on the value of the dimension, we used two approaches when computing the dimension: Computing the dimension density, as defined below, and keeping the number of modes in the unstable band of the dispersion relation (1.5) constant. For fixed  $r$ , we computed the **dimension density**, defined as

$$D(r) := \lim_{\ell \rightarrow \infty} \frac{d_f(\mathcal{A}(\ell, r))}{\ell}, \quad (6.2)$$

(for one-dimensional domains) where we are thinking of the attractor  $\mathcal{A}$  of (1.4) as being parametrized by  $\ell$  and  $r$ .

Figure 6.9 shows a plot of the Kaplan-Yorke dimension as a function of the domain size  $\ell$ , for various values of  $r$ , as well as some log-log plots of the dimension density (6.2) versus  $r$ .

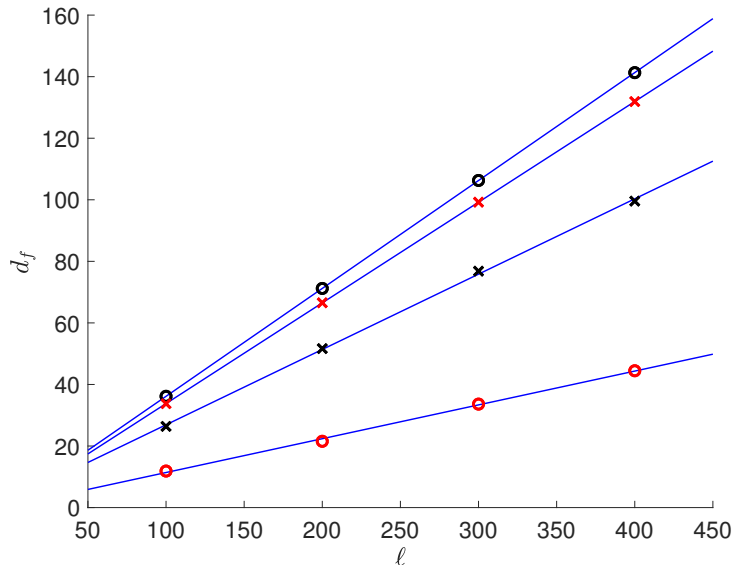


Figure 6.8: Kaplan-Yorke dimension  $d_f$  as a function of  $\ell$  for:  $\circ$   $r = 1$  with  $D = 0.351$ ,  $\times$   $r = 0.5$  with  $D = 0.328$ ,  $\times$   $r = 0.1$  with  $D = 0.244$  and  $\circ$   $r = 0.03$  with  $D = 0.110$ .

Figure 6.8 shows that the dimension is proportional to  $\ell$ , even for these relatively small values of  $\ell$ . Since the lines in Figure 6.8 fit the data very closely, this suggests that the value  $\ell = 100$  is already within the asymptotic regime for the linear scaling of the dimension  $d_f$  with  $\ell$ . The numerical scaling of the exponent with  $\ell$  is better than the  $\mathcal{O}(\ell^{19/12})$  scaling that we obtained analytically in (4.46). This suggests that the analytic results are not optimal in  $\ell$  and can still be improved.

Both the analytic results and the numerical results suggest that the dimension's dependence on  $\ell$  is constant, whenever  $\ell$  is large enough. Decreasing  $r$  does not seem to influence the fact that the dimension scales as  $\mathcal{O}(\ell)$  for fixed  $r$ . The dimension densities we computed agree with those of other authors. Xi *et al.* [XTGT00] obtained a value of  $1/3$  for  $D$  at  $r = 0.5$ . Our result of  $D \approx 0.327$  at  $r = 0.5$  is very close to this value, and we suspect that the ‘exact’ value of  $1/3$  in [XTGT00] comes from rounding an approximate value, since it was also obtained through a numerical computation.

Figures 6.9a and 6.9b show how the dimension density  $D$  scales with  $r$ . Since  $D$  is proportional to  $d_f$  and captures the ( $\ell$ -independent) behaviour of  $d_f$  for large  $\ell$ , the  $r$  scaling of  $D$  is expected to be the same as the  $r$  scaling of  $d_f$  for large  $\ell$ . For  $0.2 \leq r \leq 1$ , we have that  $D \approx \mathcal{O}(r^{0.093})$ . Comparing this with the analytic bound (4.46) we obtained

for the dimension, we see that the exponent 0.093 is bounded by the exponent  $7/48$  that was obtained analytically.

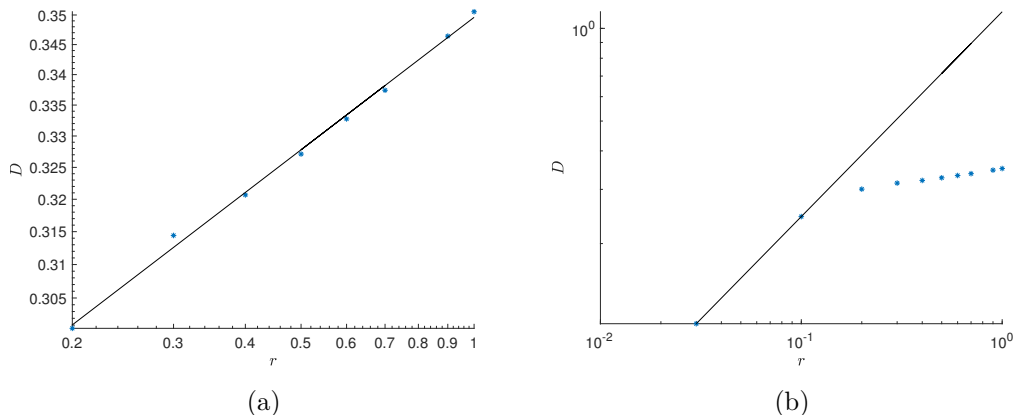


Figure 6.9: (a) Log-log plot of the dimension density  $D$  versus  $r$ , over a range  $0.2 \leq r \leq 1$ . The scaling exponent on  $r$  is approximately 0.093. (b) A log-log plot of  $D$  vs.  $r$ , over a range  $0.03 \leq r \leq 1$ . The scaling exponent on  $r$  for small  $r$  is approximately 0.665.

If we turn to Figure 6.9b however, we see that when  $r$  is decreased below 0.2, then the scaling of the exponent on  $r$  changes drastically. Taking only the last two points, we compute that the slope is approximately 0.665, suggesting that in the limit as  $r \rightarrow 0$ , we have  $d_f = \mathcal{O}(r^p)$  for  $p \geq 0.665$ ! This means that in order to capture the asymptotic scaling of  $d_f$  with  $r$ , one would have to compute the Kaplan-Yorke dimension for  $r$  values below 0.2. In fact, as we will see, even  $r = 0.2$  is not small enough yet, and in order to compute asymptotic scalings as  $r \rightarrow 0$ , one has to compute the dimension for  $r$  values below  $r = 0.01$ !

As mentioned earlier we used two approaches when computing the dimension. In the second approach, we tried to determine the scaling of the dimension  $d_f$  with  $r$  by keeping the number of unstable modes constant while varying  $r$ . The purpose of this was to minimize the effect that different system sizes could have on the results. We followed the approach in [WP09], and chose the values of  $\ell$  and  $r$  in such a way that the number of Fourier modes in the unstable band of the linear dispersion relation (1.5) remains approximately constant. Recall that the linear dispersion relation can be written as  $\omega(k) = -k^2(r - (1 - k^2)^2)$ . Therefore, now writing  $r := \epsilon^2$ , the width of the unstable band (for  $k \geq 0$ ) is

$$\Delta k = \sqrt{1 + \epsilon} - \sqrt{1 - \epsilon} = \left(1 + \frac{\epsilon}{2}\right) - \left(1 - \frac{\epsilon}{2}\right) + \mathcal{O}(\epsilon^2) = \epsilon + \mathcal{O}(\epsilon^2).$$

Then since  $k = 2\pi n/\ell$ , we have that the number of Fourier modes in the unstable band is

$$\Delta n = \frac{\ell \epsilon}{\pi}, \tag{6.3}$$

to leading order. Note that this number isn't  $\ell\epsilon/(2\pi)$ , because technically there are two unstable bands, one for positive  $k$  and one for negative  $k$ , even though we only plot the positive one.

We then computed the Kaplan-Yorke dimension for various decreasing values of  $r$ , keeping the number of modes  $\Delta n$  in the unstable band fixed at 50, 36 and 24. We chose the same  $r$  values as in [WP09]; namely, for each fixed value of  $\Delta n$ , we computed the dimension for  $r = 0.1, 0.04, 0.01, 0.001$  and  $0.0001$ , letting  $\ell$  be induced by (6.3). We computed solutions on a grid of  $2^n$  points, where  $n$  is an integer ranging from 11 to 16, and using a time step in a range between  $dt = 0.001$  and  $dt = 1$ . Initial conditions were taken as solutions that are already on the attractor of equation (1.4), and were computed over times varying from  $t = 3500$  to  $t = 10^5$ , where the computation times increased with decreasing  $r$ . In order to check that the computations converged sufficiently, we stored the approximations to the LS periodically throughout the time integration of the solutions. We then computed the standard deviation of all the approximation over the last 25% of the total integration time, and found that its value scaled with the value of the maximum LE. If the standard deviation (for the maximum exponent) was less than 5% of the maximum exponent, then we took the results to be sufficiently converged. For small values of  $r$ , it wasn't feasible to run longer time averages due to time constraints.

