"El saber de mis hijos hará mi grandeza"

UNIVERSIDAD DE SONORA

División de Ciencias Exactas y Naturales

Programa de Posgrado en Matemáticas

Numerical Solutions to the Stochastic and Deterministic Burgers' Equation by 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, February 10, 2020

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. For example, the existence and uniqueness of solutions for ordinary differential equations (ODEs) can be summed up in the Picard-Lindelöf theorem, but this same problem is far from satisfactorily solved for partial differential equations (PDEs).

There are alternatives to find solutions to differential equations, for example, computational fluid dynamics (CFD) 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 (DNS) 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 technologically 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.

In this thesis, we will present the most common spectral methods in practice

to implement them in Burgers' equation, in addition to studying their convergence, stability, and consistency under this approach. For its development we divide the work into five parts organized as follows:

1. In Chapter 1 (present chapter) we will present a brief history of Burgers' equation in its deterministic version, and in addition to its analytical solution. Later we will also present the stochastic version of Burgers' equation.

2. In Chapter 2 will study the fundamental bases of spectral methods, in addition to some results that will be useful for the study of our main problem.

3. In Chapter 3 we will show the implementation of the studied spectral methods in Chapter 2 illustrating them with some simulations, in addition, some results of the numerical analysis and their convergence will be obtained.

4. In Chapter 4 we will study the implementation of a method to solve stochastic Burgers' equation developed in [1], and we will also see some results of the numerical analysis.

5. Finally, Appendixes A and B were added presenting some results that will be useful for the analysis of the problems studied in Chapters 3 and 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 $g(x,t)$ as the x component of G, then the equation (1.2) becomes just the viscid Burgers' equation

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

Some assumptions are made, namely: $\rho=$ constant (density), $\mu=$ constant (viscosity), $p=$ 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

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

where F and A are given by

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

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

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

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

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

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 see Appendix A, and 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.

In this thesis, we will study the spectral methods focused on the formulations given above, in addition to studying the necessary bases for the analysis of convergence, stability, and consistency under this approach. Firstly, we will show the exact solution for the problem (1.3) with $g \equiv 0$ in order to compare the schemes studied and be able to make a more detailed analysis.

1.1.1 Analytical solution for Burgers' equation.

The main goal of this work is to develop and test numerical methods to solve (PDEs) of Convection-Diffusion type as (1.3). 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) with $g(x, t) \equiv 0$

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

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

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

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,

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

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

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)

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

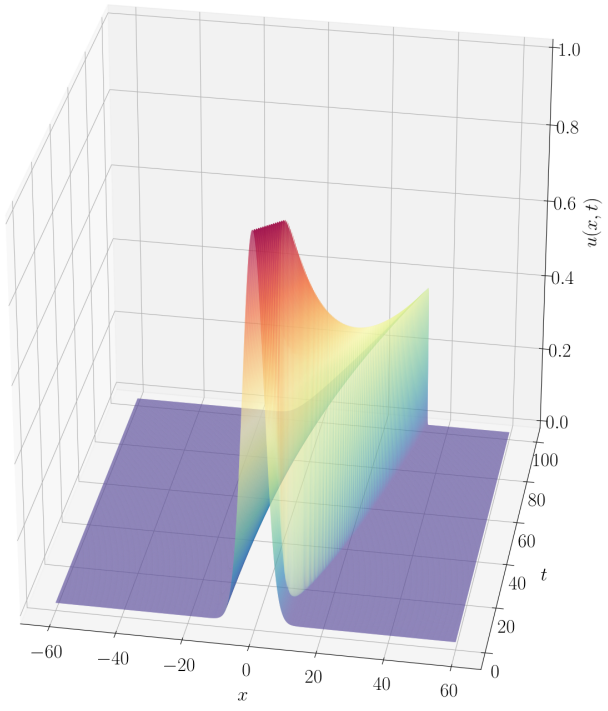


Figure 1.1: Exact solution for (1.7) with initial condition $u_0(x) = e^{-0.05x^2}$ using the equation (1.12) for $x \in [-60, 60]$, $t \in [0, 100]$, and $\alpha = 0.01$.

1.2 Stochastic Burger's 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 B.

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

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

Equation (1.14) is supplemented with Dirichlet boundary conditions

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

and the initial condition

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

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].

In the last chapter 4, we will present an interesting method to approximate the solutions to equation (1.14), which was studied and developed in [1]. The interesting thing about this method is the similarity with the spectral methods that will be developed for the deterministic case.

Chapter 2

Fundamental Theory For Spectral Methods

In this chapter, we will present the necessary elements to solve any type of (PDE) using spectral methods given by the following way

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

where u and $f(x,t)$ are defined in some Hilbert space $\mathcal{H}$, with initial condition $g(x) \in \mathcal{H}$ and $\mathcal{L}$ is some spatial differential operator. The aim is to investigate how well the spectral methods approximate the exact solution and how the level of precision improves as we refine the grid in time and space. Such behavior is known as convergence.

For this, first we will present the necessary elements to understand the implementation of these methods and finally, the theorems which are necessary for the convergence analysis.

2.1 Elements in Convergence Theory

Spectral methods involve representing the solution of the differential equation in terms of a series truncated of known, smooth functions of the independent variables. They have recently emerged as a viable alternative to finite difference and finite element methods for the numerical solution of partial differential equations. The key recent advance was the development of the fast Fourier transform algorithm for efficient implementation.

The origin of the terminology spectral is not entirely clear but probably arises from the use of Fourier expansion especially in connection with time series analysis and the fundamental frequencies of a process, namely, the spectrum.

These methods are a class of spatial discretization for differential equations. The key components for their formulation are the test functions (also called the expansion or approximating functions). A finite linear combination (truncated expansion) of suitable test functions can provide the approximate representation of the solution ensuring that the differential equation is satisfied as closely as possible. This is

achieved by minimizing the residual, i.e., the error in the differential equation produced by using the truncated expansion instead of the exact solution with respect to a suitable norm or equivalently that the residual satisfy an orthogonality condition with respect to each test function.

The choice of test functions is one of the characteristics that distinguish spectral methods from finite element and finite differences methods since they are infinitely differentiable and are defined as global functions throughout the space. The first spectral methods computations were simulations of homogeneous turbulence on periodic domains. For that type of problem, the natural choice for representing functions is the family of (periodic) trigonometric polynomials, which representation is known as Fourier series.

In this section, we will discuss the behavior of these series when used to approximate smooth functions, considering the properties of both the continuous and discrete representation, come to an understanding of the factors determining the behavior of the approximating series.

Firstly, we define the set of functions given by

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

It can be proved that ϕ_n is an orthogonal system over the interval $(0, 2\pi)$ and also

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

For a complex-valued function u defined on $(0, 2\pi)$, we define the Fourier coefficients of u by

$$\hat{u}_n = \frac{1}{2\pi}\int_0^{2\pi} u(x)e^{-inx}dx, \quad k = 0, \pm1, \pm2, \ldots. \tag{2.4}$$

The integrals in (2.4) 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.

The relation (2.4) 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

$$a_n = \frac{1}{2\pi}\int_0^{2\pi} u(x)\cos(nx)dx, \quad n = 0, \pm1, \pm2, \ldots, \tag{2.5}$$

and

$$b_n = \frac{1}{2\pi}\int_0^{2\pi} u(x)\sin(nx)dx, \quad n = 0, \pm1, \pm2, \ldots. \tag{2.6}$$

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 $\hat{u}_{-n} = \hat{u}_n$.

As it is well known that if $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ is a separable Hilbert space and, if $\{\phi_k\}_{k \in I}$ is a countable orthonormal base of $\mathcal{H}$, then each $u \in \mathcal{H}$ can be written as

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

which is well known as the Fourier expansion of u. If we choose the base $B = span\{e^{inx} : |n| \leq \infty\}$, then for $u(x) \in L^2[0, 2\pi]$ the Fourier series $F[u]$ of the function u is defined as follows

$$F[u] \equiv \sum_{|n| \leq \infty} \hat{u}_n e^{inx}, \tag{2.7}$$

which is known as the classical continuous series of trigonometric polynomials, where $\hat{u}_n$ are the Fourier coefficients given by (2.4).

In the next subsections, we will investigate when and in what sense is the Fourier series convergent, its relation with the function u, and also, how fast approaches a function or rather to the solution of a partial differential equation. For this, two types of operators will be defined, those of projection and interpolation, which define the two most commonly used spectral scheme, Galerkin and collocation. First, the projection and interpolation operators will be studied independently, and at the end of this chapter, we will show the detailed implementation of each of them.

2.1.1 Projection Operator

The operator $\mathcal{P}_N$ is defined 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 (2.7) 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 (see Appendix A) show us that

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

Conversely, if for any complex sequence $\{\hat{u}_n\}$, $n = 0, \pm 1, \ldots$, 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 $\{\hat{u}_n\}$, $n = 0, \pm 1, \pm 2, \ldots$, 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

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

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.,

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

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

$$\begin{aligned} 2\pi\hat{u}_N &= \int_0^{2\pi} u(x)e^{-inx}dx \\ &= -\frac{1}{in}(u(2\pi) - u(0)) - \frac{1}{in}\int_0^{2\pi} u'(x)e^{inx}dx, \end{aligned}$$

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 (see Appendix A), in this particular case in L^2_p, 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*

$$u(x) = \frac{3}{5 - 4\cos(x)} \tag{2.13}$$

with its expansion coefficients

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

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.

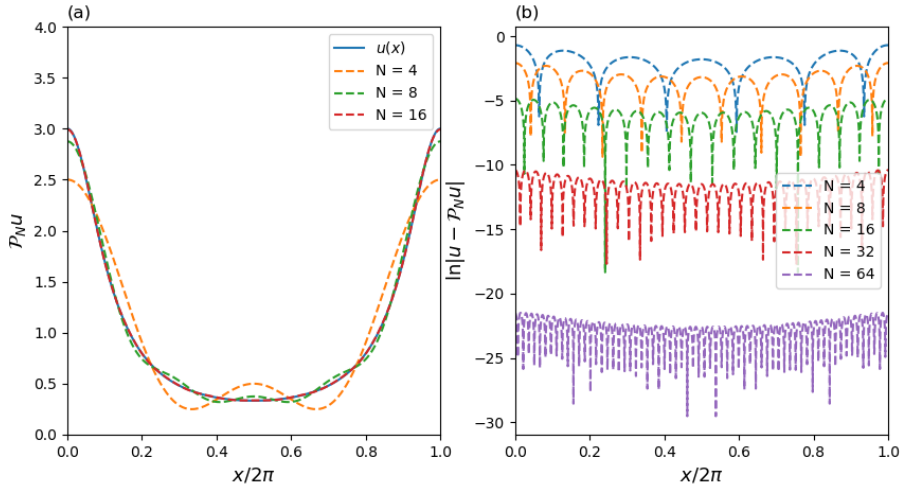


Figure 2.1: (a) Continuous Fourier series approximation of the equation (2.13). (b) The Pointwise error of approximation.

Example 2.2. *The expansion coefficients of the function*

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

are given by

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

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*

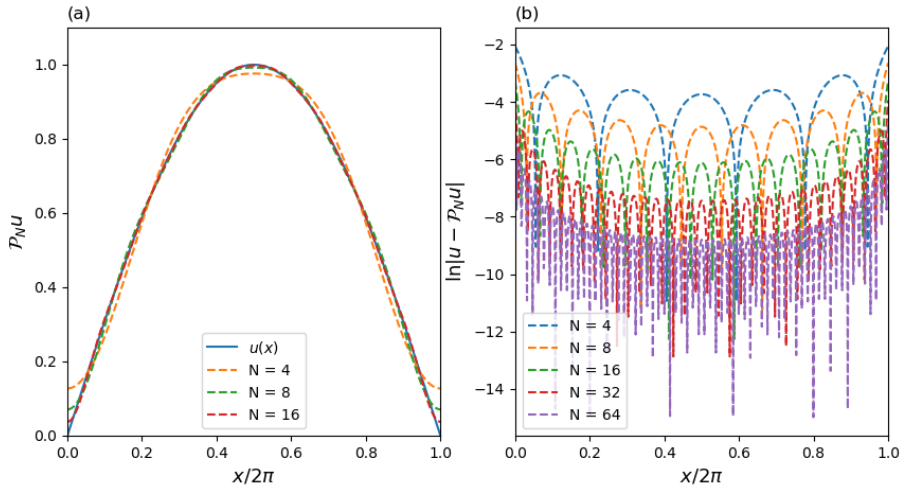


Figure 2.2: (a) Continuous Fourier series approximation of the equation (2.14). (b) The Pointwise error of approximation for increasing resolution.

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. The next example is to illustrate the above.

Example 2.3. *Set $u \in C_p^\infty[0, 2\pi]$, $\alpha > 0$, and define the following initial value problem as*

$$\begin{cases} \frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u(x,t)}{\partial x^2}, & x \in [0, 2\pi], \quad t > 0, \\ \\ u(x, 0) = u(x), & x \in [0, 2\pi], \quad t = 0. \end{cases}$$

Its possible to write the differentiation operator as $\mathcal{L} = \alpha \frac{\partial^2}{\partial x^2}$. Because u is a smooth function, it can be written as a truncated Fourier expansion as in (2.7), and then using the differentiation operator to obtain

$$\mathcal{L}\mathcal{P}_N u = \alpha \frac{\partial^2}{\partial x^2} \sum_{|n| \le \frac{N}{2}} \hat{u}_n(t) e^{inx} = -\alpha \sum_{|n| \le \frac{N}{2}} n^2 \hat{u}_n(t) e^{inx}.$$

On the other hand

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

so we get the following system of first-order ODEs

$$\sum_{|n| \le \frac{N}{2}} \frac{\partial}{\partial t} \hat{u}_n(t) e^{inx} = -\alpha \sum_{|n| \le \frac{N}{2}} n^2 \hat{u}_n(t) e^{inx},$$

which for every n has the solution given by

$$\hat{u}_n(t) = \hat{u}_n(0) e^{-n^2 \alpha t}.$$

Note that when n goes to infinity, $\hat{u}_n(t)$ tends to zero for all t. The above shows us the spectral convergence already mentioned before, and we also have to

$$\mathcal{L}(I - \mathcal{P}_N)u = -\alpha \sum_{|n| > \frac{N}{2}} n^2 \hat{u}_n(0) e^{-n^2 \alpha t} e^{inx}.$$

Therefore if N goes to infinity, then $\mathcal{P}_N \mathcal{L}(I - \mathcal{P}_N)u = 0$.

Remark. *Note that the operator in the previous example is included in the equation (1.3). Later we will see that when our main problem is considered, some of its properties are still preserved. Therefore, we will emphasize this operator to better understand its features.*

Approximation theory for smooth functions

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

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

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

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

where the interchange of the summation is allowed provided $u(x)$ has sufficient smoothness. Recall that it is possible to do the analysis with some equivalent norm, in this case it is possible to obtain the following norm denoted as $\|\cdot\|_{W_p^q[0,2\pi]}$ and defined by

$$\|u\|_{W_p^q[0,2\pi]} = \left(\sum_{|n|\leq\infty} (1 + n^{2q}) |\hat{u}_n|^2 \right)^{1/2}, \tag{2.16}$$

since for $u(x) \in C_p^q[0, 2\pi]$, $n \neq 0$ and $q > \frac{1}{2}$

$$(1 + n^{2q}) \leq \sum_{m=0}^{q} n^{2m} \leq (q + 1)(1 + n^{2q})$$

therefore is equivalent to $\|\cdot\|_{H_p^q[0,2\pi]}$.

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

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

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

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

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

The next two results that we will present show the behavior of the approximation of u and $\mathcal{L}u$ in terms of their derivatives, but now using the equivalent norm defined in (2.16).

