UNIVERSIDAD DE SONORA

DIVISIÓN DE CIENCIAS EXACTAS Y NATURALES

Programa de Posgrado en Matemáticas

Numerical Solutions for Burgers' Equation
in the Stochastic and Deterministic Version
Using Spectral Methods

T E S I S

Que para obtener el grado académico de:

Maestro en Ciencias
(Matemáticas)

Presenta:

Alan Daniel Matzumiya Zazueta

Directores de Tesis:
Dr. Daniel Olmos Liceaga
Dr. Saúl Díaz Infante Velasco

Hermosillo, Sonora, México,   June 16, 2020

"El saber de mis hijos
hará mi grandeza"

SINODALES

Dr. Daniel Olmos Liceaga
Universidad de Sonora

Dr. Saúl Díaz Infante Velasco
CONACYT-Universidad de Sonora

Dr. Francisco Javier Delgado Vences
Universidad Nacional Autónoma de México

Dr. Martín Gildardo García Alvarado
Universidad de Sonora

Contents

Chapter 1

Introduction

Differential equations, whether ordinary or partial, allow modeling phenomena that evolve with respect to space and time. Typical problems are the propagation of sound or heat, electrostatics, electrodynamics, fluid dynamics, elasticity, quantum mechanics and many others. Finding analytical solutions and the mathematical analysis of these problems, mainly for those problems that are not linear, has been a challenge of interest due to a large amount of information they can provide.

There are alternatives to find solutions to differential equations, for example, computational fluid dynamics is one of the branches of fluid mechanics that uses numerical methods and algorithms to solve and analyze fluid flow problems that perform millions of calculations to simulate the interaction of liquids and gases through complex surfaces. However, even with simplified equations and high-performance supercomputers, in many cases, only approximate results can be achieved.

The major challenge in the field of complex systems is a thorough understanding of the phenomenon of turbulence. Direct numerical simulations have contributed substantially to our understanding of the phenomena of disorderly flow that inevitably arise in the high Reynolds numbers (Re), which is a dimensionless number used in fluid mechanics to characterize the movement of the fluid indicating whether follows a laminar or turbulent flow. However, a successful theory of turbulence is still lacking which should allow predicting features of important phenomena like turbulent mixing, turbulent convection, and turbulent combustion on the basis of the fundamental dynamical equations.

Spectral methods have recently emerged as a viable alternative for the numerical solution of partial differential equations. They have proved particularly useful in fluid dynamics simulation where are now regularly used large spectral hydrodynamics codes to study turbulence, numerical weather prediction, ocean dynamics and any other problems where high accuracy is desired.

Due to the above, he has motivated the development of this thesis by studying spectral methods extensively in order to acquire the ability to use this tool and understand them from the point of view of mathematical analysis. To develop this study we will focus first on the elementary theory of these methods, encompassing enough knowledge to allow us to develop, implement and analyze under this ap-

proach a wide variety of problems that arise in the partial differential equations that evolve over time.

To understand the application of these methods, the well-known Burgers' equation has been considered, since it is an ideal problem for understanding these methods because, in addition to being a non-linear problem that presents interesting characteristics, it can be useful to develop the ability to attack more complex problems. Furthermore, in order to extend the study of the implementation of spectral methods, we are going to work with the stochastic version of this equation that will be very useful for us to know in general terms trying to solve problems of this type, which are considered of great importance for its wide field of applications and that it is still an area that is in full development due to the great difficulty in obtaining solutions.

To carry out this study, we will divide the work into six parts organized as follows

1. In chapter 1, a brief history of Burgers' equation will be presented in its deterministic version, and we will also present how to obtain the analytical solution for an initial value problem of this equation, transforming it into another linear one that can be solved using the Fourier transform. Later, the origin of the stochastic version will be discussed, in addition to its importance within mathematics and physics.

2. In chapter 2, we will study the theoretical bases of spectral methods using the well-known Fourier series as the main tool, which will allow us to study a theory of approximation of functions under two approaches, using orthogonal projections and another using interpolation techniques. These two approaches will be examined independently, studying their implementation and some theoretical results that will be useful in chapter 3.

3. In chapter 3, the spectral methods known as Fourier-Galerkin and Fourier-Collocation will be developed using the tools examined in chapter 2, verifying their convergence theory. For this, the deterministic Burgers' equation will be used, taking advantage of its linearized form to describe the methods, which will then be applied to the original nonlinear equation to describe the algorithms of its computational implementation that will allow us to perform numerical experiments and to be able to observe some characteristics interesting for your discussion.

4. In chapter 4, a spectral method used to solve stochastic partial differential equations that was studied in [1] will be disclosed in as much detail as possible. We will see that this method, which is built based on the well-known Hermite polynomials, will allow us to obtain solutions of stochastic problems by solving a deterministic type problem, and will be illustrated using the stochastic Burgers' equation developing its implementation and also numerical simulations.

5. In this last chapter, we will discuss the most relevant of each chapter, and we

will give some observations of the obtained numerical results to conclude with some ideas that can be considered to extend this work.

6. At the end of this work, an appendix A was added, which will be useful to understand in more detail the spectral method developed in chapter 4.

1.1   Brief History of Burgers' Equation

The simplest fluids (called Newtonian fluids), are described by the well-known Navier-Stokes equations, named after Claude-Louis Navier and George Gabriel Stokes. These are a set of non-linear (PDEs), which are obtained by applying the principles of conservation of mechanics and thermodynamics on a volume of fluid to obtain the so-called integral formulation of the equations. Applying certain considerations, especially those in which the tangential forces have a linear relationship with the velocity gradient (Newton's viscosity law), the differential formulation is obtained which is generally more useful for solving the problems that arise in the mechanics of fluids. For further details about the Navier-Stokes equations see [2, 3, 4, 5, 6].

Let v be a vector field, Navier-Stokes equations are given as follows

$$\begin{cases} \nabla \cdot v = 0, \\ (\rho v)_t + (\nabla \cdot \rho v)v + \nabla p - \mu \nabla^2 v - \rho G = 0. \end{cases} \tag{1.1}$$

It is well known that when ρ is considered the density, p the pressure, v the velocity and μ the viscosity of a fluid, these equations describe the dynamics of an incompressible fluid (free divergence, and $\rho_t = 0$), where G represents the gravitational effects.

In contrast to equation (1.1), this can be investigated in one spatial dimension. Simplification in (1.1) of the x component of the velocity vector, which we will call v^x, gives

$$\rho \frac{\partial v^x}{\partial t} + \rho v^x \frac{\partial v^x}{\partial x} + \rho v^y \frac{\partial v^x}{\partial y} + \rho v^z \frac{\partial v^x}{\partial z} + \frac{\partial p}{\partial x} - \mu \left(\frac{\partial^2 v^x}{\partial x^2} + \frac{\partial^2 v^x}{\partial y^2} + \frac{\partial^2 v^x}{\partial z^2} \right) - \rho G^x = 0.$$

Considering a $1D$ problem with no pressure gradient, the above equation reduces to

$$\rho \frac{\partial v^x}{\partial t} + \rho v^x \frac{\partial v^x}{\partial x} - \mu \frac{\partial^2 v^x}{\partial x^2} - \rho G^x = 0. \tag{1.2}$$

If we use now the traditional variable v rather than v^x, take α to be the kinematic viscosity, i.e, $\alpha = \frac{\mu}{\rho}$ and $G \equiv 0$, then the equation (1.2) becomes just the viscid Burgers' equation

$$\frac{\partial v(x,t)}{\partial t} + \underbrace{v(x,t)\frac{\partial v(x,t)}{\partial x}}_{\textbf{Convection}} - \underbrace{\alpha \frac{\partial^2 v(x,t)}{\partial x^2}}_{\textbf{Diffusion}} = 0. \tag{1.3}$$

Some assumptions are made, namely: $\rho = \text{constant}$ (density), $\mu = \text{constant}$ (viscosity), $p = \text{constant}$ (pressure).

Burgers' equation was introduced in 1915 by Harry Bateman [7], an English mathematician, in his paper along with its corresponding initial condition and boundary values. Later in 1939, Johannes Martinus Burgers [8, 9], a Dutch physicist, simplified the Navier-Stokes equation (1.1) by just dropping the pressure term, and in 1948 explained the mathematical modeling of turbulence with the help of the equation (1.3). The name of this equation is because Burgers became one of the leading figures in the field of fluid mechanics and, therefore, honors his contributions.

The equation (1.3) is a partial differential equation nonlinear, where the second term is known as the convective part of the equation and the third as the diffusive part. This equation appears in several areas of applied mathematics, such as fluid mechanics, nonlinear acoustics, gas dynamics, traffic flow, and many others. It is generally considered a toy model, i.e., a tool that is used to understand part of the internal behavior of the general problem.

The formulation given by the equation (1.3) is called the strong form, i.e., the partial differential equation requires that it be satisfied for each point x in its domain and for each t. This formulation can be written as follows

$$\frac{\partial v}{\partial t} + A(v) + F(v) = 0, \quad t > 0, \tag{1.4}$$

where F and A are given by

$$A(v) = -\alpha v_{xx}, \quad F(v) = \frac{1}{2}(v^2)_x, \quad x \in I.$$

Multiplying both sides of (1.4) by $\phi \in X$, for some arbitrary smooth function ϕ of compact support, such that the integral of the PDE over the space I is satisfied, we get

$$\int_I \frac{\partial v}{\partial t} \phi dx + \int_I A(v)\phi dx + \int_I F(v)\phi dx = 0, \quad \forall \phi \in X, \ \forall t > 0. \tag{1.5}$$

The formulation (1.5) is called the weak form of (1.3) and ϕ are known as the test functions. If we denote $\langle \cdot, \cdot \rangle$ as the inner product in X, then (1.5) can be written in compact form as follows

$$\left\langle \frac{\partial v}{\partial t} + A(v) + F(v), \phi \right\rangle = 0, \quad \forall \phi \in X, \ \forall t > 0. \tag{1.6}$$

Note that the two formulations, (1.5) and (1.4), are equivalent if the solution is smooth enough, however the weak formulation can adjust less regular solutions than in the strong form. In fact, the solution to (1.5) is known as the distribution solution of the original equation (1.3), since it can be shown to satisfy (1.3) in the sense of distributions. For more details of the above it is recommended to see Schwartz [10],

Lions and Magenes [11], Renardy and Rogers [12].

Proper use of these formulations allows one to recover the strong form from the weak form, therefore, an appropriate way to design a numerical method is to first choose one of the formulations satisfied by the exact solution, then restrict the choice of test functions to a space of finite dimension, to replace u with the discrete solution u_N, and possibly to replace the exact integration with quadrature rules.

1.1.1   Analytical solutions for Burgers' equation.

When we want to know the precision of an approximation, an alternative is through numerical experiments that can be compared with some other solution considered as a good reference. For this type of tests, the better is to have exact solutions to compare with an approximate one, and fortunately the equation (1.3) can be solved exactly by Hopf-Cole transformation introduced by Eberhard Hopf [13] and Julian David Cole [14] independently to convert the Burgers' equation into a linear parabolic equation and solve it exactly for any initial condition.

Now consider the problem of initial value for the equation (1.3) as follows

$$\begin{cases} u_t + u u_x = \alpha u_{xx} & x \in \mathbb{R}, \ t > 0, \ \alpha > 0 \\ u(x,0) = u_0(x) & x \in \mathbb{R}, \end{cases} \tag{1.7}$$

Hence, the transformation known as the Cole-Hopf transformation is given by

$$u = -2\alpha \frac{\varphi_x}{\varphi} \tag{1.8}$$

Operating (1.8) into each term of (1.7) we find that

$$u_t = \frac{2\alpha(\varphi_t \varphi_x - \varphi \varphi_{xt})}{\varphi^2} \ , \ u u_x = \frac{4\alpha^2 \varphi_x(\varphi \varphi_{xx} - \varphi_x^2)}{\varphi^3},$$

and

$$\alpha u_{xx} = -\frac{2\alpha^2(2\varphi_x^3 - 3\varphi \varphi_{xx}\varphi_x + \varphi^2 \varphi_{xxx})}{\varphi^3}.$$

Substituting these expressions into (1.7),

$$\frac{2\alpha(-\varphi \varphi_{xt} + \varphi_x(\varphi_t - \alpha \varphi_{xx}) + \alpha \varphi \varphi_{xxx})}{\varphi^2} = 0,$$

so we have the following,

$$-\varphi \varphi_{xt} + \varphi_x(\varphi_t - \alpha \varphi_{xx}) + \alpha \varphi \varphi_{xxx} = 0 \iff \varphi_x(\varphi_t - \alpha \varphi_{xx}) = \varphi(\varphi_{xt} - \alpha \varphi_{xxx})$$
$$\iff \varphi_x(\varphi_t - \alpha \varphi_{xx}) = \varphi(\varphi_t - \alpha \varphi_{xx})_x.$$

Therefore, if φ solves the equation $\varphi_t - \alpha \varphi_{xx} = 0$, $x \in \mathbb{R}$, then $u(x,t)$ given by the transformation (1.8) solves the Burgers equation.

To completely transform the problem (1.7) we still have to work with the initial condition function. To do this, note that (1.8) can be written as

$$u = -2\alpha(\log\varphi)_x, \tag{1.9}$$

hence, we get

$$\varphi(x,t) = e^{-\int \frac{u(x,t)}{2\alpha} dx}.$$

It is clear from (1.9) that multiplying φ by a constant does not affect u, so we can write the last equation as

$$\varphi(x,t) = e^{-\int_0^x \frac{u(y,t)}{2\alpha} dy}. \tag{1.10}$$

The initial condition on (1.7) must be transformed by using (1.9) to get

$$\varphi(x,0) = \varphi_0(x) = e^{-\int_0^x \frac{u_0(y)}{2\alpha} dy}.$$

In summary, we have reduced the problem (1.7) to this one

$$\begin{cases} \varphi_t - \alpha\varphi_{xx} = 0, & x \in \mathbb{R}, \ t > 0, \ \alpha > 0, \\ \varphi(x,0) = \varphi_0(x) = e^{-\int_0^x \frac{u_0(y)}{2\alpha} dy}, & x \in \mathbb{R}. \end{cases} \tag{1.11}$$

Parabolic Equation. The general solution of the initial value problem for the equation (1.11) is well known and can be handled by a variety of methods. An interesting method related to spectral methods is the following: one can take the Fourier transform with respect to x for both the equation and the initial condition $\varphi_0(x)$ to obtain a first-order (ODE) as follows

$$\begin{cases} \hat{\varphi}_t = \xi^2\alpha\hat{\varphi}, & \xi \in \mathbb{R}, \ t > 0, \ \alpha > 0, \\ \hat{\varphi}(\xi,0) = \hat{\varphi}_0(\xi), & \xi \in \mathbb{R}, \end{cases}$$

where $\hat{\varphi}(\xi,t) = \int_{-\infty}^{\infty} \varphi(x,t)e^{i\xi x} dx.$

Then the solution for this problem is

$$\hat{\varphi}(\xi,t) = \hat{\varphi}_0(\xi)e^{\xi^2\alpha t}.$$

To recover $\varphi(x,t)$ we have to use the inverse Fourier transformation F^{-1}, namely,

$$\varphi(x,t) = F^{-1}(\hat{\varphi}(\xi,t)) = F^{-1}(\hat{\varphi}_0 e^{\xi^2\alpha t}) = \varphi_0(x) * F^{-1}(e^{\xi^2\alpha t}),$$

where $*$ denotes the convolution product.

On the other hand

$$F^{-1}(e^{\xi^2\alpha t}) = \frac{1}{2\sqrt{\pi\alpha t}} e^{-\frac{x^2}{4\alpha t}},$$

so the initial value problem (1.11) has the analytic solution

$$\varphi(x,t) = \frac{1}{2\sqrt{\pi\alpha t}} \int_{-\infty}^{\infty} \varphi_0(\xi) e^{-\frac{(x-\xi)^2}{4\alpha t}} d\xi.$$

Finally, from (1.8), we obtain the analytic solution for the problem (1.7)

$$u(x,t) = \frac{\int_{-\infty}^{\infty} \frac{x-\xi}{t} \varphi_0(\xi) e^{-\frac{(x-\xi)^2}{4\alpha t}} d\xi}{\int_{-\infty}^{\infty} \varphi_0(\xi) e^{-\frac{(x-\xi)^2}{4\alpha t}} d\xi}. \tag{1.12}$$

However, the previous solution cannot always be calculated explicitly and we must implement some numerical integration method very efficiently to obtain a good approximation.
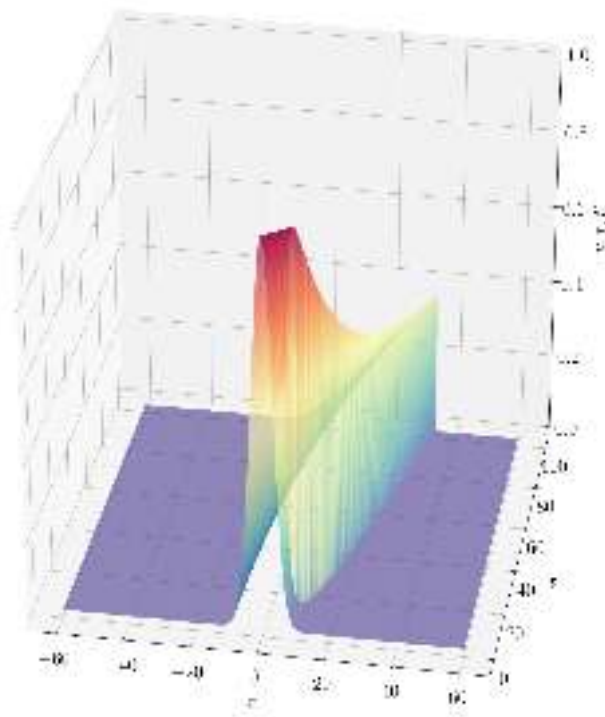

Figure 1.1: Exact solution for (1.7) with initial condition $u_0(x) = e^{-0.05x^2}$ using the equation (1.12) for $x \in [-60, 60]$, $t \in [0, 100]$, and $\alpha = 0.01$.

1.2   The Stochastic Burgers' equation

In real situations, the mathematical modeling of physical phenomena in a deterministic manner does not always produce satisfactory results, since certain hypotheses are established for their formulation, increasing uncertainty regarding spatial or temporal variables. To predict the behavior of a fluid, it is necessary to calculate the exact trajectory of each of the particles that compose it (which is an unapproachable problem).

When a fluid is in a closed container under pressure, each particle gets pushed against by all the surrounding particles. The container walls and the pressure-inducing surface (such as a piston) push against them in (Newtonian) reaction. These macroscopic forces are actually the net result of a very large number of intermolecular forces and collisions between the particles in those molecules. One fluid flow is isotropic if there is no directional preference (e.g. in fully developed turbulence); the kinetic theory of gases is also an example of isotropy if it's assumed that the molecules move in random directions and as a consequence, there is an equal probability of a molecule moving in any direction.

The equation given by (1.1) assumes that the fluid is incompressible and isotropic, where the viscous stress is given by a linear relationship with the velocity gradient (Newton's viscosity law). In addition, the collective behavior of the fluid depends only on a few macroscopic variables (such as pressure, volume, and temperature) where the internal structure of the system and the individual behavior of the particles is not relevant for thermodynamic quantities.

Sometimes, due to the large size of such a system, quantum effects can be ignored and Newton's laws may be a good approximation (in some cases, if particles move very quickly with relativistic mechanics). But it is also possible to model a fluid as a set of randomly displaced point particles that do not interact with each other, analyzed by statistical mechanics.

The information necessary to specify a physical system has to do with its entropy. When energy is degraded, Boltzmann said, it is because atoms assume a more disorderly state. And entropy is a parameter of disorder: that is the profound conception that emerges from Boltzmann's new interpretation. Oddly enough, you can create a measure for the disorder; is the probability of a particular state, defined here as the number of ways in which it can be assembled from its atoms.

When the interaction between the particles increases, their dispersion affects their positions and their velocities, which makes the entropy of the distribution increase over time until reaching a maximum (when the same system is as homogeneous and disorganized as possible). Then given a system of particles whose states X (usually position and velocity), it is possible to define a certain probability distribution that involves the various possible microstates of the system. The Maxwell-Boltzmann distribution shows how the speeds of the molecules are distributed in a Gaussian manner.

The fundamental postulate of statistical mechanics, also known as a priori equiprobability postulate, says that given an isolated system in equilibrium, the system has

the same probability of being in any of the accessible microstates. That is, a system in equilibrium has no preference for any of the microstates available for that balance. Then, in general, a system that ignores individual particles exhibits a global behavior that can be described statistically by defining macroscopic variables from a probability distribution over the microstates space.

The basic concept of entropy in information theory has a lot to do with the uncertainty that exists in any random experiment or signal, which is also called the amount of "noise" or "disorder" that a system contains or releases. In this way, we can talk about the amount of information that a signal carries. Because of this, the idea of implementing the Brownian movement, which represents the random movement observed in particles that are in a fluid medium (liquid or gas) as a result of collisions against the molecules of that fluid, gives us another way to describe complex fluids.

Over more than half a century a lot of deep mathematics was developed to tackle the rigorous understanding of turbulence and related questions in hydrodynamics problems. One of the approaches was to use stochastic analysis based on modifying the equations (as e.g. Euler, Navier-Stokes, and Burgers') adding a noise term. The idea here was to use the smoothing effect of the noise but also to discover new phenomena of stochastic nature on the other hand. In addition, this was also motivated by physical considerations, aiming at including perturbative effects, which cannot be modeled deterministically, due to too many degrees of freedom being involved, or aiming at taking into account different time scales to components of the underlying dynamics.

Because Burgers' equation given by (1.3) has a unique solution for any initial condition given, it is not a good model for turbulence. It does not display any chaos; even when a force is added to the right-hand side all solutions converge to a unique stationary solution as time goes to infinity. However, developed a parallel, theoretical, and abstract mathematical beyond its dominant presence in applications. Motivated by the intention to reinstate the Burgers' equation as a model for turbulence, the community turned its attention to the randomly forced Burgers' equation.

Several authors have suggested using the stochastic Burgers' equation as a simple model to study turbulence, [15, 16, 17, 18]. In [19] the stochastic burgers equation has been proposed to study the dynamics of the interfaces by adding a white noise (or Brownian motion) to the equation (1.3) on the right side, given as follows

$$\frac{\partial u(x,t)}{\partial t} = \alpha \frac{\partial^2 u(x,t)}{\partial x^2} + \frac{1}{2}\frac{\partial}{\partial x}(u^2(x,t)) + \frac{\partial^2 \widetilde{W}}{\partial t \partial x}. \tag{1.13}$$

This equation is a class of quasilinear stochastic PDEs (SPDEs), where $\widetilde{W}(x,t)$, $t \geq 0$, $x \in \mathbb{R}$ is a zero-mean Gaussian process. Moreover, we can write a cylindrical Wiener process W by setting

$$W(t) = \frac{\partial \widetilde{W}}{\partial x} = \sum_{j=1}^{\infty} \beta_j e_j,$$

where e_j is an orthonormal basis of $L_2(0,1)$ and β_j is a sequence of mutually independent real Brownian motions in a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ adapted to a filtration $\{\mathcal{F}_t\}_{t \geq 0}$. For more details of the above, see the Appendix A.

In the following we shall write (1.13) as follows:

$$dX(\xi,t) = \left[\alpha \partial_\xi^2 X(\xi,t) + \frac{1}{2}\partial_\xi \left(X^2(\xi,t)\right)\right]dt + dW(\xi,t), \ \ \xi \in [0,1], \ \ t > 0. \quad (1.14)$$

Equation (1.14) is supplemented with Dirichlet boundary conditions

$$X(0,t) = X(1,t) = 0, \ \ \forall t \geq 0$$

and the initial condition

$$X(\xi,0) = x(\xi), \ \ \xi \in [0,1]$$

The introduction of randomness in Burgers' equation produced a number of very interesting new directions; directions connected with dynamical systems aspects of the equation, e.g. existence and properties of invariant measures, directions related to various questions on the well-posedness of the equation in various functional settings using techniques from infinite-dimensional stochastic analysis. For further details see [19] for instance.

Also, during the past few decades, the stochastic Burgers' equation has found applications in diverse fields ranging from statistical physics, cosmology to fluid dynamics. The problem of Burgers' turbulence, that is the study of the solutions of Burgers' equation with random initial conditions or random forcing is a central issue in the study of nonlinear systems out of equilibrium. For further details see [21, 20] for instance.

A main difficulty with the multidimensional stochastic Burgers equation is that the solutions take values in a distributional space, but in the case of one-dimension, the problem of existence of solutions for stochastic Burgers equation is well understood, see [22, 23, 24, 25].

Chapter 2

The Fundamental Theory For Spectral Methods

In this chapter, we will present the elements necessary to solve partial differential equations using spectral methods given as follows

$$
\begin{cases}
\frac{\partial u}{\partial t} = \mathcal{L}u, & x \in I, \quad t > 0, \\
\\
u(x,0) = g(x), & x \in I,
\end{cases}
\tag{2.1}
$$

where u is defined in some Hilbert space $\mathcal{H}$, with initial condition $g(x) \in \mathcal{H}$ and $\mathcal{L}$ is some spatial differential operator, which allows us to represent the previous problem in another whose solution u will be given by a linear combination of already known functions.