Figure 6.10 shows the Kaplan-Yorke dimension as a function of  $r$  for fixed values of  $\Delta n$ .

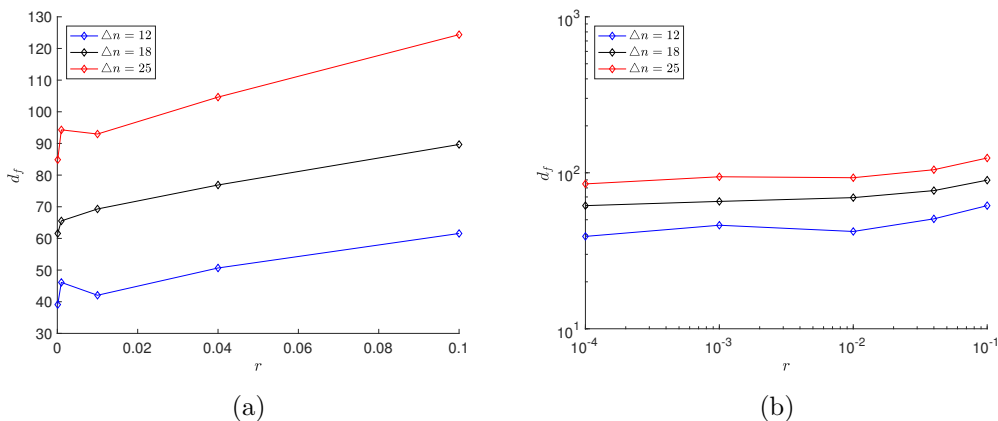


Figure 6.10: (a) Kaplan-Yorke dimension  $d_f$  versus  $r$  for fixed  $\Delta n$ . (b) Log-log plot of Kaplan-Yorke dimension  $d_f$  versus  $r$  for fixed  $\Delta n$ . In both plots,  $\Delta n \in \{24, 36, 50\}$ .

As the number of unstable modes  $\Delta n$  increases, the Kaplan-Yorke dimension of the attractor also increases. This is to be expected from the  $\ell$  scaling in Figure 6.9. In general, the value of the dimension does decrease with  $r$ ; however, for  $\Delta n = 24$  and  $\Delta n = 50$ , at  $r = 0.01$  the value of the computed value of the dimension is actually smaller than at  $r = 0.001$ , suggesting that the result of the computation is not accurate for this  $r$  value.



Even though the dimension increases with  $\Delta n$  overall, there is no clear rate of convergence visible with  $r$  yet. In fact, for larger  $r$  values, the rate of convergence seems to be somewhat slower than for smaller  $r$  values, as shown in the log-log plots in Figure 6.10b. The plots seem to suggest that the value of  $r$  is not yet small enough to capture the asymptotic convergence rate that the dimension of the attractor obeys as  $r \rightarrow 0$ . There is however, good reason to doubt the accuracy of the results in these plots, since the type of behaviour shown in them is not seen in any of our other computations. In all likelihood, the values given in the figures are not very good and the value of the dimension has not yet fully converged; one would have to compute averages over a much longer time to obtain accurate results.

Both Tanaka [Tan05a], and the power spectrum in Figure 6.1, suggest that the scale separation that characterizes Nikolaevskiy chaos does not fully develop until  $r \leq 0.01$  at least, even though the solutions themselves already look qualitatively different for  $r$  values that are slightly larger than 0.01, as can be seen in Figure 6.1 for instance. The value of  $r = 0.01$  is right around the value at which we see the slope of the Kaplan-Yorke dimension as a function of  $r$  change in Figure 6.10. This suggests that the type of chaos exhibited by solutions of equation (1.4) influences the  $r$  scaling of the attractor dimension. Analytically, we saw that for any  $r \leq 1$ , the attractor dimension scales at least as  $r^{7/48}$  for large enough system size (4.46). This means that even if the scaling changes during the transition period when  $r \approx 0.01$ , we still have an  $r$  scaling of at least  $r^{7/48}$ .

This suggests that the reason why we obtained a relatively loose lower bound for the scaling in (4.46) is because we were trying to bound the behaviour for two qualitatively different regimes of the dynamics of the Nikolaevskiy equation (1.4). As further work, it may therefore be possible to derive tighter bounds on the scaling of the attractor with  $r$ , if instead of assuming that  $r < 1$ , we assume that  $r < 0.01$  or even smaller.

# Chapter 7

## Conclusion

In this final chapter we summarize and elaborate on the main findings in the thesis. Specifically, we will review what we found about the scaling of various norms of  $u$  with  $\ell$  and with  $r$ , and comment on the extent to which these bounds are optimal. We will then state our main results, namely the scaling of the fractal dimension of the attractor  $d_f$  with the parameters in the equation. Finally, we will propose some further work that could be done to improve the results that we have so far.

### Scaling of Norms:

In Sections 3.2 and 3.3 we rigorously derived bounds on the radii of the absorbing balls of solutions in the Sobolev spaces  $\dot{H}^1$  and  $\dot{H}^2$ , and showed that they scale with the parameters  $r$  and  $\ell$ . Specifically, we found that

$$\limsup_{t \rightarrow \infty} \|\partial_x u\|_{\dot{L}^2} = \mathcal{O}(r^{5/8} \ell^2),$$

and

$$\limsup_{t \rightarrow \infty} \|\partial_x^2 u\|_{\dot{L}^2} = \mathcal{O}(r^{7/8} \ell^3).$$

Deriving the optimal scaling in  $r$  and  $\ell$  rigorously remains an open problem, but in Section 6.1 we provided numerical evidence which shows that the scaling of the  $\dot{L}^2$  norm with  $\ell$  is in fact  $\ell^{1/2}$ , as is generally believed to be the case for the Nikolaevskiy equation.

In Section 6.1 we also provided numerical evidence that the  $\dot{L}^\infty$  norm of  $u$  is  $\ell$ -independent. Using the inequality

$$\|u\|_{\dot{L}^2} \leq \ell^{1/2} \|u\|_{\dot{L}^\infty} \leq \ell \|\partial_x u\|_{\dot{L}^2},$$

and the results about the  $r$  scaling of the absorbing ball in  $\dot{H}^1$ , we can thus make the case that  $\|u\|_{\dot{L}^\infty} = \mathcal{O}(r^{5/8})$ .

### Scaling of the Dimension of the Attractor:

In Section 4.7, we found that the fractal dimension  $d_f$  of the attractor of the Nikolaevskiy equation obeys the scaling

$$d_f = \mathcal{O}(r^{7/48}\ell^{19/12})$$

for  $0 < r \leq 1$ , whenever  $\sqrt{r}\ell$  is large. In Section 6.4, we found numerical evidence that the  $\ell$  scaling of the dimension is  $d_f = \mathcal{O}(\ell)$  for fixed  $r$ , rather than the scaling of  $\ell^{19/12}$  that we derived analytically. As the analytic bounds depend on the scaling of the  $\dot{L}^2$  norm of  $u$ , it is unlikely that significantly better bounds can be derived rigorously until better bounds on the  $\dot{L}^2$  norm of  $u$  are found.

We also computed the  $r$  scaling of  $d_f$  and of the dimension density  $D$  (which we expect to be the same as that of  $d_f$ ) for  $r < 1$  in Section 6.4. We found that for  $0.2 \leq r \leq 1$  the scaling of the dimension density is  $D \approx \mathcal{O}(r^{0.092916})$ , which is bounded by the analytically derived scaling of  $d_f = \mathcal{O}(r^{7/48})$ . Furthermore, we showed that as  $r$  decreases below 0.2, the scaling exponent once more increases. However, for the range of  $r$  values that we computed, we were unable to find the asymptotic  $r$  scaling of  $d_f$  yet, as all our results indicate that one has to compute the dimension for  $r$  values well below 0.01 to hope to get into the regime of the asymptotic limit.

### Further Work:

There are two projects that immediately come to mind which could be undertaken to extend the work of this thesis: The first is to continue the numerical calculations to estimate the fractal dimension of the attractor for  $r$  values less than 0.001. For  $r$  values in this range, the power spectrum already possesses strong separation between the long and short wavelength unstable modes, and so it is reasonable to expect any values of  $r$  less than 0.001 (and in fact even less than 0.01) to give results that accurately capture the scaling of the dimension of the attractor in the Nikolaevskiy chaos regime.

Working along similar lines, a second project would be to improve the analytic estimates on the fractal dimension for very small  $r$  values. As we saw, the reason why we obtained seemingly large bounds on the dimension estimate was because our estimate included  $r$  values that were still close to one. It would thus be interesting to see if the analytic bound on the fractal dimension can be improved further, if one makes the assumption that  $r \ll 1$ .