***Theorem* 2.3.** *For any real r and any real q where $0 \leq q \leq r$, if $u(x) \in W_p^r[0, 2\pi]$, then there exists a positive constant C, independent of N, such that*

$$\|u - \mathcal{P}_{2N}u\|_{W_p^q[0,2\pi]} \leq \frac{C}{N^{r-q}}\|u\|_{W_p^r[0,2\pi]}.$$

Proof. Again by Parsevals identity yields

$$\|u - \mathcal{P}_{2N}u\|^2_{W_p^q[0,2\pi]} = 2\pi \sum_{|n|>N} (1 + |n|^{2q})|\hat{u}_n|^2.$$

First, notice that $|n| \geq 1$, and $1 + |n|^{2q} \geq 1$. Then if $q > 1/2$ we have to

$$(1 + |n|^{2q})^{-2q} \leq 1,$$

therefore

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

Since $|n| + 1 \geq N$, for any $0 \leq q \leq r$ we obtain

$$(1 + |n|^{2q}) \leq (1 + |n|)^{2q} = \frac{(1 + |n|)^{2r}}{(1 + |n|)^{2(r-q)}} \leq \frac{(1 + |n|)^{2r}}{N^{2(r-q)}}.$$

For any fixed $r > 1/2$ we have to

$$(1 + |m|)^{2r} \leq (r + 1)m^{2r} \leq (r + 1)(1 + m^{2r})$$

for some $m \in \mathbb{N}$. Then solving the following inequality

$$\left(\frac{1 + |m|}{m}\right)^{2r} \leq \left(1 + \frac{1}{m}\right)^{2r} \leq (r + 1)$$

therefore

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

Hence, for $n \geq m$ we have to

$$(1+|n|)^{2r} \leq (r+1)(1+n^{2r})$$

This immediately yields

$$\|u - \mathcal{P}_{2N}u\|^2_{W_p^q[0,2\pi]} \leq C \sum_{|n|>N} \frac{(1+n^{2r})}{N^{2(r-q)}} |\hat{u}_n|^2 \leq C \frac{\|u\|^2_{W_p^r[0,2\pi]}}{N^{2(r-q)}}.$$

□

***Theorem* 2.4.** *Let $\mathcal{L}$ be a constant coefficient differential operator*

$$\mathcal{L}u = \sum_{j=1}^{s} a_j \frac{d^j u}{dx^j}.$$

For any real r and any real q where $0 \leq q+s \leq r$, if $u(x) \in W_p^r[0,2\pi]$, then there exists a positive constant C, independent of N, such that

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

Proof. Using the definition of $\mathcal{L}$ we get

$$\begin{aligned}
\|\mathcal{L}u - \mathcal{L}\mathcal{P}_N u\|_{W_p^q[0,2\pi]} &\leq \left\| \sum_{j=1}^{s} a_j \frac{d^j u}{dx^j} - \sum_{j=1}^{s} a_j \frac{d^j \mathcal{P}_N u}{dx^j} \right\|_{W_p^q[0,2\pi]} \\
&\leq \max_{0\leq j\leq s} |a_j| \left\| \sum_{j=1}^{s} \frac{d^j}{dx^j}(u - \mathcal{P}_N u) \right\|_{W_p^q[0,2\pi]}.
\end{aligned}$$

Due to the triangle inequality and the definition of the norm W_p^q, we have to

$$\begin{aligned}
\|\mathcal{L}u - \mathcal{L}\mathcal{P}_N u\|_{W_p^q[0,2\pi]} &\leq \max_{0\leq j\leq s} |a_j| \sum_{j=1}^{s} \|u - \mathcal{P}_N u\|_{W_p^{q+s}[0,2\pi]} \\
&\leq C \|u - \mathcal{P}_N u\|_{W_p^{q+s}[0,2\pi]}.
\end{aligned}$$

Note that the above inequality is bounded by Theorem 2.3, and the result immediately follows. □

Now we can observe that it is now possible to estimate the error in L^2 using the W_p^q norm, these will give us a more practical way to analyze our main problem.

2.1.2 Interpolation Operator

The continuous Fourier series method requires the evaluation of the coefficients

$$\hat{u}_n = \frac{1}{2\pi}\int_0^{2\pi} u(x)e^{-inx}dx. \tag{2.17}$$

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.

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

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

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

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

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,

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

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 $f(x) = e^{inx}$, $|n| < N$. □

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

$$\widetilde{u}_n = \frac{1}{N\widetilde{c}_n}\sum_{j=0}^{N-1} u(x_j)e^{-inx_j}, \tag{2.19}$$

where we introduce the coefficients

$$\widetilde{c}_n = \begin{cases} 2 \text{ if } |n| = N/2, \\ \\ 1 \text{ if } |n| < N/2. \end{cases} \tag{2.20}$$

These relations define a new projection of u

$$\mathcal{I}_N u(x) = \sum_{|n|\le \frac{N}{2}} \widetilde{u}_n e^{inx} \tag{2.21}$$

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

$$u(x) = \frac{3}{3 - 4\cos(x)}. \tag{2.22}$$

Its expansion coefficients are

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

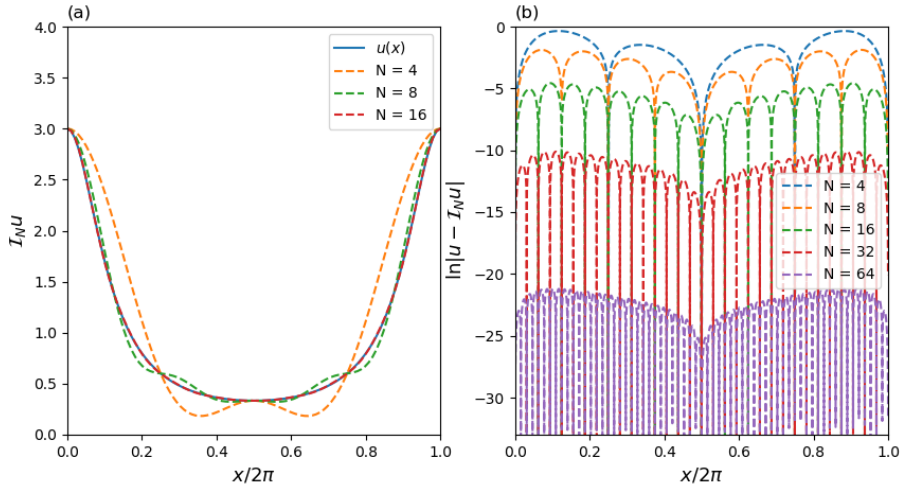


Figure 2.3: (a) Discrete Fourier series approximation of the equation (2.22). (b) Pointwise error of approximation for increasing resolution.

Example 2.5. *The expansion coefficients of the function*

$$u(x) = \sin(\frac{x}{2}), \tag{2.23}$$

are given by

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

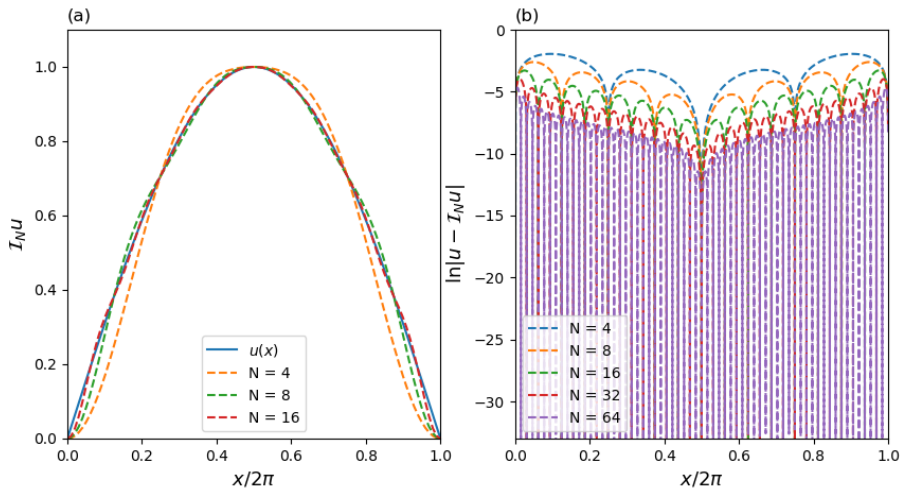


Figure 2.4: (a) Discrete Fourier series approximation of the equation (2.23). (b) Pointwise error of approximation for increasing resolution.

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* 2.6.** *Let the discrete Fourier transform be defined by Equations (2.19)-(2.21). 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 (2.19) into Equation (2.21) we obtain

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

Exchanging the order of the sum gives

$$\mathcal{I}_N u(x) = \sum_{j=0}^{N-1} u(x_j) g_j(x), \tag{2.24}$$

where

$$\begin{aligned} g_j(x) &= \sum_{|n| \le \frac{N}{2}} \frac{1}{N\widetilde{c}_n} e^{in(x-x_j)} \\ &= \frac{1}{N} \sin\left[N\frac{x-x_j}{2}\right] \cot\left[\frac{x-x_j}{2}\right] \end{aligned}$$

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 $g_j(x) \in \widetilde{B}_N$. □

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

$$\widetilde{u}_n = \frac{1}{N+1}\sum_{j=0}^{N} u(x_j)e^{-inx_j}, \tag{2.25}$$

to obtain the interpolation operator

$$\mathcal{J}_N u(x) = \sum_{|n|\leq \frac{N}{2}} \widetilde{u}_n e^{inx}. \tag{2.26}$$

Again as before, the quadrature formula is highly accurate.

***Theorem* 2.7.** *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,

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

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,

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

where

$$h_j(x) = \frac{1}{N+1}\frac{\sin(\frac{N+1}{2}(x-x_j))}{\sin(\frac{x-x_j}{2})} \tag{2.27}$$

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

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

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

where $\widetilde{c}_n$ is given by (2.20). Higher order derivatives can be obtained simply by further differentiating the basis functions.

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

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

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

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

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

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

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

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

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

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

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.

Results for the 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 2.5, we have

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

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

$$K^{-1} \|f_N\|_{L^2[0,2\pi]}^2 \leq \|f_N\|_N^2 \leq K \|f_N\|_{L^2[0,2\pi]}^2. \tag{2.33}$$

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.,

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

This phenomenon is known as aliasing.

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

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

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*

$$\widetilde{c}_n \widetilde{u}_n = \hat{u}_n + \sum_{\substack{|m| \leq \infty \\ m \neq 0}} \hat{u}_{n+Mm} \tag{2.34}$$

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

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

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

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

□

The conclusions of the previous discussion are equally valid in the number of odd or even points. An equivalent formulation of (2.34) 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

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

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

$$\begin{aligned} \sum_{|n|\leq N} \left| \sum_{\substack{|m|\leq\infty \\ m\neq 0}} \hat{u}_{n+Mm} \right|^2 &\leq \sum_{|n|\leq N} C_1 N^{-2r} \sum_{\substack{|m|\leq\infty \\ m\neq 0}} |n+Mm|^{2r} |\hat{u}_{n+Mm}|^2 \\ &\leq C_2 N^{-2r} ||u^{(r)}||^2_{L^2[0,2\pi]} \end{aligned}$$

□

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

***Theorem* 2.8.** *For any $u(x) \in W_p^r[0,2\pi]$ 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

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

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. □

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.

The above results are in the L^2 norm, but we can obtain essentially the same information about the derivatives, using the Sobolev norms. First, we need to obtain a Sobolev norm bound on the aliasing error.

Lemma 2.3. *Let $u(x) \in W_p^r[0, 2\pi]$, where $r > 1/2$. For any real q, for which $0 \le q \le r$, the aliasing error*

$$||\mathcal{A}_N||_{W_p^q[0,2\pi]} = \left(\sum_{-\infty}^{\infty} |\widetilde{c}_n \widetilde{u}_n - \hat{u}_n|^2 \right)^{1/2} \le C N^{-(r-q)} ||u||_{W_p^r[0,2\pi]}$$

Proof.

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

such that

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

The second factor is, as before, bounded by

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

provided $r > 1/2$ and $|n| \leq N$. Also, since $(1+|n|)^{2q} \leq C_2 N^{2q}$ for $|n| \leq N$ we recover

$$\begin{aligned}
&\sum_{|n|\leq N} (1+|n|)^{2q} \left| \sum_{\substack{|m|\leq\infty \\ m\neq 0}} \hat{u}_{n+Mm} \right|^2 \\
&\qquad \leq C_1 C_2 N^{-2(r-q)} \sum_{\substack{|m|\leq\infty \\ m\neq 0}} (1+|n+Mm|)^{2r} |\hat{u}_{n+Mm}|^2 \\
&\qquad \leq C_3 N^{-2(r-q)} ||u||^2_{W_p^r[0,2\pi]}
\end{aligned}$$

With this bound on the aliasing error, and the truncation error bounded by Theorem 2.3, we are now prepared to state the following Theorems. □

***Theorem* 2.9.** *Let $u(x) \in W_p^r[0,2\pi]$, where $r > 1/2$. Then for any real q for which $0 \leq q \leq r$, there exists a positive constant, C, independent of N , such that*

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

Proof. Note that the following inequality holds

$$\begin{aligned}
\|u - \mathcal{J}_{2N} u\|_{W_p^q[0,2\pi]} &\leq \|u - \mathcal{P}_{2N}\|_{W_p^q[0,2\pi]} + \|\mathcal{P}_{2N} - \mathcal{J}_{2N}\|_{W_p^q[0,2\pi]} \\
&\leq \|u - \mathcal{P}_{2N}\|_{W_p^q[0,2\pi]} + \|\mathcal{A}_N\|_{W_p^q[0,2\pi]}.
\end{aligned}$$

By using Theorems (2.3), and Lemma (2.3) are used for the first and second term on the right side respectively, the result immediately follows. □

***Theorem* 2.10.** *Let $\mathcal{L}$ be a constant coefficient differential operator*

$$\mathcal{L}u = \sum_{j=1}^{s} a_j \frac{d^j u}{dx^j}$$

For any real r and any real q where $0 \leq q + s \leq r$, if $u(x) \in W_p^r[0,2\pi]$, then there exists a positive constant, C, independent of N such that

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

Proof. Similarly as in above Theorem,

$$\begin{aligned}
\|\mathcal{L}u - \mathcal{L}\mathcal{J}_{2N} u\|_{W_p^q[0,2\pi]} &\leq \|\mathcal{L}u - \mathcal{L}\mathcal{P}_{2N} u\|_{W_p^q[0,2\pi]} + \|\mathcal{L}\mathcal{P}_{2N} u - \mathcal{L}\mathcal{J}_{2N} u\|_{W_p^q[0,2\pi]} \\
&\leq \|\mathcal{L}u - \mathcal{L}\mathcal{P}_{2N} u\|_{W_p^q[0,2\pi]} + C\|\mathcal{A}_N\|_{W_p^{q+s}[0,2\pi]},
\end{aligned}$$

and by Theorems (2.4), and (2.3) the Theorem is proved. □

2.2 Convergence Theory to Initial Value Problems

This section is one of the most important in this chapter since we will discuss the necessary tools for the convergence analysis of our main problem. First, we define the following initial value problem as follows

$$\begin{cases} \frac{\partial u}{\partial t} = Au + f(x,t), & x \in \mathcal{D}, \ \ t > 0, \\ u(x,t) = 0, & x \in \partial\mathcal{D}, \ \ t > 0, \\ u(x,0) = g(x), & x \in \mathcal{D}, \ \ t = 0, \end{cases} \tag{2.35}$$

where $\mathcal{D}$ is a spatial domain with boundary $\partial\mathcal{D}$, A is a linear (spatial) differential operator.

It is assumed that, for each t, $u(x,t)$ is an element of a Hilbert space $\mathcal{H}$ with inner product $\langle\cdot,\cdot\rangle$ and norm $||\cdot||$, and for each $t > 0$, the solution $u(t)$ belongs to the subspace $\mathcal{B}$ of $\mathcal{H}$ consisting of all functions $u \in \mathcal{H}$ satisfying $u(t) = 0$ on $\partial\mathcal{D}$. We do not require that $u(x,0) = g(x) \in \mathcal{B}$ but only that $u(x,0) \in \mathcal{H}$. The operator A is typically an unbounded differential operator whose domain is dense in, but smaller than, $\mathcal{H}$.

For notational convenience we shall assume henceforth that A is time independent, and we will often denote $u(x,t)$ by $u(t)$ when discussing u as a function of t. It is well known that if A is an infinitesimal generator of a C_0-semigroup $T(t)$ (See Appendix A), namely, the evolution operator e^{At}, and f is continuously differentiable, then the formal and unique solution of (2.35) is

$$u(t) = e^{At}u(0) + \int_0^t e^{A(t-s)}f(s)ds, \tag{2.36}$$

The above is known as the non-homogeneous Cauchy solution problem, and can be justified under the conditions that $f(t)$, $Af(t)$ and $A^2f(t)$ exist and are continuous functions of t in the norm $||\cdot||$ for all $t \geq 0$ **(see Richtmyer and Morton [29])**.

Note that the solution 2.36 is the superposition (sum) of the general solution of the homogeneous ($f(x,t) \equiv 0$) and not homogeneous problem. In fact, in both expressions, the first term represents the solution to the problem, while the second one incorporates the contribution of the second member. Finally, it should be noted that the second term of the expression, corresponding to the contribution of the second member f, is a temporary average of expressions of the form $T(\cdot, t-s) * f(\cdot, s)$, which are actually the solutions of the homogeneous equation at the instant $t - s$ that begin in the initial data $f(s)$. We see that the contribution of a second member in the equation is similar to that of the initial data averaged over time.

The semi-discrete approximations to (2.35) to be studied here are of the form

$$\frac{\partial u_N(x,t)}{\partial t} = A_N u_N(x,t) + f_N(x,t) \tag{2.37}$$

where, for each t, $u_N(x,t)$ belongs to an N-dimensional subspace $\mathcal{B}_N$ of $\mathcal{B}$, and A_N is a linear operator from $\mathcal{H}$ to $\mathcal{B}_N$ of the form

$$A_N = \mathcal{P}_N A \mathcal{P}_N$$

where $\mathcal{P}_N$ is a projection operator in some sense of $\mathcal{H}$ onto $\mathcal{B}_N$ and $f_N = \mathcal{P}_N f$. We shall assume that $\mathcal{B}_N \subset \mathcal{B}_M$ when $N < M$. To be definite, we shall also assume the initial conditions for the approximate equations (2.37) to be $u_N(0) = \mathcal{P}_N u(0)$, where $u(0) = g(x)$ is the initial condition of (2.35).