To do this, suppose that $\mathcal{H}$ is a separable Hilbert space with the inner product $\langle \cdot, \cdot \rangle$. Therefore, we can represent the function u in terms of a known orthonormal base of $\mathcal{H}$, which we will denote as $\{\phi_k\}_{k \in I}$, given as follows

$$
u = \sum_{k \in I} \langle \phi_k, u \rangle \phi_k.
$$

There is a wide variety of families of base functions, which define different spectral methods. In this chapter, we will consider the well-known Fourier basis given by

$$
\phi_n(x) = e^{inx}.
\tag{2.2}
$$

that form an orthogonal set with the standard interior product L^2 in the interval $(0, 2\pi)$, that is,

$$
\int_0^{2\pi} \phi_k(x)\overline{\phi_l(x)}dx = 2\pi \delta kl = \begin{cases} 0 & \text{if } k \neq l, \\ 2\pi & \text{if } k = l. \end{cases}
\tag{2.3}
$$

We will denote as $B = span\{e^{inx} : |n| \leq \infty\}$ the set containing the Fourier bases. Therefore, we can define the Fourier series $F[u]$ for $u(x) \in L^2[0, 2\pi]$ as follows

$$
F[u] \equiv \sum_{|n| \leq \infty} \hat{u}_n e^{inx},
\tag{2.4}
$$

where

$$\hat{u}_n = \frac{1}{2\pi} \int_0^{2\pi} u(x)e^{-inx}dx, \quad k = 0, \pm 1, \pm 2, \ldots. \qquad (2.5)$$

which is known as the classical continuous series of trigonometric polynomials, where $\hat{u}_n$ are the Fourier coefficients.

It is important noted that the integrals in (2.5) exist if u is Riemann-integrable, i.e., if u is bounded and piecewise continuous in $(0, 2\pi)$. More generally, the Fourier coefficients are defined for any function that is integrable in the Lebesgue sense. Also the relation (2.5) associates with u a sequence of complex numbers called the Fourier transform of u. It is possible as well to introduce a Fourier cosine transform and a Fourier sine transform of u, respectively, through the formulas

$$a_n = \frac{1}{2\pi} \int_0^{2\pi} u(x)\cos(nx)dx, \quad n = 0, \pm 1, \pm 2, \ldots, \qquad (2.6)$$

and

$$b_n = \frac{1}{2\pi} \int_0^{2\pi} u(x)\sin(nx)dx, \quad n = 0, \pm 1, \pm 2, \ldots. \qquad (2.7)$$

The three Fourier transforms of u are related by the formula $\hat{u}_n = a_n - ib_n$ for $n = 0, \pm 1, \pm 2, \ldots$. Moreover, if u is a real valued function, a_n and b_n are real numbers, and $\hat{u}_{-n} = \hat{u}_n$.

Based on the above, we will present two tools to build the methods that will be used in Chapter 3, and that will be developed independently in the following two sections. In the first section, we will see that with the continuous Fourier expansion we can define a projection operator on a space of finite dimension that will allow us to approximate a function and its derivatives. In the second and last section, due to the complexity of the calculation of the previous integrals, we will see that it is possible to use quadrature rules to approximate them, and thus define an interpolation operator that will give us a discrete representation for a function and its derivatives.

For these two operators, at the end of each section we will discuss the factors that determine the behavior of the series when used to approximate smooth functions, showing how fast they approach, when, and in what sense they are convergent.

2.1   Projection Operator: Continuous Fourier Expansion

We define the projection operator denoted as $\mathcal{P}_N$ as the truncated Fourier series, i.e.,

$$\mathcal{P}_N u(x) \equiv \sum_{|n| \leq \frac{N}{2}} \hat{u}_n e^{inx}. \tag{2.8}$$

We will denote to $\hat{B}_N$ as the finite subset of $B = span\{e^{inx} : |n| \leq \infty\}$ on which it is projected the function, represented as follows

$$\hat{B}_N = span\left\{ e^{inx} : |n| \leq \frac{N}{2} \right\}, \quad dim(\hat{B}_N) = N + 1.$$

Then by the orthogonality relation (2.3), it can be seen that for $u(x) \in L^2[0, 2\pi]$

$$\langle \mathcal{P}_N u, v \rangle = \langle u, v \rangle, \quad \forall v \in S_N.$$

This shows that $\mathcal{P}_N u$ is the orthogonal projection of u upon the space of the trigonometric polynomials of degree N.

Equivalently, $\mathcal{P}_N u$ is the closest element to u in $\hat{B}_N$ with respect to the inner product

$$\langle u, v \rangle = \int_0^{2\pi} u(x)\overline{v(x)}dx,$$

and this also defines the norm

$$\|u\|^2 = \int_0^{2\pi} |u(x)|^2 dx. \tag{2.9}$$

A full characterization of the functions for which the Fourier series is convergent is the framework of Lebesgue integration for convergence in mean. This convergence can be defined in $L^2(0, 2\pi)$ (square-integrable functions), also is a complex Hilbert space with inner product defined by (2.9). Then for $u \in L^2(0, 2\pi)$ the Fourier series $F(u)$ given by (2.4) is said to be convergent in mean (or L^2-convergent) to u if

$$\int_0^{2\pi} |u(x) - \mathcal{P}_N u(x)|^2 dx \to 0, \quad \text{as} \quad N \to \infty, \tag{2.10}$$

Then the Functions in $L^2(0, 2\pi)$ can be characterized in terms of their Fourier coefficients, according to the Riesz theorem, in the following sense. If $u \in L^2(0, 2\pi)$, then its Fourier series converges to u in the sense of (2.10), and by Parseval's identity show us that

$$\|u\|^2 = 2\pi \sum_{-\infty}^{\infty} |\hat{u}_n|^2. \tag{2.11}$$

Conversely, if for any complex sequence $\{\hat{u}_n\}$, $n = 0, \pm 1, \dots$, and $\sum_{n=-\infty}^{\infty} |\hat{u}_n|^2 < \infty$, there exists a unique function $u \in L^2(0, 2\pi)$ such that its Fourier coefficients are precisely the $\hat{u}_n$'s for any n. Thus, for any function $u \in L^2(0, 2\pi)$ can be written as

$$u = \sum_{n=-\infty}^{\infty} \hat{u}_n \phi_n. \tag{2.12}$$

The Riesz theorem states that the finite Fourier transform is an isomorphism between $L^2(0, 2\pi)$ and the space l^2 of complex sequences $\{\hat{u}_n\}$, $n = 0, \pm 1, \pm 2, \dots$, such that $\sum_{n=-\infty}^{\infty} |\hat{u}_n|^2 < \infty$. The above can be summed up in the following theorem.

Theorem 2.1. *If the sum of squares of the Fourier coefficients is bounded*

$$\sum_{|n| \leq \infty} |\hat{u}_n|^2 < \infty$$

then the truncated series converges in the L^2 norm

$$\|u - \mathcal{P}_N u\|_{L^2[0,2\pi]} \to 0 \quad as \quad N \to \infty.$$

If, moreover, the sum of the absolute values of the Fourier coefficients is bounded

$$\sum_{|n| \leq \infty} |\hat{u}_n| < \infty$$

then the truncated series converges uniformly

$$\|u - \mathcal{P}_N u\|_{L^\infty[0,2\pi]} \to 0 \quad as \quad N \to \infty.$$

Note that if the truncated sum converges implies that the error is dominated by the tail of the series, i.e.,

$$\|u - \mathcal{P}_N u\|_{L^2[0,2\pi]}^2 = 2\pi \sum_{|n| > \frac{N}{2}} |\hat{u}_n|^2,$$

and

$$\|u - \mathcal{P}_N u\|_{L^\infty[0,2\pi]} \leq \sum_{|n| > \frac{N}{2}} |\hat{u}_n|.$$

Thus, the error committed by replacing $u(x)$ with its Nth-order Fourier series depends solely on how fast the expansion coefficients of $u(x)$ decay.

To appreciate this, suppose that $u(x) \in L_p^2[0, 2\pi]$ and that its derivative $u'(x) \in L_p^2[0, 2\pi]$, where the subscript p indicate that the function is periodic. then for $n \neq 0$ we have to

$$2\pi \hat{u}_N = \int_0^{2\pi} u(x)e^{-inx}dx$$

$$= -\frac{1}{in}(u(2\pi) - u(0)) - \frac{1}{in}\int_0^{2\pi} u'(x)e^{inx}dx,$$

therefore

$$|\hat{u}_N| \propto \frac{1}{n}.$$

In general, if for $u(x)$ and its derivatives $(m-1)$, and its periodic extensions are all continuous, and also if its derivative mth is measurable at $[0, 2\pi]$, also known in the literature as the regularity of the function, in this particular case in L_p^2, we have to $\forall n \neq 0$, repeating the previous procedure successively, the behavior of Fourier coefficients $\hat{u}_n$ of $u(x)$ is similar, i.e.,

$$|\hat{u}_n| \propto \left(\frac{1}{n}\right)^m.$$

This is known as spectral convergence, which means that the smoother the function, the series converges faster.

This result is important since it will allow us to investigate the convergence rate of the methods, which we will define in detail later. Therefore, we will focus on periodic functions expanded in Fourier series since its rapid decay of the coefficients implies that the Fourier series truncated after just a few more terms represents an exceedingly good approximation of the function. However, in practice, this decay is not exhibited until there are enough coefficients to represent all the essential structures of the function but in general, functions can be described both through their values in physical space and through their coefficients in transform space. The following examples illustrate the previous results.

Example 2.1. *Consider the function $u(x) \in C^\infty[0, 2\pi]$ given by*

$$u(x) = \frac{1}{5 - 4\cos(x)} \tag{2.13}$$

with its expansion coefficients

$$\hat{u}_n = \frac{2^{-|n|}}{3}.$$

In Figure 2.1 we can clearly observe the convergence of the Fourier series and that in addition, the convergence of the approximation is almost uniform. This is due to the periodicity of the function and its derivatives.

15

Figure 2.1: (a) Continuous Fourier series approximation of the equation (2.13). (b) The Pointwise error of approximation.

Example 2.2. *The expansion coefficients of the function*

$$u(x) = \frac{\pi}{2}\sin(\frac{x}{2}) \tag{2.14}$$

are given by

$$\hat{u}_n = \frac{1}{(1 - 4n^2)}.$$

Note that is infinitely differentiable in $[0, 2\pi]$, but $u'(0) \neq u'(2\pi)$. In Figure 2.2 we can see that the convergence is much slower than in the Example 2.13, as expected
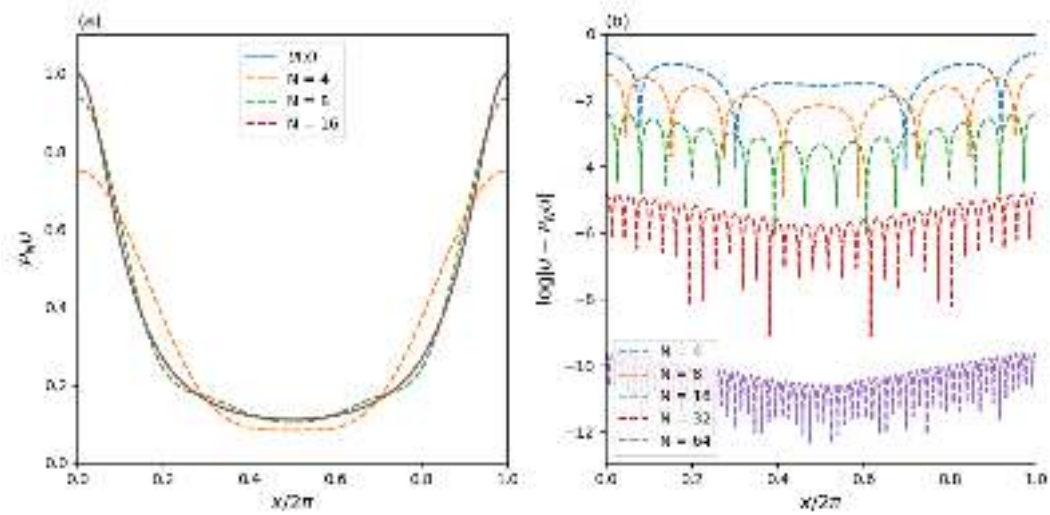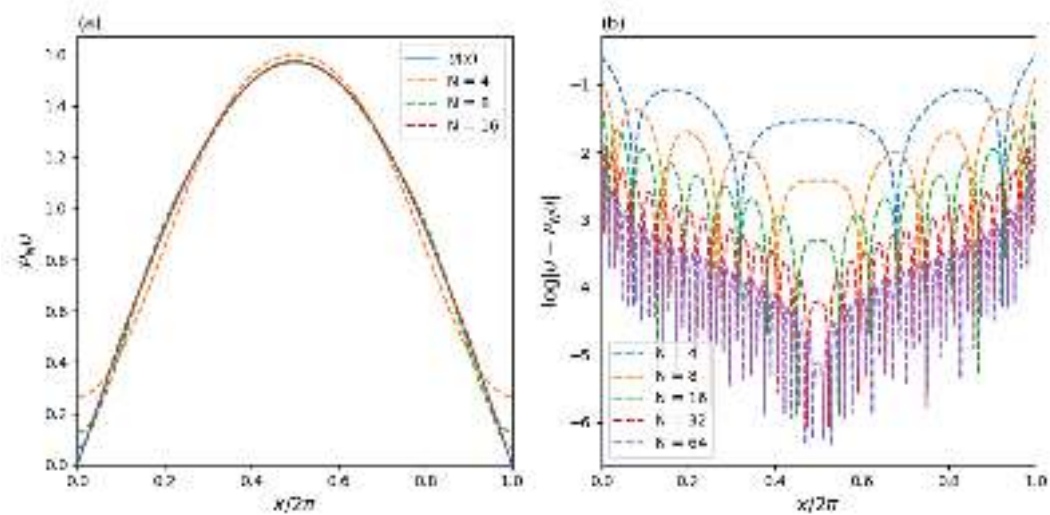

Figure 2.2: (a) Continuous Fourier series approximation of the equation (2.14). (b) The Pointwise error of approximation for increasing resolution.

2.1.1   Differentiation of the Continuous Expansion

To find solutions of partial differential equations using the spectral methods, in addition to approximating a function $u(x)$ by the finite Fourier series $\mathcal{P}_N u$, we also need to obtain its derivatives. Due to the linearity of the derivative and that these functions are exponential, we can easily obtain the derivatives of $\mathcal{P}_N u$ by simply differentiating the basis functions term by term. Therefore, if we have the following series truncated

$$\mathcal{P}_N u(x) = \sum_{|n| \leq \frac{N}{2}} \hat{u}_n e^{inx},$$

from this, we can get

$$\frac{d^q}{dx^q}\mathcal{P}_N u(x) = \sum_{|n| \leq \frac{N}{2}} \hat{u}_n \frac{d^q}{dx^q} e^{inx} = \sum_{|n| \leq \frac{N}{2}} (in)^q \hat{u}_n e^{inx}.$$

Therefore the projection and differentiation operators commute, i.e.,

$$\mathcal{P}_N \frac{d^q}{dx^q} u = \frac{d^q}{dx^q}\mathcal{P}_N u.$$

This property implies that for any differentiation operator $\mathcal{L}$ with constant coefficients,

$$\mathcal{P}_N \mathcal{L}(I - \mathcal{P}_N)u$$

vanishes, which known as the truncation error. Thus, the Fourier approximation to the equation $u_t = \mathcal{L}u$ is exactly the projection of the analytic solution.

2.1.2   Approximation theory for Continuous Expansion.

The behavior of the functions and their derivatives that we have shown is relevant when the solutions of the differential equations are approximated using spectral methods since it allows us to investigate how fast and precise they can be. In this subsection, we will present these properties based in [28] as detailed as possible some useful results for our main objective regarding the analysis of the projection operator already defined above.

When using the Fourier approximation to discretize the spatial part of the equation

$$u_t = \mathcal{L}u,$$

it is important that our approximation, both to u and to $\mathcal{L}u$, be accurate,i.e., we must consider not only the difference between u and $\mathcal{P}_N u$ if not also the distance

between $\mathcal{L}u$ and $\mathcal{L}\mathcal{P}_N u$, measured in an appropriate norm. This is because the actual rate of convergence is determined by the truncation error

$$\mathcal{P}_N \mathcal{L}(I - \mathcal{P}_N)u.$$

Thus, the error is determined not only by the behavior of the Fourier approximations of the function but also of its derivatives, as we have seen previously. Therefore, the Sobolev q-norm denoted by $H_p^q[0, 2\pi]$, It is appropriate to estimate the truncation error since it measures the smoothness of the derivatives and the function. This norm is defined as follows

$$\|u\|_{H_p^q[0,2\pi]}^2 = \sum_{m=0}^{q} \int_0^{2\pi} |u^m(x)|^2 \, dx. \tag{2.15}$$

The subscript p indicates the fact that all functions are periodic. By substituting the Fourier expansion for each derivative in (2.15), the Sobolev norm can be written as

$$\|u\|_{H_p^q[0,2\pi]}^2 = 2\pi \sum_{m=0}^{q} \sum_{|n|\leq\infty} |n|^{2m} |\hat{u}_n|^2 = 2\pi \sum_{|n|\leq\infty} \left(\sum_{m=0}^{q} |n|^{2m} \right) |\hat{u}_n|^2,$$

where the interchange of the summation is allowed provided $u(x)$ has sufficient smoothness.

Before starting the analysis, without loss of generality, we first consider the continuous Fourier series given by

$$\mathcal{P}_{2N} u(x) = \sum_{|n|\leq N} \hat{u}_n e^{inx}.$$

The first important result is the estimate in L^2 for the distance between u and its trigonometric approximation $\mathcal{P}_{2N}u$, which shows everything we've seen previously.

Theorem 2.2. *For any $u(x) \in H_p^r[0, 2\pi]$, there exists a positive constant C, independent of N, such that*

$$\|u - \mathcal{P}_{2N}u\|_{L^2[0,2\pi]} \leq CN^{-q}\|u^{(q)}\|_{L^2[0,2\pi]},$$

provided $0 \leq q \leq r$.

Proof. By Parsevals identity given by (2.11) we get

$$\|u - \mathcal{P}_{2N}u\|_{L^2[0,2\pi]}^2 = 2\pi \sum_{|n|>N} |\hat{u}_n|^2.$$

We rewrite this summation as follows

$$
\begin{aligned}
\sum_{|n|>N} |\hat{u}_n|^2 &= \sum_{|n|>N} \frac{n^{2q}}{n^{2q}} |\hat{u}_n|^2 \\
&\leq N^{-2q} \sum_{|n|>N} n^{2q} |\hat{u}_n|^2 \\
&\leq N^{-2q} \sum_{|n|\geq 0} n^{2q} |\hat{u}_n|^2 \\
&= \frac{1}{2\pi} N^{-2q} \|u^{(q)}\|^2_{L^2[0,2\pi]}.
\end{aligned}
$$

Putting all the above together and taking out the square root, we get our result.   □

Note that the smoother the function, the larger the value of q and therefore, the better the approximation, as seen before. Now let's notice the following. Suppose that $u(x)$ is analytical, so we have to

$$
u^{(q)} = \sum_{|n|\leq \infty} (in)^q \hat{u}_n e^{inx}.
$$

Since $u^{(q)} \in W_p^q$, and by (2.11)

$$
\|u^{(q)}\|_{L^2[0,2\pi]} = \sum_{|n|\leq \infty} |n|^{2q} |\hat{u}_n|^2 \leq Cq! \sum_{|n|\leq \infty} |\hat{u}_n|^2 \leq Cq! \|u\|_{L^2[0,2\pi]},
$$

and so by the previous theorem

$$
\|u - \mathcal{P}_{2N}u\|_{L^2[0,2\pi]} \leq N^{-q} \|u^{(q)}\|_{L^2[0,2\pi]} \leq C \frac{q!}{N^q} \|u\|_{L^2[0,2\pi]}.
$$

Using Stirlings formula, $q! \sim q^q e^{-q}$, and assuming that $q \propto N$, we obtain

$$
\|u - \mathcal{P}_{2N}u\|_{L^2[0,2\pi]} \leq\sim C \left(\frac{q}{N}\right)^q e^{-q} \|u\|_{L^2[0,2\pi]} \sim K e^{-cN} \|u\|_{L^2[0,2\pi]}.
$$

Thus, for an analytic function, its spectral convergence is exponential convergence.

We must not forget that the theory we have previously presented is with the assumption that the functions and their derivatives are all periodic. But it is possible to do a similar analysis considering some other class of functions, such as functions that vanish at the borders. However, these kinds of functions belong to spaces very similar to those we have studied, and it is possible to use the same results.

2.2   Interpolation Operator: Discrete Fourier Expansion

The continuous Fourier series method requires the evaluation of the coefficients

$$\hat{u}_n = \frac{1}{2\pi} \int_0^{2\pi} u(x)e^{-inx}dx. \tag{2.16}$$

In general, these integrals cannot be computed analytically, and one resorts to the approximation of the Fourier integrals by using quadrature formulas. This procedure defines a discrete transform between the set of values of u at the quadrature points and the set of approximate, or discrete, coefficients. The finite series defined by the discrete transform is actually the interpolate of u at the quadrature nodes. If the properties of accuracy (in particular the spectral accuracy) are retained by replacing the finite transform with the discrete transform, then the interpolant series can be used instead of the truncated series to approximate functions. Also, quadrature formulas differ based on the exact position of the grid points, and the choice of an even or odd number of grid points results in slightly different schemes.

2.2.1   The Even Expansion

Define an equidistant grid, consisting of an even number N of gridpoints $x_j \in [0, 2\pi)$, defined by

$$x_j = \frac{2\pi j}{N}, \quad j \in [0, \cdots, N-1].$$

The trapezoidal rule yields the discrete Fourier coefficients $\widetilde{u}_n$, which approximate the continuous Fourier coefficients $\hat{u}_n$ given as follows

$$\widetilde{u}_n = \frac{1}{N} \sum_{j=0}^{N-1} u(x_j)e^{-inx_j}. \tag{2.17}$$

The difference between the continuous and the discrete approximation is very clear since here we only need precision in the points x_j. This may somehow be an advantage in the numerical calculation because in some cases it is possible to obtain the same order of precision, as shown in the following theorem when trigonometric polynomials are involved, the trapezoidal quadrature rule is a very natural approximation.

Theorem 2.3. *For the points x_j defined as above, the quadrature formula*

$$\frac{1}{2\pi} \int_0^{2\pi} f(x)dx = \frac{1}{N} \sum_{j=0}^{N-1} f(x_j),$$

is exact for any trigonometric polynomial $f(x) = e^{inx}$, $|n| < N$.

Proof. Given a function $f(x) = e^{inx}$, It is easy to observe that

$$\frac{1}{2\pi} \int_0^{2\pi} f(x)dx = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{otherwise.} \end{cases}$$

On the other hand,

$$\frac{1}{N} \sum_{j=0}^{N-1} f(x_j) = \frac{1}{N} \sum_{j=0}^{N-1} e^{in(\frac{2\pi j}{N})}$$

$$= \frac{1}{N} \sum_{j=0}^{N-1} q^j$$

where $q = e^{i\frac{2\pi n}{N}}$. If n is an integer multiple of N , i.e., $n = mN$, then, we have to

$$\frac{1}{N} \sum_{j=0}^{N-1} e^{iNm(\frac{2\pi j}{N})} = \frac{1}{N} \sum_{j=0}^{N-1} e^{i(2\pi jm)} = 1$$

Otherwise,

$$\frac{1}{N} \sum_{j=0}^{N-1} q^j = \frac{q^N - 1}{q - 1} = 0$$

Thus, the quadrature formula is exact for any function of the form $f(x) = e^{inx}$, $|n| < N$. $\qquad\square$

Moreover, we can see that the quadrature formula is exact for $f(x) \in \hat{B}_{2N-2}$ where $\hat{B}_N$ is defined as before. Then using the trapezoid rule, the discrete Fourier coefficients become

$$\widetilde{u}_n = \frac{1}{N\widetilde{c}_n} \sum_{j=0}^{N-1} u(x_j)e^{-inx_j}, \tag{2.18}$$

where we introduce the coefficients

$$\widetilde{c}_n = \begin{cases} 2 & \text{if} \quad |n| = N/2, \\ 1 & \text{if} \quad |n| < N/2. \end{cases} \tag{2.19}$$

These relations define a new projection of u

$$\mathcal{I}_N u(x) = \sum_{|n| \leq \frac{N}{2}} \widetilde{u}_n e^{inx} \tag{2.20}$$

This is the complex discrete Fourier transform, based on an even number of quadrature points. From the above, we can see that

$$\widetilde{u}_{-N/2} = \widetilde{u}_{N/2},$$

so we have exactly N independent Fourier coefficients, corresponding to the N quadrature points. As a consequence, $\mathcal{I}_N \sin(\frac{N}{2}x) = 0$, so that the function $\sin(\frac{N}{2}x)$ is not represented in the above expansion. Therefore, the space $\hat{B}_N$ does not include $\sin(\frac{N}{2}x)$, and the correct space must be as follows

$$\widetilde{B}_N = span\left\{ \left(\cos(nx), \;\; 0 \le n \le \frac{N}{2} \right) \cup \left(\sin(nx), \;\; 1 \le n \le \frac{N}{2} - 1 \right) \right\},$$

which has dimension $\dim(\widetilde{B}_N) = N$.