Our results suggest that the full Nikolaevskiy chaos regime does not begin until  $r$  is very small, and therefore it would be interesting to investigate the dynamics of the Nikolaevskiy equation for such  $r$  values.

# Bibliography

- [Bea96] A. F. Beardon. Sums of powers of integers. *American Mathematical Monthly*, **103**:201–213, 1996.
- [Bel97] R. Bellman. *Introduction to Matrix Analysis, Second Edition*. SIAM, Philadelphia, 2nd edition, 1997.
- [BG06] J. C. Bronski and T. N. Gambill. Uncertainty estimates and  $L_2$  bounds for the Kuramoto-Sivashinsky equation. *Nonlinearity*, **19**:2023–2039, 2006.
- [Bre10] A. Bressan. *Lecture Notes on Functional Analysis*. American Mathematical Society, Providence, Rhode Island, 2010.
- [CEES93a] P. Collet, J.-P. Eckmann, H. Epstein, and J. Stubbe. Analyticity for the Kuramoto-Sivashinsky equation. *Physica D*, **67**:321–326, 1993.
- [CEES93b] P. Collet, J.-P. Eckmann, H. Epstein, and J. Stubbe. A global attracting set for the Kuramoto-Sivashinsky equation. *Communications in Mathematical Physics*, **152**:203–214, 1993.
- [Cer94] S. Cerrai. A Hille-Yosida theorem for weakly continuous semigroups. *Semigroup Forum*, **49**:349–367, 1994.
- [CH93] M. C. Cross and P. C. Hohenberg. Pattern formation outside of equilibrium. *Review of Modern Physics*, **65**:851–1112, 1993.
- [CH94] M. C. Cross and P. C. Hohenberg. Spatiotemporal chaos. *Science*, **263**(5153):1569–1569, 1994.
- [CM02] S. M. Cox and P. C. Matthews. Exponential time differencing for stiff systems. *Journal of Computational Physics*, **176**:430–455, 2002.
- [Cul66] W. J. Culver. On the existence and uniqueness of the real logarithm of a matrix. *Proceedings of the American Mathematical Society*, **17**:1146–1151, 1966.
- [EP98] S. V. Ershov and A. B. Potapov. On the concept of stationary Lyapunov basis. *Physica D*, **118**:167–198, 1998.
- [Eva15] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, Providence, Rhode Island, 2nd edition, 2015.
- [FR08] M. Frankel and V. Roytburd. Stability for a class of nonlinear pseudo-differential equations. *Applied Mathematics Letters*, **21**:425–430, 2008.

- [GF18] D. Goluskin and G. Fantuzzi. Bounds on mean energy in the Kuramoto–Sivashinsky equation computed using semidefinite programming. *Preprint: arXiv:1802.08240*, 2018.
- [GO04] L. Giacomelli and F. Otto. New bounds for the Kuramoto-Sivashinsky equation. *Communications on Pure and Applied Mathematics*, **58**(3):297–318, 2004.
- [GPT<sup>+</sup>07] F. Ginelli, P. Poggi, A. Turchi, H. Chaté, R. Livi, and A. Politi. Characterizing dynamics with covariant Lyapunov vectors. *Physical Review Letters*, **99**:130601, 2007.
- [Gre96] H. S. Greenside. Spatiotemporal chaos in large systems: The scaling of complexity with size. *Preprint: arXiv:chao-dyn/9612004*, 1996.
- [HN86] J. M. Hyman and B. Nicolaenko. The Kuramoto-Sivashinsky equation: A bridge between PDEs and dynamical systems. *Physica D*, **18**(1–3):113–126, 1986.
- [Hun96] B. R. Hunt. Maximum local Lyapunov dimension bounds the box dimension of chaotic attractors. *Nonlinearity*, **9**(4):845–852, 1996.
- [IS14] X. Ioakim and Y.-S. Smyrlis. Analyticity for Kuramoto-Sivashinsky type equations and related systems. *Procedia IUTAM*, **11**:69–80, 2014.
- [Joh75] F. John. *Partial Differential Equations*. Springer, New York, 2nd edition, 1975.
- [Kre89] E. Kreyszig. *Introductory Functional Analysis with Applications*. Wiley, New Delhi, 1989.
- [KT76] Y. Kuramoto and T. Tsuzuki. Persistent propagation of concentration waves in dissipative media far from thermal equilibrium. *Progress of Theoretical Physics*, **55**(2):356–369, 1976.
- [KT05] A.-K. Kassam and L. N. Trefethen. Fourth-order time-stepping for stiff PDEs. *SIAM Journal on Scientific Computing*, **26**(4):1214–1233, 2005.
- [KY79] J. L. Kaplan and J. A. Yorke. Chaotic behavior of multidimensional difference equations. *Functional Differential Equations and the Approximation of Fixed Points.*, **730**:204–227, 1979.
- [LeV07] R. LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations*. SIAM, Philadelphia, 2007.
- [Li04] Y. Charles Li. *Chaos in Partial Differential Equations*. International Press, Missouri, 2004.
- [LMRT75] L. E. LaQuey, S. M. Mahajan, P. H. Rutherford, and W. M. Tang. Nonlinear saturation of the trapped-ion mode. *Physical Review Letters*, **34**(7):391–394, 1975.
- [Lor63] E. N. Lorenz. Deterministic nonperiodic flow. *Journal of Atmospheric Science*, **20**:130–141, 1963.

- [Man85] P. Manneville. Liapounov exponents for the Kuramoto-Sivashinsky model. In *Macroscopic Modelling of Turbulent Flows*, pages 319–326, Berlin, Heidelberg, 1985. Springer.
- [MC00] P. C. Matthews and S. M. Cox. One-dimensional pattern formation with Galilean invariance near a stationary bifurcation. *Physical Review E*, **62**:R1473–R1476, 2000.
- [Mei17] J. D. Meiss. *Differential Dynamical Systems, Revised Edition*. SIAM, Philadelphia, 2017.
- [Mol93] T. C. A. Molteno. Fast  $o(n)$  box-counting algorithm for estimating dimensions. *Physical Review E*, **48**(5):R3263, 1993.
- [Nik89] V. N. Nikolaevskii. Dynamics of viscoelastic media with internal oscillators. *Lecture Notes on Engineering*, **39**:210–221, 1989.
- [NST85] B. Nicolaenko, B. Scheurer, and R. Temam. Some global dynamical properties of the Kuramoto-Sivashinsky equation: Nonlinear stability and attractors. *Physica D*, **16**:155–183, 1985.
- [Ose68] V. I. Oseledets. A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. *Transactions of the Moscow Mathematical Society*, **19**:179–210, 1968.
- [OY08] W. Ott and J.A. Yorke. When Lyapunov exponents fail to exist. *Physical Review E*, **78**:056203, 2008.
- [PC89] T. S. Parker and L. O. Chua. *Practical Numerical Algorithms for Chaotic Systems*. Springer, New York, 1989.
- [Per00] L. Perko. *Differential Equations and Dynamical Systems*. Springer, New York, 3rd edition, 2000.
- [Poo09] K.-F. Poon. *Dynamics of the Nikolaevskiy and related equations*. PhD thesis, Simon Fraser University, 2009.
- [PP16] A. Pikovsky and A. Politi. *Lyapunov Exponents*. Cambridge University Press, Cambridge, 2016.
- [Rob01] J. C. Robinson. *Infinite-Dimensional Dynamical Systems*. Cambridge University Press, Cambridge, 2001.
- [RSK11] P. N. Ryabov, D. I. Sineshchikov, and M. B. Kochadov. Application of the Kudryashov method for finding exact solutions of the high order nonlinear evolution equations. *Applied Mathematics and Computation*, **218**(7):3965–3972, 2011.
- [SBH99] P. So, E. Barreto, and B. R. Hunt. Box-counting dimension without boxes: Computing  $D_0$  from average expansion rates. *Physical Review E*, **60**(1):378–85, 1999.