The fundamental problem of the numerical analysis of initial value problems is to find conditions under which $u_N(x,t)$ converges to $u(x,t)$ as $N \to \infty$ for some time interval $0 \le t \le T$ and to estimate the error $||u - u_N||$. To do this, we will denote $u(t)$ and $u_N(t)$ the solutions to the problem (2.35) and its approximation given by (2.37) respectively, then we will proceed to define the concept of well-posed problem, as well as the definitions of stability, consistency, and convergence as follows.

Definition 2.2.1. A problem is well defined if, for each initial condition $g \in C^r$ and for each time $T_0 > 0$ exists a unique solution $u(x,t)$ such that

$$\|u(t)\|_{L^2(\mathcal{D})} \le Ce^{\alpha t}\|g\|_{\mathcal{H}^p(\mathcal{D})} \quad 0 \le t \le T_0,$$

for $p \le r$, some real α, and some positive constant C. For $p = 0$ it is said that it is strongly defined, i.e, for the norm $\mathcal{L}^2$.

From the above we can say that if the problem (2.35) is well posed, the evolution operator is a bounded linear operator from $\mathcal{H}$ to $\mathcal{B}$. Boundedness implies that the domain of the evolution operator can be extended in a standard way from the domain of A to the whole space $\mathcal{H}$ **(Richtmyer and Morton [29] (1967, p. 34))**.

Definition 2.2.2. An approximation scheme is convergent if

$$\|u(t) - u_N(t)\|_{L^2(\mathcal{D})} \to 0 \quad \text{as} \quad N \to \infty$$

for all $t \in [0,T]$, $u(0) \in \mathcal{H}$, , and $u_N(0) \in B_N$.

Also for any spatial operator $\mathcal{L}$ we define the following.

Definition 2.2.3. An approximation scheme is consistent if

$$\|\mathcal{P}_N \mathcal{L}(I - \mathcal{P}_N)u)\|_{L^2(\mathcal{D})} \to 0 \quad \text{as} \quad N \to \infty$$

$$\|\mathcal{P}_N u(0) - u_N(0)\|_{L^2(\mathcal{D})} \to 0 \quad \text{as} \quad N \to \infty$$

for all $u(0) \in \mathcal{H}$ and $u_N(0) \in B_N$.

Definition 2.2.4. An approximation scheme is stable if

$$\|e^{\mathcal{L}_N t}\|_{L^2(\mathcal{D})} \leq K(t), \quad \forall N$$

where $K(t)$ is a finite function of t independet of N and with the norm associated with the operator norm

$$\|e^{\mathcal{L}_N t}\|_{L^2(\mathcal{D})} = \sup_{u \in B} \frac{\|e^{\mathcal{L}_N t} u\|_{L^2(\mathcal{D})}}{\|u\|_{L^2(\mathcal{D})}}$$

$K(t)$ is independent of N and bounded for any $t \in [0, T]$.

Note that the stability of the approximation is closely related to the question of well-posed for the partial differential equation, it also guarantees that the solution remains bounded as N approaches infinity.

The classical and well-known Lax-Richtmyer equivalence theorem relates the above definitions as follows.

Theorem* 2.11. *(Lax-Richtmyer equivalence theorem) *A consistent approximation to a linear well-posed partial differential equation is convergent if and only if it is stable.*

Proof. To show that stability implies convergence we use (2.35) and (2.37) to obtain

$$\frac{\partial(u - u_N)}{\partial t} = A_N(u - u_N) + Au - A_N u + f - f_N$$

Thus,

$$\begin{aligned} u(t) - u_N(t) =& e^{A_N t}[u(0) - u_N(0)] \\ &+ \int_0^t e^{A_N(t-s)}[Au(s) - A_N u(s) + f(s) - f_N(s)]ds. \end{aligned}$$

Under the assumption of stability and consistency of the approximation, and using the triangle inequality we obtain the estimate

$$\begin{aligned} \|u(t) - u_N(t)\|_{L^2(\mathcal{D})} \leq& K(t)\|u(0) - u_N(0)\|_{L^2(\mathcal{D})} \\ &+ \int_0^t K(t-s)\left(\|Au(s) - A_N u(s)\|_{L^2(\mathcal{D})} + \|f(s) - f_N(s)\|_{L^2(\mathcal{D})}\right) ds. \end{aligned}$$

Thus, if $u(t)$ belongs to the dense subspace of $\mathcal{H}$ satisfying the Definition 2.2.3 and if $f(t)$ belongs to the dense subspace of $\mathcal{H}$ satisfying $\|f - \mathcal{P}_N f\| \to 0$ as $N \to \infty$, then $\|u(t) - u_N\| \to 0$ as $N \to \infty$. Since all solutions $u(t)$ of (2.35) can be approximated arbitrarily well by functions satisfying consistency, the proof that stability

implies convergence is completed.

Conversely, to show that convergence implies stability, we first observe that, for any $u \in \mathcal{H}$, $\|e^{A_N t}u\|$ is bounded for all N and each fixed t. In fact, convergence implies

$$0 \leq \left| \|e^{A_N t}u\| - \|e^{At}u\| \right| \leq \|e^{A_N t}u - e^{At}u\| \to 0, \qquad N \to \infty,$$

while well-posedness requires that $||e^{At}u||$ is finite. However, $\max ||e^{A_N t}u||$ may depend on u and on t, so stability is not yet proved. To complete the proof we use the fact that $\mathcal{H}$ is a Hilbert space. The principle of uniform boundedness implies that if $\|e^{A_N t}u\|$ is bounded as $N \to \infty$ for each t and $u \in \mathcal{H}$, then $\|e^{A_N t}\|$ is bounded as $N \to \infty$ for each t. This proves stability and completes the proof of the equivalence theorem. □

Note that the above Theorem, the study of the convergence of discrete approximations to the solutions of lineal initial value problems is reduced to the study of the stability of the discrete approximations.

The intention of the proof of previous theorem is to give us an idea of how to investigate under what conditions it is possible to adapt or extend the above to nonlinear problems. To do this, in the space H_p^q defined as in (2.15) we will define the operator A defined in (2.35) in via bilinear as $a(\cdot,\cdot) : H_0^1(\mathcal{D}) \times H_0^1(\mathcal{D}) \to \mathbb{R}$, where $H_0^1(\mathcal{D})$ is defined in Appendix A, satisfying

$$a(u,u) \leq C_0 \|u\|_{H^1}^2, \quad for \;\; u \in H_0^1(\mathcal{D}), \tag{2.38}$$

and

$$|a(u,v)| \leq C_1 \|u\|_{H^1} \|v\|_{H^1}, \quad \text{for} \;\; u, v \in H_0^1(\mathcal{D}). \tag{2.39}$$

Also for $u \in D(A) = H_p^2 \cup H_0^1(\mathcal{D})$, we define $(Au, v) = a(u,v)$ for any $v \in H_0^1(\mathcal{D})$, and $Au \in L^2(\mathcal{D})$.

It is well known that these conditions are satisfied in general by a parabolic equation, and we also by Lemma A.1, we have to $-A$ is the infinitesimal generator of an analytic semigroup $\left\{e^{-tA}\right\}_{t \geqq 0}$. Also the usual estimates for analytic semigroups hold for $\alpha \geq 0$, $t > 0$,

$$\|A^\alpha e^{-tA}\| \leq C_\alpha t^{-\alpha}. \tag{2.40}$$

Now we consider the next approximation in the weak form as in (1.5). First, let $\left\{V_N\right\}$ be sequence of finite-dimensional subspaces of H_0^1, and $a_N(\cdot,\cdot)$, $V_N \times V_N \to \mathbb{R}$ be discretization of the bilinear form $a(\cdot,\cdot)$ satisfying

$$a_N(\varphi,\varphi) \leq C_0' \|\varphi\|_{H^1}^2 \;\; \text{for} \;\; \varphi \in V_N, \tag{2.41}$$

$$|a_N(\varphi,\psi)| \leqq C_1'\|\varphi\|_{H^1}\|\psi\|_{H^1} \quad \text{for} \quad \varphi,\psi \in V_N. \tag{2.42}$$

We consider the following approximation scheme for (2.35):

$$\begin{cases} u_N(t): \ [0,T] \to V_N, \\ \frac{du_N}{dt} + A_N u_N = f_N(t), \\ u_N(0) = \mathcal{P}_N g, \end{cases} \tag{2.43}$$

where A_N is defined by $A_N : V_N \to V_N$ as

$$(A_N v, \varphi) = a_N(v,\varphi) \ \text{ for any } \ v,\varphi \in V_N,$$

and $f_N(t) \in C([0,T],V_N)$ is an approximation of $f(t)$, and $\mathcal{P}_N$ is an operator (a projection operator in some sense) chosen in each application,

$$\mathcal{P}_N : X = L^2(\mathcal{D}) \to V_N.$$

Conditions (2.41) and (2.42) guarantee that the numerical scheme (2.43) has some smoothing properties as shown in the following lemma proved in [31].

Lemma 2.4. *There exist a constant $\delta \in (\pi/2,\pi)$, such that $\sigma(A_N) \supset \sum_\delta = \{\xi \in \mathbb{C}, |arg(\xi)| < \delta\}$. Moreover, we have the following estimates for $0 \leqq \alpha \leqq 1, \lambda \in \sum_\delta$,*

$$\|A_N^\alpha e^{-A_N t}\varphi\|_{0,w} \leqq \frac{C_\alpha}{t^\alpha}\|\varphi\|_{0,w}, \quad \varphi \in V_N, \tag{2.44}$$

where C_α is independent of N and φ.

Now we will consider the nonlinear evolutionary equation in $L^2(\mathcal{D})$ as follows

$$\begin{cases} \frac{du}{dt} + Au + F(t,u) = 0, \\ u(0) = g, \end{cases} \tag{2.45}$$

where A is the operator considered as before, and F be a nonlinear function defined in the space $H_0^1(\mathcal{D})$. We will assume that this system has a unique solution that satisfies some regularity requirements specified below.

Thus, an approximation scheme is given by:

$$\begin{cases} u_N(t) : [0,T] \to V_N, \\ \frac{du_N}{dt} + A_N u_N + F_N(t,u_N) = 0, \\ u_N(0) = P_N g \in V_N, \end{cases} \tag{2.46}$$

where $F_N(t,v) : [0,T] \times V_N \to V_N$ is the approximation of the nonlinear term.

Using the variation of constants formula as in (2.36), we can write (2.45) and (2.46) in integral forms

$$u(t) = e^{-At}a - \int_0^t e^{-A(t-s)}F(s,u(s))ds, \tag{2.47}$$

$$u_N(t) = e^{-A_N t}P_N a - \int_0^t e^{-A_N(t-s)}F_N(s,u_N(s))ds \tag{2.48}$$

It is natural to solve above equations using a fixed point theorem to find a weak solution, and then to show that, under appropriate assumptions, exist a unique weak solution and is the classical, i.e., prove that the operator $G(u)$ is a contraction, where G is defined as

$$G(u)(t) = e^{-At}a - \int_0^t e^{-A(t-s)}F(s,u(s))ds \tag{2.49}$$

For example, one sufficiently condition to uniqueness and existence solution is, if F be a Lipschitz continuous function with respect to u on any bounded subset of X, with Lipschitz constant independent of t, i.e.,

$$\|F(t,u) - F(t,v)\| \leq L\|u-v\|, \quad t \in [0,T], \quad \text{for all } \ u,v \in X$$

In general, (2.47) makes sense under some smoothness conditions for u and F. In the next chapter, we will investigate what conditions are sufficient for $u_N(t)$ to convergence to $u(t)$, and also obtain its estimate.

Remark. *Although it seems strange, the above says that our numerical method should converge in some sense for linear problems and the approximation for the nonlinear term should be consistent and stable. Keep in mind that it is a very weak form of the stability assumption that corresponds to the local condition of Lipschitz in F necessary for the existence of a local solution.*

2.3 Fourier Spectral Methods

According to the projection and interpolate operators discussed in above section, we will proceed to study the methods that arise. for this, we will define the following initial value problem defined in some space of Hilbert $\mathcal{H}$ on the interval $I = [0, 2\pi]$. Set $u(x,t) \in C^1[0,T] \times \mathcal{H}[I]$ be a periodic function in I, and define the following problem

$$\begin{cases} \frac{\partial u}{\partial t} = \mathcal{L}u + f(x,t), & x \in [0,2\pi], \ \ t > 0, \\ \\ u(x,0) = u_0(x), & x \in [0,2\pi], \ \ t = 0. \end{cases} \tag{2.50}$$

In the following subsections, we will present the schemes that we will implement in the next chapter to our main problem and that we will then analyze independently. For the moment we will approach the methods in a general way using the

problem given by (2.50).

The methods to be formulated in the following sections are based on the methods of residual weights. The main idea of these methods is to look for solutions $\bar{u}$ in a subspace of finite dimension $\hat{B}_N$ of some Hilbert $\mathcal{H}$ space satisfying the following: We will define the R_N function, which we will call as the residual function and is given by

$$R_N(\bar{u}) = \frac{\partial \bar{u}}{\partial t} - \mathcal{L}\bar{u} - f(x,t).$$

These subspace are denoted as before, i.e.,

$$\hat{B}_N = span\left\{\phi_n(x) = e^{inx} : |n| \leq N\right\},$$

then for some functions $\hat{u}_n$, $\bar{u}$ can be expressed by

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

The key to these methods to approximate the solutions to the problem (2.50) is as follows. For some functions $W_n(x)$, called weight functions, the following must be satisfied

$$\int_I R_N(x,t) W_n(x) dx = 0, \quad \forall |n| \leq N.$$

Now, if we define $W_n(x) = \varphi_n(x)$, where the φ_n functions will be called test functions, the above can be write equivalently as follows

$$\langle R_N(\bar{u}), \varphi_n \rangle = 0, \quad \forall |n| \leq N$$

We can observe that when the residue tends to zero, the approximation $\bar{u}$ approximates the exact solution of the problem (2.50).

The choice of trial functions ϕ_n defines the spectral method, in this case, the family of functions defined in $\hat{B}_N$ defines the Fourier spectral methods.

2.3.1 Fourier-Galerkin Method

The main objective of the Fourier-Galerkin method is to approximate the function u using the projection operator $\mathcal{P}_N$ described in the previous sections to satisfy the problem given by (2.50) and minimize the error by orthogonal projection ($R_N(x) = 0$) of the original problem.

This method is defined by the choice of test functions as $W_n = \frac{\partial u_N}{\partial \hat{u}_n} = \phi_n$. So, the problem to solve is the following

$$\int_I R_N(x,t)\overline{\phi_n(x)}dx = \int_I \left(\frac{\partial u_N}{\partial t} - \mathcal{L}u_N - f_N(x,t)\right)\overline{\phi_n(x)}dx = 0,$$

where $f_N = \mathcal{P}_N f$, which must be satisfied for every $t \in [0,T]$, and for all $|n| \leq N$. Moreover, due to the orthogonality property of the test and trial functions we can write the problem as a set of $2N+1$ ODEs given by

$$\frac{d\hat{u}_n(t)}{dt} = \frac{d}{dt}\langle u_N(t), \phi_i(x)\rangle = \langle \mathcal{L}u_N(t), \phi_i(x)\rangle + \langle f_N(t), \phi_i(x)\rangle = F(t, \hat{u}_n)$$

and its corresponding initial conditions are

$$u_N(x,0) = \sum_{|n|\leq N} \hat{u}_n(0)e^{inx}, \quad \hat{u}_n(0) = \frac{1}{2\pi}\int_0^{2\pi} u_0(x)e^{-inx}dx$$

In general, Fourier Galerkin scheme for the problem given by (2.50) can be formulated as follows

$$\begin{cases} u_N(t) : [0,T] \to V_N, \text{ such that for every } \phi \in V_N \\ \left\langle \frac{\partial u_N(t)}{\partial t} - \mathcal{L}u_N(t) - f_N(t), \phi \right\rangle = 0, \\ u_N(0) = \mathcal{P}_N u_0(x), \end{cases} \tag{2.51}$$

The orthogonal projection is given by the restriction $(u - \mathcal{P}_N u, \phi) = 0$, in the space $V_N = \hat{B}_N \cap \mathcal{H}$ and $\mathcal{P}_N : L^2_p(I) \to V_N$, $\forall \phi \in V_N$. So the approximations have the following form,

$$V_N = \left\{ u : u(x,t) = \sum_{|n|\leq N} \hat{u}_n(t)e^{inx} \right\}$$

In general, the coefficients $\hat{u}_n(t)$ of the approximation are not equal to Fourier coefficients; only if we obtain the exact solution of the problem will they be equal. If the residual is smooth enough, this requirement implies that the residual itself is small. In particular, if the residual itself lives in the space V_N , the orthogonal complement must be zero.