In the same way, as in the previous subsection using the discrete expansion for Examples 2.13 and 2.14, we can observe the same behavior as with continuous expansion, but now we have that the error at each point x_j of the grid is zero.

Example 2.3. *Consider the $C_p^\infty[0, 2\pi]$ function*

$$u(x) = \frac{1}{5 - 4\cos(x)}. \tag{2.21}$$

Its expansion coefficients are
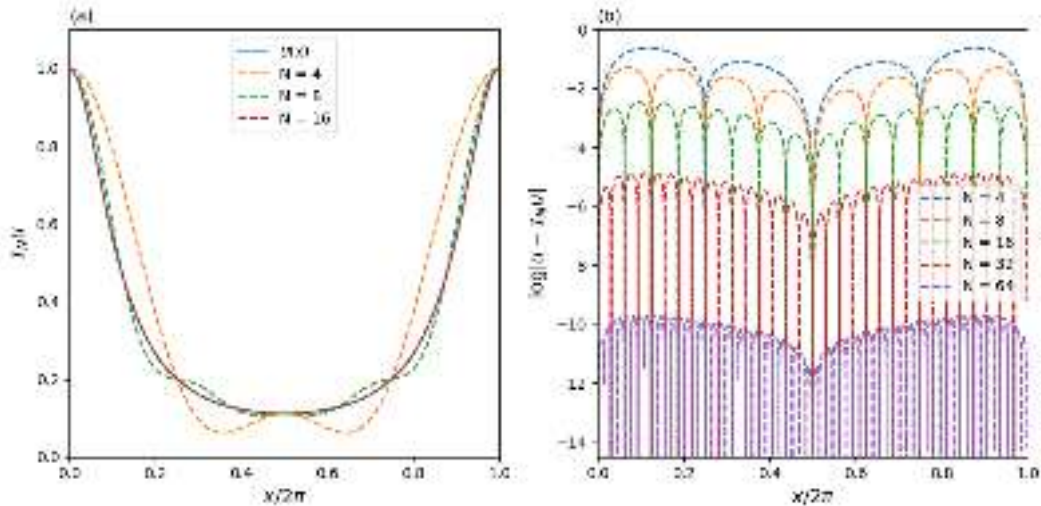
$$\hat{u}_n = \frac{2^{-|n|}}{3}.$$

Figure 2.3: (a) Discrete Fourier series approximation of the equation (2.21). (b) Pointwise error of approximation for increasing resolution.

Example 2.4. *The expansion coefficients of the function*

$$u(x) = \frac{\pi}{2}\sin(\frac{x}{2}), \tag{2.22}$$

are given by

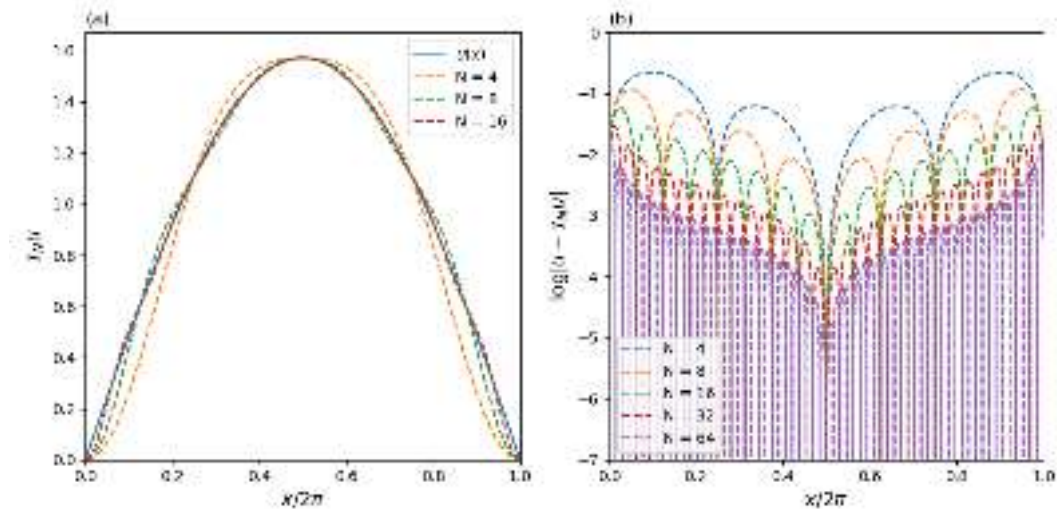$$\hat{u}_n = \frac{1}{(1 - 4n^2)}.$$

Figure 2.4: (a) Discrete Fourier series approximation of the equation (2.22). (b) Pointwise error of approximation for increasing resolution.

Therefore, we can see that the discrete expansion is, in fact, an interpolation operator as mentioned. This can be shown in the following theorem.

Theorem 2.4. *Let the discrete Fourier transform be defined by Equations (2.18)-(2.20). For any periodic function, $C_p^0[0, 2\pi]$, we have*

$$\mathcal{I}_N u(x_j) = u(x_j), \quad \forall x_j = \frac{2\pi j}{N}, \quad j = 0, \ldots, N - 1.$$

Proof. Substituting Equation (2.18) into Equation (2.20) we obtain

$$\mathcal{I}_N u(x) = \sum_{|n| \leq \frac{N}{2}} \left(\frac{1}{N \widetilde{c}_n} \sum_{j=0}^{N-1} u(x_j) e^{-in x_j} \right) e^{inx}.$$

Exchanging the order of the sum gives

$$\mathcal{I}_N u(x) = \sum_{j=0}^{N-1} u(x_j) g_j(x), \tag{2.23}$$

23

where

$$g_j(x) = \sum_{|n| \leq \frac{N}{2}} \frac{1}{N \widetilde{c}_n} e^{in(x - x_j)}$$

$$= \frac{1}{N} \sin \left[N \frac{x - x_j}{2} \right] \cot \left[\frac{x - x_j}{2} \right]$$

by summing as a geometric series. It is easily verified that $g_j(x_i) = \delta_{ij}$

We still need to show that $g_j(x) \in \widetilde{B}_N$. Clearly, $g_j(x) \in \hat{B}_N$ as $g_j(x)$ is a polynomial of degree $\leq N/2$. However, since

$$\frac{1}{2} e^{-i \frac{N}{2} x_j} = \frac{1}{2} e^{i \frac{N}{2} x_j} = \frac{(-1)^j}{2},$$

and, by convention $\widetilde{u}_{-N/2} = \widetilde{u}_{N/2}$, we do not get any contribution from the term $\sin(\frac{N}{2} x)$, hence $g_j(x) \in \widetilde{B}_N$. □

2.2.2   The Odd Expansion

Similarly, we define a grid with an odd number of grid points as follows

$$x_j = \frac{2\pi}{N + 1} j, \quad j \in [0, \ldots, N],$$

and using the trapezoidal rule we get

$$\widetilde{u}_n = \frac{1}{N + 1} \sum_{j=0}^{N} u(x_j) e^{-inx_j}, \tag{2.24}$$

to obtain the interpolation operator

$$\mathcal{J}_N u(x) = \sum_{|n| \leq \frac{N}{2}} \widetilde{u}_n e^{inx}. \tag{2.25}$$

Again as before, the quadrature formula is highly accurate.

Theorem 2.5. *For the points x_j defined as above, the quadrature formula*

$$\frac{1}{2\pi} \int_0^{2\pi} f(x) dx = \frac{1}{N + 1} \sum_{j=0}^{N} f(x_j),$$

is exact for any $f(x) = e^{inx}$, $|n| < N$, i.e., for all $f(x) \in \widetilde{B}_{2N}$.

Proof. Given a function $f(x) = e^{inx}$, It is easy to observe that

$$\frac{1}{2\pi} \int_0^{2\pi} f(x)dx = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{otherwise.} \end{cases}$$

On the other hand,

$$\frac{1}{N+1} \sum_{j=0}^{N} f(x_j) = \frac{1}{N+1} \sum_{j=0}^{N} e^{in(\frac{2\pi j}{N+1})}$$

$$= \frac{1}{N+1} \sum_{j=0}^{N} q^j$$

where $q = e^{i\frac{2\pi n}{N+1}}$. If n is an integer multiple of $N+1$, i.e., $n = (N+1)m$, then we have to

$$\frac{1}{N+1} \sum_{j=0}^{N} e^{i(N+1)m(\frac{2\pi j}{N+1})} = \frac{1}{N+1} \sum_{j=0}^{N} e^{i(2\pi jm)} = 1$$

Otherwise,

$$\frac{1}{N+1} \sum_{j=0}^{N} q^j = \frac{q^{N+1} - 1}{q - 1} = 0$$

Thus, the quadrature formula is exact for any function of the form $f(x) = e^{inx}$, $|n| < N$. □

The scheme may also be expressed through the use of a Lagrange interpolation polynomial,

$$\mathcal{J}_N u(x) = \sum_{j=0}^{N} u(x_j)h_j(x) \tag{2.26}$$

where

$$h_j(x) = \frac{1}{N+1} \sum_{|k| \leq \frac{N}{2}} e^{ik(x-x_j)} = \frac{1}{N+1} \frac{\sin(\frac{N+1}{2}(x - x_j))}{\sin(\frac{x-x_j}{2})} \tag{2.27}$$

One easily shows that $h_j(x_l) = \delta_{jl}$ and that $h_j(x) \in \hat{B}_N$.

2.2.3 Differentiation of the discrete expansions

Similarly, as in continuous expansion, we require compute derivatives of the discrete approximation. In the following subsections, we assume that our function u and all its derivatives are continuous and periodic on $[0, 2\pi]$.

We consider the case of an even number of grid points. Using expansion coefficients given the values of the function $u(x)$ at the points x_j , differentiating the basis functions in the interpolant yields

$$\frac{d}{dx}\mathcal{I}_N u(x) = \sum_{|n|\leq N/2} in\widetilde{u}_n e^{inx}, \ \ \widetilde{u}_n = \frac{1}{N\widetilde{c}_n}\sum_{j=0}^{N-1} u(x_j)e^{-inx_j}, \qquad (2.28)$$

where $\widetilde{c}_n$ is given by (2.19). Higher order derivatives can be obtained simply by further differentiating the basis functions.

Similarly, for the case of an odd number of grid points

$$\frac{d}{dx}\mathcal{J}_N u(x) = \sum_{|n|\leq N/2} in\widetilde{u}_n e^{inx}, \ \ \widetilde{u}_n = \frac{1}{N+1}\sum_{j=0}^{N} u(x_j)e^{-inx_j}, \qquad (2.29)$$

The procedure for differentiating using expansion coefficients can be described as follows: first, we transform the point values $u(x_j)$ in physical space into the coefficients $\widetilde{u}_n$ in mode space. We then differentiate in mode space by multiplying $\widetilde{u}_n$ by in, and return to physical space.

There are other ways to obtain these derivatives, which may have greater advantage and be more efficient to calculate. In the literature, it can commonly find the use of differentiation matrices, for which there is a great variety. We will present some matrices that have been studied in [28], [26], and we will observe the difference between the cases of an even and odd number of grid points.

Differentiation Matrix. Recall that to the case of an even number of grid points, the interpolation operator can be written as

$$\mathcal{I}_N u(x) = \sum_{j=0}^{N-1} u(x_j)g_j(x),$$

where g_j are the Lagrange interpolation polynomials given by

$$g_j(x) = \frac{1}{N}\sin\left[N\frac{x-x_j}{2}\right]\cot\left[\frac{x-x_j}{2}\right].$$

26

Then, by differentiating the interpolation directly, it can get an approximation to the derivative of $u(x)$ at the points x_j as follows

$$\frac{d}{dx}\mathcal{I}_N(x)\Big|_{x_l} = \sum_{j=0}^{N-1} u(x_j)\frac{d}{dx}g_j(x)\Big|_{x_l} = \sum_{j=0}^{N-1} D_{lj}u(x_j),$$

where D_{lj} are the differentiation matrix entries given by

$$D_{ij} = \frac{d}{dx}g_j(x)\Big|_{x_i} = \begin{cases} \frac{(-1)^{i+j}}{2}\cot\left[\frac{x_i-x_j}{2}\right] & i \neq j, \\ 0 & i = j, \end{cases} \qquad (2.30)$$

it is also well known that D is circulant and skew-symmetric matrix. In the same way, the entries of the second order differentiation matrix $D^{(2)}$ gives us

$$D_{ij}^{(2)} = \frac{d^2}{dx^2}g_j(x)\Big|_{x_i} = \begin{cases} -\frac{(-1)^{i+j}}{2}\left[\sin\left[\frac{x_i-x_j}{2}\right]\right]^{-1} & i \neq j, \\ -\frac{N^2+2}{12} & i = j. \end{cases} \qquad (2.31)$$

The approximation of higher derivatives follows exactly the same route, and similarly to obtain the entries of the differentiation matrix $\widetilde{D}$ for the interpolation based on an odd number of points given by

$$\widetilde{D}_{ij} = \begin{cases} -\frac{(-1)^{i+j}}{2}\left[\sin\left[\frac{x_i-x_j}{2}\right]\right]^{-2} & i \neq j, \\ 0 & i = j. \end{cases} \qquad (2.32)$$

It is also known that $\widetilde{D}$ is a circulant, skew-symmetric matrix. The advantage of this method is that the differentiation matrix takes us from physical space to physical space, and the act of differentiation is hidden in the matrix itself.

It is interesting to observe that the differentiation operator for the interpolation based on an odd number of grid points, takes elements of $\widetilde{B}_N$ out of $\widetilde{B}_N$ and then

$$\mathcal{I}_N\frac{d^2}{dx^2}\mathcal{I}_N \neq \left(\mathcal{I}_N\frac{d}{dx}\right)^2\mathcal{I}_N.$$

But for the interpolation based on an odd number of grid points the differentiation operator remain in $\hat{B}_N$ when takes elements of $\hat{B}_N$, and thus,

$$\mathcal{J}_N\frac{d^2}{dx^2}\mathcal{J}_N = \left(\mathcal{J}_N\frac{d}{dx}\right)^2\mathcal{J}_N$$

Moreover, for all values of q we have

$$\widetilde{D}^{(q)} = \mathcal{J}_N\frac{d^q}{dx^q}\mathcal{J}_N = \widetilde{D}^q$$

allowing us to calculate approximate high derivatives by just multiplying the D matrix as many times as necessary.

For the above, and for some interesting properties that we will see later about interpolation operator based on an odd number of grid points, it has been decided to use it for the study of this work.

2.2.4 Approximation theory for Discrete Expansion.

Based on the theory developed in [28] with respect to the interpolation operator analysis for the case of an even number of grid points, we will adapt the results for the case of an odd number of grid points in the most detailed way possible.

First of all, we can define a discrete version of the inner product L^2 as follows

$$\langle f_N, g_N \rangle_N = \frac{1}{N+1} \sum_{j=0}^{N} f_N(x_j) \bar{g}_N(x_j),$$

and the associated norm

$$\|f_N\|_N^2 = \langle f_N, f_N \rangle_N$$

where f_N , $g_N \in \hat{B}_N$ and there are an odd number of grid points x_j , $j = 0, \ldots, N$. Note also that the interpolant $\mathcal{J}_N u$ of a continuous function u and for all $v \in \hat{B}_N$, satisfies trivially the identity

$$\langle \mathcal{J}_N u, v \rangle_N = \langle u, v \rangle_N.$$

Moreover, as a consequence of the exactness of the quadrature rule for trigonometric functions, as have seen in Theorem 2.3, we have

$$\langle f_N, g_N \rangle_N = \frac{1}{2\pi} \int_0^{2\pi} f_N \bar{g}_N dx, \quad \|f_N\|_{L^2[0,2\pi]} = \|f_N\|_N \tag{2.33}$$

Hence, in $\hat{B}_N$, the continuous and discrete inner product are the same.

The situation is different when we discuss an even number of grid points. If $f_N, g_N \in \tilde{B}_N$ and we have an even number of grid points x_j , the discrete inner product

$$\langle f_N, g_N \rangle_N = \frac{1}{N} \sum_{j=0}^{N-1} f_N(x_j) \bar{g}_N(x_j), \quad \|f_N\|_N^2 = \langle f_N, f_N \rangle_N$$

is not equal to the continuous inner product. However, using the fact that $f_N \in L^2[0, 2\pi]$ it can be shown that there exists a $K > 0$ such that

$$K^{-1} \|f_N\|_{L^2[0,2\pi]}^2 \leq \|f_N\|_N^2 \leq K \|f_N\|_{L^2[0,2\pi]}^2. \tag{2.34}$$

Something very interesting and useful in the use of discrete expansion to approximate functions and their derivatives is that the behavior is very similar to that shown in the previous subsection for continuous expansion. We will see that the approximation theory for the discrete expansion yields essentially the same results as for the continuous expansion. The proofs are based on the fact that the Fourier

coefficients of the discrete approximation are sufficiently close to those of the continuous approximation.

Recall that the interpolation operator associated with an odd number of grid points is given by

$$\mathcal{J}_{2N} u = \sum_{|n| \leq N} \widetilde{u}_n e^{inx},$$

with expansion coefficients

$$\widetilde{u}_n = \frac{1}{2N+1} \sum_{j=0}^{2N} u(x_j) e^{-inx_j}, \quad x_j = \frac{2\pi j}{2N+1}.$$

First we observe the following, the interpolation operator associated with an odd number of grid points are based on the points x_j, for which the $(n + Mm)$th mode, where $M = 2N + 1$, is indistinguishable from the nth mode, i.e.,

$$e^{i(n+Mm)x_j} = e^{inx_j} e^{i2\pi m j} = e^{inx_j}$$

This phenomenon is known as aliasing.

Moreover, due to the orthogonality relation as before seen we have to

$$\frac{1}{M} \sum_{j=0}^{M-1} e^{-inx_j} = \left\{ \begin{array}{ll} 1 & \text{if } n = Mm, \ m = 0, \pm 1, \pm 2, \ldots, \\ 0 & \text{otherwise.} \end{array} \right.$$

The relationship between the discrete expansion coefficients $\widetilde{u}_n$, and the continuous expansion coefficients $\hat{u}_n$, is given in the following lemma.

Lemma 2.1. *Consider $u(x) \in W_p^q[0, 2\pi]$, where $q > 1/2$. For $|n| \leq N$ we have*

$$\widetilde{c}_n \widetilde{u}_n = \hat{u}_n + \sum_{\substack{|m| \leq \infty \\ m \neq 0}} \hat{u}_{n+Mm} \tag{2.35}$$

Proof. Substituting the continuous Fourier expansion into the discrete expansion yields

$$\widetilde{c}_n \widetilde{u}_n = \frac{1}{M} \sum_{j=0}^{M-1} \sum_{|l| \leq \infty} \hat{u}_l e^{i(l-n)x_j}$$

To interchange the two summations we must ensure uniform convergence, i.e., $\sum_{|l| \leq \infty} |\hat{u}_l| < \infty$. This is satisfied, since if $q > 1/2$ then, as before there is $m \in \mathbb{N}$ such that for $l \geq m$ we have to

$$(1 + |l|)^{2q} \leq 2q(1 + l^{2q})$$

taking m as follows

$$\frac{1}{m} \leq (2q)^{\frac{1}{2q}} - 1$$

Therefore

$$\sum_{|l| \leq \infty} |\hat{u}_l| = \sum_{|l| \leq \infty} (1 + |l|)^q \frac{|\hat{u}_l|}{(1 + |l|)^q}$$

$$\leq \left(2q \sum_{|l| \leq \infty} (1 + l^{2q}) |\hat{u}_l|^2 \right)^{1/2} \left(\sum_{|l| \leq \infty} (1 + |l|)^{-2q} \right)^{1/2},$$

where the last expression follows from the Cauchy-Schwarz inequality. As $u(x) \in W_p^q[0, 2\pi]$ the first part is clearly bounded. Furthermore, the second term is a p-series and then converges provided $q > 1/2$, ensuring boundedness.

Interchanging the order of summation and using orthogonality of the exponential function at the grid yields the desired result

$$\tilde{c}_n \tilde{u}_n = \frac{1}{M} \sum_{j=0}^{M-1} \sum_{|l| \leq \infty} \hat{u}_l e^{i(l-n)x_j} = \sum_{|l| \leq \infty} \frac{1}{M} \sum_{j=0}^{M-1} \hat{u}_l e^{i(l-n)x_j}$$

$$= \sum_{|m| \leq \infty} \frac{1}{M} \sum_{j=0}^{M-1} \hat{u}_{n+Mm} e^{i(n+Mm)x_j}$$

$$= \frac{1}{M} \sum_{j=0}^{M-1} \hat{u}_n e^{inx_j} + \sum_{\substack{|m| \leq \infty \\ m \neq 0}} \frac{1}{M} \sum_{j=0}^{M-1} \hat{u}_{n+Mm} e^{i(n+Mm)x_j}$$

$$= \hat{u}_n + \sum_{\substack{|m| \leq \infty \\ m \neq 0}} \hat{u}_{n+Mm}$$

$\square$

The conclusions of the previous discussion are equally valid in the number of odd or even points. An equivalent formulation of (2.35) is

$$\mathcal{J}_N u = \mathcal{P}_N u + \mathcal{A}_N u$$

It is orthogonal to the truncation error, $u - \mathcal{P}_N u$, so that

$$\|u - \mathcal{J}_N u\|^2 = \|u - P_N u\|^2 + \|\mathcal{A}_N u\|^2$$

Hence, the error due to the interpolation is actually always larger than the error due to the truncation of the Fourier series.

Rather than deriving the estimates of the approximation error directly, we shall use the results obtained in the previous section and then estimate the difference between the two different expansions, which we recognize as the aliasing error given by

$$
\|\mathcal{A}_N\|_{L^2[0,2\pi]} = \left\| \sum_{|n|<N} \left(\sum_{\substack{|m|\leq\infty \\ m\neq 0}} \hat{u}_{n+Mm} \right) \right\|_{L^2[0,2\pi]}
$$

As before, we first consider the behavior of the approximation in the L^2-norm. We will first show that the bound on the aliasing error, $\mathcal{A}_N$, in equation above is of the same order as the truncation error. The error caused by truncating the continuous expansion is essentially the same as the error produced by using the discrete coefficients rather than the continuous coefficients.

Lemma 2.2. *For any $u(x) \in W_p^r[0,2\pi]$, where $r > 1/2$, the aliasing error*

$$
\|\mathcal{A}_N\|_{L^2[0,2\pi]} = \left(\sum_{|n|\leq\infty} |\widetilde{c}_n \widetilde{u}_n - \hat{u}_n|^2 \right)^{1/2} \leq CN^{-r}\|u^r\|_{L^2[0,2\pi]}
$$

Proof. From Lemma 2.1 we have

$$
|\widetilde{c}_n \widetilde{u}_n - \hat{u}_n|^2 = \left| \sum_{\substack{|m|\leq\infty \\ m\neq 0}} \hat{u}_{n+Mm} \right|^2
$$

To estimate this, we first note that

$$
\left| \sum_{\substack{|m|\leq\infty \\ m\neq 0}} \hat{u}_{n+Mm} \right|^2 = \left| \sum_{\substack{|m|\leq\infty \\ m\neq 0}} |n+Mm|^r \hat{u}_{n+Mm} \frac{1}{|n+Mm|^r} \right|^2
$$

$$
\leq \left(\sum_{\substack{|m|\leq\infty \\ m\neq 0}} |n+Mm|^{2r} |\hat{u}_{n+Mm}|^2 \right) \left(\sum_{\substack{|m|\leq\infty \\ m\neq 0}} \frac{1}{|n+Mm|^{2r}} \right)
$$

using the Cauchy-Schwartz inequality. Since $M = 2N+1$ and $|n| \leq N$, we have to $N(2m-1) = 2Nm - N \leq |n+Mm|$. Hence, bounding of the second term is ensured by

$$
\sum_{\substack{|m|\leq\infty \\ m\neq 0}} \frac{1}{|n+Mm|^{2r}} \leq \frac{2}{N^{2r}} \sum_{m=1}^{\infty} \frac{1}{(2m-1)^{2r}} = C_1 N^{-2r},
$$

provided $r > 1/2$. Here, the constant C_1 is a consequence of the fact that the power series converges, and it is independent of N.

Summing over n, we have

$$\sum_{|n|\leq N}\left|\sum_{\substack{|m|\leq\infty\\m\neq 0}}\hat{u}_{n+Mm}\right|^2 \leq \sum_{|n|\leq N}C_1 N^{-2r}\sum_{\substack{|m|\leq\infty\\m\neq 0}}|n+Mm|^{2r}|\hat{u}_{n+Mm}|^2$$

$$\leq C_2 N^{-2r}\|u^{(r)}\|^2_{L^2[0,2\pi]}$$

$\square$

We are now in a position to state the error estimate for the discrete approximation.