- [Siv77] G. I. Sivashinsky. Nonlinear analysis of hydrodynamic instability in laminar flames—I. Derivation of basic equations. *Acta Astronautica*, **4**:1177–1206, 1977.
- [Siv80] G. I. Sivashinsky. On flame propagation under conditions of stoichiometry. *SIAM Journal of Applied Mathematics*, **39**(1):67–82, 1980.
- [SN79] I. Shimada and T. Nagashima. A numerical approach to ergodic problem of dissipative dynamical systems. *Progress of Theoretical Physics*, **61**(6):1605–1616, 1979.
- [Str15] S. H. Strogatz. *Nonlinear Dynamics and Chaos*. Westview Press, Colorado, 2nd edition, 2015.
- [Tan05a] D. Tanaka. Bifurcation scenario to Nikolaevskiy turbulence in small systems. *Journal of the Physical Society of Japan*, **74**(8):2223–2225, 2005.
- [Tan05b] D. Tanaka. Critical exponents of Nikolaevskii turbulence. *Physical Review E*, **71**, 2005.
- [TB97] L. N. Trefethen and D. Bau. *Numerical Linear Algebra*. SIAM, Philadelphia, 3rd edition, 1997.
- [Tem98] R. Temam. *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*. Springer, New York, 2nd edition, 1998.
- [Tre00] L. N. Trefethen. *Spectral Methods in MATLAB*. SIAM, Philadelphia, 2000.
- [TT96] M. I. Tribelsky and K. Tsuboi. New scenario for transition to turbulence? *Physical Review Letters*, **76**(10):1631–1634, 1996.
- [TV96] M. I. Tribelsky and M. G. Velarde. Short-wavelength instability in systems with slow long-wavelength dynamics. *Physical Review E*, **54**(5):4973–4981, 1996.
- [TYG<sup>+</sup>11] K. A. Takeuchi, H.-L. Yang, F. Ginelli, G. Radons, and H. Chaté. Hyperbolic decoupling of tangent space and effective dimension of dissipative systems. *Physical Review E*, **84**:1–19, 2011.
- [WH99] R. W. Wittenberg and P. Holmes. Scale and space localization in the Kuramoto-Sivashinsky equation. *Chaos*, **9**(2):452–465, 1999.
- [Wit02] R. W. Wittenberg. Dissipativity, analyticity and viscous shocks in the (de)stabilized Kuramoto-Sivashinsky equation. *Physics Letters A*, **300**:407–416, 2002.
- [Wit14] R. W. Wittenberg. Optimal parameter-dependent bounds for the Kuramoto-Sivashinsky equation. *Discrete and Continuous Dynamical Systems A*, **34**(12):5325–5357, 2014.
- [WP09] R. W. Wittenberg and K.-F. Poon. Anomalous scaling on a spatiotemporally chaotic attractor. *Physical Review E*, **79**:056225, 2009.
- [XTGT00] H.-W. Xi, R. Toral, J. D. Gunton, and M. I. Tribelsky. Extensive chaos in the Nikolaevskii model. *Physical Review E*, **62**(1):17–20, 2000.



[YHK93] M. Yamaguti, M. Hata, and J. Kigami. *Mathematics of Fractals*. American Mathematical Society, Providence, Rhode Island, 1993.

# Appendix A

## Table of Notation

- $A$  : A spatial differential operator of the form  $A = \partial_x^n$  for  $n \in \mathbb{N}$  or  $A = |\partial_x|$ .
- $\mathcal{A}$  : The attractor of (1.4), as well as a general attractor.
- $b(u, v, w)$  : The trilinear form (2.14)  $b(u, v, w) := \int_0^\ell u \partial_x v w \, dx$ .
- $B$  : A general Banach space.
- $\mathcal{B}$  : The absorbing set of (1.4), as well as of a general dynamical system.
- $\mathcal{D}$  : The space of test functions.
- $d_f(X)$  : The fractal dimension of the set  $X$ .
- $f(x), f : H \rightarrow H$  : The right-hand side a dynamical system, may contain spatial differential operators.
- $\mathcal{G}_u$  : The operator (4.31) linearised about a solution to (1.4).
- $H$  : A general Hilbert space.
- $\dot{H}^n$  : The Sobolev space  $H^n(0, \ell)$  with periodic, zero mean elements,  $n \in \mathbb{N}$ .
- $K$  : The constant, independent of  $r$  or  $\ell$ , in the  $\dot{L}^2$  bound on  $u$  (3.1).
- $\ell$  : The spatial domain size.
- $\mathcal{L}$  : The linear operator of (1.4),  $\mathcal{L} = \partial_x^6 + 2\partial_x^4 + (1 - r)\partial_x^2$ .
- $\Lambda$  : A general linear operator.
- $\dot{L}^2$  : The Lebesgue space  $L^2(0, \ell)$  with periodic, zero mean elements.
- $\mathcal{N}(\cdot)$  : The nonlinear term in (1.4),  $\mathcal{N}(u) := u \partial_x u$ .
- $\omega(k)$  : The linear dispersion relation (1.5) of (1.4).
- $r$  : The parameter in the coefficient of the second-order term in (1.4).

- $R$ : The radius of the  $\dot{L}^2$  absorbing ball of (1.4).
- $S(k)$ : The time-averaged power spectrum (6.1).

# Appendix B

## Lemmas and Tools

### B.1 ODE Theory

In Section 2.2, the system of ODEs (2.11) obtained by Galerkin projections are solved to approximate the solution to the full PDE (1.4). Here we state some results which allow us to obtain solutions to ODEs. We don't give the proofs since they are somewhat lengthy, but proofs can be found in any book on ODE theory.

**Theorem B.1.1** (Existence and Uniqueness for ODEs). *Consider the dynamical system*

$$\frac{dx}{dt} = f(x), \quad x(0) = x_0. \quad (\text{B.1})$$

*If for any bounded set  $B$  containing  $x_0$  there exists a constant  $L(B)$  such that  $f$  satisfies*

$$|f(x) - f(y)| \leq L(B)|x - y|$$

*for any  $x, y \in B$ , then there exists a  $T = T(x_0)$  such that (B.1) has a unique solution on  $[0, T]$ .*

**Lemma B.1.2** (Maximal Interval of Existence). *A solution  $x(t)$  to (B.1) has a finite maximal interval of existence  $[0, \tilde{T})$  if and only if  $|x(t)| \rightarrow \infty$  as  $t \rightarrow \tilde{T}$ .*

Theorem B.1.1 is stated in [Rob01, p. 45]. For a proof of Theorem B.1.1 and Lemma B.1.2 see for instance [Per00, pp. 70 & 87].

### B.2 Functional Analysis and Linear Algebra

Convergence of the Galerkin solutions (2.9) to a solution of (1.4) is proved in a weak sense. In this section some functional analysis and general convergence results are stated. Proofs are given when the proof is illuminating, or when the result in question is less general and applies specifically to the spaces in which we are working in the thesis.

**Theorem B.2.1** (Alaoglu Compactness Theorem). *Let  $X$  be a reflexive Banach space with dual  $X^*$ . Then if  $x_n$  is a bounded sequence in  $X$ , there exists a subsequence  $x_{n_j}$  of  $x_n$  and an element  $x \in X$  such that*

$$\lim_{j \rightarrow \infty} \langle f, x_{n_j} \rangle = \langle f, x \rangle \quad (\text{B.2})$$

for every  $f \in X^*$ , where  $\langle f, x \rangle$  denotes the action of  $f$  on  $x$ . That is, the sequence  $x_n$  contains a weakly convergent subsequence in  $X$ .

For a proof of Theorem B.2.1 see [Rob01, p. 106].

**Lemma B.2.2** (Ehrling's Lemma [Rob01, p. 215]). *Suppose we have three Banach spaces  $X, Y$  and  $H$ , such that  $X \subset\subset H \subset Y$  (here  $\subset\subset$  denotes a compact embedding). Then for each  $\eta > 0$  there exists a constant  $c_\eta$  such that*

$$\|u\|_H \leq \eta \|u\|_X + c_\eta \|u\|_Y \quad \text{for all } u \in X.$$

*Proof.* Suppose not; then exists  $\eta > 0$  such that for every  $n \in \mathbb{N}$ , there exists some  $u_n \in X$  such that

$$\|u_n\|_H > \eta \|u_n\|_X + n \|u_n\|_Y.$$

Note that  $u_n \neq 0$  under these assumptions. Now let  $v_n := u_n / \|u_n\|_X$ ; then since  $X$  is continuously embedded in  $H$ , it follows that there exists some constant  $C > 0$  such that

$$C \geq \frac{\|u_n\|_H}{\|u_n\|_X} > \eta + n \|v_n\|_Y.$$

But this implies that  $v_n \rightarrow 0$  in  $Y$ . Furthermore,  $v_n$  is bounded in  $X$ , since  $\|v_n\|_X = 1$  for all  $n$ . Thus by the compact embedding of  $X$  in  $H$ ,  $v_n$  contains a subsequence (which we will relabel as  $v_n$ ) that also converges in  $H$ , and since 0 is a common element of both  $Y$  and  $H$ , it must also tend to 0 in  $H$ , which contradicts the fact that

$$\|v_n\|_H > \eta > 0.$$

□

**Corollary B.2.3.** *For  $p > 1$  we also have*

$$\|u\|_H^p \leq \eta \|u\|_X^p + c_\eta \|u\|_Y^p \quad \text{for all } u \in X, \quad (\text{B.3})$$

for some  $c_\eta > 0$ .

The proof of the above result follows from the triangle inequality.

The following result is very important for Section 2.4 in that it allows one to obtain convergence of the nonlinear term in (2.10) to the one in (1.4) in a distributional sense. We will give the full proof of this theorem as it given in [Rob01].