2.3.2 Fourier-Collocation Methods

This method is similar to Fourier-Galerkin, The difference is that we seek an approximation in the space of $2N$th-order polynomials $\widetilde{B}_N$ for $u_N(x,t) \in C^1[0,T] \times S_N[I]$, where $S_N = \widetilde{B}_N \cap \mathcal{H}$. We define a set of even points number $x_j \in [0, 2\pi)$ given by

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

Recall that the space $\widetilde{B}_N$ is defined as

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

Then we can write the residual function as follows

$$R_N(x_j, t) = \frac{du_N(x_j, t)}{dt} - \mathcal{L}u_N(x_j, t) - \mathcal{J}_N f(x_j, t),$$

which must be satisfied for each x_j, $j = 0, 1, \ldots, 2N$, and as before the method require again that the residual satisfies

$$\int_I R_N(x_j, t) W_j(x) dx = 0, \ \text{ for } \ j = 0, 1, \ldots, 2N.$$

Using $W_j(x) = \delta(x - x_j)$, we have that R_N must be zero in the grid-points, i.e,

$$R_N(x_j, t) = 0, \ \text{ for } \ j = 0, 1, \ldots, 2N.$$

leading $2N + 1$ ODEs to determine the point values $u_N(x_j, t)$ as follows

$$\frac{du_N(x_j, t)}{dt} = \mathcal{L}u_N(x_j, t) + f(x_j, t), \ \text{ for } \ j = 0, 1, \ldots, 2N.$$

Recall that we can expressing $u_N(x, t)$ in terms of Lagrange polynomials $g_j(x)$, i.e., the approximation can be written as follows

$$u_N(x, t) = \sum_{j=0}^{2N} u_N(x_j, t) g_j(x)$$

where $g_j(x)$ satisfies $g_j(x_i) = \delta_{ij}$, and then we have that $\mathcal{J}_N u(x_j) = u(x_j)$ for every j.

Therefore we can formulating the Fourier Collocation method as follows

$$\begin{cases} u_N(x_j, t) : [0, T] \to S_N, \ \text{ such that for every } \ x_j \\ \frac{\partial u_N}{\partial t}(x_j) - \mathcal{L}u_N(x_j) = f(x_j), \\ u_N(x_j, 0) = u_0(x_j). \end{cases} \tag{2.52}$$

We can write the approximation to solution of problem given by (2.50) as

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

where $\widetilde{u}_n(t)$ are given by

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

Then, u_N must satisfies the problem

$$\begin{cases} u_N(t) : [0,T] \to S_N, \text{ such that for every } \phi \in S_N \\ \left\langle \frac{\partial u_N(t)}{\partial t} - \mathcal{L}u_N(t) - f_N(t), \phi \right\rangle = 0, \\ u_N(0) = \mathcal{J}_N u_0(x), \end{cases} \tag{2.53}$$

where $f_N = \mathcal{J}_N f$.

2.4 Semi-Bounded Operator

In this section, we will discuss the linear and nonlinear parts of the equation (1.3) separately, but first, we will analyze a type of operator of interest.

A special case of a well posed problem is if $\mathcal{L}$ is semi-bounded in the Hilbert space scalar product, i.e., $\mathcal{L} + \mathcal{L}^* \leq \beta I$ for some constant β and $\mathcal{L}^*$ is the adjoint operator. For example, given the solution $u(x,t) \in (\mathcal{H}, \langle \cdot, \cdot \rangle)$, where $\mathcal{H}$ is some Hilbert space, and supposes that is the scalar product in $L^2[0, 2\pi]$, which satisfies the following problem with initial condition suitable $u(0)$

$$\frac{\partial u}{\partial t} = \mathcal{L}u. \tag{2.54}$$

Lets show that equation above with a semi-bounded operator $\mathcal{L}$ is well posed. To show this, we estimate the derivative of the norm by considering

$$\frac{d}{dt}\|u\|^2 = \frac{d}{dt}\langle u, u\rangle = \frac{d}{dt}\int_0^{2\pi} u\bar{u}dx = \int_0^{2\pi} \frac{du}{dt}\bar{u}dx + \int_0^{2\pi} u\frac{d\bar{u}}{dt}dx, \tag{2.55}$$

thus, we have to

$$\frac{d}{dt}\|u\|^2 = \langle \frac{du}{dt}, u\rangle + \langle u, \frac{du}{dt}\rangle.$$

By using adjoint operator definition

$$\begin{aligned} \frac{d}{dt}\|u\|^2 &= \langle \mathcal{L}u, u\rangle + \langle u, \mathcal{L}u\rangle = \langle u, \mathcal{L}^*u\rangle + \langle u, \mathcal{L}u\rangle \\ &= \langle u, (\mathcal{L} + \mathcal{L}^*)u\rangle. \end{aligned}$$

Since $\mathcal{L} + \mathcal{L}^* \leq \beta I$, we have $\frac{d}{dt}\|u\|^2 \leq \beta\|u\|^2$, and so $\frac{d}{dt}\|u\| \leq \beta\|u\|$, which means that the norm is bounded, i.e.,

$$\|u(t)\| \leq e^{\beta t}\|u(0)\|.$$

Therefore the problem is well posed.

To illustrate the above, let's consider the following examples of well posed problems.

Example 2.6. *Let $u \in C_p^\infty[0,2\pi]$, $\alpha > 0$, and considering the following initial value problem*

$$\begin{cases} u_t = \alpha u_{xx}, \quad x \in [0,2\pi], \quad t > 0, \\ \\ u(x,0) = u_0(x), \quad x \in [0,2\pi]. \end{cases} \tag{2.56}$$

We define $A = \alpha \frac{\partial^2}{\partial x^2}$. We will show that the operator A is self-adjoint as follows

$$\langle Au, v\rangle_{L^2[0,2\pi]} = \alpha \int_0^{2\pi} \frac{\partial^2 u}{\partial x^2} \bar{v} dx.$$

Integration by parts yields

$$\begin{aligned} \alpha \left[\bar{v} \frac{\partial u}{\partial x}\Big|_0^{2\pi} - \int_0^{2\pi} \frac{\partial u}{\partial x}\frac{\partial \bar{v}}{\partial x} dx \right] &= -\alpha \left[\int_0^{2\pi} \frac{\partial u}{\partial x}\frac{\partial \bar{v}}{\partial x} dx \right] \\ &= -\alpha \left[u \frac{\partial \bar{v}}{\partial x}\Big|_0^{2\pi} - \int_0^{2\pi} u \frac{\partial^2 \bar{v}}{\partial x^2} dx \right] \\ &= \alpha \int_0^{2\pi} u \frac{\partial^2 \bar{v}}{\partial x^2} dx \\ &= \langle u, A^* v\rangle_{L^2[0,2\pi]}. \end{aligned}$$

From the above, we can observe that $A = A^*$, and we also have

$$\langle Au, u\rangle_{L^2[0,2\pi]} = -\alpha \int_0^{2\pi} \left(\frac{\partial u}{\partial x} \right)^2 dx \leq 0,$$

so that the operator A is semi-bounded, and following the steps as in (2.55) we get

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|u\|^2 &= \langle Au, u\rangle_{L^2[0,2\pi]} \\ &= -\alpha \int_0^{2\pi} \left(\frac{\partial u}{\partial x} \right)^2 dx \leq 0. \end{aligned}$$

Therefore the problem is well-posed, i.e.,

$$\|u(x,t)\| \leq \|u(x,0)\|$$

Example 2.7. *Let $u \in H_p^1([0,2\pi])$ a periodic function such that is the solution of the following initial value problem*

$$\begin{cases} \frac{\partial u(x,t)}{\partial t} = -u(x,t)\frac{\partial u(x,t)}{\partial x}, \\ \\ u(x,0) = u_0(x), \quad x \in [0,2\pi]. \end{cases} \tag{2.57}$$

Multiplying for u and integration over the space

$$\frac{1}{2}\frac{d}{dt}\int_0^{2\pi} u^2(x,t) dx = -\int_0^{2\pi} u^2(x,t)\frac{\partial u(x,t)}{\partial x} dx = -\frac{1}{3}u^3(x,t)\Big|_0^{2\pi} = 0.$$

So we have to

$$\|u(x,t)\| = \|u(x,0)\|,$$

and therefore the problem is well posed.

Remark. *The last equation in above example means that the energy is conserved. This problem is often referred to as the Burgers' equation without viscosity, and is in some sense the simplest nonlinear conservation law, also it expresses that u is conserved with a flux density given by $f(u) = \frac{u^2}{2}$.*

The following result ensures the stability of an approximation using Fourier-Galerkin when it is shown that the operator $\mathcal{L}$ is semi-bounded.

***Theorem* 2.12.** *Given the problem $\frac{\partial u}{\partial t} = \mathcal{L}u$, where the operator $\mathcal{L}$ is semi-bounded in the usual $L^2[0, 2\pi]$ scalar product. Then the Fourier-Galerkin method is stable.*

Proof. First, we show that $\mathcal{P}_N = \mathcal{P}_N^*$. We begin with the simple observation that,

$$\langle u, \mathcal{P}_N v\rangle = \langle \mathcal{P}_N u, \mathcal{P}_N v\rangle + \langle (I - \mathcal{P}_N)u, \mathcal{P}_N v\rangle.$$

The second term on the right side is the scalar product of the projection of u on the complement of the space $\hat{B}_N$ with the projection of v on the space $\hat{B}_N$, and hence

$$\langle u, \mathcal{P}_N v\rangle = \langle \mathcal{P}_N u, \mathcal{P}_N v\rangle.$$

By the same argument, we find that

$$\langle \mathcal{P}_N u, v\rangle = \langle \mathcal{P}_N u, \mathcal{P}_N v\rangle.$$

Therefore, $\langle \mathcal{P}_N u, v\rangle = \langle u, \mathcal{P}_N v\rangle$, i.e. $\mathcal{P}_N = \mathcal{P}_N^*$. Now, the Fourier-Galerkin method involves seeking the trigonometric polynomial u_N such that

$$\frac{\partial u_N}{\partial t} = \mathcal{P}_N \mathcal{L} \mathcal{P}_N u_N = \mathcal{L}_N u_N,$$

and we can observe that if $\mathcal{L}$ is semi-bounded, then

$$\begin{aligned}\mathcal{L}_N + \mathcal{L}_N^* &= \mathcal{P}_N \mathcal{L} \mathcal{P}_N + \mathcal{P}_N \mathcal{L}^* \mathcal{P}_N \\ &= \mathcal{P}_N(\mathcal{L} + \mathcal{L}^*)\mathcal{P}_N \leq 2\alpha \mathcal{P}_N\end{aligned}$$

Following Equation (2.55) this leads to the stability estimate

$$\|u_N(t)\| \leq e^{\alpha t}\|u_N(0)\|,$$

which means that it is stable. □

Note the following, if we considering the approximation using $\mathcal{J}_N$, and the fact that the Fourier coefficients of the discrete approximation are sufficiently close to those of the continuous approximation, then

$$\begin{aligned}\langle u, \mathcal{J}_N v\rangle_N &= \langle \mathcal{J}_N u, \mathcal{J}_N v\rangle_N + \langle (I - \mathcal{J}_N)u, \mathcal{J}_N v\rangle_N \\ &= \langle \mathcal{J}_N u, \mathcal{J}_N v\rangle,\end{aligned}$$

where the last equation is because by Theorem 2.7, also using the same argument as above we have to

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

so we have $\mathcal{J}_N = \mathcal{J}_N^*$, and therefore the above theorem holds to Fourier-Collocation method using the operator $\mathcal{J}_N$.

By Example 2.3 it is clearly that the approximation to problem given in Example 2.6 is consistent, and also is well defined and stable. Therefore, by Theorem 2.11 the approximation is convergent.

Furthermore, recall by Example 2.3 that the solution to approximate problem is

$$u_N(x,t) = \sum_{|n|\le \frac{N}{2}} \hat{u}_n(t) e^{inx}$$

where $\hat{u}_n(t)$ is

$$\hat{u}_n(t) = \hat{u}_n(0) e^{-n^2 t}$$

Therefore estimating convergence we have to

$$\|u - u_N\| = \left(\sum_{|n|>\frac{N}{2}} |\hat{u}_n(0) e^{-n^2 t}|^2 \right)^{1/2} \le e^{-N^2 t} \sum_{|n|>\frac{N}{2}} |\hat{u}_n(0)|^2 \le e^{-N^2 t} \|u(0)\|.$$

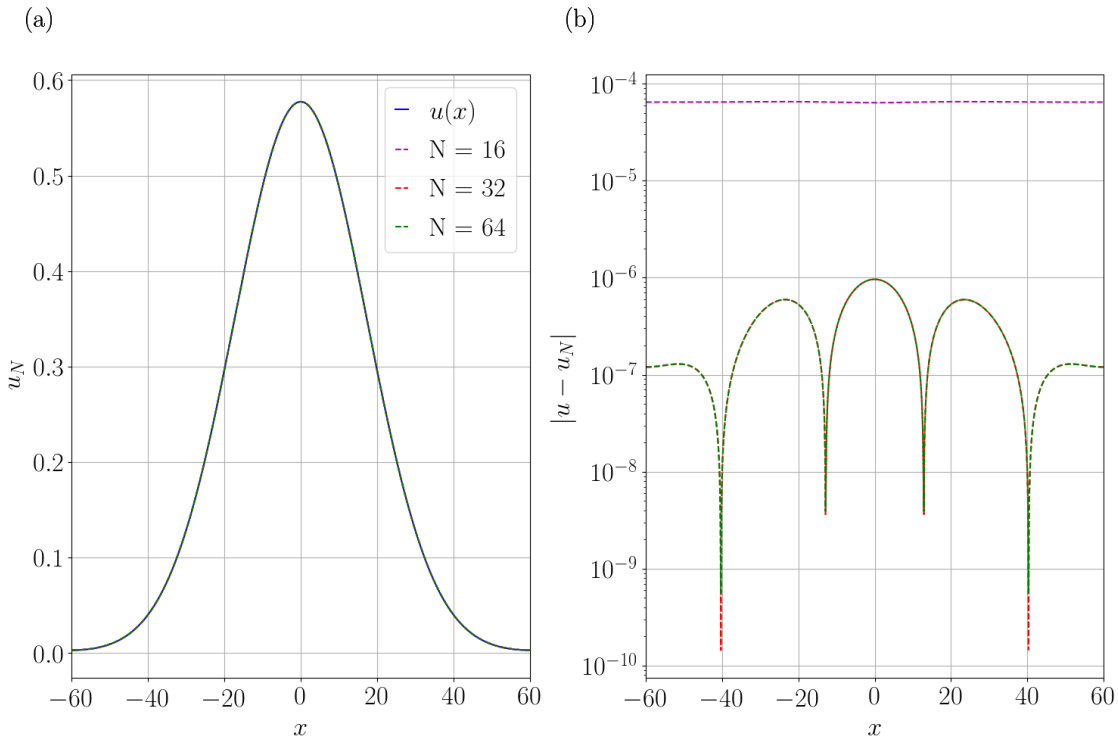


Figure 2.5: (a) Different approximations to the problem given by (2.56) using Galerkin method at the time $T = 100$ with initial condition $u_0(x) = e^{-0.005x^2}$, $x \in [-60, 60]$, and $\alpha = 1.0$. (b) Pointwise error of approximation.

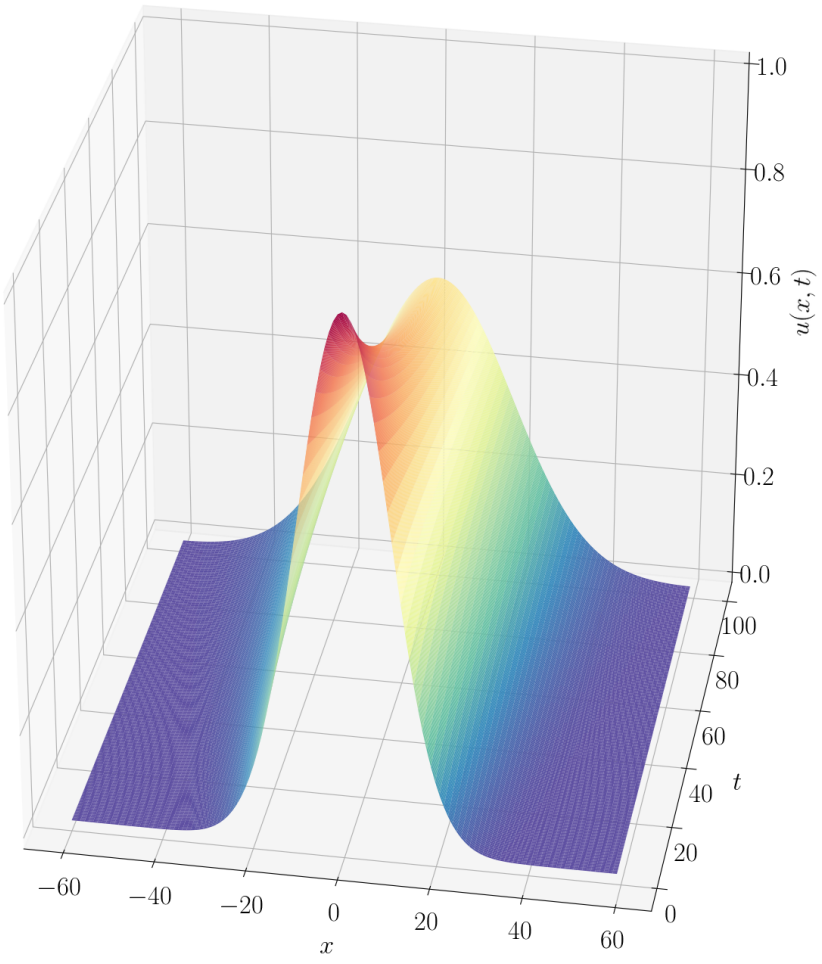


Figure 2.6: Numerical solution for problem given by (2.56) using Galerkin method with $N = 128$, $u_0(x) = e^{-0.005x^2}$, $x \in [-60, 60]$, $t \in [0, 100]$, and $\alpha = 1.0$.