Theorem 2.6. _For any $u(x) \in W^r_p[0,2\pi]$ with $r > 1/2$, there exists a positive constant C, independent of N, such that_

$$\|u - \mathcal{J}_{2N}u\|_{L^2[0,2\pi]} \leq CN^{-r}\|u^{(r)}\|_{L^2[0,2\pi]}$$

Proof. Lets write the difference between the function and its discrete approximation

$$\|u - \mathcal{J}_{2N}u\|_{L^2[0,2\pi]} = \|(\mathcal{P}_{2N} - \mathcal{J}_{2N})u + u - \mathcal{P}_{2N}u\|_{L^2[0,2\pi]}$$

$$\leq \|(\mathcal{P}_{2N} - \mathcal{J}_{2N})u\|_{L^2[0,2\pi]} + \|u - \mathcal{P}_{2N}u\|_{L^2[0,2\pi]}$$

Thus, the error has two components. The first one, which is the difference between the continuous and discrete expansion coefficients, is the aliasing error, which is bounded in Lemma 2.2. The second, which is the tail of the series, is the truncation error, which is bounded by the result of Theorem 2.2. The desired result follows from these error bounds. $\square$

Theorem above confirms that the approximation errors of the continuous expansion and the discrete expansion are of the same order, as long as $u(x)$ has at least half a derivative. Furthermore, the rate of convergence depends, in both cases, only on the smoothness of the function being approximated.

Chapter 3

Solutions for Burgers' Equation in the Deterministic Version

This chapter is one of the most important, since we will begin with the development of the main objective of this work. We will construct two spectral methods known as Fourier-Galerkin and Fourier-Collocation using the projection and interpolation operators respectively, considering an initial value problem that will be defined using our objective equation presented in (1.3).

For this, to illustrate what we have studied in the Chapter (2), we will consider the function space $H_p^q[0, 2\pi]$ with the norm defined in (2.15) for some q that will be specified later. We will assume that the problems have solutions $u(x, t) \in H_p^q[\mathcal{D}]$ for every $t \in I$, where $\mathcal{D} = [x_L, x_R]$ for some fixed reals x_L, x_R and $I = [0, T]$ with $T > 0$. So, given some real $\alpha \geq 0$ and an initial condition function $u_0(x) \in H_p^q[\mathcal{D}]$, our initial value problem is as follows

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{1}{2}(u^2)_x = \alpha \frac{\partial^2 u}{\partial x}, & 0 < t \leq T, \ x \in I \\ \\ u(x, 0) = u_0(x), & x \in I \end{cases} \tag{3.1}$$

The analytical solution of the previous problem was given in (1.12), and we observe that it is not easy to evaluate it directly. Therefore, it is necessary to choose a suitable method of numerical integration to calculate the integrals involved, since they have exponential behavior that is noticeably affected when the parameter α is very small. However, precision problems can also arise because arithmetic operations can generate considerable errors if they are not performed correctly, and therefore this equation is not a good choice for finding solutions to the problem.

In the next section, we will present the aforementioned spectral methods considering the linear problem obtained by using the transformation given by (1.8), And this will allow us to approximate the solutions of the problem (3.1) More easily and with excellent precision. In addition, this will give us the advantage of having solutions that can be considered exact and use them to compare them with those that will be obtained in the numerical experiments that we will describe in the second section when implementing these methods for the nonlinear problem (3.1), Since we will observe that the precision will be much less because we will need to use numerical methods to solve in the variable t, decreasing with respect to the parameter α.

3.1   Fourier Spectral Methods

In the following sections we will work with the problem obtained by the transformation of the problem (1.8) given by

$$u(x,t) = -2\alpha \frac{\partial_x \varphi(x,t)}{\varphi(x,t)} = -2\alpha \left(\log \varphi(x,t) \right)_x \tag{3.2}$$

with $\alpha > 0$, and φ solves the following initial value problem

$$\begin{cases} \frac{\partial \varphi}{\partial t} = \alpha \varphi_{xx}, \ \ 0 < t \leq T, \ \ x \in I \\[2mm] \varphi(x,0) = \varphi_0(x) = e^{-\int_0^x \frac{u_0(y)}{2\alpha} dy}, \ \ x \in I \end{cases} \tag{3.3}$$

3.1.1   Fourier-Galerkin

In chapter 2 we saw that for a function $\varphi \in H_p^2[0, 2\pi]$ it can be written as

$$\varphi(x,t) = \sum_{|n| \leq \infty} \hat{\varphi}_n(t) \phi_n(x), \ \ \hat{\varphi}_n(t) = \frac{1}{2\pi} \langle \varphi(x,t), \phi_n(x) \rangle \tag{3.4}$$

where $\phi_n(x) = e^{inx}$, and let's assume that this function is the solution of the problem (3.3).

The Fourier-Galerkin method will be constructed using the project operator described in (2.8) to project and force the function φ to satisfy the problem (3.3) on a space that will be defined a continuation.

Let V_N the space of trigonometric polynomials of degree $2N + 1$ given by $V_N = B_N \cap H_p^2[0, 2\pi]$, where $B_N = span\{\phi_n(x) : |n| \leq N\}$. Then for $\varphi \in V_N$ we can obtain its expansion as

$$\varphi_N(x,t) = \sum_{|n| \leq N} \hat{\varphi}_n(t) \phi_n(x), \ \ \hat{\varphi}_n(t) = \frac{1}{2\pi} \langle \varphi(x,t), \phi_n(x) \rangle \tag{3.5}$$

Substituting (3.4), and (3.5) into the equation (3.3) to taking the difference as follows

$$\left[\frac{\partial}{\partial t} \varphi_N - \alpha \frac{\partial^2}{\partial x} \varphi_N \right] - \left[\frac{\partial \varphi}{\partial t} - \alpha \frac{\partial^2 \varphi}{\partial x} \right] = R_N(x,t)$$

where $R_N(x,t)$ is a residual function, and since the second term is exactly zero because φ is the exact solution we have to

$$\frac{\partial}{\partial t} \varphi_N - \alpha \frac{\partial^2}{\partial x} \varphi_N = R_N(x,t)$$

In the Fourier-Galerkin method, it is desired that the remainder R_N belong to the orthogonal space of V_N, that is, for every $\varphi, \phi \in V_N$ such that $\langle \varphi - \mathcal{P}_N \varphi, \phi \rangle = 0$. This is achieved by forcing for each $|n| \leq N$ the following condition

$$\langle R_N, \phi_n \rangle = \left\langle \frac{\partial \varphi_N}{\partial t} - \alpha \frac{\partial^2 \varphi_N}{\partial x^2}, \phi_n \right\rangle = 0, \quad \phi_n \in V_N, \ \ \forall t > 0 \tag{3.6}$$

or equivalently

$$\int_{\mathcal{D}} \frac{\partial}{\partial t} \varphi_N(x,t) \overline{\phi_n(x)} dx = \alpha \int_{\mathcal{D}} \frac{\partial^2}{\partial x^2} \varphi_N(x,t) \overline{\phi_n(x)} dx, \quad \phi_n \in V_N, \ \ \forall t > 0$$

Using the orthogonality $\langle \phi_k, \phi_n \rangle = 2\pi \delta_{kn}$, for each n fixed we have to

$$\left\langle \frac{\partial \varphi_N}{\partial t}, \phi_n \right\rangle = \left\langle \sum_{|k| \leq N} \frac{d\hat{\varphi}_k(t)}{dt} \phi_k, \phi_n \right\rangle = 2\pi \frac{d\hat{\varphi}_n(t)}{dt}$$

Note that the spatial interval is not the desired one, for this we are going to define the following linear transformation from $\mathcal{D}_p = [0, 2\pi]$ to $\mathcal{D} = [x_L, x_R]$ given by $x = Pz + x_L$ to escalate the problem, where $P = \frac{x_R - x_L}{2\pi}$ and $z \in \mathcal{D}_p$. Then, using the chain rule we can rewritten the derivatives as

$$\frac{\partial^2 \varphi_N}{\partial x^2} = P^2 \frac{\partial^2 \varphi_N}{\partial x^2}$$

and using $\frac{\partial^2}{\partial x^2} \phi_n(x) = -P^2 n^2 \phi_n(x)$ we have to

$$\alpha \left\langle \frac{\partial^2 \varphi_N}{\partial x^2}, \phi_n \right\rangle = -2\pi \alpha P^2 n^2 \hat{\varphi}_n(t), \ \ |n| \leq N$$

Therefore, we have the next ODE

$$\frac{d\hat{\varphi}_n(t)}{dt} = -\alpha P^2 n^2 \hat{\varphi}_n(t), \ \ |n| \leq N$$

which is solved using the projection of the initial condition given by

$$\varphi_N(x, 0) = \sum_{|n| \leq N} \hat{\varphi}_n(0) \phi_n(x), \ \ \hat{\varphi}_n(0) = \frac{1}{2\pi} \langle \varphi_0(x), \phi_n(x) \rangle \tag{3.7}$$

We will denote $\lambda_n = \alpha P^2 n^2$, so the exact solution for the above system is given as follows

$$\hat{\varphi}_n(t) = \hat{\varphi}_n(0) e^{-\lambda_n t}, \ \ |n| \leq N$$

Therefore, using (3.5) the solution can be expressed as

$$\varphi_N(x, t) = \sum_{|n| \leq N} \hat{\varphi}_n(0) e^{-\lambda_n t} \phi(x)$$

We can notice that λ_n is actually an eigenvalue of the problem (3.3), which is associated with the eigenvector $\varphi_n(0)$. Therefore, we should note that what we are actually solving is an eigenvalues problem to obtain a representation of the solution in terms of its eigenvectors.

To see this more clearly, we will represent the solution by configuring the following vectors

$$\hat{\varphi}_N(t) = [\hat{\varphi}_{-N}, \hat{\varphi}_{-N+1}(t), \ldots, \hat{\varphi}_N(t)]^T, \quad \hat{\varphi}_N(0) = [\hat{\varphi}_{-N}(0), \hat{\varphi}_{-N+1}(0), \ldots, \hat{\varphi}_N(0)]^T$$

and the matrix

$$\mathcal{L}_N = \begin{bmatrix} \lambda_{-N} & 0 & \ldots & 0 & 0 \\ 0 & \lambda_{-N+1} & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \ldots & \lambda_{N-1} & 0 \\ 0 & 0 & \ldots & 0 & \lambda_N \end{bmatrix},$$

then we can express the solution system as

$$\hat{\varphi}_N(t) = e^{-\mathcal{L}_N t} \hat{\varphi}_N(0),$$

where $e^{-\mathcal{L}_N}$ is the inverse of the exponential matrix of $\mathcal{L}_N$ given by

$$e^{\mathcal{L}_N} = \begin{bmatrix} e^{\lambda_{-N}} & 0 & \ldots & 0 & 0 \\ 0 & e^{\lambda_{-N+1}} & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \ldots & e^{\lambda_{N-1}} & 0 \\ 0 & 0 & \ldots & 0 & e^{\lambda_N} \end{bmatrix}.$$

From the above, we can notice that the solution approaches very fast to zero when t tends to infinity. To show this, we can see that the solution is bounded as follows

$$\|\varphi_N(x,t)\|^2 = \sum_{|n| \leq N} |\hat{\varphi}_n(t)|^2 = \sum_{|n| \leq N} |\hat{\varphi}_n(0)|^2 e^{-2\lambda_n t}$$

$$\leq e^{-2t} \sum_{|n| \leq N} |\hat{\varphi}_n(0)|^2 = e^{-2t} \|\varphi(0)\|^2$$

showing us that the coefficients vanishing with exponential speed, which tells us that the convergence must have the same behavior. We can verify this with the following estimate

$$\|\varphi(x,t) - \varphi_N(x,t)\|^2 = \sum_{|n| > N} |\hat{\varphi}_n(0)|^2 e^{-\lambda_n t} \leq e^{-\lambda_N t} \sum_{|n| > N} |\hat{\varphi}_n(0)|^2$$

$$\leq e^{-\lambda_N t} \sum_{|n| \leq N} |\hat{\varphi}_n(0)|^2 = e^{-\lambda_N t} \|\varphi_N(0)\|^2$$

Therefore, the rate of convergence is exponential, which verifies the theory studied in chapter 2, ensuring an excellent approximation of the solution of the problem (3.1), which is obtained using (1.8) to obtain

$$u_N(x,t) = -2\alpha \frac{\partial_x \varphi_N(x,t)}{\varphi_N(x,t)} = -2\alpha \frac{\displaystyle\sum_{|n|\leq N} in\hat{\varphi}_n(0)e^{-\lambda_n t}\phi_n(x)}{\displaystyle\sum_{|n|\leq N} \hat{\varphi}_n(0)e^{-\lambda_n t}\phi_n(x)} \qquad (3.8)$$

The previous expression is much more practical to calculate solutions to the problem (3.1), we only have to handle the arithmetic operations correctly, and with low approximation orders we will have excellent results.

3.1.2　Fourier-Collocation

The main idea of this method that we will see next is very similar to the Fourier-Galerkin method, except that we will use the interpolation operator described in (2.25) for an odd number of points in the grid. For this, we must use another polynomial space, which we already defined in chapter 2 as $\widetilde{B}_N$ given by

$$\widetilde{B}_N = span\left\{\left(cos(nx),\ \ 0 \leq n \leq \frac{N}{2}\right) \cup \left(sin(nx),\ \ 1 \leq n \leq \frac{N}{2}-1\right)\right\}.$$

Now we will look for a solution to the problem (3.3) in the space given by $S_N = \widetilde{B}_N \cap H_p^2[0, 2\,pi]$, using the discrete expansion for the function φ as follows

$$\mathcal{J}_N\varphi(x,t) = \sum_{|n|\leq \frac{N}{2}} \widetilde{\varphi}_n(t)e^{inx},\ \ \widetilde{\varphi}_n(t) = \sum_{j=0}^{2N} \varphi(x_j,t)e^{-inx_j} \qquad (3.9)$$

where x_j are given by

$$xj = \frac{2\pi j}{2N+1},\ \ j = 0,1,\ldots,2N.$$

Remember that by (2.26) the previous expansion can be written equivalently and also conveniently as

$$\mathcal{J}_N\varphi(x,t) = \sum_{j=0}^{2N} \varphi(x_j,t)\psi_j(x)$$

where $\psi_j(x_i) = \delta_{ij}$ for $i,j = 0,1,\ldots,2N$, and satisfies $\mathcal{J}_N\varphi(x_j,t) = \varphi(x_j,t)$ for each j.

Using the previous expansion in the equation (3.3) we can obtain a residual function as we obtained it in the Fourier-Galerkin method given as follows

$$R_N(x,t) = \frac{\partial \mathcal{J}_N\varphi(x,t)}{\partial t} - \alpha\frac{\partial}{\partial x^2}\mathcal{J}_N\varphi(x,t),$$

and similarly, we must force the orthogonality, for which we must remember by (2.33) that in this case, the discrete product coincides with the continuum, therefore, we have to

$$\langle R_N, \psi_j \rangle_N = \int_I R_N(x,t)\overline{\psi}_j(x)dx = 0, \ \ \text{for} \ j = 0, 1, \dots, 2N.$$

Since $\mathcal{J}_N\varphi(x_j, t) = \varphi(x_j, t)$ for every $j = 0, 1 \dots, 2N$, the orthogonality can be satisfied by solving the following problem

$$\frac{d\mathcal{J}_N\varphi(x_j, t)}{dt} = \alpha \mathcal{J}_N \frac{\partial}{\partial x^2} \mathcal{J}_N\varphi(x_j, t), \ \ j = 0, 1 \dots, 2N \tag{3.10}$$

which is a system of $2N + 1$ ordinary differential equations that can be solved using the initial condition given by

$$\mathcal{J}_N\varphi(x_j, t) = \varphi_0(x_j), , \ \ j = 0, 1 \dots, 2N$$

A convenient way to solve the above problem is to solve for each fixed j the following system of differential equations for the coefficients $\widetilde{\varphi}_n$ given by

$$\frac{d\widetilde{\varphi}_n(t)}{dt} = \alpha n^2 \widetilde{\varphi}_n(t), \ \ |n| \leq 2N$$

which is exactly the same that we obtained in the Fourier-Galerkin method, with the solution given by

$$\widetilde{\varphi}_n(t) = \widetilde{\varphi}_n(0)e^{-\alpha n^2 t}, \ \ |n| \leq 2N$$

Finally, after solving the previous problem for each j we can express the solution with the expansion given by (3.9), which is basically the same solution that we have found using the Fourier-Galerkin method.

When we need to use some numerical method to solve in the variable t, it is better to express the system of differential equations by configuring the following vector

$$\varphi_N(t) = [\varphi(x_0, t), \varphi(x_1, t), \dots, \varphi(x_{2N}, t)]^T$$

and using the differentiation matrix given by (2.32) to obtain

$$\frac{d}{dt}\varphi_N(t) = \alpha D_{2N}^{(2)}\varphi_N(t)$$

In this way we can calculate derivatives directly in real space, being a great advantage that we will discuss in the next section when implementing numerical methods based on the previous representation.

3.2 Numerical Experiments

This section intends to be the most important in the chapter since it will take advantage of everything that has been previously studied and expand in the best possible way the implementation of the spectral methods studied in the previous section. For this, we will focus on the nonlinear problem defined in (3.1) directly, because after understanding these tools you want to have the ability to attack more complex problems such as (1.1) which it is nonlinear, and in general, in most problems, it is not always possible to obtain analytical solutions, and then an option is to approximate them with numerical methods.

According to what was studied in the previous section, we will start solving our problem with each method considering $\alpha > 0$, and later we will study the case when $\alpha = 0$. To find solutions we are going to use Euler's method, either explicit or implicit, to solve in the variable t, and due to the importance of carrying out numerical studies, we will try to give a detailed methodology about each method implemented to construct an algorithm that can be useful in the realization of computational codes. Finally, with the aim of giving a more detailed analysis, we will present some simulations of the results obtained.

3.2.1 Numerical Solutions for Burgers' Equation with Viscosity

Simulations for Fourier-Galerkin Method

Following the ideas in the previous section for the Fourier-Galerkin method, we will assume that the expansion of the solution of (3.1) is as follows

$$u(x,t) = \sum_{|n| \leq \infty} \hat{u}_n(t)\phi_n(x), \quad \hat{u}_n(t) = \frac{1}{2\pi} \langle u(x,t), \phi_n(x) \rangle,$$

and as before $\phi_n(x) = e^{inx}$.

Similarly, we consider the space $V_N = \hat{B}_N \cap H_p^2(\mathcal{D})$ where we will look for its truncated expansion given by

$$u_N(x,t) = \sum_{|n| \leq N} \hat{u}_n(t)\phi_n(x), \quad \hat{u}_n(t) = \frac{1}{2\pi} \langle u(x,t), \phi_n(x) \rangle,$$

which we are going to force to satisfy the following

$$\langle R_N, \phi_n \rangle = \left\langle \frac{\partial u_N}{\partial t} + \frac{1}{2}(u_N^2)_x - \alpha \frac{\partial^2 u_N}{\partial x^2}, \phi_n \right\rangle = 0, \quad \phi_n \in V_N, \ \forall t > 0 \qquad (3.11)$$

or equivalently

$$\int_I \frac{\partial}{\partial t} u_N(x,t)\overline{\phi_n(x)}dx = \alpha \int_I \frac{\partial^2}{\partial x^2} u_N(x,t)\overline{\phi_n(x)}dx - \int_I \frac{1}{2}\frac{\partial}{\partial x}\left[u_N(x,t)\right]^2 \overline{\phi_n(x)}dx$$

39

Since $u_N(x,t)$ and its derivatives are periodic in space, integrating by parts we obtain that

$$\int_I \frac{\partial}{\partial t} u_N(x,t)\overline{\phi_n(x)}dx = -\alpha \int_I \frac{\partial}{\partial x} u_N(x,t)\frac{\partial}{\partial x}\overline{\phi_n(x)}dx + \int_I \frac{1}{2}\left[u_N(x,t)\right]^2 \frac{\partial}{\partial x}\overline{\phi_n(x)}dx$$

Then writing again as an inner product

$$\left\langle \frac{\partial u_N}{\partial t}, \phi_n \right\rangle = \left\langle -\alpha\frac{\partial u_N}{\partial x} + \frac{1}{2}[u_N]^2, \frac{\partial \phi_n}{\partial x} \right\rangle$$

and using the orthogonality $\langle \phi_l, \phi_n \rangle = 2\pi\delta_{ln}$, for each n fixed we have to

$$\left\langle \frac{\partial u_N}{\partial t}, \phi_n \right\rangle = \left\langle \sum_{|l|\leq N} \frac{d\hat{u}_l(t)}{dt}\phi_l, \phi_n \right\rangle = 2\pi\frac{d\hat{u}_n(t)}{dt}$$

and for the other terms

$$\left\langle \frac{\partial u_N}{\partial x} - \frac{1}{2}[u_N]^2, \frac{\partial \phi_n}{\partial x} \right\rangle = \alpha\left\langle -i\sum_{|l|\leq N} l\hat{u}_l(t)\phi_l, in\phi_n \right\rangle$$

$$+ \left\langle \frac{1}{2}\left(\sum_{|p|\leq N}\hat{u}_p(t)\phi_p\right)\left(\sum_{|q|\leq N}\hat{u}_q(t)\phi_q\right), in\phi_n \right\rangle$$

$$= -\alpha n\left\langle \sum_{|l|\leq N} l\hat{u}_l(t)\phi_l, \phi_n \right\rangle - \frac{in}{2}\left\langle \sum_{|p|\leq N}\sum_{|q|\leq N}\hat{u}_p(t)\hat{u}_q(t)\phi_{p+q}, \phi_n \right\rangle$$

$$= -2\pi\alpha n^2\hat{u}_n(t) - in\pi\sum_{p+q=n}\hat{u}_p(t)\hat{u}_q(t), \quad |q|,|p| \leq N.$$

Therefore, using the linear transformation from $\mathcal{D}_p = [0,2\pi]$ to $\mathcal{D} = [x_L, x_R]$ given by $x = Pz + x_L$ to escalate the problem, where $P = \frac{x_R - x_L}{2\pi}$ and $z \in \mathcal{D}_p$, we have the following system of nonlinear ordinary differential equations

$$\frac{d\hat{u}_n(t)}{dt} = -\alpha P^2 n^2\hat{u}_n(t) - \frac{in}{2}P\sum_{p+q=n}\hat{u}_p(t)\hat{u}_q(t), \quad |q|,|p| \leq N, \qquad (3.12)$$

or equivalently

$$\frac{d\hat{u}_n(t)}{dt} = -\alpha P^2 n^2\hat{u}_n(t) - \frac{in}{2}P\sum_{|q|\leq N}\hat{u}_{n-q}(t)\hat{u}_q(t), \quad |n| \leq N,$$

which is solved using the initial condition of the original projected problem on the space V_N, that is,

$$\mathcal{P}_N u_0 = \sum_{|n|\leq N} a_n\phi_n(x), \quad a_n = \frac{1}{2\pi}\langle u_0(x), \phi_n(x)\rangle.$$

This problem can be treated in different ways and, in this work, we will develop it with the following approach. First, let's discard the terms $\hat{u}_q$ such that $|q| > N$ on the right side of the equation above, to get the following

$$\frac{d\hat{u}_n(t)}{dt} = -\alpha P^2 n^2 \hat{u}_n(t) - \frac{in}{2} P \sum_{|q| \leq N} \hat{u}_{n-q}(t)\hat{u}_q(t), \ \ |n-q| \leq N,$$

which is a system of $2N + 1$ equations.

There is a wide variety of numerical methods to solve the above problem, and since this is not linear, it is not easy to justify choosing an appropriate method, since advanced knowledge of numerical analysis is required to investigate its characteristics. The way to choose a candidate method is by investigating its numerical stability, which guarantees that the numerical solution does not explode towards infinity and that it is at least bounded.