**Theorem B.2.4** (Strong Convergence from Weak Convergence [Rob01, p. 214]). *Let  $X \subset \subset H \subset Y$  be Banach spaces, where  $X$  is reflexive, and  $H$  is continuously embedded in  $Y$ . If  $u_n$  is a sequence that is uniformly bounded in  $L^2(0, T, X)$ , and  $\partial_t u_n$  is uniformly bounded in  $L^2(0, T, Y)$ , then there is a subsequence of  $u_n$  that converges strongly in  $L^2(0, T, H)$ .*

*Proof.* First since  $u_n$  is uniformly bounded in  $L^2(0, T; X)$  there exists some subsequence, which we will also call  $u_n$  such that  $u_n \rightharpoonup u$  for some  $u \in L^2(0, T; X)$  (Theorem B.2.1). In order to show that  $u_n \rightarrow u$  in  $L^2(0, T; H)$ , we will first show that if  $u_n \rightarrow u$  in  $L^2(0, T; Y)$  then we also have  $u_n \rightarrow u$  in  $L^2(0, T; H)$ . Of course we haven't yet proved the convergence in  $L^2(0, T; Y)$ . For this purpose define  $v_n := u_n - u$ . Then by the corollary to Ehrling's Lemma B.2.3 we have that for each  $\eta > 0$  there exists  $c_\eta$  such that

$$\|v_n\|_H^2 \leq \eta \|v_n\|_X^2 + c_\eta \|v_n\|_Y^2,$$

which, if we integrate over time from 0 to  $T$  and use the fact that  $v_n$  is uniformly bounded in  $L^2(0, T; X)$ , becomes

$$\|v_n\|_{L^2(0, T; H)}^2 \leq \eta C + c_\eta \|v_n\|_{L^2(0, T; Y)}^2.$$

Here  $C$  depends on the uniform bound on the  $u_n$ , as well as on the norm of  $u$ , but not on  $\eta$ . But since  $v_n$  tends to zero in  $L^2(0, T; Y)$ , we have that

$$\limsup_{n \rightarrow \infty} \|v_n\|_{L^2(0, T; H)}^2 \leq \eta C,$$

which is true for any  $\eta > 0$ , and so we must have that

$$\lim_{n \rightarrow \infty} \|v_n\|_{L^2(0, T; H)}^2 = 0.$$

Thus under the assumption that  $u_n \rightarrow u$  in  $L^2(0, T, Y)$ , we also have that  $u_n \rightarrow u$  in  $L^2(0, T, H)$ .

To continue, we first note that  $\partial_t v_n \in L^2(0, T; Y)$  implies that  $v_n \in H^1(0, T; Y)$ , so we have that  $v_n \in C([0, T], Y)$ , and we have the uniform bound

$$\max_{0 \leq t \leq T} \|v_n(t)\|_Y \leq \|v_n\|_{H^1(0, T; Y)} \leq M.$$

As an aside, we should note here that we have assumed that  $u$  has a weak time derivative in  $L^2(0, T; Y)$  and that  $\partial_t u_n$  converges to this derivative. By the fundamental theorem of calculus [Rob01, p. 192] we also have that

$$v_n(t) = v_n(\beta) - \int_t^\beta \partial_\tau v_n(\tau) d\tau,$$

which we can integrate with respect to  $\beta$  from  $t$  to  $t + s$  to get

$$v_n(t) = \frac{1}{s} \left[ \int_t^{t+s} v_n(\beta) d\beta - \int_t^{t+s} \int_t^\beta \partial_\tau v_n(\tau) d\tau d\beta \right] = a_n + b_n.$$

Now we find a more useful form for  $b_n$ , which we will achieve by switching the order of integration of the double integral. We find

$$\begin{aligned} b_n &= -\frac{1}{s} \int_t^{t+s} \int_t^\beta \partial_\tau v_n(\tau) d\tau d\beta = -\frac{1}{s} \int_t^{t+s} \int_\tau^{t+s} \partial_\tau v_n(\tau) d\beta d\tau \\ &= -\frac{1}{s} \int_t^{t+s} (t+s-\tau) \partial_\tau v_n(\tau) d\tau = -\frac{1}{s} \int_t^{t+s} (t+s-\beta) \partial_\tau v_n(\beta) d\beta, \end{aligned}$$

where, in the last step, we relabelled  $\tau$  as  $\beta$ . In this form, we can now bound  $b_n$  by

$$|b_n| \leq \frac{1}{s} \int_t^{t+s} |t+s-\beta| |\partial_\tau v_n(\beta)| d\beta \leq \int_t^{t+s} |\partial_\tau v_n(\beta)| d\beta,$$

since for  $\beta \in [t, t+s]$ ,  $|t+s-\beta| \leq s$ . So now using the triangle inequality property of norms, as well as Hölder's inequality we get that

$$\|b_n\|_Y \leq \int_t^{t+s} \|\partial_\tau v_n(\beta)\|_Y d\beta \leq s^{1/2} \|\partial_\tau v_n\|_{L^2(0,T;Y)} \leq s^{1/2} M.$$

So choosing  $s$  small enough gives us that  $\|b_n\|_Y \leq \epsilon/2$ . Next we claim that

$$a_n := \frac{1}{s} \int_t^{t+s} v_n(\beta) d\beta$$

converges weakly in  $X$ . To see this let  $\chi$  be the characteristic function on  $[t, t+s]$  and let  $\phi \in X^*$  be arbitrary, then we have by the weak convergence of  $v_n$  to 0 in  $L^2(0, T; X)$  that

$$0 \leftarrow \int_0^T \langle v_n, \chi\phi \rangle d\beta = \int_t^{t+s} \langle v_n, \phi \rangle d\beta = \left\langle \int_t^{t+s} v_n(\beta) d\beta, \phi \right\rangle = \langle a_n, \phi \rangle.$$

So  $a_n$  converges weakly to 0 in  $X$ , which means that it is bounded in  $X$ , which means that it (or a subsequence of it to be precise) converges strongly in  $H$ , and since  $H$  is continuously embedded in  $Y$ , it also converges strongly in  $Y$ . Thus we have  $\|a_n\|_Y \rightarrow 0$  as  $n \rightarrow \infty$ . Thus for large enough  $n$  this is smaller than  $\epsilon/2$  and therefore we have that

$$\|v_n(t)\|_Y \leq \|a_n\|_Y + \|b_n\|_Y \leq \epsilon.$$

Thus  $v_n(t)$  converges strongly in  $Y$  too. Now since  $\|v_n(t)\|_Y$  is uniformly bounded and converges to some limit for almost every  $t$  (if we choose not to redefine it on a set of measure zero), we have by the dominated convergence theorem that

$$\lim_{n \rightarrow \infty} \int_0^T \|v_n(t)\|_Y^2 dt = \int_0^T \lim_{n \rightarrow \infty} \|v_n(t)\|_Y^2 dt = 0.$$

Hence  $v_n$  converges to zero in  $L^2(0, T; Y)$  and so by the first part of the proof,  $v_n$  converges to zero in  $L^2(0, T; H)$ . Thus we have

$$\|u_n - u\|_{L^2(0,T;H)} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

□

The next theorem comes from linear algebra. We use it in Section 4.4, to derive an evolution equation for an  $n$ -dimensional volume in phase space.

**Theorem B.2.5** (Jacobi's Formula). *Let  $M(t)$  be an  $n \times n$ , non-singular matrix which is differentiable in  $t$  with  $t$  derivative  $\frac{d}{dt}M(t)$ . Then*

$$\frac{d}{dt} \det M(t) = \det M(t) \operatorname{Tr} \left( M(t)^{-1} \frac{d}{dt} M(t) \right),$$

where  $\det$  is the determinant and  $\operatorname{Tr}$  is the trace.

*Proof.* For a proof see for instance [Bel97]. □

### B.3 Young's-type Inequalities

We give a few inequalities that derive from Young's inequality as well as from the interpolation inequality between higher and lower derivatives. We will be needing many different variations of these inequalities, and with the exception of the basic ones, proofs are difficult to find in the literature. We therefore give the proofs of most of these inequalities. These inequalities are very important in our estimates, because they allow us to bound derivatives in terms of higher (and lower) derivatives without introducing additional parameter-dependent factors.

**Proposition B.3.1** (Young's Inequality [Eva15, p. 708]). *For any positive real numbers  $f, g > 0$ , and any constant  $\epsilon > 0$ , one has*

$$fg \leq \frac{\epsilon}{2} f^2 + \frac{1}{2\epsilon} g^2. \tag{B.4}$$

**Proposition B.3.2** (Cauchy-Schwarz Inequality). *Let  $(\cdot, \cdot)$  be the inner product of some Hilbert space  $H$ , and let  $\|\cdot\|$  be the norm induced by this inner product. Then for any  $x, y \in H$  one has*

$$(x, y) \leq \|x\| \|y\|. \tag{B.5}$$

We will mainly use the above inequality with the  $\dot{L}^2$  inner product.