The above shows the spectral convergence, which in effect is exponential convergence. In figure 2.6 we can observe this behavior, showing that only a few coefficients $\hat{u}_n$ are required to obtain a good approximation like the one shown in Figure 2.6.

However, we still cannot establish convergence for the problem given in 2.7 because it is nonlinear. At the moment we will only make some observations that will be useful in the next chapter.

Note that this problem is the Burgers' equation with $\alpha = 0$, known also as a simple model for the propagation of fluids that can be understood as a simplified

one-dimensional version of the known Euler equations for an ideal or perfect incompressible fluid. As we will see, the equation is a model example of systems in which solutions can generate singularities in finite time. As we will see, it is in this case of non-avoidable discontinuities of the solution, also called shock.

We define the next curves $x = x(t)$ that start from a point x_0, and satisfies the following equation:

$$\begin{cases} x'(t) = u(x(t), t), \quad t > 0, \\ \\ x(0) = x_0. \end{cases} \tag{2.58}$$

When the solution $u = u(x, t)$ is locally Lipschitz in the variable x and, say, it continues in time t, for every $x_0 \in \mathbb{R}$ the previous equation admits a unique local solution that we will call characteristic curve. Differentiating $u(x(t), t)$ with respect to time t, and using the equation above along with the one in the Example 2.7, we obtain the following

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

In effect, the value of the solution u along a characteristic curve, that is, $u(x(t), t)$ is independent of the time t. Therefore, the solutions are constant along the characteristic curves, and we also have to

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

So the solution to the problem (2.58) is given by

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

We see therefore that the characteristic curves are actually straight, and also the solution $u(x, t)$ can be written as

$$u(x, t) = u_0(x_0), \quad x_0 = x - u_0(x_0)t$$

Note the following. Let x_0, x_1 be start points such that $x_0 < x_1$, then for some t we have to

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

which tells us that two characteristic curves that start from x_0 and x_1 respectively, are cut at a point (x, t), where the time t is given by

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

for some $c \in (x_0, x_1)$. Therefore, if (x,t) is a point where two characteristic curves are cut that start from x_0 and x_1 respectively, at this point the solution cannot be continuous. This can be easily seen since the solution is constant along the characteristic curves, and then we would have to $u(x,t) = u_0(x_0) = u_0(x_1)$, which is impossible if $u_0(x_0) \neq u_0(x_1)$.

If we consider the minimum time for when the discontinuity or shock occurs denoted by T_c as follows

$$Tc = \min_{x\in\mathbb{R}} \left[\frac{-1}{u_0'(x)} \right] \tag{2.59}$$

Therfore, the solution $u \in H_p^1[0, 2\pi]$ over the time interval $[0, Tc]$, and we can approximate the solution using projection operator as follows

$$\begin{cases} \frac{\partial}{\partial t} u_N(x,t) + \frac{1}{2}\frac{\partial}{\partial x}(\mathcal{P}_N(u_N^2)(x,t)) = 0, \;\; t \in (0Tc], \;\; x \in [0, 2\pi] \\ \\ u_N(x,0) = \mathcal{P}_N u_0(x), \;\; x \in [0, 2\pi]. \end{cases} \tag{2.60}$$

It is also possible to get stability rewritten the above equation as

$$\frac{\partial}{\partial t} u_N + \frac{1}{2}\frac{\partial}{\partial x}(u_N^2) = \frac{1}{2}\frac{\partial}{\partial x}(I - \mathcal{P}_N)(u_N^2).$$

Multiplying by u_N, and integrating over whole space

$$\frac{1}{2}\frac{d}{dt}\int_0^{2\pi} u_N^2(x,t)dx = \frac{1}{2}\int_0^{2\pi} u_N \partial_x (I - \mathcal{P})(u_N^2)dx = -\frac{1}{2}\int_0^{2\pi} \frac{\partial}{\partial x} u_N (I - \mathcal{P})(u_N^2)dx,$$

where the last integral is obtained by integration by parts, which vanishes and hence we have the following stability estimating

$$\|u_N(t)\| = \|u_N(0)\|$$

Furthermore, by setting $(u_N - u)^2 = |u_N|^2 - |u|^2 - 2u(u_N - u)$ we can get the next convergence estimating as follows

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\int_0^{2\pi} (u_N - u)^2 dx &= \frac{d}{dt}\int_0^{2\pi} (\frac{|u_N|^2}{2} - \frac{|u|^2}{2} - u(u_N - u))dx \\ &= \frac{1}{2}\int_0^{2\pi} u_N \partial_x (I - \mathcal{P})(u_N^2)dx - \int_0^{2\pi} \partial_t (u(u_N - u))dx = I_1 + I_2. \end{aligned}$$

Recall that I_1 vanishes, and the term I_2 it can be decompose it into two terms as

$$I_2 = \int_0^{2\pi} \partial_t (u(u_N - u))dx = \int_0^{2\pi} u_t(u_N - u)dx + \int_0^{2\pi} u(\partial_t u_N - \partial_t u)dx.$$

Using original equation and its approximation to convert time derivatives to spatial ones, we find

$$\begin{aligned} I_2 &= -\int_0^{2\pi} uu_x(u_N-u)dx - \int_0^{2\pi} u\partial_x(\frac{u_N^2}{2}-\frac{u^2}{2})dx + \frac{1}{2}\int_0^{2\pi} u\partial_x(I-\mathcal{P}_N)(u_N^2)dx \\ &= -\int_0^{2\pi} uu_x(u_N-u)dx + \int_0^{2\pi} u_x(\frac{u_N^2}{2}-\frac{u^2}{2})dx - \frac{1}{2}\int_0^{2\pi} u_x(I-\mathcal{P}_N)(u_N^2)dx \\ &= \int_0^{2\pi} u_x(\frac{u_N^2}{2}-\frac{u^2}{2}-u(u_N-u))dx - \frac{1}{2}\int_0^{2\pi} u_x(I-\mathcal{P}_N)(u_N^2)dx, \end{aligned}$$

then the following inequality holds

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\int_0^{2\pi} |u_N(x,t)-u(x,t)|^2dx \leq & \frac{|u_x(\cdot,t)|_\infty}{2}\int_0^{2\pi} |u_N(x,t)-u(x,t)|^2dx \\ & -\frac{1}{2}\int_0^{2\pi} u_N^2(I-\mathcal{P}_N)(u_x)dx \end{aligned}$$

Estimating of the last integral gives us

$$|e_N(t)| = \int_0^{2\pi} |u_N^2(I-\mathcal{P}_N)(u_x)|dx \leq CN^{1-q}\|u_N\|^2 \leq CN^{1-q}\|u_N(0)\|^2$$

Therefore, from above we have the following convergence estimating

$$\|u_N(x,t)-u(x,t)\|^2 \leq e^{U'_\infty(t;0)}\int_0^{2\pi} |u_N(x,0)-u(x,0)|^2dx + \int_0^{2\pi} e^{U'_\infty(t;s)}u_N^2|e_N(s)|ds$$

where

$$U'_\infty(t;s) = \int_{\tau=s}^{t} |u_x(\cdot,\tau)|_\infty d\tau$$

Moreover, since $u_N(0) = \mathcal{P}_N u(0)$ we have to

$$\|u_N(x,t)-u(x,t)\|^2 \leq e^{\int_0^t |u_x(\cdot,s)|_\infty ds}\left[N^{-2q}\|u(0)\|_{H^q}^2 + N^{1-q}\max_{s\leq t}\|u(s)\|_{H^s}\right].$$

Note that the previous estimate depends on the solution derivative. Due to this, the convergence rate of the u_N approximation may be slower as shown in the Figure 2.7 Where we can observe that the error decreases in smaller order compared to the behavior already seen in the Figure 2.5. Recall that Tc is the shock time given by (2.59) where a discontinuity occurs in the solution that depends on the initial condition $u_0(x)$. Therefore, a higher order of approximation will be required, that is, high degrees N of polynomials to obtain a good approximation.

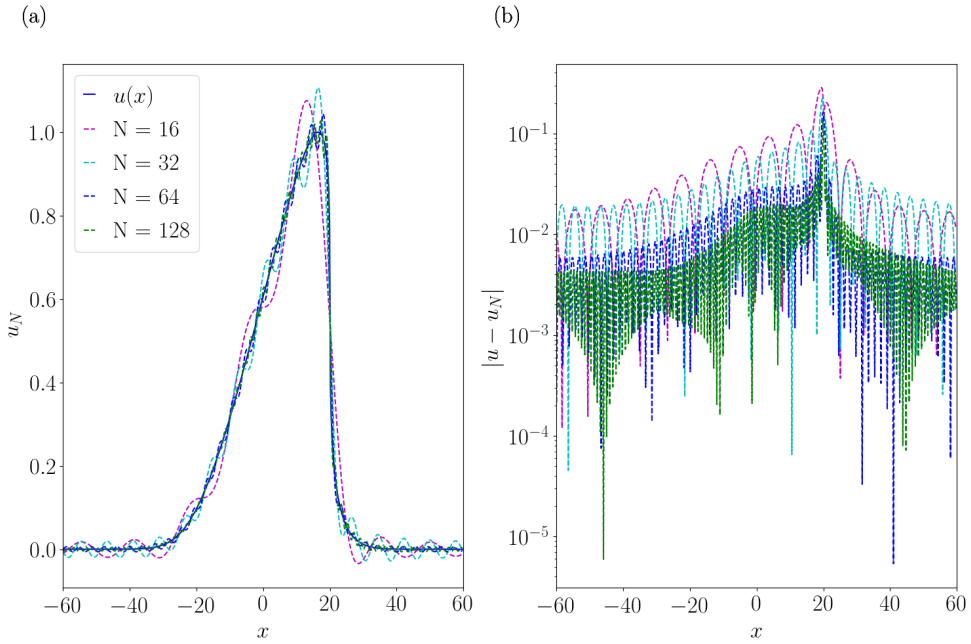


Figure 2.7: (a) Exact solution for the problem defined in the example (2.7), and its approximations using the scheme given by (2.60) at the time Tc with initial condition $u_0(x) = e^{-0.005x^2}$, $x \in [-60, 60]$. (b) Pointwise error of approximation.

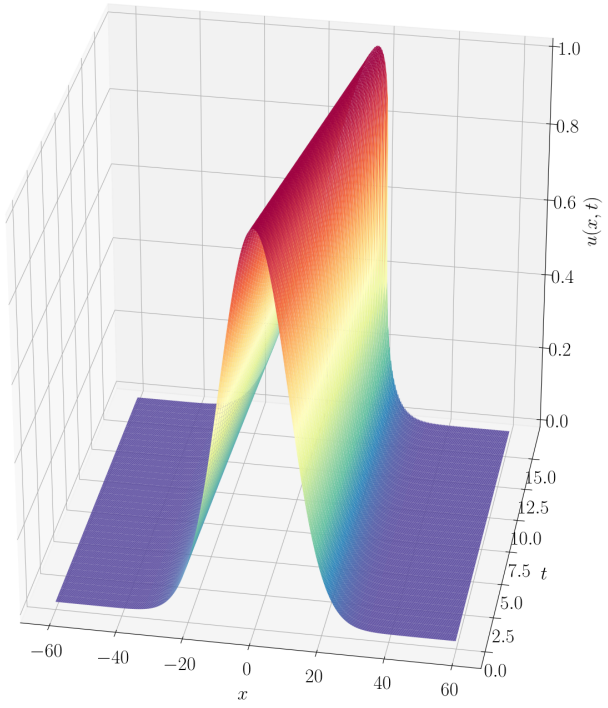


Figure 2.8: Numerical approximation for the problem defined in the Example 2.7 using the scheme given by (2.60) with $N = 512$, $u_0(x) = e^{-0.005x^2}$, $x \in [-60, 60]$, and $t \in [0, Tc]$.

The evolution of the profile of this solution simulates in a simplified way that of a marine wave that approaches the coast. As it do it profile curls until it breaks. Once it has broken, it slides to the shore without further deformation.

Perfect or ideal fluids, while constituting interesting mathematical models, are still unrealistic to the extent that each fluid has a certain degree of viscosity. Something similar occurs within the framework of the Burgers equation. The presence of a viscosity term introduces an additional regularization effect in the nonlinear equation. The important thing, in this case, is that this regularization effect is effective for all time $t > 0$ to avoid shocks. To observer this, recall by (1.12) that the solution to problem (1.3) is given by

$$u(x,t) = -2\alpha \frac{[G_x(\cdot,t) * g_\alpha](x)}{[G(\cdot,t) * g_\alpha](x)}, \tag{2.61}$$

where G is known as the fundamental solution or Gauss kernel, and is given by

$$G(x,t) = (4\pi t)^{-1/2} \exp(-x^2/4t)$$

and g_α defined by

$$g_\alpha(x) = e^{-\int_{-\infty}^{x} \frac{u_0(s)}{2\alpha} ds}$$

From the previous expression for the solution u of the viscous Burgers' equation, we can observe that it is globally defined for all $\alpha > 0$, and it is also of class $C^\infty(\mathbb{R} \times (0,\infty))$ for all $\alpha > 0$.

This means that the introduction of the term viscosity or diffusion into the Burgers' equation, however small $\alpha > 0$, makes the solutions regular. Recall that in the Burgers' equation, in the absence of viscosity, shocks occur for a finite time Tc and the solution was no longer continuous. The regularizing effect of the viscosity term $\alpha > 0$ is thus evident. Simultaneously, the solutions become global in time.

From (2.61) It can also be seen that the limit when $\alpha \to 0$ of the solution to the problem given in Example 2.7, is a solution of the Burgers equation in the absence of viscosity terms. In fact, this process of passing to the limit is actually a criterion for selecting the solution of the Burgers equation in the absence of viscosity that has real physical meaning. It is the so-called entropy solution.

From the modeling point of view, it is a natural procedure since the Burgers equation without viscosity is a simple ideal or perfect fluid model in the absolute absence of viscosity. However, in practice, every fluid has a certain degree of viscosity. It is therefore natural that the only relevant solutions of the Burgers equation without viscosity are those that can be obtained as limits of the evanescent viscosity procedure just described. A classic and relevant result in the field of non-linear hyperbolic systems due to Kruzkov guarantees that the entropy solution thus obtained is unique (see [32]). Furthermore, is recommended to see [33], [34].

Chapter 3

Numerical Solution to Burger's Equation

In this chapter, we will use the tools studied in the previous chapter to implement them in Burgers' equation. In the first section, we will present the numerical solution using the Fourier Galerkin method, and in the second section the Fourier Collocation method, also for each scheme we will establish consistency, stability, and convergence. Finally, in the last section we will present numerical results.

First note the following, for $\alpha > 0$ we define $u(x,t) = \frac{1}{\sqrt{\alpha}} v(\sqrt{\alpha}x, t)$ and $f(x,t) = \frac{1}{\sqrt{\alpha}} g(\sqrt{\alpha}x, t)$. Substituting in Burgers' equation given by (1.3) we have to

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x} + \frac{1}{2}(u^2)_x = f(x,t) \tag{3.1}$$

In the next sections, the previous equation will be studied, since it is equivalent to Burgers' equation.

In the space W_p^q defined by the norm given by (2.16), we define the initial value problem for (3.1) in the interval $I = [0, 2\pi]$ as follows

$$\begin{cases} \frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x} + \frac{1}{2}(u^2)_x = f(x,t), \;\; t > 0, \;\; x \in I \\ \\ u(x,0) = u_0(x), \;\; t = 0, \;\; x \in I \end{cases} \tag{3.2}$$

In the following sections of this chapter, we will use the above formulation to develop the method, and to establish convergence we will use the formulation given as in (2.45) by configuring $Au = -\langle u_{xx}, v\rangle$ and $F(t,u) = \langle \frac{1}{2}(u^2)_x - f, v\rangle$, $u, v \in W_p^1$ writing it as follows

$$\begin{cases} \frac{du}{dt} + Au + F(t,u) = 0, \;\; t > 0, \;\; u(t) : [0,T] \to W_p^1 \\ \\ u(0) = u_0(x) \end{cases} \tag{3.3}$$

Note that the above formulation is the weak form described in (1.6).