In general, explicit methods are not suitable for nonlinear problems and are commonly handled by implicit methods. To show the implementation, and in addition to looking at its most relevant characteristics, we will solve the problem using Euler's implicit method. First, note that when $n = 0$ we have $\hat{u}_0(t) = \hat{u}_0(0)$, and that the nonlinear term involves the term $\hat{u}_n$ only when $q = 0$. So, using the integrating factor $e^{\lambda_n t}$ with $\lambda_n = \alpha P^2 n^2 + \frac{in}{2} P\hat{u}_0(0)$ we can obtain the following formulation

$$\frac{d\left[e^{\lambda_n t}\right]\hat{u}_n(t)}{dt} = -\frac{in}{2}Pe^{\lambda_n t}\left[\hat{u}_0(0)\hat{u}_n(t) + \sum_{\substack{|q| \leq N \\ q \neq 0,n}} \hat{u}_{n-q}(t)\hat{u}_q(t)\right],$$

and its solution using implicit Euler

$$\hat{u}_n^{j+1} = e^{-\lambda_n \Delta t}\left[\hat{u}_n^j - \Delta t\frac{in}{2}P\sum_{\substack{|q| \leq N \\ q \neq 0,n}} \hat{u}_{n-q}^j\hat{u}_q^j\right] - \Delta t\frac{in}{2}P\hat{u}_0(0)\hat{u}_n^{j+1},$$

and then solving for the term $\hat{u}_n^{j+1}$ to obtain

$$\hat{u}_n^{j+1} = \frac{e^{-\lambda_n \Delta t}}{1 + \Delta t\frac{in}{2}P\hat{u}_0(0)}\left[\hat{u}_n^j - \Delta t\frac{in}{2}P\sum_{\substack{|q| \leq N \\ q \neq 0,n}} \hat{u}_{n-q}^j\hat{u}_q^j\right].$$

Furthermore, using this formulation repeatedly for each j gives us

$$\hat{u}_n^{j+1} = \left[\frac{e^{-\lambda_n \Delta t}}{1 + \Delta t\frac{in}{2}P\hat{u}_0(0)}\right]^{j+1}\hat{u}_n^0 - \Delta t\frac{in}{2}P\sum_{k=1}^{j+1}\left[\frac{e^{-\lambda_n \Delta t}}{1 + \Delta t\frac{in}{2}P\hat{u}_0(0)}\right]^k \sum_{\substack{|q| \leq N \\ q \neq 0,n}} \hat{u}_{n-q}^{j+1-k}\hat{u}_q^{j+1-k}$$

$$(3.13)$$

41

$$\hat{u}_N^{j+1} = \mathcal{L}_N^{j+1}\hat{u}_N^0 - \Delta t \frac{in}{2} P \sum_{k=1}^{j+1} \mathcal{L}_N^k B_N(\hat{u}_N^{j+1-k})$$

Now, we will proceed to describe the steps necessary to implement the method based on the above. The simulations we are going to show were based on the formulation (3.13), taking advantage of the fast Fourier transformation to calculate the coefficients $\hat{u}_n$, since it allows us to go from physical space to the space of Fourier and conversely. Also, it is worth mentioning that the codes used in this work can be found at https://github.com/alanmatzumiya/Maestria.git.

To carry out these numerical experiments it was necessary to follow the following steps:

1. Set the intervals to consider, $[x_L, x_R]$ for physical space and $[t_0, t_f]$ for time, and give an initial condition function u_0. It is also necessary to choose the value of N, and another for the parameter α.

2. We proceed to calculate the points in the spatial grid where we want to obtain the solution, and for this, we define the points $z_i \in [0, 2\pi]$ as

$$z_i = \frac{2\pi i}{N}, \quad i = 0, 1, \ldots, N$$

which are used to obtain the points $x_i \in [x_L, x_R]$

$$x_i = P z_i + x_L, \quad P = \frac{x_R - x_L}{2\pi}$$

To obtain the points $t_j \in [t_0, t_f]$, we define the sequence $j = 0, 1, \ldots, M$, for some positive M, and calculate the following

$$t_j = t_0 + j\Delta t, \quad \Delta t = \frac{t_f - t_0}{M}$$

The values $|n| = 0, 1, \ldots, N$ must be established to calculate in, n^2, which are required to approximate the derivatives.

3. Calculate the coefficients $\hat{u}_0$ of the initial condition function u_0, using the fast Fourier transformation based on the already fixed spatial grid. So, it is possible to start the recursive rule given by (3.13) for every t_j.

4. Finally, the approximation obtained is evaluated using the equation given by

$$u_N(x, t_j) = \sum_{|n| \leq N} \hat{u}_n(t_j) e^{inx}.$$

For the numerical study, we will establish the following initial condition

$$u_0(x) = e^{0.05x^2}, \quad x \in [x_L, x_R] \tag{3.14}$$

To compare results, we will use the approximation given by (3.8) as the exact solution. In the figure 3.1 shows the maximum distance over every $t \in [0, 100]$ between the exact solution and its approximations given by (3.13) for $N = 2^m$, $m = 4, \ldots, 12$, $\Delta t = 1.0 \times 10^{-5}$, and different values of α. Furthermore, in Tables 3.1 and 3.2, we can see the numerical values of these distances for different configurations of N and Δt. Similarly, in Tables 3.3 and 3.4 but for $\alpha = 0.005$.
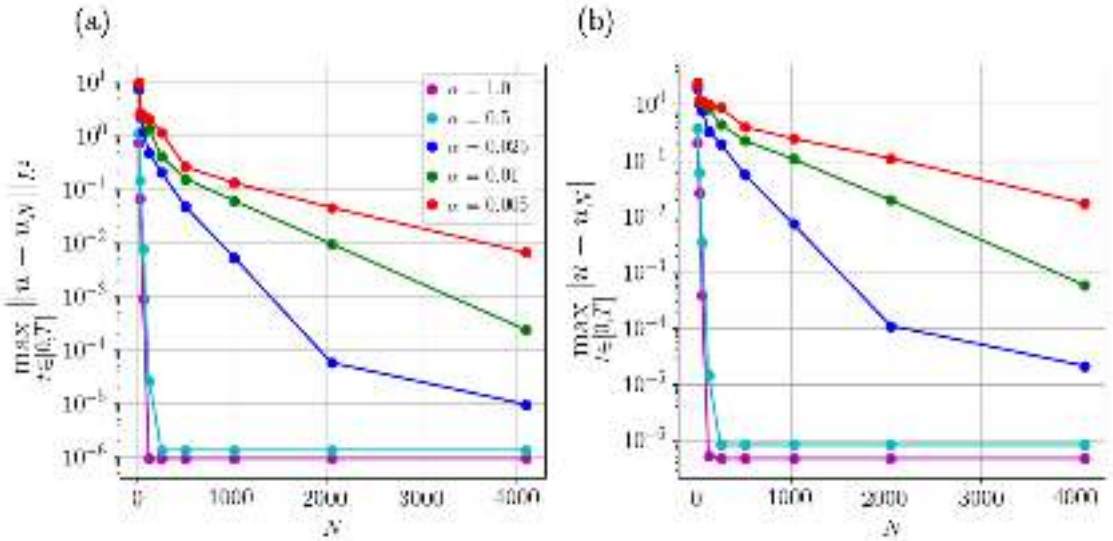

Figure 3.1: (a) L^2-norm between the exact solution and its approximations using Galerkin method. (b) Max norm between the exact solution and its approximations.

Figure 3.2: Numerical solution for (3.1) using (3.13) with $\alpha = 1.0$, $N = 2048$, and $\Delta t = 1.0 \times 10^{-5}$.
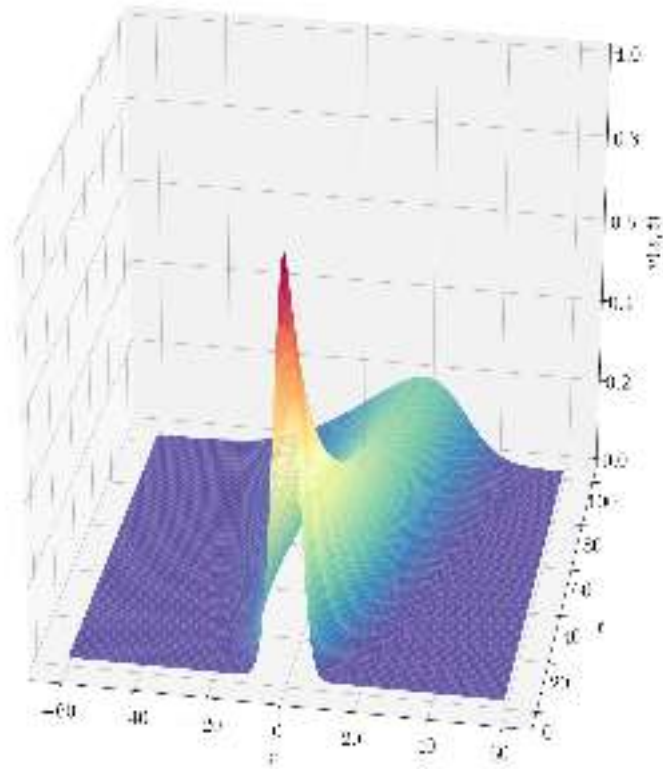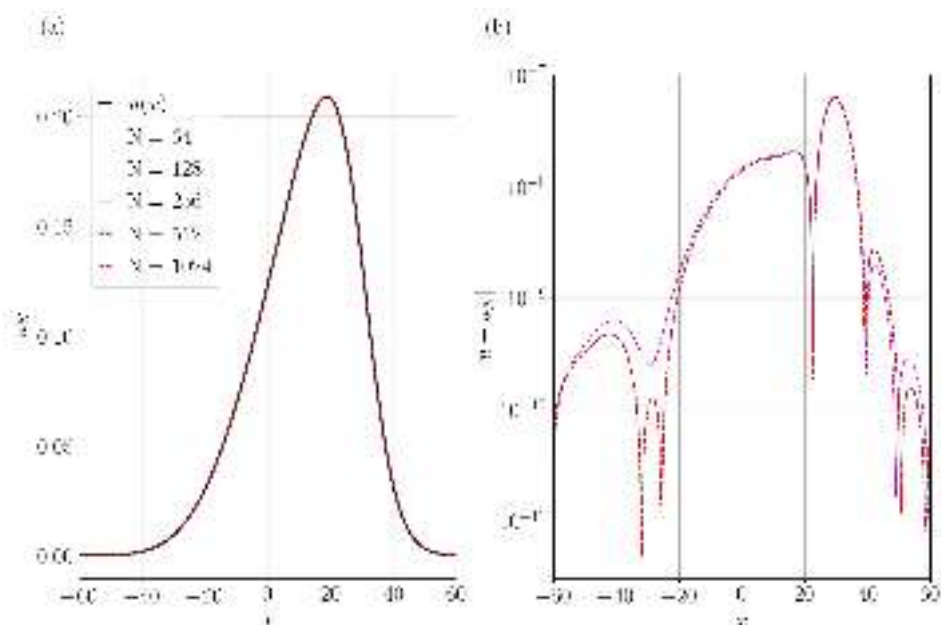

Figure 3.3: Numerical solution for (3.1) using (3.13) at the time $T = 100$ with $\alpha = 1.0$, and $\Delta t = 1.0 \times 10^{-5}$. (b) Point-wise error of approximation

Approximation	Error			
N	$\Delta t = 1 \times 10^{-2}$	$\Delta t = 1 \times 10^{-3}$	$\Delta t = 1 \times 10^{-4}$	$\Delta t = 1 \times 10^{-5}$
16	0.72504	0.72504	0.72504	0.72504
32	6.90249×10^{-2}	6.88052×10^{-2}	6.87838×10^{-2}	6.87816×10^{-2}
64	1.23827×10^{-3}	8.85367×10^{-4}	8.80521×10^{-4}	8.80410×10^{-4}
128	9.43454×10^{-4}	9.41793×10^{-5}	9.41148×10^{-6}	9.41827×10^{-7}
256	9.43454×10^{-4}	9.41793×10^{-5}	9.41109×10^{-6}	9.36411×10^{-7}
512	9.43454×10^{-4}	9.41793×10^{-5}	9.41109×10^{-6}	9.36411×10^{-7}
1024	$*$	9.41793×10^{-5}	9.41109×10^{-6}	9.36411×10^{-7}
2048	$*$	$*$	9.41109×10^{-6}	9.36411×10^{-7}

Table 3.1: Error using L^2-norm with $\alpha = 1.0$

Approximation Max	Error			
N	$\Delta t = 1 \times 10^{-2}$	$\Delta t = 1 \times 10^{-3}$	$\Delta t = 1 \times 10^{-4}$	$\Delta t = 1 \times 10^{-5}$
16	0.203363	0.203333	0.203331	0.20333
32	2.64192×10^{-2}	2.6248×10^{-2}	2.62491×10^{-2}	2.62492×10^{-2}
64	6.93001×10^{-4}	4.11641×10^{-4}	3.85563×10^{-4}	3.82972×10^{-4}
128	4.74934×10^{-4}	4.73649×10^{-5}	4.74295×10^{-6}	5.16105×10^{-7}
256	4.74936×10^{-4}	4.7368×10^{-5}	4.72569×10^{-6}	4.64922×10^{-7}
512	4.74936×10^{-4}	4.7368×10^{-5}	4.72569×10^{-6}	4.64922×10^{-4}
1024	$*$	4.7368×10^{-5}	4.72569×10^{-6}	4.64922×10^{-7}
2048	$*$	$*$	4.72569×10^{-6}	4.64922×10^{-7}

Table 3.2: Error using Max norm with $\alpha = 1.0$

Figure 3.4: Numerical solution for (3.1) using (3.13) with $\alpha = 0.005$, $N = 2048$, and $\Delta t = 1.0 \times 10^{-5}$.
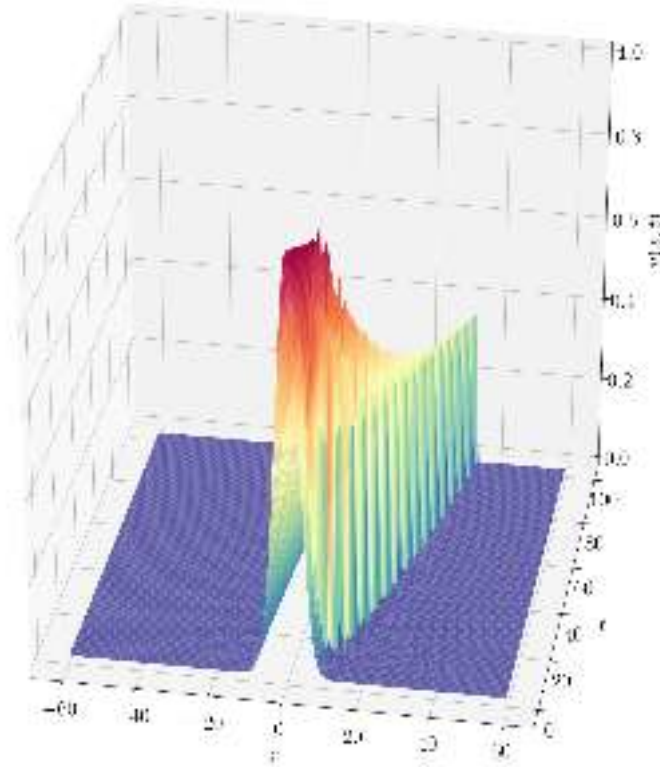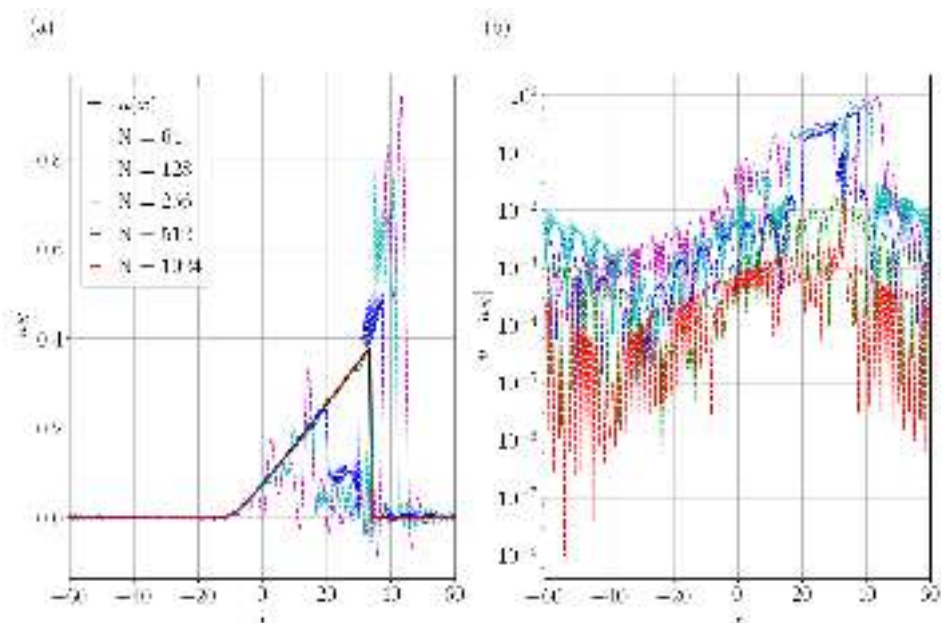

Figure 3.5: Numerical solution for (3.1) using (3.13) at the time $T = 100$ with $\alpha = 1.0$, and $\Delta t = 1.0 \times 10^{-5}$. (b) Point-wise error of approximation

Approximation	Error			
N	$\Delta t = 1 \times 10^{-2}$	$\Delta t = 1 \times 10^{-3}$	$\Delta t = 1 \times 10^{-4}$	$\Delta t = 1 \times 10^{-5}$
16	9.95328	9.91901	9.91597	9.91567
32	2.72607	2.70558	2.70347	2.70326
64	2.50343	2.45988	2.45543	2.45497
128	2.16142	2.06992	2.05918	2.05795
256	1.3658	1.19385	1.17602	1.17412
512	0.339826	0.265843	0.262164	0.261805
1024	0.161405	0.133743	0.131882	0.131699
2048	6.50292 $\times 10^{-2}$	4.70602 $\times 10^{-2}$	4.57371 $\times 10^{-2}$	4.56090 $\times 10^{-2}$
4096	*	7.26917 $\times 10^{-3}$	6.64157 $\times 10^{-3}$	6.60753 $\times 10^{-3}$

Table 3.3: Error using L^2-norm with $\alpha = 0.005$

Approximation Max	Error			
N	$\Delta t = 1 \times 10^{-2}$	$\Delta t = 1 \times 10^{-3}$	$\Delta t = 1 \times 10^{-4}$	$\Delta t = 1 \times 10^{-5}$
16	2.50002	2.48992	2.48891	2.48881
32	1.21263	1.20544	1.2047	1.20463
64	1.21269	1.17736	1.17517	1.17495
128	1.10164	1.03493	1.03093	1.03048
256	0.954369	0.881472	0.873392	0.87259
512	0.665071	0.418735	0.398664	0.396931
1024	0.241841	0.244188	0.244437	0.244461
2048	0.133067	0.104675	0.109151	0.109596
4096	*	2.37273 $\times 10^{-2}$	1.75687 $\times 10^{-2}$	1.69531 $\times 10^{-2}$

Table 3.4: Error using Max norm with $\alpha = 0.005$

Simulations for Fourier-Collocation Method

Following the same ideas from the previous section for Fourier-Collocation, we will seek solutions in the space given by $S_N = \widetilde{B}_N \cap H_p^2[0, 2\pi]$, and using the discrete expansion to the problem (3.1) as follows

$$\mathcal{J}_N u(x,t) = \sum_{|n| \leq \frac{N}{2}} \widetilde{u}_n(t) e^{inx}, \;\; \widetilde{u}_n(t) = \sum_{j=0}^{2N} u(x_j, t) e^{-inx_j}$$

or equivalently

$$\mathcal{J}_N u(x,t) = \sum_{j=0}^{2N} u(x_j, t) \psi_j(x)$$

where x_j are given by

$$x_j = \frac{2\pi j}{2N+1}, \;\; j = 0, 1, \ldots, 2N.$$

For this case, the Fourier-Collocation method will be given by the following problem

$$R_N(x_j, t) = \frac{\partial u_N}{\partial t}(x_j, t) - \frac{\partial^2 u_N}{\partial x^2}(x_j, t) + \frac{1}{2} \frac{\partial [u_N]^2}{\partial x}(x_j, t) = 0 \qquad (3.15)$$

which is a system of $2N + 1$ ordinary differential equations that can be solved using the initial condition given by

$$\mathcal{J}_N u(x_j, t) = u_0(x_j), , \;\; j = 0, 1 \ldots, 2N$$

By Setting a vector as follows

$$u_N(t) = (u_N(x_0, t), u_N(x_1, t), \ldots, u_N(x_{2N}, t))^T,$$

and the above system of ordinary differential equations can be written as follows

$$\frac{du_N(t)}{dt} = \alpha D_N^2 u_N(t) - \frac{1}{2} D_N u_N^2(t)$$

where D_N is the matrix given by (2.32), that represents discrete Fourier differentiation, which can be rewrite as

$$D_N = C^{-1} \Lambda_N^2 C \qquad (3.16)$$

where $\Lambda_N = diag\{ik\}_{|k| \leq N}$, C represents the discrete Fourier transform and C^{-1} the inverse.

Note that the situation here is very different compared to the systems obtained with Fourier-Galerkin. The representation of the non-linear term is more practical to handle with implicit methods, for example, an implicit approach is as follows

$$u_N(t_{i+1}) = C^{-1}\Lambda_N C u_N(t_i) - \frac{\Delta t}{2} D_N u_N^2(t_{i+1})$$

It is possible to formulate the above problem as

$$W(u_N(t_{i+1})) = C^{-1}\Lambda_N C u_N(t_i) - u_N(t_{i+1}) - \frac{\Delta t}{2} D_N u_N^2(t_{i+1}),$$

and then we must find $u_N(t_{i+1})$ that satisfies $W(u_N(t_{i+1})) = 0$, for example, using the Newton-Raphson method. However, this formulation may require many calculations because the differentiation matrix increases as a function of N, and therefore the number of operations. Here we are going to develop a numerical solution that is more practical to implement, implicitly approaching as we did with Fourier-Galerkin on the linear term as follows

$$\left[I_N + \frac{\Delta t}{2} u_0 \Lambda_N\right] u_N(t_{i+1}) = e^{-\Delta t \Lambda_N^0} \left[u_N(t_i) - \frac{\Delta t}{2} D_N u_N^2(t_i)\right]$$

where $\Lambda_N^0 = diag\{\alpha k^2 + \frac{ik}{2}u_0\}_{|k|\leq N}$, and solving for $u_N(t_{i+1})$ gives us

$$u_N(t_{i+1}) = \left[I_N + \frac{\Delta t}{2} u_0 \Lambda_N\right]^{-1} e^{-\Delta t \Lambda_N^0} \left[u_N(t_i) - \frac{\Delta t}{2} D_N u_N^2(t_i)\right] \qquad (3.17)$$

which is very similar to 3.13, except for the nonlinear term.

This formulation will be used for its implementation, and we will present some numerical results that were obtained using the same information that was used with Fourier-Galerkin, the difference is that the calculations are performed in real space using the differentiation matrix D_N, and in this case, we have the advantage of approximating the second derivative with $D_N^2 = D_N \cdot D_N$.

The description of the following results is as follows. In the figure 3.6 shows the maximum distance over every $t \in [0, 100]$ between the exact solution and its approximations given by (3.17) for $N = 2^m$, $m = 4, \ldots, 12$, $\Delta t = 1.0 \times 10^{-5}$, and different values of α. Furthermore, in Tables 3.5 and 3.6, we can see the numerical values of these distances for different configurations of N and Δt. Similarly, in Tables 3.7 and 3.8 but for $\alpha = 0.005$.
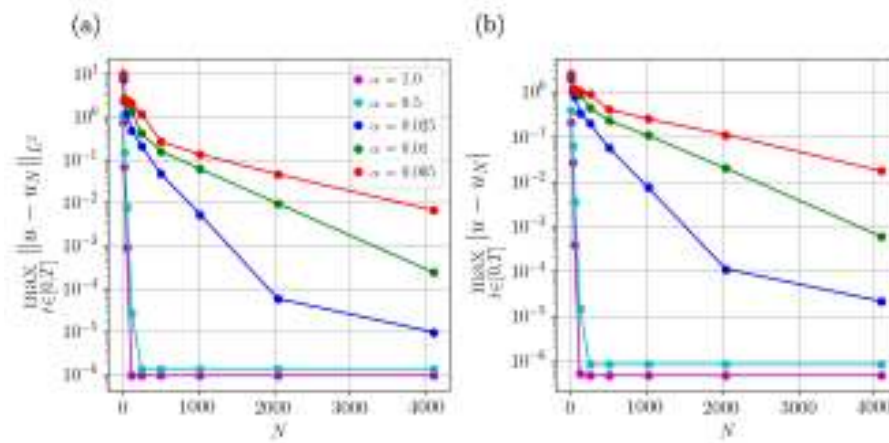
Figure 3.6: (a) L^2-norm between the exact solution and its approximations using Collocation method. (b) Max norm between the exact solution and its approximations.