**Proposition B.3.3** (Bounds on Derivatives Identity a [Rob01, p. 426]). *If  $u \in \dot{H}^2$  we have*

$$\|\partial_x u\|_{L^2}^2 \leq \|u\|_{L^2} \left\| \partial_x^2 u \right\|_{L^2}. \tag{B.6}$$



*Proof.* Assume first that  $u \in \dot{C}_p^2$ ; then by the density of  $\dot{C}_p^2$  in  $\dot{H}^2$ , we can conclude that the following also holds for functions in the larger space:

$$\begin{aligned} 0 &\leq \|\partial_x u\|_{\dot{L}^2}^2 = \int_0^\ell (\partial_x u)^2 dx = u \partial_x u \Big|_0^\ell - \int_0^\ell u \partial_x^2 u dx = - \int_0^\ell u \partial_x^2 u dx \\ &= \left| - \int_0^\ell u \partial_x^2 u dx \right| \leq \left( \int_0^\ell u^2 dx \right)^{1/2} \left( \int_0^\ell (\partial_x^2 u)^2 dx \right)^{1/2} = \|u\|_{\dot{L}^2} \|\partial_x^2 u\|_{\dot{L}^2}, \end{aligned}$$

where we have used the Cauchy-Schwarz inequality (B.5).  $\square$

In the results in the rest of this section,  $n \in \mathbb{N}$  is arbitrary.

**Corollary B.3.4.** *For any  $u \in \dot{H}^{n+1}$  one has*

$$\|\partial_x^n u\|_{\dot{L}^2}^2 \leq \|\partial_x^{n-1} u\|_{\dot{L}^2} \|\partial_x^{n+1} u\|_{\dot{L}^2}.$$

*Proof.* Apply (B.6) to  $v = \partial_x^{n-1} u$ .  $\square$

**Proposition B.3.5** (Bounds on Derivatives b). *For any  $u \in \dot{H}^{n+2}$ , and for any  $\epsilon > 0$  we have*

$$\|\partial_x^n u\|_{\dot{L}^2}^2 \leq \frac{1}{\epsilon} \|\partial_x^{n-1} u\|_{\dot{L}^2}^2 + \frac{\epsilon^2}{4} \|\partial_x^{n+2} u\|_{\dot{L}^2}^2. \quad (\text{B.7})$$

*Proof.* We apply Young's inequality repeatedly. Let  $\epsilon > 0$ ; then for any  $\delta > 0$  we have

$$\begin{aligned} \|\partial_x^n u\|_{\dot{L}^2}^2 &\leq \|\partial_x^{n-1} u\|_{\dot{L}^2} \|\partial_x^{n+1} u\|_{\dot{L}^2} \leq \frac{1}{2\epsilon} \|\partial_x^{n-1} u\|_{\dot{L}^2}^2 + \frac{\epsilon}{2} \|\partial_x^{n+1} u\|_{\dot{L}^2}^2 \\ &\leq \frac{1}{2\epsilon} \|\partial_x^{n-1} u\|_{\dot{L}^2}^2 + \frac{\epsilon}{2} \|\partial_x^n u\|_{\dot{L}^2} \|\partial_x^{n+2} u\|_{\dot{L}^2} \\ &\leq \frac{1}{2\epsilon} \|\partial_x^{n-1} u\|_{\dot{L}^2}^2 + \frac{\epsilon}{4\delta} \|\partial_x^n u\|_{\dot{L}^2}^2 + \frac{\epsilon\delta}{4} \|\partial_x^{n+2} u\|_{\dot{L}^2}^2. \end{aligned}$$

Letting  $\delta = \epsilon/2$ , we have

$$\|\partial_x^n u\|_{\dot{L}^2}^2 \leq \frac{1}{2\epsilon} \|\partial_x^{n-1} u\|_{\dot{L}^2}^2 + \frac{1}{2} \|\partial_x^n u\|_{\dot{L}^2}^2 + \frac{\epsilon^2}{8} \|\partial_x^{n+2} u\|_{\dot{L}^2}^2,$$

which implies

$$\|\partial_x^n u\|_{\dot{L}^2}^2 \leq \frac{1}{\epsilon} \|\partial_x^{n-1} u\|_{\dot{L}^2}^2 + \frac{\epsilon^2}{4} \|\partial_x^{n+2} u\|_{\dot{L}^2}^2. \quad \square$$

**Corollary B.3.6.** *For  $u \in \dot{H}^{n+3}$  and any  $\epsilon > 0$  we have*

$$\|\partial_x^{n+2} u\|_{\dot{L}^2}^2 \leq \frac{\epsilon^2}{4} \|\partial_x^n u\|_{\dot{L}^2}^2 + \frac{1}{\epsilon} \|\partial_x^{n+3} u\|_{\dot{L}^2}^2. \quad (\text{B.8})$$

Furthermore (B.7) and (B.8) also hold for the absolute value,  $|\partial_x|$  (3.7), of the derivatives.

*Proof.* The proof parallels that of Proposition B.3.5.  $\square$

**Proposition B.3.7** (More Bounds on Derivatives c). *For  $u \in \dot{H}^{n+4}$ , and for any  $\epsilon > 0$  one has*

$$\left\| \partial_x^{n+2} u \right\|_{L^2}^2 \leq \frac{1}{\epsilon} \left\| \partial_x^{n+4} u \right\|_{L^2}^2 + \frac{\epsilon}{2} \left\| \partial_x^n u \right\|_{L^2}^2. \quad (\text{B.9})$$

*Proof.* By (B.8) we have

$$\left\| \partial_x^{n+2} u \right\|_{L^2}^2 \leq \frac{\epsilon^2}{4} \left\| \partial_x^n u \right\|_{L^2}^2 + \frac{1}{\epsilon} \left\| \partial_x^{n+3} u \right\|_{L^2}^2;$$

applying (B.4) and (B.6) to this, we get

$$\left\| \partial_x^{n+2} u \right\|_{L^2}^2 \leq \frac{\epsilon^2}{4} \left\| \partial_x^n u \right\|_{L^2}^2 + \frac{1}{2\sigma\epsilon} \left\| \partial_x^{n+4} u \right\|_{L^2}^2 + \frac{\sigma}{2\epsilon} \left\| \partial_x^{n+2} u \right\|_{L^2}^2,$$

for an arbitrary  $\sigma > 0$ . Letting  $\sigma = \epsilon$ , one gets

$$\frac{1}{2} \left\| \partial_x^{n+2} u \right\|_{L^2}^2 \leq \frac{\epsilon^2}{4} \left\| \partial_x^n u \right\|_{L^2}^2 + \frac{1}{2\epsilon^2} \left\| \partial_x^{n+4} u \right\|_{L^2}^2.$$

Relabelling  $\epsilon^2$  as  $\epsilon$  gives the result.  $\square$

**Corollary B.3.8** (More Bounds on Derivatives d). *For  $u \in \dot{H}^{n+5}$ , and for any  $\sigma > 0$  one has*

$$\left\| \partial_x^{n+3} u \right\|_{L^2}^2 \leq 2\sigma^2 \left\| \partial_x^{n+5} u \right\|_{L^2}^2 + \frac{1}{\sigma^3} \left\| \partial_x^n u \right\|_{L^2}^2. \quad (\text{B.10})$$

*Proof.* We start with inequality (B.9), and then apply (B.8) to the  $\left\| \partial_x^{n+1} u \right\|_{L^2}^2$  term to get

$$\begin{aligned} \left\| \partial_x^{n+3} u \right\|_{L^2}^2 &\leq \frac{1}{\epsilon} \left\| \partial_x^{n+5} u \right\|_{L^2}^2 + \frac{\epsilon}{2} \left\| \partial_x^{n+1} u \right\|_{L^2}^2 \\ &\leq \frac{1}{\epsilon} \left\| \partial_x^{n+5} u \right\|_{L^2}^2 + \frac{\epsilon\eta^2}{8} \left\| \partial_x^{n+3} u \right\|_{L^2}^2 + \frac{\epsilon}{\eta} \left\| \partial_x^n u \right\|_{L^2}^2. \end{aligned}$$

Then, choosing  $\epsilon = 4/\eta^2$  and subsequently letting  $\sigma = \eta/2$ , we get the result.  $\square$

**Proposition B.3.9** (More Bounds on Derivatives e). *For  $u \in \dot{H}^{n+4}$ , one has*

$$\left\| \partial_x^{n+3} u \right\|_{L^2}^2 \leq \frac{1}{2\mu} \left\| \partial_x^{n+4} u \right\|_{L^2}^2 + \left( \frac{\alpha}{2\eta} + \frac{\mu}{2} \right) \left\| \partial_x^{n+2} u \right\|_{L^2}^2 - \alpha \left\| \partial_x^{n+1} u \right\|_{L^2}^2 + \frac{\alpha\eta}{2} \left\| \partial_x^n u \right\|_{L^2}^2, \quad (\text{B.11})$$

for any  $\alpha, \mu, \eta > 0$ .