Note that A can be extended as a positive definite linear self-adjoint operator on W_p^0. Therefore, we can define the powers of A. Furthermore, we have $D(A^k) = W_p^{2k}$ and there exist constants C_1 and C_2, such that

$$C_1 \|u\|_{W_p^{2k}} \leq \|A^k u\|_{W_p^0} \leq C_2 \|u\|_{W_p^{2k}} \;\text{ for any }\; u \in D(A^k) \tag{3.4}$$

3.1 Fourier Galerkin

Following the Fourier Galerkin method described in the previous chapter, in section 2.3, define V_N as the space of trigonometric polynomials of degree N given by $V_N = \hat{B}_N \cap W_p^1(I)$, and the projection $\mathcal{P}_N : L_p^2(I) \to V_N$, such that satisfies $\langle v - \mathcal{P}_N v, \phi \rangle$ for $v, \phi \in V_N$. Therefore, the Fourier Galerkin scheme for the problem given by (3.2) can be formulated as follows

$$\begin{cases} u_N(t) : [0,T] \to V_N, \text{ such that for every } \phi \in V_N \\ (\frac{\partial u_N}{\partial t} + \frac{1}{2}(u_N^2)_x - \frac{\partial^2 u_N}{\partial x^2} - f, \phi) = 0 \\ u_N(0) = \mathcal{P}_N u_0(x) \end{cases} \tag{3.5}$$

or equivalently of dynamical form given by (2.46)

$$\begin{cases} u_N(t) : [0,T] \to V_N, \\ \frac{\partial u_N}{\partial t} + A_N u_N + F_N(t, u_N) = 0 \\ u_N(0) = \mathcal{P}_N u_0(x) \end{cases} \tag{3.6}$$

where A_N is defined via bilinear as $a_N(v,\phi) = -\int_I \frac{\partial^2 v}{\partial x^2}\phi dx$, and $F_N(t,v) = \mathcal{P}_N(\frac{1}{2}(v^2)_x - f(t))$ for $v, \phi \in V_N$.

Recall that every $u_N \in V_N$ are given by the following projection

$$u_N(x,t) = \sum_{|n| \le \frac{N}{2}} \hat{u}_n e^{inx}$$

Then scheme given by (3.5) can be handled as follows

$$\frac{1}{2\pi}\int_0^{2\pi} \left(\frac{\partial u_N}{\partial t} + \frac{1}{2}\frac{\partial}{\partial x}\left(u_N^2\right) - \frac{\partial^2 u_N}{\partial x^2} - f_N \right) e^{-inx} dx = 0, \quad \forall |n| \le \frac{N}{2}$$

where $f_N = \mathcal{P}_N f(x,t) = \sum_{|n| \le \frac{N}{2}} \hat{f}_n(t) e^{inx}$.

We define the following linear transformation from $I = [0, 2\pi]$ to $I_0 = [a, b]$ given by $T(z) = Pz + a$ to escalate the problem, where $P = \frac{b-a}{2\pi}$ and $z \in I$. Therefore, using the previous series each term becomes

$$\frac{\partial u_N(x,t)}{\partial t} = \sum_{|n| \le \frac{N}{2}} \frac{d\hat{u}_n(t)}{dt} e^{inx}, \quad \frac{\partial^2 u_N(x,t)}{\partial x^2} = -P^2 \sum_{|n| \le \frac{N}{2}} n^2 \hat{u}_n(t) e^{inx}$$

and the Fourier coefficients of u_N^2 are given by the convolution sum given as follows

$$\frac{\partial}{\partial x}\left(u_N^2\right) = \sum_{|n|\leq \frac{N}{2}} in \left(P \sum_{|k|\leq \frac{N}{2}} \hat{u}_n(t)\hat{u}_{n-k}(t) \right) e^{inx}$$

or equivalently in a compact form

$$\frac{\partial}{\partial x}\left(u_N^2\right) = \sum_{|n|\leq \frac{N}{2}} \hat{b}_n(t) e^{inx}$$

where $\hat{b}_n(t)$ is defined as

$$\hat{b}_n(t) = in \left(P \sum_{|k|\leq \frac{N}{2}} \hat{u}_n(t)\hat{u}_{n-k}(t) \right)$$

Substituting each term in the previous integral, and due to the orthogonality property of ϕ, we obtain a set of ODE's for $\hat{u}_n(t)$ as follows

$$\frac{d\hat{u}_n(t)}{dt} = -P^2 n^2 \hat{u}_n(t) - \frac{1}{2}\hat{b}_n(t) + \hat{f}_n(t), \quad \forall |n| \leq \frac{N}{2} \tag{3.7}$$

which is equivalent to a set of $N+1$ ordinary differential equations, which allows us to determine the coefficients $\hat{u}_n(t)$ with the following initial conditions

$$u_N(x,0) = \sum_{|n|\leq \frac{N}{2}} \hat{u}_n(0) e^{inx}, \quad \hat{u}_n(0) = \frac{1}{2\pi}\int_0^{2\pi} u(0) e^{-inx} dx$$

A suitable way to solve the system of differential equations above is to use a semi-implicit method, such as backward Euler for nonlinear term which is given as follows

$$\hat{u}_n(t_{j+1}) = \hat{u}_n(t_j) + \Delta t \left[\alpha P^2 n^2 \hat{u}_n(t_j) - \frac{1}{2}\hat{b}_n(t_{i+1}) + \hat{f}_n(t_i) \right], \tag{3.8}$$

But it is also possible to express the system in an equivalent way to notice the following. Multiplying both sides of the equation (3.7) by $e^{P^2 n^2 t}$, we obtain the following equivalent system

$$\frac{d}{dt}\left[e^{P^2 n^2 t} \hat{u}_n(t) \right] = e^{P^2 n^2 t} \left[-\frac{1}{2} P \hat{b}_n(t) + \hat{f}_n(t) \right], \quad \forall |n| \leq \frac{N}{2}$$

Then, forward Euler approximation is defined as follows. For M subintervals in $[0,T]$, of sizes equal to $\Delta t = \frac{T}{M}$, we obtain the discrete times $t_i = i\Delta t$ for $i = 0, 1, \ldots, M$. Therefore, the above equation becomes

$$\hat{u}_n(t_{i+1}) = e^{-P^2 n^2 \Delta t} \left[\hat{u}_n(t_i) - \frac{1}{2}\Delta t \hat{b}_n(t_i) + \hat{f}_n(t_i) \right], \quad \forall |n| \leq \frac{N}{2}$$

Note the following, substituting $\hat{u}_n(t_i)$ in $\hat{u}_n(t_{i+1})$ successively for each i, we get

$$\hat{u}_n(t_{i+1}) = e^{-(i+1)P^2n^2\Delta t}\hat{u}_n(t_0) + \Delta t \sum_{j=0}^{i} e^{-(i-j)P^2n^2\Delta t}\left[-\frac{1}{2}\hat{b}_n(t_j) + \hat{f}_n(t_j)\right] \quad (3.9)$$

If we look at equation (2.45), we can see that the numerical solution of the first term is solved exactly. Therefore the error is produced only by the nonlinear term.

3.1.1 Numerical Analysis to Fourier Galerkin

Now we are ready to establish some results of the problem analysis described above. Based on the definitions given by (2.2.2-2.2.4) we will establish the stability and consistency of the method, and finally, we will study the convergence.

Lemma 3.1. *Let $u_N(t) \in C([0,T], V_N)$ and $f_N(t) \in C([0,T], V_N)$ defined in $I = [0, 2\pi]$. Then Fourier Galerkin scheme given by (3.5) is stable and bounded by*

$$\|u_N(t)\|_{L^2(I)} \leq K(t)\|u_N(0)\|_{L^2(I)} + \sup_{0\leq t\leq T} \|f_N(t)\|_{L^2(I)}$$

for any $t \in [0,T]$ and $K(t)$ is independent of N.

Proof. Let u_N be the solution for (3.5). So u_N satisfies the following equation

$$\frac{\partial u_N}{\partial t} = \frac{\partial^2 u_N}{\partial x^2} - \frac{1}{2}(u_N^2)_x + f_N(x,t)$$

Multiplying both sides of the equation above by u_N and integrating over $[0, 2\pi]$ we have to

$$\frac{1}{2}\int_0^{2\pi} \frac{\partial u_N^2}{\partial t}dx = \int_0^{2\pi} u_N \frac{\partial^2 u_N}{\partial x^2}dx - \frac{1}{2}\int_0^{2\pi} u_N(u_N^2)_x dx + \int_0^{2\pi} u_N f(x,t)dx$$

The integration of the first two terms of the right side by parts and, the use of the Cauchy-Schwarz inequality (A.1.1) for the last term gives us

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|u_N\|^2 &= \int_0^{2\pi} u_N \frac{\partial^2 u_N}{\partial x^2}dx - \int_0^{2\pi} (u_N)^2 \frac{\partial}{\partial x}u_N dx + \int_0^{2\pi} u_N f_N(x,t)dx \\ &\leq -\int_0^{2\pi} \left(\frac{\partial u_N}{\partial x}\right)^2 dx + \frac{1}{3}u_N^3\Big|_0^{2\pi} + \|u_N\|_{L^2(I)}\|f_N\|_{L^2(I)}\end{aligned}$$

Using the Poincaré inequality (A.1.2) for the first term on the right side gives

$$\frac{d}{dt}\|u_N\|_{L^2(I)}^2 \leq -\|u_N\|_{L^2(I)}^2 + \|u_N\|_{L^2(I)}\|f_N\|_{L^2(I)}$$

or equivalently

$$\frac{d}{dt}\|u_N\|_{L^2(I)} \leq -\|u_N\|_{L^2(I)} + \|f_N\|_{L^2(I)}$$

Finally, solving the equation we obtain

$$\begin{aligned}\|u_N\|_{L^2(I)} &\leq e^{-t}\left(\|u_N(0)\|_{L^2(I)} + \int_0^t e^s \|f_N\|_{L^2(I)} ds\right) \\ &\leq e^{-t}\|u_N(0)\|_{L^2(I)} + \sup_{0\leq t\leq T}\|f_N\|_{L^2(I)}\end{aligned}$$

□

Remark. *In the previous Lemma we can see that the solution u_N vanishes if f too as t goes to infinity, or rather, if $f \equiv 0$. Furthermore, the problem seems converge to lineal problema.*

Lemma 3.2. *Let $u(x,t) \in W_p^q(I)$ and $f(x,t) \in W_p^{q-1}(I)$ where $I = [0, 2\pi]$, and consider Fourier-Galerkin scheme to Burgers' equation given by (3.5). Then for $q > 2$ the scheme is consistent and its estimate is given by*

$$\|\mathcal{P}_N\mathcal{L}(I-\mathcal{P}_N)\|_{L_p^2(I)} \leq \frac{K}{N^{q-1}}\left(\|u\|_{W_p^q(I)} + \|u\|^2_{W_p^q(I)} + \|f\|^2_{W_p^{q-1}(I)}\right)$$

where K is independent of N.

Proof. Set $\mathcal{L}u = Au + Bu + f(x,t)$, where $Au = u_{xx}$ and $Bu = \frac{1}{2}(u^2)_x$. Then we have the following

$$\begin{aligned}\|\mathcal{P}_N\mathcal{L}(I-\mathcal{P}_N)u\|_{L_p^2(I)} \leq \|\mathcal{P}_N A(I-\mathcal{P}_N)u\|_{L_p^2(I)} + \|\mathcal{P}_N B(I-\mathcal{P}_N)u\|_{L_p^2(I)} \\ + \|\mathcal{P}_N(I-\mathcal{P}_N)f\|_{L_p^2(I)}\end{aligned}$$

By (), we have to first term for all $q \geq 2$

$$\|\mathcal{A}(I-\mathcal{P}_N)u)\|_{L_p^2(I)} \leq \frac{C}{N^q}\|u\|_q \tag{3.10}$$

For the second term, by theorem 2.4 we have to

$$\|B(I-\mathcal{P}_N)u\| = \|(u^2)_x - \mathcal{P}_N(u^2)_x\| \tag{3.11}$$

$$\leq \frac{C}{N^{q-1}}\|u^2\|_q \tag{3.12}$$

$$\leq \frac{C}{N^{q-1}}\|u\|_q^2 \tag{3.13}$$

In the same way as above for the last term, we finally get

$$\|\mathcal{P}_N\mathcal{L}(I-\mathcal{P}_N)u)\|_{\mathcal{L}^2(\mathcal{D})} \leq \frac{K}{N^{q-1}}\left(\|u\|_q + \|u\|_q^2 + \|f\|_{q-1}\right) \tag{3.14}$$

□

***Theorem* 3.1.** *Assume for $q \geq 2$, $u(t) \in C([0,T], W_p^q)$, $f(t) \in C([0,T], W_p^{q-1})$ and consider Fourier-Galerkin scheme given by (3.5). Then when N tends to infinity, $u_N(t) \in C([0,T], \hat{B}_N)$ converges to the unique solution $u(t)$. In addition, its estimate is given by*

$$\|u_N(t) - u(t)\|_{L_p^2[0,2\pi]} \leq \frac{K(T)}{N^{q-1}} \left(|||u(t)|||_{W_p^q}^2 + \|u(0)\|_{W_p^q} + |||f(t)|||_{W_p^{q-1}} \right),$$

where $K(T)$ is independent of N

Proof. Set $e_N(t) = u(t) - u_N(t)$. Using variation of constants formula given by (2.47) and (2.48), we get

$$\begin{aligned} e_N(t) &= e^{-At}u_0 - e^{-A_N t}\mathcal{P}_N u_0 - \frac{1}{2}\int_0^t (e^{-A(t-s)}u_x^2 - e^{-A_N(t-s)}\mathcal{P}_N(u_x^2))ds \\ &\quad - \frac{1}{2}\int_0^t (e^{-A_N(t-s)}\mathcal{P}_N(u_x^2 - (u_N^2)_x)ds + \int_0^t (e^{-A(t-s)}f(s) - e^{-A_N(t-s)}\mathcal{P}_N f(s))ds \\ &= I_1 + I_2 + I_3 + I_4 \end{aligned}$$

Note that the restriction of A to V_N gives A_N. Therefore, we can replace A_N by A, or rather $e^{-A_N t}$ by e^{-At} to estimating I_1 as follows

$$\begin{aligned} \|e^{-At}u_0 - e^{-A_N t}P_N u_0\| &= \|e^{-At}u_0 - e^{-At}P_N u_0\| \\ &= \|e^{-At}\left(u_0 - P_N u_0\right)\| \end{aligned}$$

and by (2.40) and Theorem 2.3, we have to

$$\|I_1\| \leq \frac{C}{N^{q-1}}\|u(0)\|_q$$

Similarly to estimating I_2, using Theorem (2.3) to $(u^2)_x$

$$\begin{aligned} \|I_2\| &\leq \int_0^t \|e^{-A(t-s)}\left(u_x^2 - P_N(u_x^2)\right)\|ds \leq C\int_0^t \|u_x^2 - P_N(u_x^2)\|ds \\ &\leq \frac{C}{N^{q-1}}\int_0^t \|u^2\|_q ds \leq \frac{C}{N^{q-1}}\int_0^t \|u\|_q^2 ds \leq \frac{C}{N^{q-1}}\sup_{0\leq t\leq T}\|u\|_q^2 \end{aligned}$$

To estimate I_3 we observe that if $v = P_N(u^2 - u_N^2)_x$ and $w = P_N(u^2 - u_N^2)$, then we have to $\|v\| = \|A^{1/2}w\|$. Then $\|A^{-1/2}v\|^2 = (A^{-1/2}v, A^{-1/2}v) = \|w\|^2$. Furthermore, if $\|u + u_N\|_\infty = M$ we have

$$\begin{aligned} \|u^2 - u_N^2\| = \|(u+u_N)(u-u_N)\| &= \left(\int_0^{2\pi} |u+u_N|^2|u-u_N|^2 dx\right)^{1/2} \\ &\leq M\left(\int_0^{2\pi}(u-u_N)^2 dx\right)^{1/2} = M\|u - u_N\| \end{aligned}$$

Therefore, using the above the estimate of I_3 gives

$$\|I_3\| \leq \int_0^t \|A^{1/2}e^{-A(t-s)}A^{-1/2}P_N(u^2-u_N^2)_x\|ds \leq C\int_0^t (t-s)^{-1/2}\|A^{-1/2}P_N(u^2-u_N^2)_x\|ds$$
$$\leq C\int_0^t (t-s)^{-1/2}\|u^2-u_N^2\|ds \leq \beta\int_0^t (t-s)^{-1/2}\|u-u_N\|ds$$

Finally estimating I_4

$$\|I_4\| \leq \int_0^t \|e^{-A(t-s)}(f-f_N)\|ds \leq \frac{C}{N^{q-1}}\int_0^t \|f\|_{q-1}ds \leq \frac{C}{N^{q-1}}\sup_{0\leq t\leq T}\|f\|_{q-1}$$

Setting $\alpha = \frac{C}{N^{q-1}}\left(|||u|||_q^2 + \|u_0\|_q + |||f|||_q\right)$, and we have

$$\|e_N(t)\| \leq \alpha + \beta\int_0^t \|e_N(s)\|ds$$

Therefore, by Gronwall inequality given by (A.2) we get

$$\|e_N(t)\| \leq C\alpha e^{C\beta^2 t} \leq C\frac{e^{C\beta^2 T}}{N^q}(|||u|||_q^2 + \|u_0\|_q + |||f|||_{q-1})$$

□

3.2 Fourier Collocation

We will consider Fourier-Collocation approach for (3.2) using the operator $\mathcal{J}_N$ given by (2.26), with an odd number of points on the grid given as follows

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

As before described in the previous chapter, Fourier-Collocation scheme is given by

$$\begin{cases} \text{Find } u_N(t) : [0,T] \to V_N, \text{ such that for every } j \in [0,1,\ldots,N] \\ \frac{\partial u_N}{\partial t}(x_j,t) - \frac{\partial^2 u_N}{\partial x^2}(x_j,t) + \frac{1}{2}\mathcal{J}_N\left(u_N^2\right)_x(x_j,t) = f(x_j,t) \\ u_N(x_j,0) = u_0(x_j). \end{cases} \tag{3.15}$$

V_N is defined as before in the previous section such that for $u_N \in V_N$ it has the form

$$u_N(x,t) = \sum_{|n|\leq \frac{N}{2}} \hat{u}_n(t)e^{inx} = \sum_{j=0}^{N} u_N(x_j,t)g_j(x)$$

where $h_j(x)$ is the Lagrange interpolation polynomial for an odd number of points that satisfies $h_j(x_i) = \delta_{ji}$.