Figure 3.7: Numerical solution for (3.1) using (3.17) with $\alpha = 1.0$, $N = 2048$, and $\Delta t = 1.0 \times 10^{-5}$.
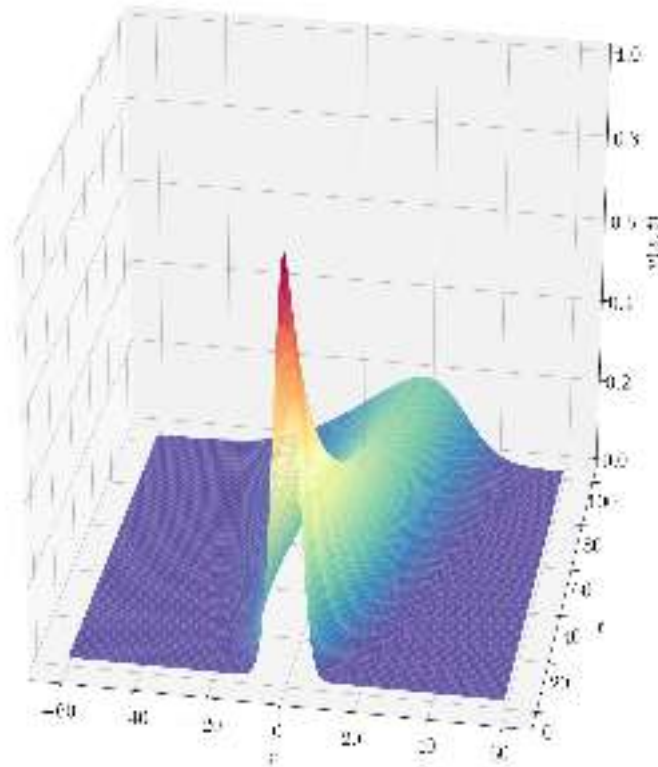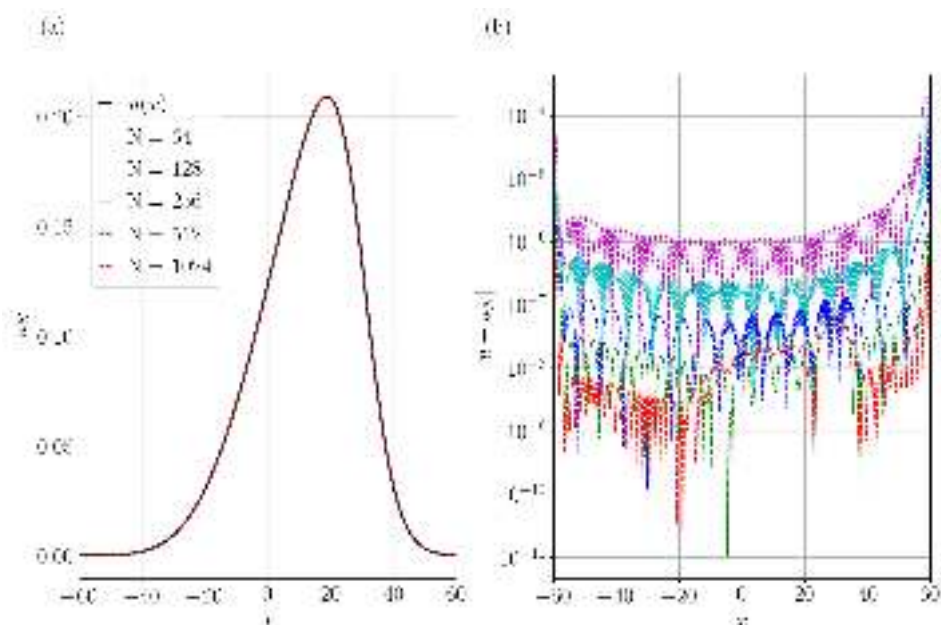

Figure 3.8: Numerical solution for (3.1) using (3.17) at the time $T = 100$ with $\alpha = 1.0$, and $\Delta t = 1.0 \times 10^{-5}$. (b) Point-wise error of approximation

51

Approximation	Error			
N	$\Delta t = 1 \times 10^{-2}$	$\Delta t = 1 \times 10^{-3}$	$\Delta t = 1 \times 10^{-4}$	$\Delta t = 1 \times 10^{-5}$
16	0.721112	0.721112	0.721112	0.721112
32	4.71797×10^{-2}	4.72892×10^{-2}	4.73004×10^{-2}	4.73015×10^{-2}
64	1.17954×10^{-3}	7.35344×10^{-4}	7.27561×10^{-4}	7.27283×10^{-4}
128	9.43454×10^{-4}	1.75152×10^{-4}	1.74583×10^{-4}	1.74574×10^{-4}
256	9.43454×10^{-4}	1.15509×10^{-4}	1.14669×10^{-4}	1.14659×10^{-4}
512	9.43454×10^{-4}	9.41793×10^{-5}	7.78847×10^{-5}	7.78707×10^{-5}
1024	$*$	9.41793×10^{-5}	5.32213×10^{-5}	5.32019×10^{-5}
2048	$*$	$*$	3.56779×10^{-5}	3.56498×10^{-5}
4096	$*$	$*$	2.24122×10^{-5}	$*$

Table 3.5: Error using L^2-norm with $\alpha = 1.0$

Approximation Max	Error			
N	$\Delta t = 1 \times 10^{-2}$	$\Delta t = 1 \times 10^{-3}$	$\Delta t = 1 \times 10^{-4}$	$\Delta t = 1 \times 10^{-5}$
16	0.317617	0.317617	0.317617	0.317617
32	1.95279×10^{-2}	1.96812×10^{-2}	1.96965×10^{-2}	1.96981×10^{-2}
64	6.21793×10^{-4}	2.9813×10^{-4}	2.80086×10^{-4}	2.78934×10^{-4}
128	4.74952×10^{-4}	1.64746×10^{-4}	1.6473×10^{-4}	1.64728×10^{-4}
256	4.74936×10^{-4}	1.52482×10^{-4}	1.52467×10^{-4}	1.52465×10^{-4}
512	4.74936×10^{-4}	1.47249×10^{-4}	1.47234×10^{-4}	1.47232×10^{-4}
1024	$*$	1.45032×10^{-4}	1.45017×10^{-4}	1.45016×10^{-4}
2048	$*$	$*$	1.43941×10^{-4}	1.4394×10^{-4}
4096	$*$	$*$	1.43411×10^{-4}	$*$

Table 3.6: Error using Max norm with $\alpha = 1.0$

Figure 3.9: Numerical solution for (3.1) using (3.17) with $\alpha = 0.005$, $N = 2048$, and $\Delta t = 1.0 \times 10^{-5}$.
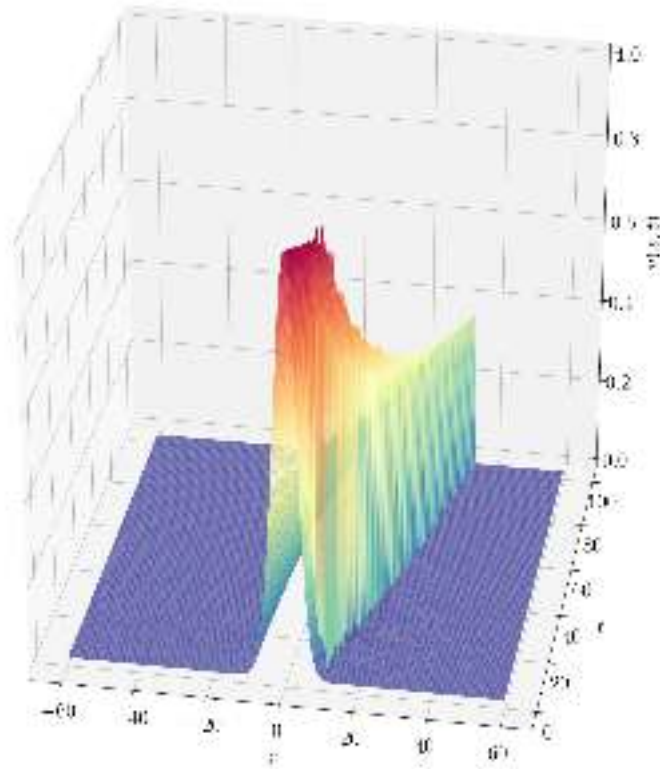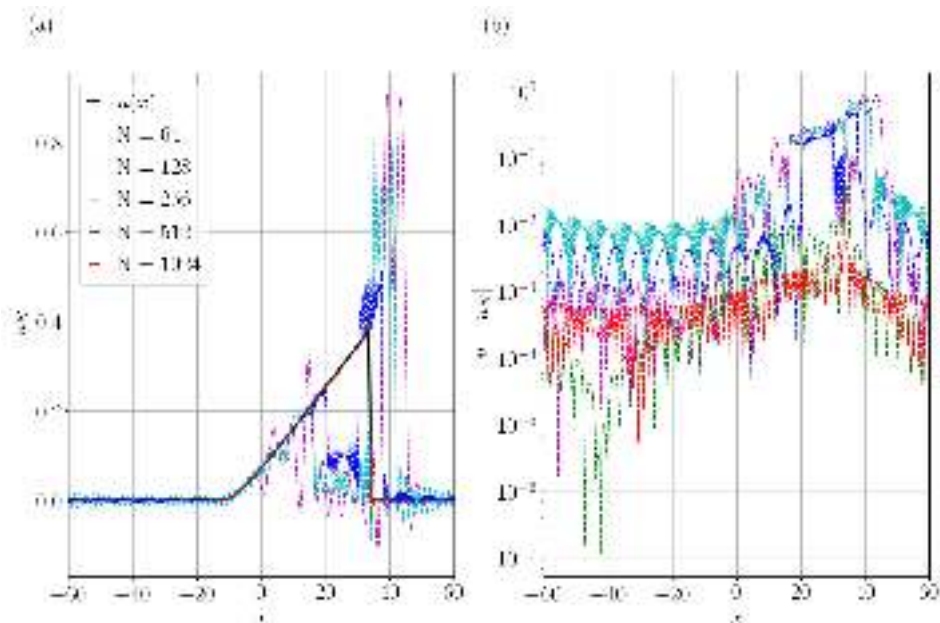

Figure 3.10: Numerical solution for (3.1) using (3.17) at the time $T = 100$ with $\alpha = 0.005$, and $\Delta t = 1.0 \times 10^{-5}$. (b) Point-wise error of approximation.

Approximation	Error			
N	$\Delta t = 1 \times 10^{-2}$	$\Delta t = 1 \times 10^{-3}$	$\Delta t = 1 \times 10^{-4}$	$\Delta t = 1 \times 10^{-5}$
16	1.36189	1.35883	1.35852	1.35849
32	2.67506	2.65305	2.65078	2.65055
64	2.50365	2.45855	2.45432	2.45387
128	2.15795	2.0632	2.05589	2.05497
256	1.362	1.18393	1.16697	1.16532
512	0.350775	0.304595	0.300865	0.300499
1024	0.168462	0.140332	0.13803	0.137804
2048	6.56161 $\times 10^{-2}$	4.63808 $\times 10^{-2}$	4.49226 $\times 10^{-2}$	4.47813 $\times 10^{-2}$
4096	*	7.66246 $\times 10^{-3}$	6.9909 $\times 10^{-3}$	6.9909 $\times 10^{-3}$

Table 3.7: Error using L^2-norm with $\alpha = 0.005$

Approximation Max	Error			
N	$\Delta t = 1 \times 10^{-2}$	$\Delta t = 1 \times 10^{-3}$	$\Delta t = 1 \times 10^{-4}$	$\Delta t = 1 \times 10^{-5}$
16	0.695784	0.695659	0.695646	0.695645
32	1.20278	1.19418	1.19329	1.1932
64	1.22454	1.18903	1.18507	1.18467
128	1.11999	1.0238	1.01754	1.01701
256	0.927954	0.877058	0.872508	0.87201
512	0.664133	0.415288	0.39714	0.395563
1024	0.247742	0.259451	0.260605	0.26072
2048	0.126824	0.103297	0.107102	0.10748
4096	*	2.04624 $\times 10^{-2}$	1.76513 $\times 10^{-2}$	1.76513 $\times 10^{-2}$

Table 3.8: Error using Max norm with $\alpha = 0.005$

3.2.2   Numerical Solutions for Burgers' Equation without Viscosity

To finish the section, we will study the problem (3.1) without viscosity, that is, when $\alpha = 0$. So the problem is the following

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{1}{2}(u^2)_x = 0, \ \ 0 < t \leq T, \ \ x \in \mathcal{D} \\[2mm] u(x,0) = u_0(x), \ \ x \in \mathcal{D} \end{cases} \tag{3.18}$$

This problem, which seems much simpler, turns out to be very interesting, because it presents relevant characteristics regarding the physical interpretation of the problem in general, that is, the case with non-zero viscosity.

The previous equation interprets the conservation of energy and is considered as a non-linear conservation law. We can understand this if we consider the function u as the speed of a fluid that conserves its energy with a flow density given by $f(u) = \frac{u^2}{2}$.

The above can be shown considering that $u \in H_p^1[\mathcal{D}]$, and multiplying by u and integrating both sides of the equation on the domain $\mathcal{D}$ to prove that $\|u\|$ does not change over time, that is,

$$\frac{1}{2}\frac{d}{dt}\int_0^{2\pi} u^2(x,t)dx = -\int_0^{2\pi} u^2(x,t)\frac{\partial u(x,t)}{\partial x}dx = -\frac{1}{3}u^3(x,t)\Big|_0^{2\pi}.$$

and because u is periodic, we have to

$$\frac{1}{2}\frac{d}{dt}\|u(x,t)\|^2 = 0.$$

This means that the energy is conserved with the same initial energy since we must bear in mind that this represents the kinetic energy of a fluid that travels with a velocity given by u_0.

In any problem, these characteristics are of utmost importance, this is because we must consider the behavior of the solutions when they must be approached with some numerical method. In the literature, when a solution remains bounded in time, the problem is said to be well defined. In the analysis context, this assures us the existence of a continuously differentiable solution u, and that besides being unique, it is possible to approximate it.

However, approximating the solutions to this problem using spectral methods can be more complicated if its characteristics are not well understood. Next, we will see that under a certain condition it is possible to find a single analytical solution and that otherwise, it may lose its uniqueness when developing discontinuities.

First, let's define the curves $x = x(t)$ that start at a point $x_0 \in \mathbb{R}$, and satisfy the following problem

$$\begin{cases} x'(t) = u(x(t), t), \;\; t > 0, \\ \\ x(0) = x_0. \end{cases}$$

The solutions for each x_0 are known as characteristic curves, which pass through the solution $u = u(x(t), t)$. It is well known in the theory of differential equations that when $u(x, t)$ is locally Lipschitz in the variable x, and continues in the variable t, the above equation admits a single solution $x(t)$ for each $x_0 \in \mathbb{R}$. So, assuming the above we have that for a solution $x(t)$ corresponding to a fixed $x_0 \in \mathbb{R}$, it satisfies the following

$$\frac{d}{dt}[u(x(t), t)] = x'(t)u_x(x(t), t) + u_t(x(t), t)$$
$$= u(x(t), t)u_x(x(t), t) - u(x(t), t)u_x(x(t), t) = 0$$

which tells us that the function $u(x(t), t)$ is independent of the variable t, remaining constant along the characteristic curve. Therefore, we have to

$$u(x(t), t) = u(x(0), 0) = u_0(x_0)$$

Furthermore, we have that the solution $x(t)$ will be given by the following curve

$$x(t) = x_0 + u_0(x_0)t, \;\; t > 0.$$

which allows us to write the solution $u(x, t)$ as follows

$$u(x, t) = u_0(x_0), \;\; x_0 = x - u_0(x_0)t \tag{3.19}$$

Note that the Lipschitz condition is necessary for the uniqueness of the above solution, since, conversely, if we set two different starting points x_0, x_1 such that $x_0 < x_1$, then its curves characteristics intersect for some t, that is,

$$x_0 + u_0(x_0)t = x_1 + u_0(x_1)t,$$

and using the mean value theorem we have that for some $c \in (x_0, x_1)$ the time t is given by

$$t = \frac{x_1 - x_0}{u_0(x_0) - u_0(x_1)} = -\frac{1}{u_0'(c)},$$

The above tells us that when these two curves intersect, the solution $u(x, t)$ cannot be continuous in the time t given by the previous equation. This is because $u_0(x_0) \neq u_0(x_1)$, and since the solution is constant over time we would have that $u(x, t) = u_0(x_0) = u_0(x_1)$, which is impossible.

Therefore, the time t for which two curves intersect represents a discontinuity, and that can be calculated by

$$Tc = \min_{x \in \mathbb{R}} \left[\frac{-1}{u_0'(x)} \right] \tag{3.20}$$

and we can observe that the continuity of the solution u is assured if $T_c < 0$, which depends only on the initial condition u_0.

This type of information allows us to know the criteria that must be considered when implementing numerical methods, but also the physical interpretation of the problem can be useful. For example, the solution to this problem can be considered to simulate the evolution of the profile of a sea wave that deforms as it approaches the coast until it breaks. But when this occurs, the deformation of the wave stops precisely at the time T_c.

We must consider that the problem without viscosity supposes that the fluid behaves in an ideal or perfect way, traveling as if they were separate sheets without rubbing. But in real life, fluids always have a certain degree of viscosity, that is, the fluid sheets can rub and cause energy dissipation, and for these cases, we could consider the problem with sufficiently small values of α. So a question that naturally arises is what the behavior of the solutions looks like when α approaches zero.

In chapter 1 we obtained the solution of the problem (3.1), which is given by (1.12), and from this equation we can see that the solutions are infinitely differentiable for any value of $\alpha > 0$. Instinctively, we can notice that the solution approaches the solution of the equation without viscosity when α approaches zero. In fact, the solution obtained as a limit when α approaches zero is known as the entropy solution, which has been studied in [33], [34] and in [32] has been proved that this solution is unique.

In order to illustrate what we have previously discussed, we will consider the problem (3.18) with the following initial condition function

$$u_0(x) = e^{-0.005x^2}, \quad x \in \mathcal{D} \tag{3.21}$$

where $\mathcal{D} = [x_L, x_R]$, and considering the interval $I = [0, T_c]$ for a value of T_c given by (3.20).

In the following simulations, we will use the Fourier-Galerkin method given in (3.13) to obtain approximations with different values of α, and we will see how they approximate the exact solution of the problem (3.21) which was given in (3.19).
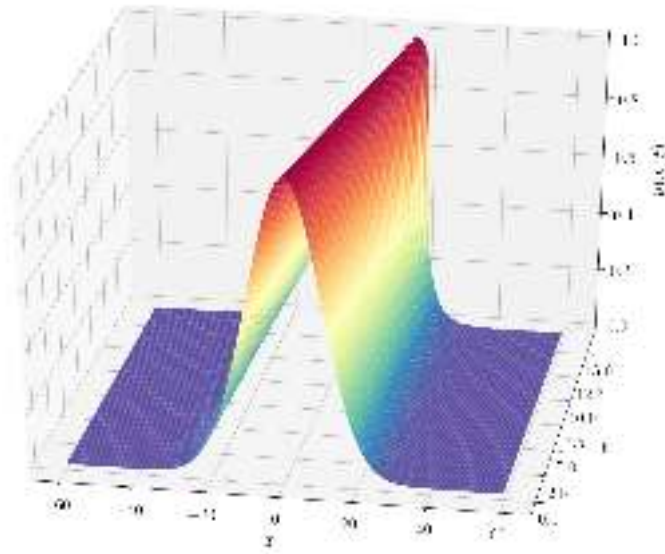
Figure 3.11: Numerical solution for (3.1) using (3.13) with $\alpha = 1.0 \times 10^{-5}$, $N = 256$, and $\Delta t = 1.0 \times 10^{-3}$.

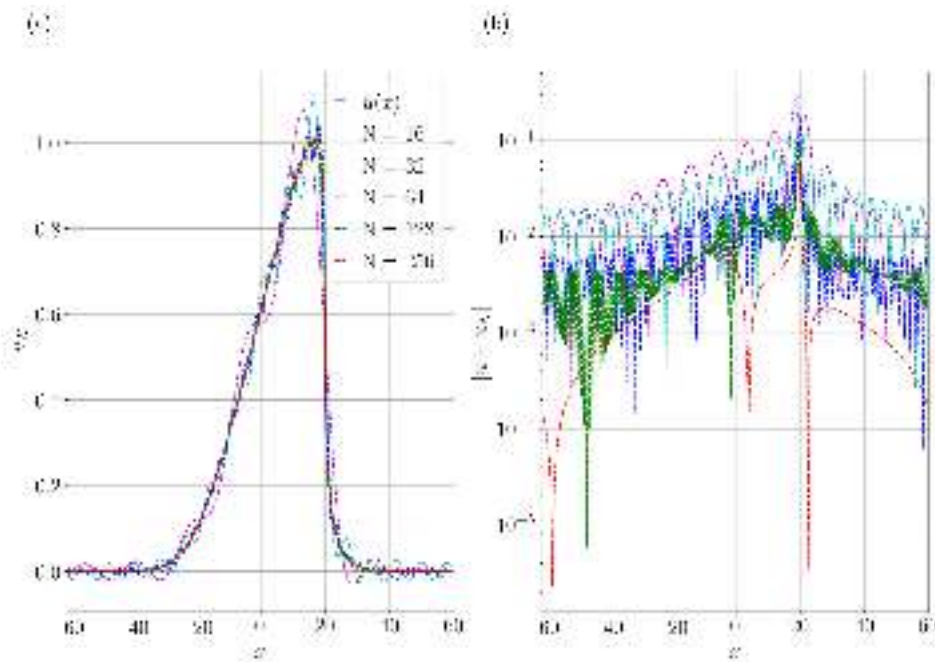Figure 3.12: Numerical solution for (3.1) using (3.13) at the time T_c with $\alpha = 1.0 \times 10^{-5}$, and $\Delta t = 1.0 \times 10^{-3}$. (b) Point-wise error of approximation.
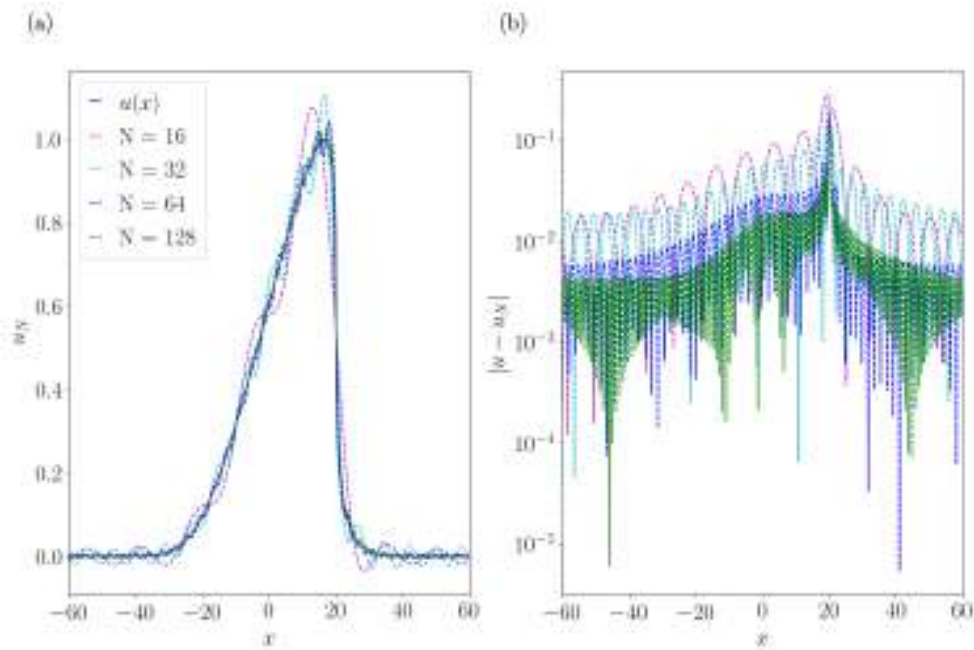
Figure 3.13: (a) Exact solution for (3.1), and its approximations using (3.13) at the time Tc with initial condition $u_0(x) = e^{-0.005x^2}$, $x \in [-60, 60]$. (b) Pointwise error of approximation.
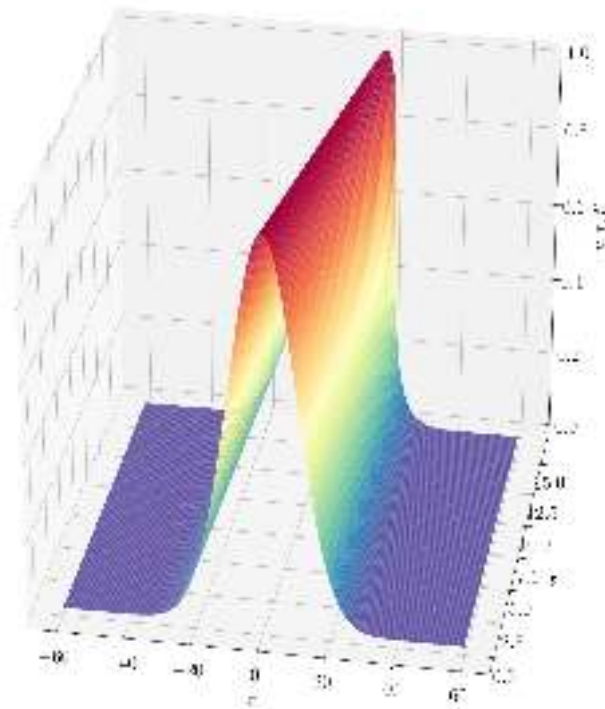

Figure 3.14: Numerical approximation for 3.1 using (3.13) with $N = 512$, $u_0(x) = e^{-0.005x^2}$, $x \in [-60, 60]$, and $t \in [0, Tc]$.