*Proof.* Applying (B.4) with  $\epsilon = \eta$  to (B.6) and then multiplying the expression by  $\alpha$  and moving the middle term to the right we have

$$0 \leq \frac{\alpha}{2\eta} \left\| \partial_x^{n+2} u \right\|_{\dot{L}^2}^2 - \alpha \left\| \partial_x^{n+1} u \right\|_{\dot{L}^2}^2 + \frac{\alpha\eta}{2} \left\| \partial_x^n u \right\|_{\dot{L}^2}^2.$$

Then if we add this to

$$\left\| \partial_x^{n+3} u \right\|_{\dot{L}^2}^2 \leq \frac{1}{2\mu} \left\| \partial_x^{n+4} u \right\|_{\dot{L}^2}^2 + \frac{\mu}{2} \left\| \partial_x^{n+2} u \right\|_{\dot{L}^2}^2,$$

we have the result. □

## B.4 Other Inequalities

The first inequality that we shall prove in this section is Poincaré's inequality, which allows one to bound the  $\dot{L}^2$  norm of a function directly by that of its derivative, but which has the disadvantage of introducing a factor of the domain size  $\ell$ .

There are many different forms of Poincaré's inequality. Here we will give the version for one-dimensional, periodic functions, which can be represented by a Fourier series of the form (2.2). The constant  $\ell/2\pi$  that is introduced in (B.12) is optimal, which can be seen by substituting the lowest Fourier mode into the inequality.

**Lemma B.4.1** (Poincaré's Inequality). *For any  $u \in \dot{H}^1$ , one has*

$$\|u\|_{\dot{L}^2} \leq \frac{\ell}{2\pi} \|\partial_x u\|_{\dot{L}^2}. \tag{B.12}$$

*Proof.* We use Parseval's identity, and the fact that the wavenumber  $|k| \geq \frac{2\pi}{\ell}$  (ignoring the zero mode). Then  $\|\partial_x u\|_{\dot{L}^2}^2$  can be bounded below as follows:

$$\|\partial_x u\|_{\dot{L}^2}^2 = \ell \sum_k |k|^2 |\hat{u}_k|^2 \geq \ell \sum_k \left(\frac{2\pi}{\ell}\right)^2 |\hat{u}_k|^2 = \left(\frac{2\pi}{\ell}\right)^2 \|u\|_{\dot{L}^2}^2.$$

Moving the  $2\pi/\ell$  term to the other side and taking the square root gives the result. □

We will quickly state the most common version of Gronwall's inequality. For a proof see for instance [Rob01, p. 54].

**Lemma B.4.2.** *Let  $x(t)$  satisfy the differential inequality*

$$\frac{dx}{dt} \leq a(t)x$$

for some function  $a(t)$ . Then

$$x(t) \leq x(t_0)e^{\int_{t_0}^t a(s) ds}. \quad (\text{B.13})$$

The following inequalities are generalizations of the well-known Gronwall inequality. We will use them to bound quantities in time; in particular, we will use them to get uniform bounds on solutions and absorbing balls in Sobolev spaces.

**Lemma B.4.3** (Gronwall Lemma). *Suppose that  $a, b \in \mathbb{R}$ , with  $a > 0$ . Then if*

$$\frac{dx}{dt} \leq -ax + b, \quad x(t_0) = x_0,$$

then

$$\limsup_{t \rightarrow \infty} x(t) \leq \frac{b}{a}. \quad (\text{B.14})$$

*Proof.* We use the integrating factor  $e^{at}$  to solve the inequality.

$$\begin{aligned} \frac{d}{dt} (xe^{at}) &\leq be^{at} \\ \implies x(t) &\leq x_0 e^{-a(t-t_0)} + e^{-at} \int_{t_0}^t be^{as} ds \\ &= \frac{b}{a} - \left( \frac{b}{a} - x_0 \right) e^{-a(t-t_0)}. \end{aligned}$$

Taking the lim sup as  $t \rightarrow \infty$  gives the result. □

We will now prove a short result analogous to Poincaré's inequality, but involving two distinct norms. We use this result on many occasions in the text.

**Proposition B.4.4.** *Let  $u \in \dot{H}^1$ ; then*

$$\|u\|_{\dot{L}^\infty} \leq \ell^{1/2} \|\partial_x u\|_{\dot{L}^2}. \quad (\text{B.15})$$

*Proof.* By the fundamental theorem of calculus we have

$$|u(x)| \leq \int_{x_0}^x |\partial_s u(s)| ds \leq \int_0^\ell |\partial_s u(s)| ds,$$

where  $x_0 \in [0, \ell]$  is a point such that  $u(x_0) = 0$ . Then by the Cauchy-Schwarz inequality (B.5) we have

$$|u(x)| \leq \|1\|_{\dot{L}^2} \|\partial_x u\|_{\dot{L}^2} = \ell^{1/2} \|\partial_x u\|_{\dot{L}^2}.$$

Taking the supremum over  $x$  of the have inequality gives the result. □

## B.5 Regularity

The following results relate the decay rate of the Fourier coefficients of a function to its degree of differentiability. In this thesis, they are used to show analyticity of solutions to equation (1.4).

**Lemma B.5.1** (Analyticity and Decay of Derivatives [Joh75, p. 65]). *Suppose  $u \in \dot{L}^2$  satisfies*

$$|\partial_x^n u(x)| \leq M \frac{(n+1)!}{\alpha^{n+1}}, \quad x \in [0, \ell], \quad (\text{B.16})$$

for some constant  $M$  and for all  $n \in \mathbb{N}$ . Here  $\alpha$  is a positive constant which, as we shall later see, can be thought of as the radius of analyticity of  $u$ . Then  $u$  is analytic on  $[0, \ell]$ .

*Proof.* We will show that  $u$  is analytic in a radius  $d < \alpha$  about  $y$  for any  $y \in [0, \ell]$ . First define, for any  $x$  with  $|x - y| \leq d$ ,

$$\phi(s) := u(y + s(x - y)),$$

then we have

$$\left| \frac{\phi^{(n)}(s)}{n!} \right| = \left| \frac{1}{n!} \partial_x^n u(y + s(x - y))(x - y)^n \right| \leq M \frac{1}{n!} \frac{(n+1)!}{\alpha^{n+1}} d^n = \frac{M(n+1)}{\alpha} \left( \frac{d}{\alpha} \right)^n.$$

Then  $u(x) = \phi(1)$ , while by Taylor's theorem, we have

$$\phi(1) = \sum_{n=1}^{N-1} \frac{1}{n!} \phi^{(n)}(0) + \underbrace{\frac{1}{(N-1)!} \int_0^1 (1-s)^{N-1} \phi^{(N)}(s) ds}_{r_N}.$$

But the remainder  $r_N$  satisfies

$$|r_N| \leq \frac{1}{(N-1)!} \int_0^1 (1-s)^{N-1} |\phi^{(N)}(s)| ds \leq \frac{1}{N!} \frac{M(N+1)!}{\alpha} \left( \frac{d}{\alpha} \right)^N,$$

so since  $d < \alpha$ , we have

$$\lim_{N \rightarrow \infty} r_N = 0.$$

So  $u$  has a Taylor series that converges at all points in  $[0, \ell]$ , which means that  $u$  is (real) analytic. □

**Corollary B.5.2.** *For any infinitely differentiable function  $u$  defined on  $[0, \ell]$ , if we have*

$$|\partial_x^n u| \leq C_u \frac{n!}{\alpha^{n+1}}, \quad (\text{B.17})$$

for all  $n \geq n_0$ , for some  $n_0 > 0$  and a constant  $C_u$  depending on  $u$ , then the above argument shows that  $u$  must be analytic on  $[0, \ell]$ .

**Theorem B.5.3** (Paley-Wiener [Joh75]). *Suppose that we are given a function  $u \in \dot{L}^2(0, \ell)$  satisfying*

$$\|e^{\alpha A} u\|_{L^2} \leq M, \quad (\text{B.18})$$

for some constants  $M$  and  $\alpha > 0$  where  $A = |\partial_x|$  is defined in (3.7). Then  $u$  is analytic in a strip of radius  $\alpha$  about  $[0, \ell]$ .

*Proof.* We begin by noting that  $e^{\alpha A}$  acts on  $u$  in the following way:

$$\begin{aligned} e^{\alpha A} u &= \sum_k (1 + \alpha |\partial_x| + \alpha^2 |\partial_x|^2 / 2 + \alpha^3 |\partial_x|^3 / 6 \dots) \hat{u}_k e^{ikx} \\ &= \sum_k (1 + \alpha |k| + \alpha^2 |k|^2 / 2 + \alpha^3 |k|^3 / 6 \dots) \hat{u}_k e^{ikx} \\ &= \sum_k e^{\alpha |k|} \hat{u}_k e^{ikx}. \end{aligned}$$

But then by Parseval's identity we have

$$\sum_k |e^{\alpha |k|} \hat{u}_k|^2 = \frac{1}{\ell} \|e^{\alpha A} u\|_{L^2}^2 \leq \frac{M^2}{\ell}.$$

Thus in order for the above sum to converge we must have  $|\hat{u}_k| \leq M e^{-\alpha |k|} / \ell^{1/2}$  for each  $k$ . Thus the Fourier coefficients decay exponentially fast in  $k$ , and so the function must be infinitely differentiable. To see why it is analytic, note that if  $n \geq 1$ , then we have

$$|\partial_x^n u| \leq \sum_{k \neq 0} |k|^n |\hat{u}_k| \leq 2 \frac{M}{\ell^{1/2}} \sum_{k > 0} |k|^n e^{-\alpha |k|}.$$

But this series is dominated by the integral  $\int_0^\infty y^n e^{-\alpha y} dy$ . Thus

$$|\partial_x^n u| \leq \frac{2M}{\ell^{1/2}} \int_0^\infty y^n e^{-\alpha y} dy,$$

so with  $s = \alpha y$ , we have

$$|\partial_x^n u| \leq \frac{2M}{\ell^{1/2}} \frac{1}{\alpha^{n+1}} \int_0^\infty s^n e^{-s} ds = \frac{2M n!}{\ell^{1/2}} \frac{1}{\alpha^{n+1}}, \quad (\text{B.19})$$

since  $\int_0^\infty s^n e^{-s} ds$  is the Gamma function.