Note that (3.15) can be written equivalently

$$\begin{cases} u_N(t) : [0,T] \to V_N, \text{ such that for every } \phi \in V_N \\ \left\langle \frac{\partial u_N}{\partial t}, \phi \right\rangle_N - \left\langle \frac{\partial^2 u_N}{\partial x^2}, \phi \right\rangle_N + \frac{1}{2} \left\langle \mathcal{J}_N (u_N^2)_x, \phi \right\rangle_N = \langle f, \phi \rangle_N, \quad t > 0, \quad x \in I \\ u_N(0) = \mathcal{J}_N u_0(x), \quad t = 0, \quad x \in I \end{cases} \tag{3.16}$$

The above equation can be put into the dynamical form given by (2.46) taking $A_N = A$ and $F_N(t,u) = \frac{1}{2}\mathcal{J}_N(u^2)_x - \mathcal{J}_N f(t)$ for $u \in V_N$.

Similarly, as in the Galerkin method, we require the residual to be zero, but now only at points x_j, i.e.,

$$R_N(x_j,t) = \frac{\partial u_N}{\partial t}(x_j,t) - \frac{\partial^2 u_N}{\partial x^2}(x_j,t) + \frac{1}{2}\mathcal{J}_N \left(u_N^2\right)_x (x_j,t) - f(x_j,t) = 0$$

The above defines N ordinary differential equations for $u_N(x_j,t)$ with initial conditions $u_N(x_j,0) = u_0(x_j)$, but firstly, we need to define again the linear transformation from $I_0 = [0,2\pi]$ to $I = [a,b]$ given by $T(z) = Pz + a$, where $P = \frac{b-a}{2\pi}$ and $z \in I_0$ is a scale factor.

If Setting as follows

$$u_N(t) = (u_N(x_0,t), u_N(x_1,t), \ldots, u_N(x_N,t))^T,$$
$$f_N(t) = (f(x_0,t), f(x_1,t), \ldots, f(x_N,t))^T,$$

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

$$\frac{du_N(t)}{dt} + \frac{1}{2} D_N u_N^2(t) - D_N^2 u_N(t) - f_N(t) = 0,$$

where D_N is the matrix given by (2.32) already scaled by factor P, that represents discrete Fourier differentiation.

Solving the above equation using backward Euler for nonlinear term gives us

$$u_N(t_{i+1}) = u_N(t_i) + \Delta t \left[D_N^2 u_N(t_i) - \frac{1}{2} D_N u_N^2(t_{i+1}) + f_N(t_i) \right]. \tag{3.17}$$

3.2.1 Numerical Analysis to Fourier Collocation

Lemma 3.3. *Let $u(x) \in L^2[0,2\pi]$ and $f(x,t) \in \mathcal{H}_0^1$. Then the scheme Fourier Galerkin is stable and its estaming is given by*

$$\|u_N(t)\|_N \leq \|u_N(0)\|_N + \sup_{0 \leq t \leq T} \|f_N(t)\|_N \tag{3.18}$$

where $C(t)$ is independent of N and bounded for any $t \in [0,T]$.

Proof. Using the scheme given by (3.15) and multiplying by u_N gives us

$$\frac{1}{2}\frac{\partial u_N^2(x_j,t)}{\partial t} = -\frac{1}{2}u_N(x_j,t)\left(u_N^2\right)_x(x_j,t) + \frac{\partial^2 u_N}{\partial x^2}(x_j,t) + u_N(x_j,t)f(x_j,t)$$

Firstly, note the following

$$\left(u_N^2\right)_x = \mathcal{J}_N\frac{\partial}{\partial x}\mathcal{J}_N u_N^2(x,t) \in V_N$$

and

$$\frac{\partial^2 u_N}{\partial x^2} = \mathcal{J}_N\frac{\partial^2}{\partial x^2}\mathcal{J}_N u_N(x,t) \in V_N$$

summing over all collocation points we obtain

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\sum_{j=0}^{N} u_N^2(x_j,t) = &-\frac{1}{2}\sum_{j=0}^{N} u_N(x_j,t)\left(u_N^2\right)_x(x_j,t)\\ &+\sum_{j=0}^{N} u_N(x_j,t)\frac{\partial^2 u_N}{\partial x^2}(x_j,t) + \frac{N+1}{2\pi}\sum_{j=0}^{N} u_N(x_j,t)f(x_j,t)\end{aligned}$$

by Theorem 2.5 the quadrature rule is exact and we have

$$\begin{aligned}\frac{N+1}{2\pi}\frac{1}{2}\frac{d}{dt}\|u_N(t)\|^2 = &-\frac{N+1}{2\pi}\int_0^{2\pi} u_N(x,t)\mathcal{J}_N\frac{\partial}{\partial x}\mathcal{J}_N u_N^2(x,t)dx\\ &+\frac{N+1}{2\pi}\int_0^{2\pi} u_N(x,t)\mathcal{J}_N\frac{\partial^2}{\partial x^2}\mathcal{J}_N u_N(x,t)dx\\ &+\frac{N+1}{2\pi}\int_0^{2\pi} u_N(x,t)\mathcal{J}_N f(x,t)dx\end{aligned}$$

Integration by part gives us

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|u_N(t)\|^2 &= -\int_0^{2\pi}\mathcal{J}_N\frac{\partial u_N(x,t)}{\partial x}\mathcal{J}_N\frac{\partial u_N(x,t)}{\partial x}dx + \int_0^{2\pi} u_N(x,t)\mathcal{J}_N f(x,t)dx\\ &\le -\left\|\frac{\partial u_N}{\partial x}\right\|^2 + \|u_N(t)\|\|f_N(t)\|\end{aligned}$$

By Poincare inequality, and similarly as in Theorem 3.1

$$\|u_N(t)\| \le -e^{-t}\|u_N(0)\|^2 + \sup_{0\le t\le T}\|f_N\|$$

Since that the continuous and discrete norms are uniformly equivalent, discussed in (2.33), the lemma is proved. □

Lemma 3.4. *Let $u \in \mathcal{H}_p^q$ and consider the Fourier Galerkin scheme to Burgers equation given by (3.16). Then for $q > 2$ the scheme is consistent. Furthermore, we have the next estimating*

$$\|\mathcal{I}_N \mathcal{L}(I - \mathcal{J}_N)u\|_{\mathcal{L}^2_w(\mathcal{D})} \leq \frac{K}{N^{q-1}} \left(\|u^{(q)}\| + \|u\|_q^2 + \|f\|_q \right)$$

where K is independent of N.

Proof. We observe the following

$$\begin{aligned} \|\mathcal{L}(I - \mathcal{J}_N)u\| &= \|\mathcal{L}(I - \mathcal{J}_N)u + \mathcal{L}\mathcal{P}_N u - \mathcal{L}\mathcal{P}_N u\| \\ &= \|\mathcal{L}(I - \mathcal{P}_N)u + \mathcal{L}(\mathcal{P}_N - \mathcal{J}_N)u\| \\ &\leq \|\mathcal{L}(I - \mathcal{P}_N)u\| + \|\mathcal{A}_{\mathcal{N}} u\| \end{aligned}$$

Using Lemma 3.2 and Theorem 2.3 we have to

$$\|\mathcal{L}(I - \mathcal{J}_N)u\| \leq \frac{K}{N^{q-1}} \left(\|u^{(q)}\| + \|u\|_q^2 + \|f\|_q \right)$$

□

***Theorem* 3.2.** *Assume for $q \geq 2$, $u(t) \in C([0,T],\, W_p^q)$, $f(t) \in C([0,T],\, W_p^{q-1})$. Then the approximation at $u_N(t) \in C([0,T], V_N)$ given by (3.16) converges to solution $u(t)$ given by (3.3). Furthermore, its estimate is given by*

$$\|u_N(t) - u(t)\|_{L_2[0,2\pi]} \leq \frac{K(T)}{N^{q-1}} \left(|||u(t)|||_{W_p^q}^2 + \|u_0\|_{W_p^q} + |||f(t)|||_{W_p^{q-1}} \right)$$

where $K(T)$ is independent of N.

Proof. Note the following

$$\begin{aligned} \|u - u_N\|_{L_2[0,2\pi]} &\leq \|\mathcal{J}_N u - u_N\|_{L_2[0,2\pi]} + \|u - \mathcal{J}_N u\|_{L_2[0,2\pi]} \\ &\leq \|\mathcal{J}_N u - \mathcal{P}_N u\|_{L_2[0,2\pi]} + \|\mathcal{P}_N u - u_N\|_{L_2[0,2\pi]} + \|u - \mathcal{J}_N u\|_{L_2[0,2\pi]} \end{aligned}$$

The first term on right side is the aliasing error, which is bounded by Theorem 2.3, the last two terms are bounded by Theorems 3.1 and 2.10. Therefore from above we have proved the theorem. □

3.3 Numerical Results

In this section, we will describe some numerical experiments performed to illustrate the results obtained from the numerical analysis for each scheme described in the previous sections independently. It is worth mentioning that the simulations were carried out in the Python programming language using the Fourier fast transform tool.

3.3.1 Galerkin Simulations

For the following numerical simulations that we will describe, we will use the discretization of the problem already described in (3.8) but with $f \equiv 0$, which is given as follows

$$\hat{u}_n(t_{j+1}) = \hat{u}_n(t_j) + \Delta t \left[\alpha p^2 n^2 \hat{u}_n(t_j) - p\widehat{G}_n(t_{j+1})\right], \tag{3.19}$$

where $\widehat{G}_n$ is evaluated by looking for a fixed point and is given by

$$\widehat{G}_n(t_{j+1}) = in \left[\sum_{|k|\leq \frac{N}{2}} \hat{u}_n(t_{j+1})\hat{u}_{n-k}(t_{j+1})\right].$$

The discretization of the time variable t over the interval $[0, T]$ is given by

$$t_j = j\Delta t, \;\; j = 0, 1, \ldots, T.$$

Finally to evaluate the numerical solution the following expression will be used

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

For the numerical study, we will use the analytical solution of (1.3) given by (1.12), establishing the following initial condition

$$u_0(x) = e^{0.05x^2}, \quad x \in [x_L, x_R] \tag{3.20}$$

In the figure 3.1 shows the maximum distance over every $t \in [0, 100]$ between the exact solution and its approximations given by (3.19) 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 $\alpha = 0.005$.

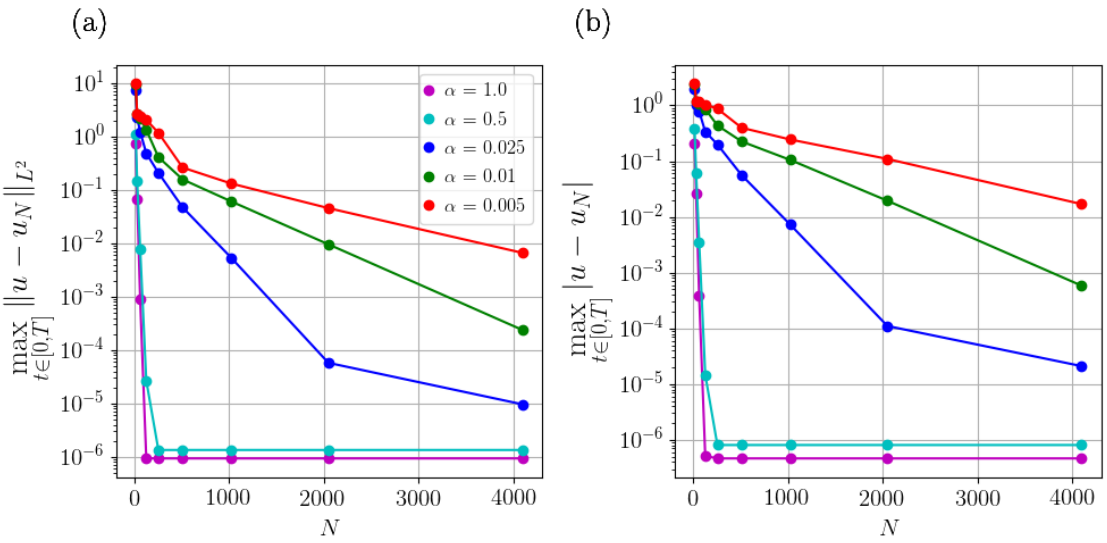


Figure 3.1: (a) L^2-norm between the exact solution and its approximations using Galerkin method. (b) Max norm between the exact solution and its approximations.

Figure 3.2: Numerical solution for (1.3) using (3.19) with $\alpha = 1.0$, $N = 2048$, and $\Delta t = 1.0 \times 10^{-5}$.

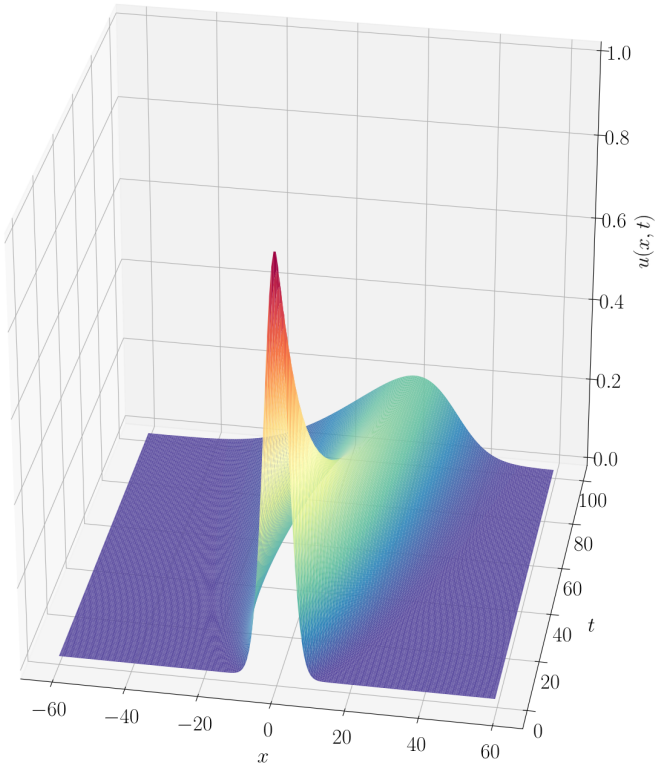
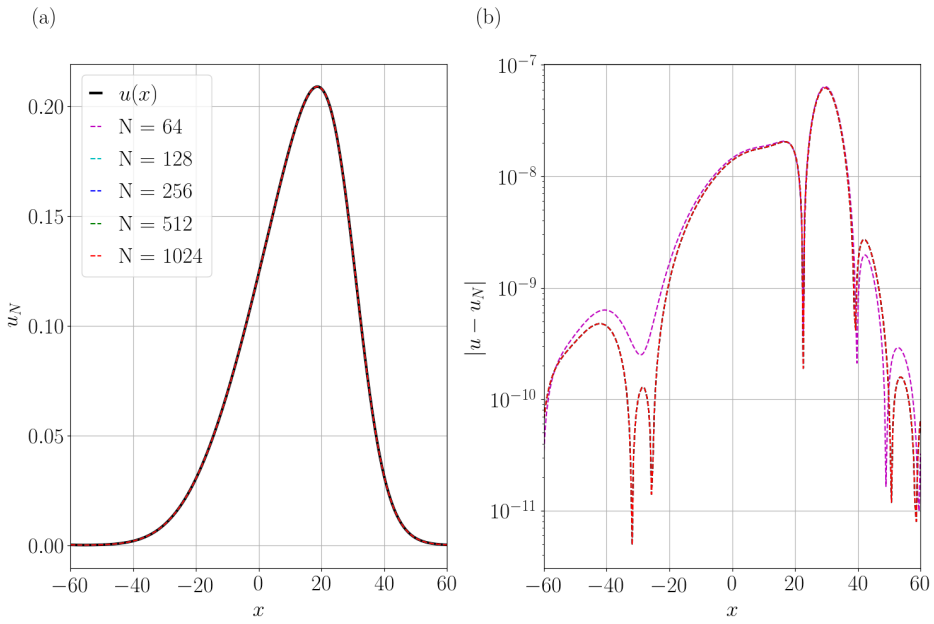


Figure 3.3: Numerical solution for (1.3) using (3.19) at the time $T = 100$ with $\alpha = 1.0$, and $\Delta t = 1.0 \times 10^{-5}$. (b) Point-wise error of approximation



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 (1.3) using (3.19) with $\alpha = 0.005$, $N = 2048$, and $\Delta t = 1.0 \times 10^{-5}$.