Approximation	Distance		
N	$\Delta t = 1 \times 10^{-2}$	$\Delta t = 1 \times 10^{-3}$	$\Delta t = 1 \times 10^{-4}$
16	0.285531	0.285732	0.285752
32	0.222737	0.223260	0.223312
64	0.160385	0.162782	0.163025
128	0.129297	0.133322	0.133733
256	0.083291	0.091449	0.092320

Table 3.9: Distance between exact solution for (3.18) and the approximation for (3.1) with $\alpha = 1.0 \times 10^{-5}$.
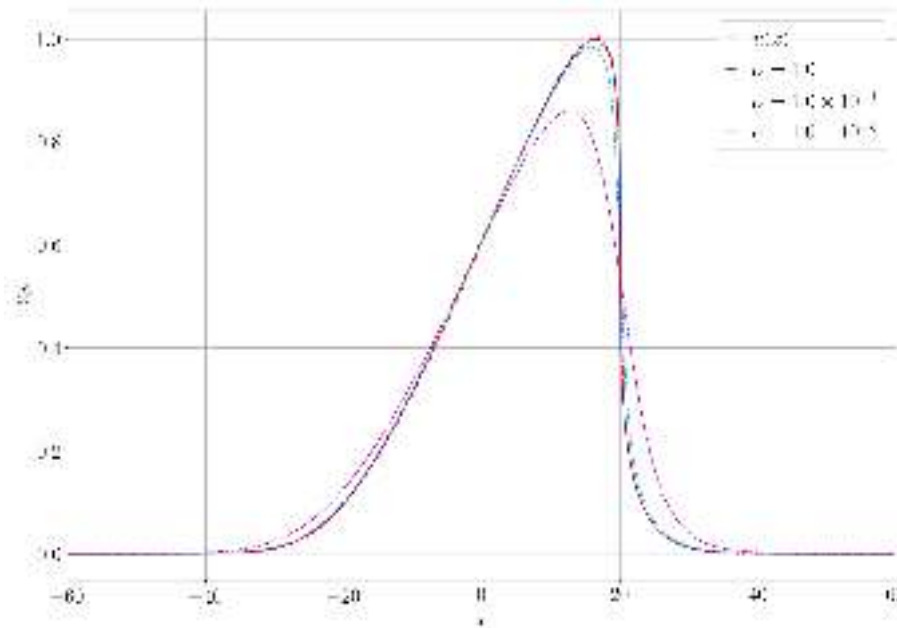

Figure 3.15: Exact solution for (3.18) and different approximations using (3.13) with $N = 256$, and $\Delta t = 1.0 \times 10^{-3}$.

Chapter 4

Numerical Solutions for Burgers' equation in the Stochastic Version

The objective of this chapter is to show how stochastic partial differential equations can be solved using spectral methods, for which it is much more difficult to find analytical or even numerical solutions. For this, we will dedicate the first section of this chapter to describe in the best possible way the details about a spectral method based on the Hermite series that has been studied in [1], which associates a stochastic partial differential equation with the well-known equation Fokker-Planck-Kolmogorov, which is a partial differential equation that describes the time evolution of the probability density function of the velocity of a particle under the influence of random forces.

We will use some results and definitions from the appendix A since the theoretical tools involved in the development of this method are outside the limitations of this work. This appendix contains the concepts necessary to understand the basic ideas of the method implementation that we will use in the second and last section considering the stochastic Burgers' equation that we have presented in (1.14). Finally, we will describe and show some numerical simulations.

4.1 Spectral Approximation for Fokker-Plank-Kolmogorov Equation

The method that we are going to describe in this section will be developed in a space of Hilbert $\mathcal{H}$ with interior product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$, where we will define a Gaussian measure μ with zero mean. Based on [1], the Fokker-Planck-Kolmogorov equation is presented as follows

$$\frac{\partial u}{\partial t} = \frac{1}{2}Tr(D^2 u) + \langle A(x), Du \rangle_{\mathcal{H}} + \langle B(x), Du \rangle_{\mathcal{H}}, \; x \in D(A) \qquad (4.1)$$

where Tr is the trace operator, $A : D(A) \subset \mathcal{H} \to \mathcal{H}$ is a linear differential operator, $B : D(B) \subset \mathcal{H} \to \mathcal{H}$ is a nonlinear operator, and D represents the Frechet derivative.

The main idea of the method is to solve the previous equation associating the following stochastic differential equation

$$dX_t = AX_t dt + B(X_t)dt + dW_t \qquad (4.2)$$

where W_t is a process Q-Wiener as defined in (A.1.10), and the solution for the equation (4.1) is defined as follows

$$u(x,t) = \mathbb{E}\left[u_0(X_t^x)\right] \tag{4.3}$$

where $u_0 : \mathcal{H} \to \mathbb{R}$, and X_t^x is the solution of the equation (4.2).

Following our reference, the solution to the problem (4.1) is represented by an expansion known as Fourier-Hermite which is given by the following series

$$u(x,t) = \sum_{n \in J} u_n(t) H_n(x), \quad x \in \mathcal{H}, \quad t \in [0,T], \tag{4.4}$$

where $u_n : [0,T] \to \mathbb{R}$ and $H_n(x)$ are the Hermite functionals defined in A.6, and J as in A.7.

The above expansion can be justified by the Lemmas A.1 and A.2, and also is known as the deterministic Wiener-Chaos descomposition. Similarly, as we have seen in the previous chapter, it must satisfy the problem (4.1). For this, define the following operator

$$\mathcal{L}u = \frac{1}{2}Tr(D^2 u) + \langle Ax, Du\rangle_{\mathcal{H}}, \quad x \in \mathcal{H} \tag{4.5}$$

which represents the linear part of (4.1), and by Lemma A.2 satisfies the following

$$\mathcal{L}u = -\sum_{n \in J} u_n(t)\lambda_n H_n(x). \tag{4.6}$$

So, substituting the expansion on the left side of the equation (4.1) we get

$$\frac{\partial u}{\partial t} = \sum_{n \in J} \dot{u}_n(t) H_n(x), \tag{4.7}$$

and for the non-linear term

$$\langle B(x), Du\rangle_{\mathcal{H}} = \left\langle B(x), D_x \sum_{n \in J} u_n(t) H_n(x) \right\rangle_{\mathcal{H}}$$

$$= \sum_{n \in J} u_n(t) \left(B(x), D_x H_n(x)\right)_{\mathcal{H}} \tag{4.8}$$

Therefore, by (4.6 - 4.8) the equation (4.1) can be written as

$$\sum_{n \in J} \dot{u}_n(t) H_n(x) = -\sum_{n \in J} u_n(t)\lambda_n H_n(x) + \sum_{n \in J} u_n(t) \left(B(x), D_x H_n(x)\right)_{\mathcal{H}}$$

To develop the above, in the space $\mathcal{H}$ define the Gaussian measure $\mu(dx) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$. So, multiplying the previous equation by $H_m(x)$, $m \in J$ and integrating over $\mathcal{H}$ with respect to the measure $\mu(dx)$ we have to

$$\sum_{n \in J} \dot{u}_n(t) \int_{\mathcal{H}} H_m(x)H_n(x)\mu(dx) = -\sum_{n \in J} u_n(t)\lambda_n \int_{\mathcal{H}} H_m(x)H_n(x)\mu(dx)$$
$$+ \sum_{n \in J} u_n(t) \int_{\mathcal{H}} H_m(x) \left(B(x), D_x H_n(x)\right)_{\mathcal{H}} \mu(dx)$$

and also using the orthogonality of the system $\{H_m(x)\}$, we get the following infinite system of coupled ordinary differential equations

$$\dot{u}_m(t) = -u_m(t)\lambda_m + \sum_{n \in \mathcal{J}} u_n(t)C_{n,m}, \quad n, m \in \mathcal{J} \tag{4.9}$$

where $C_{n,m}$ is given by

$$C_{n,m} = \int_{\mathcal{H}} H_m(x) \left(B(x), D_x H_n(x)\right)_{\mathcal{H}} \mu(dx) \tag{4.10}$$

We need to truncate and solve the above system to get approximations of the solution to the equation (4.1), which will be done in the next section focusing on the Burgers' equation given by (1.14).

4.2   Numerical Solution for Burgers' Equation

Let $\mathcal{H} = L^2(0,1)$. Let's consider an initial value problem for (1.14) as follows

$$dX(\xi,t) = \left[\alpha \partial_\xi^2 X(\xi,t) + \frac{1}{2}\partial_\xi \left(X^2(\xi,t)\right)\right] dt + dW_t(\xi,t), \quad \xi \in [0,1] \tag{4.11}$$

The boundary condition and its initial condition are respectively

$$X(0,t) = X(1,t) = 0, \quad t > 0$$
$$X(\xi,0) = x(\xi), \quad x \in \mathcal{H},$$

where W is a cylindrical Wiener process on $\mathcal{H}$ as was given by A.1.10, associated to a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \geq 0})$, and as usually $\alpha > 0$ is the viscosity coefficient.

Using the above problem, we will proceed to develop the implementation of the method described in the previous section.

4.2.1   Method Description and Its Implementation

We will denote the functions space that vanishes at the borders as $H_0^1(0,1)$. Setting $A = \alpha \partial_\xi^2$ and $B = \frac{1}{2}\partial_\xi(x^2)$, $x \in \mathcal{H}$, with its domains $D(A) = H^2(0,1) \cap H_0^1(0,1)$ and

$D(B) = H_0^1(0,1)$ respectively, then by (4.2), the equation (4.11) can be rewritten as

$$dX = [AX + B(X)]dt + dW_t$$
$$X(0) = x, \quad x \in \mathcal{H}$$

where A have eigenfunctions in $\mathcal{H}$ given by

$$e_k(\xi) = \sqrt{2}\sin{(k\pi\xi)}, \quad \xi \in [0,1], \quad k \in \mathbb{N}$$

Note that the operator A satisfies $Ae_k = -\alpha\pi^2 k^2 e_k$ for $k \in \mathbb{N}$, then if we set $\Lambda = (-A)^{-1}$ we have that $\Lambda^{-1/2}e_k = \sqrt{2\alpha}\pi|k|e_k$.

Therefore, as in (4.9) we need to solve the following system

$$\dot{u}_m(t) = -u_m(t)\lambda_m + \sum_{n\in\mathcal{J}} u_n(t)C_{n,m}, \ n,m \in \mathcal{J} \tag{4.12}$$

We need to calculate the value of the constants $C_{n,m}$, then we need to calculate expressions such as $B(x)$, $D_x H_n(x)$. Note that x can be written as $x = \sum_k \beta_k e_k$, with $\beta_k := \langle x, e_k \rangle_{\mathcal{H}}$. Then we have

$$B(x) = \frac{1}{2}\partial_\xi \left(\sum_k \beta_k e_k\right)^2 = \frac{1}{2}\partial_\xi \left[\sum_l \sum_k \beta_l\beta_k e_l e_k\right] = \frac{1}{2}\sum_l \sum_k \beta_l\beta_k(e_l e_k' + e_l' e_k)$$

and for $D_x H_n(x)$ we have

$$D_x H_n(x) = \sum_{j=1}^{\infty} \prod_{i=1.i\neq j}^{\infty} P_{n_i}(\langle x, \Lambda^{-1/2}e_i\rangle_{\mathcal{H}})P_{n_j}'(\langle x, \Lambda^{-1/2}e_j\rangle_{\mathcal{H}})\Lambda^{-1/2}e_j$$

Therefore, $C_{n,m}$ given by (4.10) gives

$$
\begin{aligned}
C_{n,m} =& \frac{1}{2}\int_{\mathcal{H}} H_m(x)\mu(dx)\sum_{j=1}^{\infty}\prod_{i=1,i\neq j}^{\infty} P_{n_i}(\langle x, \Lambda^{-1/2}e_i\rangle_{\mathcal{H}})P_{n_j}'(\langle x, \Lambda^{-1/2}e_j\rangle_{\mathcal{H}})\sqrt{2\alpha}\pi|j| \\
& \cdot \sum_l \sum_k \beta_l\beta_k(e_l e_k' + e_l' e_k) \\
=& \frac{1}{2}\int_{\mathcal{H}} \mu(dx)\sum_{j=1}^{\infty}\sqrt{2\alpha}\pi|j|P_{m_j}(\langle x, \Lambda^{-1/2}e_j\rangle_{\mathcal{H}})P_{n_j}'(\langle x, \Lambda^{-1/2}e_j\rangle_{\mathcal{H}}) \\
& \cdot \prod_{i=1,i\neq j}^{\infty} P_{n_i}(\langle x, \Lambda^{-1/2}e_i\rangle_{\mathcal{H}})P_{m_i}(\langle x, \Lambda^{-1/2}e_i\rangle_{\mathcal{H}}) \\
& \cdot \sum_l \sum_k \beta_l\beta_k(e_l e_k' + e_l' e_k)
\end{aligned}
$$

So, to obtain a truncated approximation of the solution, the following set of indices is considered

$$J^{M,N} = \{\gamma = (\gamma_i, \ 1 \le \gamma_i \le M) \mid \gamma_i \in \{0,1,\cdots,N\}\} \tag{4.13}$$

this is the set of M-tuple which can take values in the set $\{0,1,\cdots,N\}$.

For $N_1 \in \mathbb{N}$ define as the set $S_{N_1} = \{n_1, n_2, \cdots, n_{N_1} : n_i \in J^{M,N}, i = 1, \cdots, N_1\}$. Then for $n, m \in S_N$ we have

$$\bar{C}_{n,m} = \frac{1}{2} \sum_{j=1}^{\infty} \sqrt{2\alpha}\pi|j| \int_{\mathbb{R}^{\mathcal{M}}} P_{m_j}(\xi_j) P'_{n_j}(\xi_j)\mu(d\xi_j)$$

$$\cdot \prod_{i=1,i\neq j}^{M} P_{m_i}(\xi_i)P_{n_i}(\xi_i)\mu(d\xi_i) \sum_{l=1}^{M}\sum_{k=1}^{M}\beta_l\beta_k(e_l e'_k + e'_l e_k)$$

and for $m_1, m_2, \cdots, m_M \in J^{M,N}$ the system (4.9) give us

$$\dot{u}_{m_i}(t) = -u_{m_i}(t)\lambda_{m_i} + \sum_{j=1}^{M} u_{n_j}(t)C_{n_j,m_i}, \ \ 1 \le i \le M \tag{4.14}$$

The solutions of the previous system can be calculated in terms of their eigenvectors by establishing the following vector

$$U^M(t) = \begin{pmatrix} u_{m_1}(t) & u_{m_2}(t) & \ldots & u_{m_M}(t) \end{pmatrix}^T$$

and for its derivatives

$$\dot{U}^M(t) = \begin{pmatrix} \dot{u}_{m_1}(t) & \dot{u}_{m_2}(t) & \ldots & \dot{u}_{m_M}(t) \end{pmatrix}^T$$

So, we can now write the system (4.14) as

$$\dot{U}^M(t) = AU^M(t) \tag{4.15}$$

where the matrix A is given by

$$A = \begin{pmatrix} -\lambda_1 + C_{1,1} & C_{2,1} & \ldots & C_{M-1,1} & C_{M,1} \\ C_{1,2} & -\lambda_2 + C_{2,2} & \ldots & C_{M-1,2} & C_{M,2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{1,M-1} & C_{2,M-1} & \ldots & -\lambda_{M-1} + C_{M-1,M-1} & C_{M,M-1} \\ C_{1,M} & C_{2,M} & \ldots & C_{M-1,M} & -\lambda_M + C_{M,M} \end{pmatrix}$$

where $\lambda_i = \lambda_{mi}$ and $C_{i,j} = C_{n_i,m_j}$ para $1 \le i,j \le M$.

Then, if A has M real and distint eigenvalues η_i and M eigenvectors V_i, then the solution to (4.15) is given by

$$U^M(t) = \sum_{j=1}^{M} c_i V_i e^{\eta_i t} \tag{4.16}$$

In the case when some eigenvalue is complex, we can write it together with its eigenvector as follows

$$V = a + ib, \quad \eta = \beta + i\mu$$

to get the solutions

$$e^{\beta t}(a\cos(\mu t) - b\sin(\mu t)), \;\; e^{\beta t}(a\sin(\mu t) + b\cos(\mu t))$$

which are real and different.

Then we can write the approximation of the solution of (4.1) as

$$u_M(x,t) = \sum_{n \in J^{M,N}} u_n(t)H_n(x) = U^M(t)H^M(x), \;\; x \in \mathcal{H}, \;\;, t \in [0,T]. \qquad (4.17)$$

Also, if $u(\xi, t) = \mathbb{E}\left[X_t(\xi)\right]$, then satisfies the problem given by

$$\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial \xi^2} + \partial_\xi \left[u(\xi,t)\right]^2$$

with the initial condition $u(\xi, 0) = \mathbb{E}\left[X_0\right]$.

4.2.2 A Functional to obtain Initial Conditions

The interesting thing about the equation (4.1), is that there is no standard way to define an initial condition. For this problem, a functional is defined that acts in the initial condition, and because there are different ways of defining this functional, the method may change. For this work, the following functional was chosen

$$u_0^{z_0}(g) := g(z_0), \;\; \text{for fixed} \;\; z_0 \in [0,1].$$

To construct the initial condition, the following set of points is considered

$$P = \{z_i, \; 0 \le i \le p \; : \; z_0 = 0, \; z_p = 1\}$$

Then for each point $z_i \in P$ such that $X_0(z_i) = X(0, z_i)$ set $u_0(x)$ as the evaluation functional $z_i \longrightarrow X_t^x(z_i)$. Then from (4.3) we obtain

$$u(0,x) = \mathbb{E}[u_0^{z_i}(X_0^x)] = X^x(0, z_i) = x(z_i) \qquad (4.18)$$

For other hand

$$u(0,x) = \sum_{n \in \mathcal{J}^{M,N}} u_n(0)H_n(x)$$

multiplying for $H_m(x)$ and integrating over space $\mathcal{L}^2(\mathcal{H}, \mu)$

$$u_m(0) = \int_{\mathcal{H}} x(z_i)H_m(x)\mu(dx)$$

Note that in the direction of the eigenfunction e_k the expression x can be written as $(x, e_k)_{\mathcal{H}} e_k$, then we can write $H_m(x) x(z_i)$ in the direction e_k as $P_{m_k}(\xi_k)(x, e_k)_{\mathcal{H}} e_k(z_i)$ with $\xi_k = (x, \Lambda^{-\frac{1}{2}} e_k) = \|\lambda_k\|(x, e_k)_{\mathcal{H}}$ and P_{m_k} is given by (A.4). Then we have

$$
\begin{aligned}
u_m^{z_i}(0) &= \int_{\mathcal{H}} x(z_i) H_m(x) \mu(dx) \\
&= \int_{\mathbb{R}^N} \sum_{k=1}^{\infty} P_{m_k}(\xi_k)(x, e_k)_{\mathcal{H}} e_k(z_i) \mu(d\xi_1, d\xi_2, \cdots) e_k \\
&= \int_{\mathbb{R}^N} \sum_{k=1}^{\infty} P_{m_k}(\xi_k) \frac{\xi_k}{\lambda_k} e_k(z_i) \mu(d\xi_1, d\xi_2, \cdots) e_k \\
&= \sum_{k=1}^{\infty} \frac{e_k}{\lambda_k} \int_{\mathbb{R}} P_{m_k}(\xi_k) \xi_k(z_i) \mu(d\xi_k)
\end{aligned}
$$

truncating the above expression we have

$$
u_m^{z_i}(0) \approx \sum_{k=1}^{M} \frac{e_k}{\lambda_k} \int_{\mathbb{R}} P_{m_k}(\xi_k) \xi_k(z_i) \mu(d\xi_k) \tag{4.19}
$$

Setting the equation (4.19) for each element from $u_m^{z_i}$ as $u_{m_j}^{z_i}(0) = u_j(0)$, $1 \leq j \leq M$ and by (4.16) evaluated for $t = 0$, then the initial condition can be written as

$$
\begin{pmatrix} u_1(0) \\ u_2(0) \\ \vdots \\ u_{M-1}(0) \\ u_M(0) \end{pmatrix} = \begin{pmatrix} V1 & V2 & \ldots & V_{M-1} & V_M \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{M-1} \\ c_M \end{pmatrix}
$$

and the constants c_j are calculated as

$$
\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{M-1} \\ c_M \end{pmatrix} = \begin{pmatrix} V1 & V2 & \ldots & V_{M-1} & V_M \end{pmatrix}^{-1} \begin{pmatrix} u_1(0) \\ u_2(0) \\ \vdots \\ u_{M-1}(0) \\ u_M(0) \end{pmatrix}
$$

4.2.3   Numerical Experiments

Now, we will proceed to describe the steps necessary to implement the method based on the above:

1. Set the intervals to consider, $[0, 1]$ for physical space and $[0, t_f]$ for time with t_f fixed, and give the functional u_0 that interacting with the initial condition $x(\xi)$. It is also necessary to choose the values of N, M, N_1, and another for the parameter α.

2. We proceed to calculate the points $\xi_i \in [0, 1]$ for $i = 0, 1, \ldots, p$ with p positive entire as

$$\xi_i = \xi_0 + i\Delta\xi, \quad i = 0, 1, \ldots, p$$

   and to obtain the points $t_j \in [t_0, t_f]$, we define the sequence $j = 0, 1, \ldots, T$, for some positive entire T, and calculate the following

$$t_j = t_0 + j\Delta t, \quad \Delta t = \frac{t_f}{T}$$

3. Calculate the values of $\mathcal{J}^{N,M}$ to obtain the matrix $\bar{C}_{n,m}$. Setting the values λ_i to proceed with the calculation of the eigenvalues η_i and their respective eigenvectors V_i of the matrix A.

4. Using the functional u_0 to calculate the constants c_k for each $k = 1, \ldots, M$, and then obtain the system solution given by (4.16).

5. Finally, the approximation obtained is evaluated using the equation given by

$$u_M(x, t) = \sum_{k=1}^{M} u_k(t) H_k(x).$$

The following simulations that will be shown were performed using a discretization of 2048 points in the spatial variable ξ over the interval $[0, 1]$, and 1024 points in the variable time t over the interval $[0, 10]$. In addition, the values of the parameters $N = 5$, $M = 11$, and $\alpha = 1.0 \times 10^{-2}$ were considered. With this information, the solutions obtained were calculated using the following initial condition and its truncated Chebysheb expansion given by

$$x(\xi) = \sin(\pi\xi), \quad y(\xi) = \sum_{k=0}^{N} c_k T_k(\xi),$$

This numerical experiment consists of illustrating an interesting result given in [35], which tells us that the solutions obtained from two close initial conditions also remain close. This behavior allows characterizing what is known as the stability of the approximation, and it is described by means of continuity with respect to the

initial conditions of the numerical approximations of the equation (4.1).

To understand this better, let us denote by Ψ_t^x the solution of (4.1) obtained by

$$u(x,t) = \mathbb{E}\left[\varphi(X_t^x)\right].$$

where $\varphi : \mathcal{H} \to \mathbb{R}$ is Lipschitz and X_t^x is the solution to (4.2) with initial conditions $X_0 = x \in \mathcal{H}$. So, as before, its expansion is given by

$$\Psi_t^x = \sum_{n \in \mathcal{J}} u_n(t) H_n(x), \quad x \in \mathcal{H}, \quad t \in [0, T].$$

Therefore, following our reference, the above is summarized as follows: Given two different initial conditions $x, y \in \mathcal{H}$, then we have the following estimate

$$\|\Psi_t^x - \Psi_t^y\|_{(L^2(\mathcal{H},\mu))^2}^2 \leq \exp(Ct) \int_{\mathcal{H} \times \mathcal{H}} \|x - y\|_{\mathcal{H}}^2 \mu(dx)\mu(dy) + f(t)\|x - y\|_{\mathcal{H}},$$

for some C finite and $f(t)$ is given by

$$f(t) = \sum_{n \in J} [u_n^y(t)]^2 + \int_{\mathcal{H}} \mathbb{E}^2\left[\varphi(X_t^y)\right]\mu(dy).$$

From the above, we can see that if $\|x - y\|_{\mathcal{H}} \leq \delta$, so we have to $\|\Psi_t^x - \Psi_t^y\| \leq G(t)\delta$. As we have already mentioned, this continuity defines a type of stability for approximations, which is of utmost importance in this field since characterizations of this type are still under construction and are essential for the analysis of a numerical method.