Thus by Corollary B.5.2, and the results about analytic continuation outlined below,  $u$  is analytic in a strip of radius  $\alpha$  about  $[0, \ell]$ .

□

**Analytic Continuation:**

The above argument shows that for any point  $y \in [0, \ell]$ , the radius of analyticity of  $u$  about  $y$  on the real line is  $\alpha$ . So we have a series

$$u(x) = \sum_{n=0}^{\infty} a_n(x-y)^n,$$

that converges whenever  $|x-y| < \alpha$ . But of course, nothing stops us from evaluating the above series at some complex number  $z \in \mathbb{C}$  instead of at  $x$ . In fact, as long as  $|z-y| < \alpha$ , then the series with the same coefficients  $a_n$  also converges, and must thus represent some complex-analytic function, which coincides with  $u$  on  $[0, \ell]$ . Thus we can take the function that is represented by  $\sum_n a_n(z-y)^n$  as the analytic continuation of  $u$  to the complex plane. We will call this function  $u$ , and so we have that  $u$  is analytic in a neighbourhood of radius  $\alpha$  about the set  $[0, \ell] \subset \mathbb{C}$ .

### Regularity in Time:

The following theorem is also very important in that it gives us some regularity in time, and allows us to conclude that the limit as  $t \rightarrow 0$  of a solution gives the initial condition. The result proved here is a minor generalization of the temporal regularity result of [Eva15, p. 305].

**Theorem B.5.4** (Temporal Regularity). *Suppose  $u \in L^2(0, T; \dot{H}^n)$  and additionally  $u_t \in L^2(0, T; \dot{H}^{-n})$ .*

*Then*

$$u \in C([0, T]; \dot{L}^2),$$

*after possibly being redefined on a set of measure zero.*

*Proof.* The norm on  $C([0, T]; \dot{L}^2)$  is given by

$$\|u\|_{C([0, T]; \dot{L}^2)} := \max_{0 \leq t \leq T} \|u\|_{\dot{L}^2}.$$

We begin by extending  $u$  in time outside of the interval  $[0, T]$  by zero, and then defining  $u^\epsilon$  to be the mollification of  $u$  in the sense of [Eva15, p. 716]. Then  $u^\epsilon \rightarrow u$  in  $L^2(0, T; \dot{H}^n)$  and  $u_t^\epsilon \rightarrow u_t$  in  $L^2(0, T; \dot{H}^{-n})$  as  $\epsilon \rightarrow 0$ . Thus for each  $\epsilon, \delta > 0$ , we have that

$$\frac{d}{dt} \left\| u^\epsilon(t) - u^\delta(t) \right\|_{\dot{L}^2}^2 = 2 \langle u_t^\epsilon(t) - u_t^\delta(t), u^\epsilon(t) - u^\delta(t) \rangle,$$

where  $\langle \cdot, \cdot \rangle$  denotes the action of an element in  $\dot{H}^{-n}$  on an element in  $\dot{H}^n$ , since the mollifications are smooth in time. Now we integrate the above expression over time from  $s$  to  $t$  to get

$$\left\| u^\epsilon(t) - u^\delta(t) \right\|_{\dot{L}^2}^2 = \left\| u^\epsilon(s) - u^\delta(s) \right\|_{\dot{L}^2}^2 + 2 \int_s^t \langle u_t^\epsilon(\tau) - u_t^\delta(\tau), u^\epsilon(\tau) - u^\delta(\tau) \rangle d\tau.$$

Now using the fact that  $\langle a, b \rangle \leq \|a\|_{\dot{H}^{-n}} \|b\|_{\dot{H}^n}$  for any  $a \in \dot{H}^{-n}$  and  $b \in \dot{H}^n$ , we get

$$\left\| u^\epsilon(t) - u^\delta(t) \right\|_{\dot{L}^2}^2 \leq \left\| u^\epsilon(s) - u^\delta(s) \right\|_{\dot{L}^2}^2 + \int_0^T \|u_t^\epsilon(\tau) - u_t^\delta(\tau)\|_{\dot{H}^{-n}}^2 + \|u^\epsilon(\tau) - u^\delta(\tau)\|_{\dot{H}^n}^2 d\tau,$$

where we have used Young's inequality on the integrand. Now since  $u^\epsilon \rightarrow u$  in  $\dot{L}^2$  for a.e.  $t$ , we can choose  $s$  so that

$$\lim_{\epsilon, \delta \rightarrow 0} \left\| u^\epsilon(s) - u^\delta(s) \right\|_{\dot{L}^2}^2 = 0.$$

But then we have that

$$\begin{aligned} & \lim_{\epsilon, \delta \rightarrow 0} \sup_{0 \leq t \leq T} \left\| u^\epsilon(t) - u^\delta(t) \right\|_{\dot{L}^2}^2 \\ & \leq \lim_{\epsilon, \delta \rightarrow 0} \left( \left\| u^\epsilon(s) - u^\delta(s) \right\|_{\dot{L}^2}^2 + \int_0^T \|u_t^\epsilon(\tau) - u_t^\delta(\tau)\|_{\dot{H}^{-n}}^2 + \|u^\epsilon(\tau) - u^\delta(\tau)\|_{\dot{H}^n}^2 d\tau \right) \\ & = 0. \end{aligned}$$

Thus the  $u^\epsilon$  form a Cauchy sequence in  $C([0, T]; \dot{L}^2)$  and hence converge to a limit in  $C([0, T]; \dot{L}^2)$ ; so we define this limit to be  $u$  (after possibly redefining  $u$  on a set of measure zero). Thus we have the result.  $\square$

## B.6 The Nonlinear Term

We will often need to bound the nonlinear term. The trilinear form (2.14) associated with the nonlinear term can be bounded in several different ways, and satisfies some commutative properties, which we will state here.

Recall that the trilinear form associated with the nonlinear term  $\mathcal{N}$  of equation (1.4) is given by

$$b(u, v, w) := \int_0^\ell u \partial_x v w \, dx.$$

We will assume that  $u, v$  and  $w$  satisfy periodic boundary conditions and that they are regular enough such that the above expression is well defined.

**Proposition B.6.1.** *The trilinear form satisfies the commutative property*

$$b(u, v, w) + b(w, u, v) + b(v, w, u) = 0. \tag{B.20}$$



*Proof.* We integrate by parts and use the periodic boundary conditions on  $u, v$  and  $w$  and the product rule to get

$$\begin{aligned} b(u, v, w) &= - \int_0^\ell \partial_x(uw) v \, dx \\ &= - \int_0^\ell w \partial_x u v \, dx - \int_0^\ell v \partial_x w u \, dx = -b(w, u, v) - b(v, w, u). \end{aligned}$$

Moving all terms to the left gives the result.  $\square$

**Corollary B.6.2.** *Evaluating (B.20) with  $w = v$  and with  $v = w = u$ , gives*

$$b(u, v, v) = -\frac{1}{2}b(v, u, v),$$

and

$$b(u, u, u) = 0,$$

respectively.

**Proposition B.6.3.** *The trilinear form (2.14) satisfies:*

$$b(u, v, w) \leq \ell^{1/2} \|u\|_{\dot{L}^2} \left\| \partial_x^2 v \right\|_{\dot{L}^2} \|w\|_{\dot{L}^2} \quad \text{for all } u, w \in \dot{L}^2, v \in \dot{H}^2 \quad (\text{B.21})$$

and

$$b(u, v, w) \leq \ell^{1/2} \|u\|_{\dot{L}^2} \|\partial_x v\|_{\dot{L}^2} \|\partial_x w\|_{\dot{L}^2} \quad \text{for all } u \in \dot{L}^2, v, w \in \dot{H}^1. \quad (\text{B.22})$$

*Proof.* By (B.15) we have that for a function  $u \in \dot{H}^1$

$$\|u\|_{\dot{L}^\infty} \leq \ell^{1/2} \|\partial_x u\|_{\dot{L}^2}.$$

Thus we get that

$$b(u, v, w) \leq \ell^{1/2} \left\| \partial_x^2 v \right\|_{\dot{L}^2} \int_0^\ell uw \, dx,$$

and

$$b(u, v, w) \leq \ell^{1/2} \|\partial_x w\|_{\dot{L}^2} \int_0^\ell u \partial_x v \, dx.$$

Applying the Cauchy-Schwarz inequality to the integrals gives the results.  $\square$