This behavior is shown in the following figures, which were obtained using codes that were created following the previous steps, and which can be found in https://github.com/alanmatzumiya/Paper.git. In figure 4.1 it shows us the two initial conditions for which the equation (4.1) will be solved by associating it with (1.14), and in figure 4.2 we can see that the solutions keep the distance. Finally, the figure 4.3 shows the distances for each instant of time t, showing that they are actually getting closer as time passes.
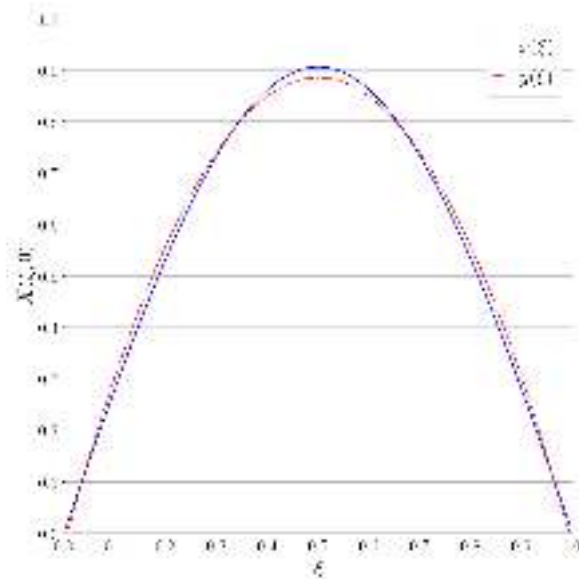
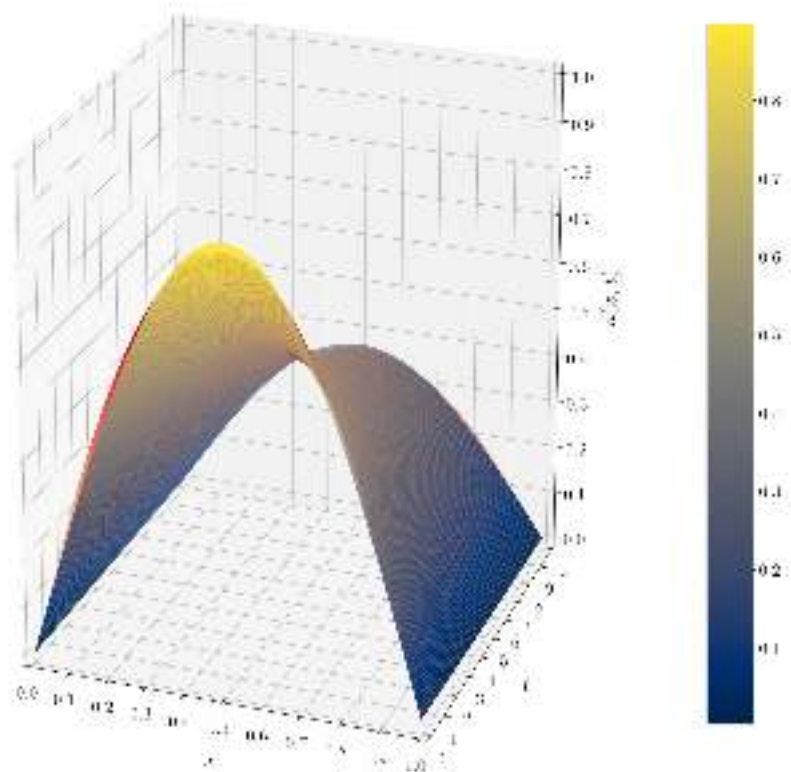Figure 4.1: Initial condition for (4.11) and its approximation.

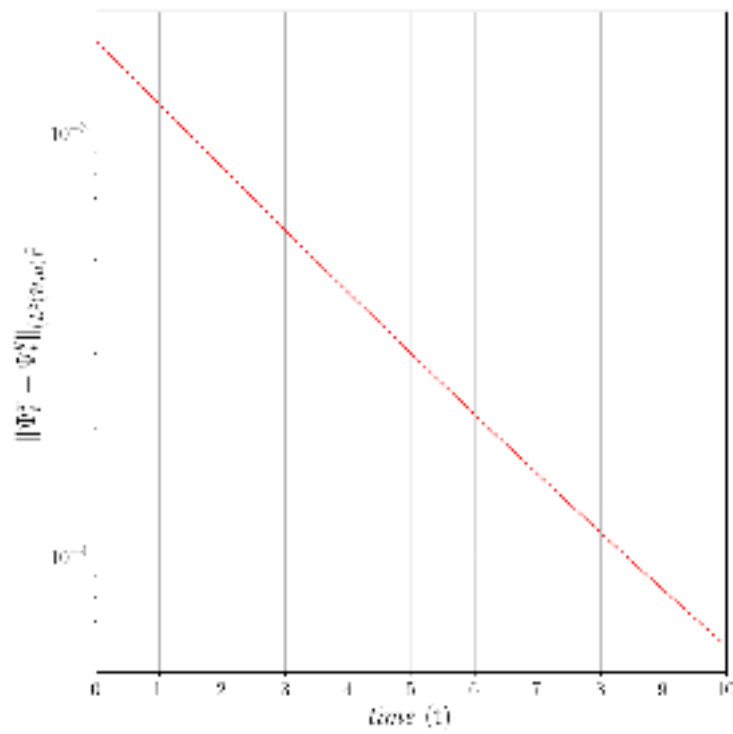Figure 4.2: Numerical solutions for (4.11) with initial conditions $x(\xi)$ and $y(\xi)$.

Figure 4.3: Distance between the numerical solutions for equation (4.11) with initial conditions $x(\xi)$, and $y(\xi)$.

Chapter 5

Discussion and Conclusions

With what has been developed throughout this thesis work, it has given us a satisfactory understanding of the knowledge necessary to apply spectral methods in solving initial value problems of nonlinear partial differential equations, which with practice it will be possible to acquire the ability to study and attack other more complex problems.

The practice that we have obtained when developing Chapter 3 with the help of the tools examined in Chapter 2, can be useful to construct and study some other spectral method using another family of polynomials that allows us to approximate more precisely a function, and to construct practical methods for its implementation solving this same problem or another one of interest. Although, as expected based on theory, we were able to obtain good approximations using Fourier methods to solve linear problems. But, the situation was different when we considered the nonlinear problem since it was not possible to obtain the solutions directly, forcing us to develop more carefully the numerical methods that finally gave us a good numerical approximation and with practical mathematical expressions for its implementation.

However, we noted that very small steps in time were required to ensure that the methods worked correctly, which greatly increased the computation time, but this was due to the presence of the aliasing error caused by the nonlinear term, which manifested itself unfavorably with small values of α, and they should be handled more carefully, possibly considering another class of polynomials. Even so, we were able to satisfactorily detect the problems that can arise when using these tools, and it is interesting to delve into this with the help of the appropriate tools that allow us to characterize and know the convergence of these methods.

Regarding the spectral method that we developed in the Chapter 4, the problem that arises is due to the number of calculations that are necessary to obtain the numerical approximations, but nevertheless, the implementation is easy to understand and can be studied with more detail to propose a more efficient and less expensive algorithm to calculate, mainly attacking the calculation of the integrals involved in the process. For this type of problem, it is worth investigating an implementation that allows you to write parallel code in some programming language to optimize memory and computation times.

Even so, it was possible to successfully carry out some numerical experiments that allowed us to observe a theoretical result that characterized the stability of this method, and perhaps could allow us to investigate other characteristics of interest, such as its convergence. Thanks to the practice developed in this chapter we have acquired the ability to implement a spectral method to carry out numerical studies that may be useful for its analysis, which could also be interesting to implement these same techniques in other problems that are considered important.

Therefore, we can conclude that spectral methods are a good option to study partial differential equations, either deterministic or stochastic since in addition to giving us the possibility of obtaining good precision in their approximations provided they are implemented correctly, they are also an excellent alternative to investigate the nature of these problems, be they physical or mathematical, opening different paths, such as some of those already mentioned, that may be interesting for developing future research.

Appendix A

Elements of Probability

A.1 Basic Definitions

Definition A.1.1. A probability measure $\mathbb{P}$ on a measurable space $(\Omega, \mathcal{F})$ is a function from $\mathcal{F}$ to $[0, 1]$ such that

- $\mathbb{P}(\emptyset) = 0$, and $\mathbb{P}(\Omega) = 1$.

- If $\{A_n\}_{n \geq 1} \in \mathcal{F}$ and $A_i \cap A_j \neq \emptyset$ if $i \neq j$, then $\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.

A σ-algebra on a set X is a collection of subsets of X that includes the empty subset and is closed under complement and under countable unions. Denote $\sigma(D) = \cap\{H : H \text{ is a } \sigma\text{-algebra of } \Omega, D \subseteq H\}$. We call $\sigma(D)$ a σ-algebra generated by D.

Definition A.1.2. A triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space if

- Ω is a sample space which is a collection of all samples.

- $\mathcal{F}$ is a σ-algebra on Ω.

- $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$.

On a given probability space $(\Omega, \mathcal{F}, \mathbb{P})(\Omega = \mathbb{R})$, if a cumulative distribution function of a random variable X is normal, i.e.,

$$\mathbb{P}(X < x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy, \quad \sigma > 0, \tag{A.1}$$

then the random variable X is called a Gaussian (normal) random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Here X is completely characterized by its mean μ and its standard deviation σ. We denote $X \sim \mathcal{N}(\mu, \sigma^2)$. The probability density function of X is

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

When $\mu = 0$ and $\sigma = 1$, we call X a standard Gaussian (normal) random variable.

Definition A.1.3. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be a complete probability space if for all $B \in \mathcal{F}$ with $\mathbb{P}(B) = 0$ and all $A \subseteq B$ one has $A \in \mathcal{F}$.

Definition A.1.4. If $(\Omega, \mathcal{F}, \mathbb{P})$ is a given probability space then a function $Y : \Omega \to \mathbb{R}^n$ is called $\mathcal{F}$-measurable if $Y^{-1}(U) = \{w \in \Omega : Y(w) \in U\} \in \mathcal{F}$ holds for all open sets $U \in \mathbb{R}^n$. If $X : \Omega \to \mathbb{R}^n$ is a function, then $\sigma(X)$ is the smallest σ-algebra on Ω containing all the sets $X^{-1}(U)$ for all open sets U in $\mathbb{R}^n$.

Definition A.1.5. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a given complete probability space. A random variable X is an $\mathcal{F}$-measurable function $X : \Omega \to \mathbb{R}^n$.

Its well known that every random variable induces a probability measure μ_X (distribution of X) on $\mathbb{R}_n$ given by

$$\mu_X(B) = \mathbb{P}(X^{-1}(B)).$$

If $\int_\Omega |X(w)| d\mathbb{P}(w) < \infty$, the expectation of X w.r.t. $\mathbb{P}$ is defined by

$$\mathbb{E}[X] = \int_\Omega X(w) d\mathbb{P}(w) = \int_{\mathbb{R}^n} x d\mu_X(x).$$

Also the p-th moment of X is defined as (if the integrals are well defined)

$$\mathbb{E}[X^p] = \int_\Omega X^p d\mathbb{P}(w) = \int_{\mathbb{R}^n} x^p d\mu_X(x).$$

The centered moments are defined by $\mathbb{E}[|X - \mathbb{E}[X]|]$, $p = 1, 2, \ldots,$. When $p = 2$, the centered moment is also called the variance.

Definition A.1.6. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $T \subseteq \mathbb{R}$ be time. A collection of random variables X_t , $t \in T$ with values in $\mathbb{R}$ is called a stochastic process. If time is an interval, $\mathbb{R}^+$ or $\mathbb{R}$, it is called a stochastic process with continuous time. For any fixed $w \in \Omega$, one can regard $X_t(w)$ as a function of t (called a sample function of the stochastic process).

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a filtration refers to an increasing sequence of σ-algebras:

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_n \subseteq \cdots .$$

A natural filtration (w.r.t. X) is the smallest σ-algebra that contains information of X. It is generated by X and $\mathcal{F}_n^X = \sigma(X_1, \ldots, X_n)$ with $\mathcal{F}_0^X = \{\emptyset, \Omega\}$. If $\lim_{n\to\infty} \mathcal{F}_n \subseteq \mathcal{F}$, then we call $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n\geq 1}, \mathbb{P})$ a filtered probability space. A stochastic process $\{X_n\}$ on a filtered probability space is an adapted process if X_n is $\mathcal{F}_n$-measurable for each n.

Definition A.1.7. A family of sub-σ-algebras $\mathcal{F}_t \subseteq \mathcal{F}$ indexed by $t \in [0, \infty)$ is called a filtration if it is increasing $\mathcal{F}_s \subseteq \mathcal{F}_t$ when $0 \leq s \leq t < \infty$.

Definition A.1.8. A collection of random variables is called a Gaussian process, if the joint distribution of any finite number of its members is Gaussian. In other words, a Gaussian process is a $\mathbb{R}^d$-valued stochastic process with continuous time (or with index) t such that $(X(t_0), X(t_1), \ldots, X(t_n))^T$ is a $(n+1)$-dimensional Gaussian random vector for any $0 \leq t_0 < t_1 < \cdots < t_n$. The Gaussian process is denoted as $X = \{X(t)\}_{t \in I}$ where I is a set of indexes.

76

Definition A.1.9. A continuous time stochastic process $W(t)$ is called a standard Brownian motion if

- $W(t)$ is almost surely continuous in t, and $W(0) = 0$.

- $W(t)$ has independent increments $W(t_{i+1}) - W(t_i)$ for all $t_n \geq 0$, $i = 0, 1, \ldots, n$.

- $W(t) - W(s) \sim \mathcal{N}(0, t - s)$, i.e., obeys the normal distribution with mean zero and variance $t - s$.

Set $x \in D \subset \mathbb{R}^d$, we define infinite dimensional Gaussian processes as follows

$$W^Q(x, t) = \sum_{j=1}^{\infty} \sqrt{q_j} e_j(x) W_j(t),$$

where $W_j(t)$ are mutually independent Brownian motions. Here $q_j \geq 0$, $j \in \mathbb{N}^d$ and $\{e_j(x)\}$ is an orthonormal basis in $L^2(D)$. The following expansion is usually considered in literature:

$$\dot{W}^Q(x, t) = \sum_{j=1}^{\infty} \sqrt{q_j} e_j(x) \dot{W}_j(t).$$

where $\dot{W}_j(t) = \frac{d}{dt} W$, is formally the first-order derivative of $W_j(t)$ in time. When $q_j = 1$ for all j, we have a space-time white noise, and if $\sum_{j=1}^{\infty} \sqrt{q_j}$ is called a Q-Wiener process.

The Brownian motion and white noise can also be defined in terms of orthogonal expansions. Suppose that $\{e_j(t)\}_{j \geq 1}$ is a complete orthonormal system in $L^2([0, T])$, then the Brownian motion $W(t)$ can be defined by

$$W(t) = \sum_{j=1}^{\infty} \beta_j \int_0^t e_j(s) ds, \ \ t \in [0, T], \tag{A.2}$$

where β_j are mutually independent standard Gaussian random variables for each j, and it can be also checked that is indeed a standard Brownian motion. Correspondingly, the white noise is defined by

$$\dot{W}(t) = \sum_{j=1}^{\infty} \beta_j e_j(t), \ \ t \in [0, T]. \tag{A.3}$$

Definition A.1.10. Let $\{e_j\}_{j \geq 1}$ be a complete orthonormal system of a separable Hilbert space $\mathcal{H}$, and $T \in \mathbb{R}^+$, and $\{\beta_j(t)\}_{j \geq 1}$ be an independent and identically distributed sequence of Brownian Motions. Then a cylindrical Wiener process W in $\mathcal{H}$ is given by

$$W(t) = \sum_{j=1}^{\infty} \beta_j e_j(t)$$

A.2   Some Important Results in Probability Spaces

It is well known that the Hermite Polynomials $\{P_k(\dot{})\}_{k\in\mathbb{N}}$ with $P_0 = 1$ is a complete orthonormal system for $L_2(\mathbb{R}, \mu_1(dx))$ with $\mu_1(dx) = \frac{1}{\sqrt{2\ pi}}e^{-\frac{x^2}{2}}\,dx$, which are defined by

$$P_k(x) = \frac{(-1)^k}{(k!)^{1/2}}e^{\frac{x^2}{2}}\frac{d^k}{dx^k}e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R} \tag{A.4}$$

A functional $\Phi : \mathcal{H} \longrightarrow \mathbb{R}$, is said to be a smooth simple functional (or a cylinder functional) if there exists a C^∞-function φ on $\mathbb{R}^n$ and n-continuous linear functional $l_1, \cdots, l_n$ on $\mathcal{H}$ such that for $h \in \mathcal{H}$

$$\Phi(h) = \phi(h_1, \cdots, h_n) \text{ where } h_i = l_i(h), \ i = 1, \cdots, n.$$

Let $\mathbb{H} = L_2(\mathcal{H}, \mu)$ denote the Hilbert space of Borel measurable functionals on the probability space with inner product

$$\langle\Phi, \Psi\rangle_{\mathbb{H}} = \int_{\mathcal{H}} \Phi(v)\Psi(v)\mu(dx), \ \ \Phi, \Psi \in \mathbb{H}, \tag{A.5}$$

and the norm $\|\Phi\|_{\mathbb{H}} = \langle\Phi, \Phi\rangle_{\mathbb{H}}^{1/2}$. In $\mathcal{H}$ we choose a basis system $\{\varphi_k\}$ such that $\varphi_k \in \mathcal{H}$.

Let Λ an operator such that $Tr < +\infty$. We will denote as $\mathcal{H}_0$ the Hilbert subspace of $\mathcal{H}$ with inner product $\langle g, h\rangle_0 = \langle\Lambda^{1/2}g, \Lambda^{1/2}h\rangle_{\mathcal{H}}$ for $g, h \in \Lambda^{1/2}\mathcal{H}$, which is the completion of $\Lambda^{1/2}\mathcal{H}$ with respect to the norm $\|g\|_0 = \langle g, g\rangle^{1/2}$, and also is dense in $\mathcal{H}$. Then, by using the above can be defined The Hermite functional as follows

$$H_n(h) = \prod_{i=1}^{\infty} P_{n_i}(l_i(h)), h \in \mathcal{H}_0, n \in J \tag{A.6}$$

with

$$l_i(h) = \left\langle h, \Lambda^{-1/2}\varphi_i\right\rangle_{\mathcal{H}}, \quad i = 1, 2 \cdots$$

and the set $\mathcal{J}$ is given by

$$\mathcal{J} = \{\alpha = (\alpha_i, i \geq 1)|\alpha_i \in \mathbb{N} \cup 0, |\alpha| := \sum_{i=0}^{\infty}\alpha_i < \infty\} \tag{A.7}$$

Lemma A.1. *For $h \in \mathcal{H}$ let $l_i(h) = (h, \Lambda^{-1/2}\varphi_i)_{\mathcal{H}}, \ i = 1, 2, \cdots$. Then the set $\{H_n\}$ of all Hermite polynomials on $\mathcal{H}$ forms a complete orthonormal system for $\mathbb{H}$. Hence the set of all functionals are dense in $\mathbb{H}$. Moreover, we have the direct sum decomposition:*

$$\mathbb{H} = \bigoplus_{j=0}^{\infty} K_j,$$

where K_j is the subspace of $\mathbb{H}$ spanned by $\{H_n : |n| = j\}$. This result can be found in [24].

Lemma A.2. *Let $H_n(h)$ be a Hermite polynomial functional given by (A.6), and suppose that the operator $-A$ have eigenfunctions e_k with eigenvalues λ_k. Then the operator given in (4.5) satisfies the following*

$$\mathcal{L}H_n(h) = -\lambda_n H_n(h) \qquad (A.8)$$

for any $n \in \mathcal{J}$ and $H_n \in \mathcal{H}$, where

$$\lambda_n = \sum_{k=1}^{\infty} n_k \lambda_k$$

Bibliography

[1] Delgado Vences, Francisco & Flandoli, Franco. (2016). A spectral-based numerical method for Kolmogorov equations in Hilbert spaces. Infinite Dimensional Analysis, Quantum Probability and Related Topics. 10.1142/S021902571650020X.

[2] Acheson, D. J. Elementary Fluid Dynamics. Clarendon Press, 2001.

[3] Batchelor, G. K. An Introduction to Fluid Dynamics, By G.K. Batchelor. 1967.

[4] Landau, Lev D., and Lifsic Evgenij M. Fluid Dynamics. Pergamon, 1987.

[5] Currie, Iain G. Fundamental Mechanics of Fluids. McGraw-Hill, 1974.

[6] Temam, Roger. Navier-Stokes Equations: Theory and Numerical Analysis. North-Holland, 1984.

[7] Bateman, H. (1915) Some Recent Researches on the Motion of Fluids. Monthly Weather Review, 43, 163-170. http://dx.doi.org/10.1175/1520-0493(1915)43¡163:SRROTM¿2.0.CO;2

[8] J M Burgers,Trans. R. Netherlands Acad. Sci. Amster-dam17, 1 (1939)

[9] J M Burgers,Adv. Appl. Mech.1, 171 (1948)

[10] Schwartz, Laurent. Theorie Des Distributions. Hermann, 1966.

[11] Lions, J.-L, and Enrico Magenes. Non-Homogeneous Boundary Value Problems and Applications. Springer-Verlag, 1972.

[12] Renardy, Michael, and Robert C. Rogers. An Introduction to Partial Differential Equations. Springer, 1993.

[13] E Hopf,Commun. Pure Appl. Math.3, 201 (1950)

[14] J D Cole,Quart. Appl. Math.9, 225 (1951)

[15] Chambers, D. H., et al. KarhunenLove Expansion of Burgers Model of Turbulence. Physics of Fluids, vol. 31, no. 9, 1988, pp. 25732582., doi:10.1063/1.866535.

[16] H. CHOI, R. TEMAM, P. MOIN, J. KIM, Feedback control for unsteady flow and its application to Burgers equation, Center for Turbulence Research, Stanford University, CTR Manuscript 131. To appear on J. Fluid Mechanics (1992)

[17] DAH-TENC JENG Forced Model Equation for Turbulence, The Physics of Fluids 12, 10, 2006-2010 (1969)

[18] I. HOSOKAWA, K. YAMAMOTO, Turbulence in the randomly forced one dimensional Burgers flow, J. Stat. Phys. 13, 245 (1975)

[19] M. KARDAR, M. PARISI, J. C. ZHANG Dynamical scaling of growing interfaces, Phys. Rev. Lett. 56, 889 (1986)

[20] J. Bec; K. Khanin. Burgers turbulence. Phys. Rep. 447 (2007), 12, 166.

[21] E. Weinan. Stochastic hydrodynamics. Current developments in mathematics, 2000, 109147, Int. Press, Somerville, 2001.

[22] L. Bertini ;N. Cancrini ;G. JonaLasinio. The stochastic Burgers equation. Comm. Math. Phys. 165 (1994), 2, 211232.

[23] P. Catuogno ;C. Olivera. Strong solution of the stochastic Burgers equation. Appl. Anal. 93 (2014), 3, 646652.

[24] G. Da Prato ; A. Debussche ; R. Temam. Stochastic Burgers equation. NoDEA Nonlinear Differential Equations Appl. 1 (1994), 4, 389402.

[25] M. Gubinelli ; N. Perkowski Lectures on singular stochastic PDEs , Ensaios Matemticos, 29. Sociedade Brasileira de Matemtica, Rio de Janeiro, 2015.

[26] Canuto, Claudio and Hussaini, M Yousuff and Quarteroni, Alfio Maria and Thomas Jr, A and others, Spectral methods in fluid dynamics. Springer Science & Business Media (2012) 2017.06.28

[27] F T Nieuwstadt and J A Steketee,Selected papers of J M Burgers(Springer Science+Business Media, BV,Dordrecht, 1995)

[28] Hesthaven, Jan S. and Gottlieb, Sigal and Gottlieb, David, Spectral Methods for Time-Dependent Problems. Cambridge Monographs on Applied and Computational Mathematics (2007) 10.1017/CBO9780511618352

[29] I., E., Richtmyer, R. D., and Morton, K. W. (1968). Difference Methods for Initial-Value Problems. Mathematics of Computation, 22(102), 465. doi: 10.2307/2004698

[30] KATO, Tosio. Fractional powers of dissipative operators. J. Math. Soc. Japan 13 (1961), no. 3, 246–274. doi:10.2969/jmsj/01330246. https://projecteuclid.org/euclid.jmsj/1261062709

[31] Lasiecka, I. (1984). Convergence Estimates for Semidiscrete Approximations of Nonselfadjoint Parabolic Equations. SIAM Journal on Numerical Analysis, 21(5), 894909. doi:10.1137/0721058

[32] S N Krukov, "FIRST ORDER QUASILINEAR EQUATIONS IN SEVERAL INDEPENDENT VARIABLES", MATH USSR SB, 1970, 10 (2), 217243

[33] Eitan Tadmor. 1989. Convergence of spectral methods for nonlinear con-
     servation laws. SIAM J. Numer. Anal. 26, 1 (February 1989), 3044.
     DOI:https://doi.org/10.1137/0726003

[34] Yvon Maday and Eitan Tadmor. 1989. Analysis of the spectral vanishing viscos-
     ity method for periodic conservation laws. SIAM J. Numer. Anal. 26, 4 (August
     1989), 854870. DOI:https://doi.org/10.1137/0726047

[35] Delgado Vences, Francisco & Matzumiya-Zazueta, Alan & Saul, Diaz-
     Infante. (2019). INITIAL CONDITIONS CONTINUITY OF A NU-
     MERICAL APPROXIMATION FOR KOLMOGOROV EQUATIONS *.
     10.13140/RG.2.2.21593.47203.