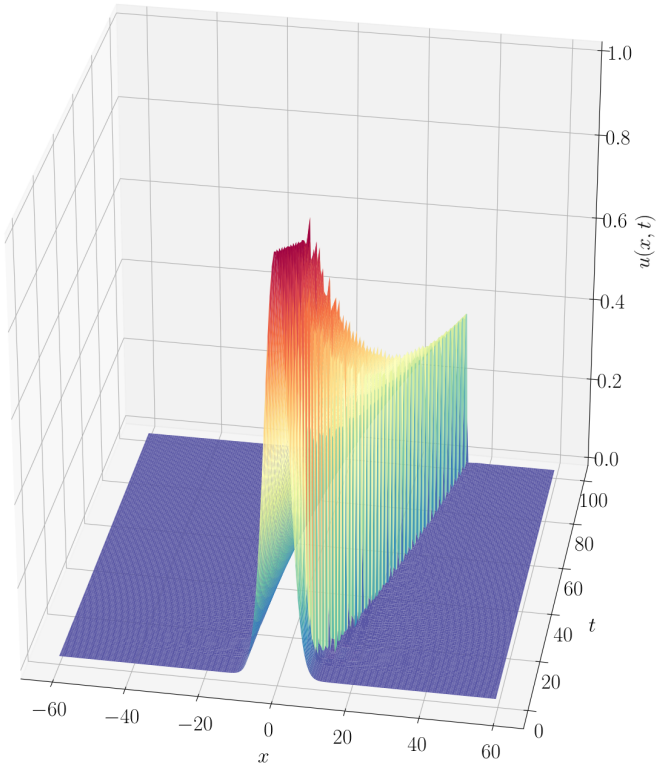
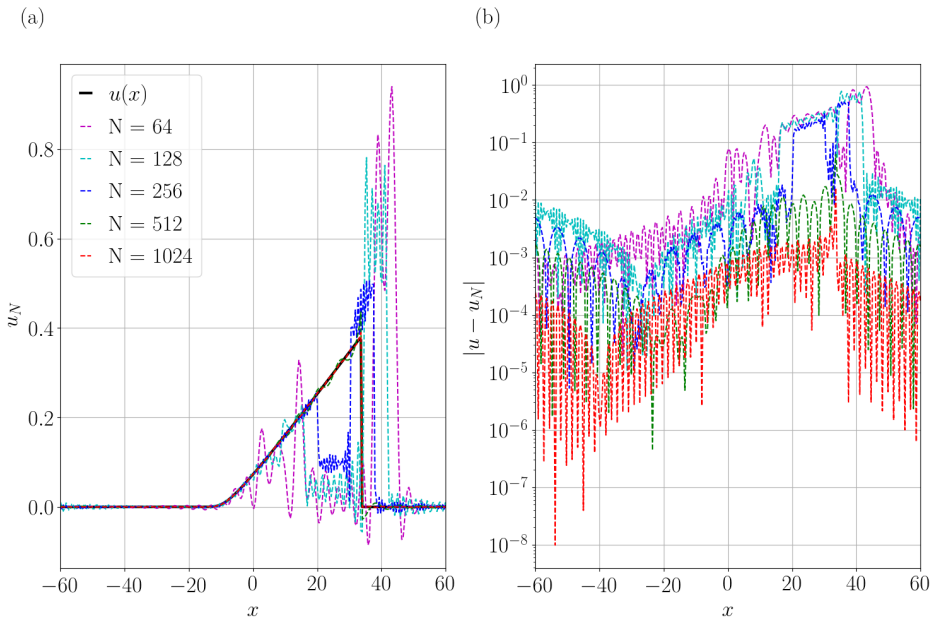


Figure 3.5: Numerical solution for (1.3) using (3.19) at the time $T = 100$ with $\alpha = 1.0$, and $\Delta t = 1.0 \times 10^{-5}$. (b) Point-wise error of approximation



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	6.50292×10^{-2}	4.70602×10^{-2}	4.57371×10^{-2}	4.56090×10^{-2}
4096	*	7.26917×10^{-3}	6.64157×10^{-3}	6.60753×10^{-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	*	2.37273×10^{-2}	1.75687×10^{-2}	1.69531×10^{-2}

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

3.3.2 Collocation Simulations

For the following numerical simulations that we will describe, we will use the discretization of the problem already described, which was the following

$$u_N(t_{i+1}) = u_N(t_i) + \Delta t \left[p^2 D_N^2 u_N(t_i) - \frac{1}{2} p D_N u_N^2(t_{i+1}) \right]. \tag{3.21}$$

The discretization of the time variable t over the interval $[0, T]$ is given by

$$t_i = i\Delta t, \quad i = 0, 1, \ldots, T,$$

and for the spatial variable x over the interval $[x_L, x_R]$ is given by $x_j = pz_n + x_L$, where $p = \frac{x_R - x_L}{2\pi}$, and

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

Finally to evaluate the numerical solution the following expression will be used

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

For the numerical study, we will use the analytical solution of (1.3) given by (1.12), establishing the following initial condition

$$u_0(x) = e^{0.05x^2}, \quad x \in [x_L, x_R] \tag{3.22}$$

In the figure 3.6 shows the maximum distance over every $t \in [0, 100]$ between the exact solution and its approximations given by (3.21) 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 $\alpha = 0.005$.

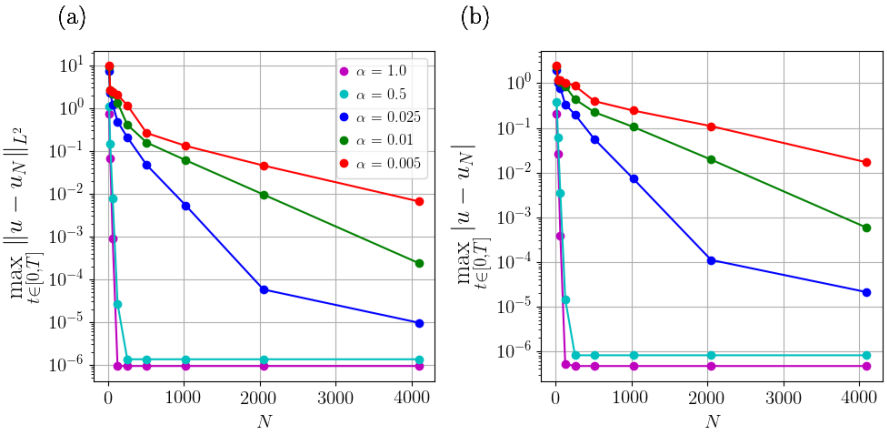


Figure 3.6: (a) L^2-norm between the exact solution and its approximations using Collocation method. (b) Max norm between the exact solution and its approximations.

Figure 3.7: Numerical solution for (1.3) using (3.21) with $\alpha = 1.0$, $N = 2048$, and $\Delta t = 1.0 \times 10^{-5}$.

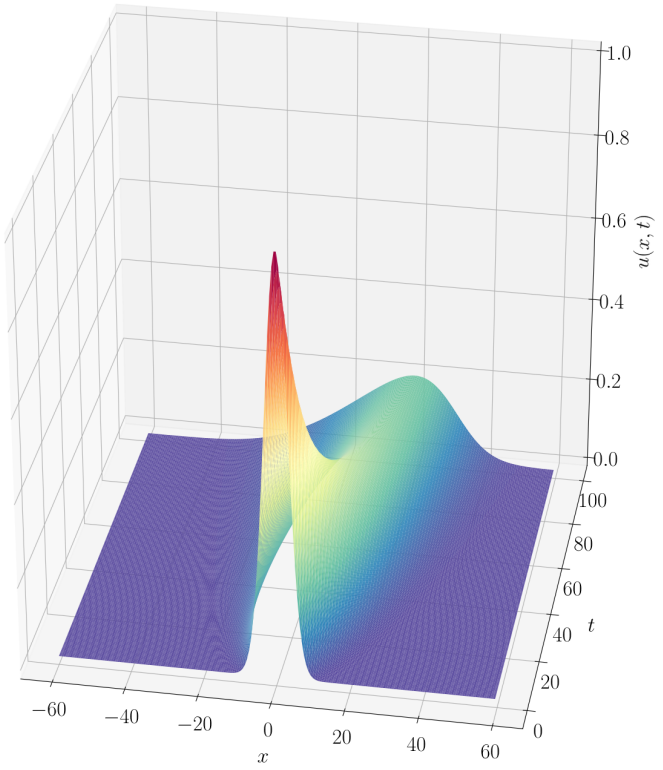
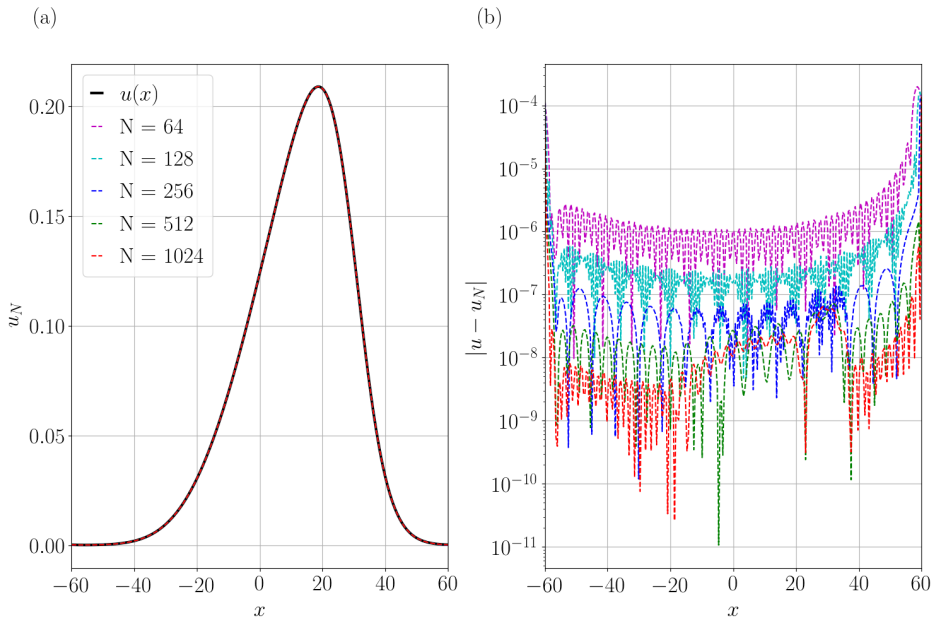


Figure 3.8: Numerical solution for (1.3) using (3.21) at the time $T = 100$ with $\alpha = 1.0$, and $\Delta t = 1.0 \times 10^{-5}$. (b) Point-wise error of approximation



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	0.721112	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	0	9.41793×10^{-5}	5.32213×10^{-5}	5.32019×10^{-5}
2048	0	0	3.56779×10^{-5}	3.56498×10^{-5}
4096	0	0	2.24122×10^{-5}	0

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	0	1.45032×10^{-4}	1.45017×10^{-4}	1.45016×10^{-4}
2048	0	0	1.43941×10^{-4}	1.4394×10^{-4}
4096	0	0	1.43411×10^{-4}	0

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

Figure 3.9: Numerical solution for (1.3) using (3.21) with $\alpha = 0.005$, $N = 2048$, and $\Delta t = 1.0 \times 10^{-5}$.

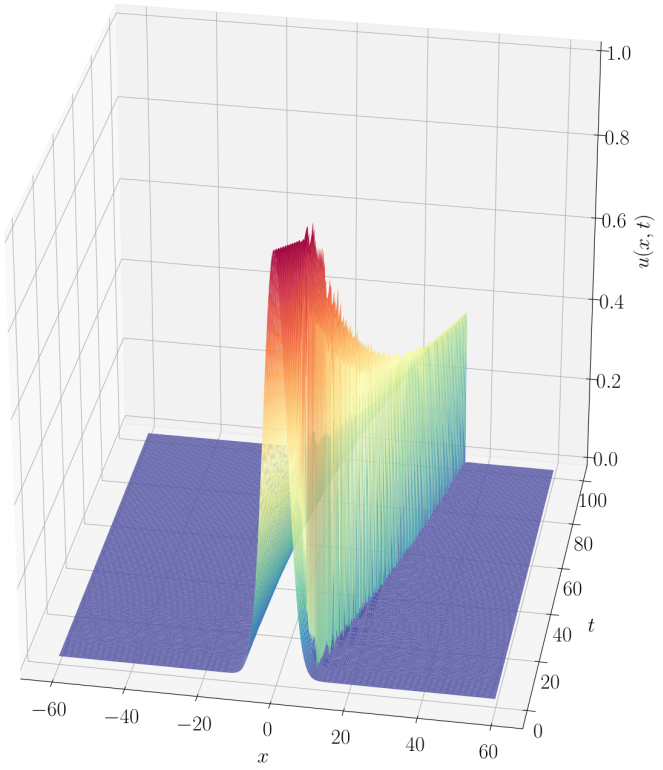
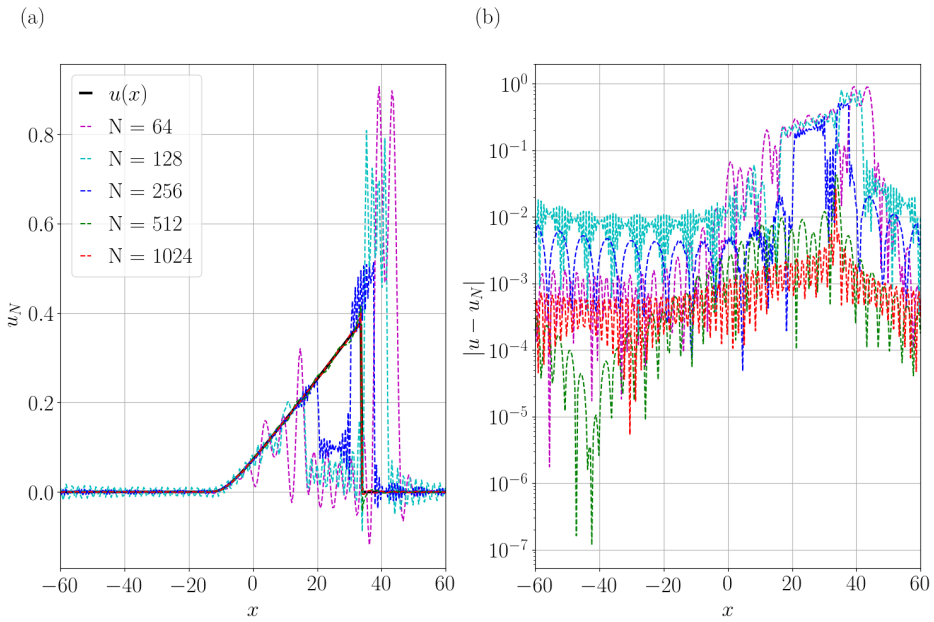


Figure 3.10: Numerical solution for (1.3) using (3.21) at the time $T = 100$ with $\alpha = 0.005$, and $\Delta t = 1.0 \times 10^{-5}$. (b) Point-wise error of approximation.



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	6.56161×10^{-2}	4.63808×10^{-2}	4.49226×10^{-2}	4.47813×10^{-2}
4096	0	7.66246×10^{-3}	6.9909×10^{-3}	0

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	0	2.04624×10^{-2}	1.76513×10^{-2}	0

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

3.3.3 Numerical Solutions for Small Viscosity Coefficients

At the end of the previous chapter, it was mentioned that when α tends to zero the solution of equation (1.3) approaches inviscid Burgers' equation given by (2.57). In order to illustrate this, we will consider the following initial condition

$$u_0(x) = e^{-0.005x^2}, \quad x \in [x_L, x_R]$$

As before, we will discretize the time variable t over the interval $[0, T_c]$, where T_c is given as in (2.59), as follows

$$t_i = i\Delta t, \;\; i = 0, 1, \ldots, T_c$$

Recall that the solution for (2.57) is given by the characteristic curves as follows

$$u(x,t) = u_0(x_0), \;\; x_0 = x - u_0(x_0)t, \tag{3.23}$$

which is resolved for every t and every $x_0 \in [x_L, x_R]$.

In the following simulations, the Galerkin method given by (3.19) was used to obtain numerical solutions with small values of α and compare them with the exact solution given above corresponding to the problem without viscosity.

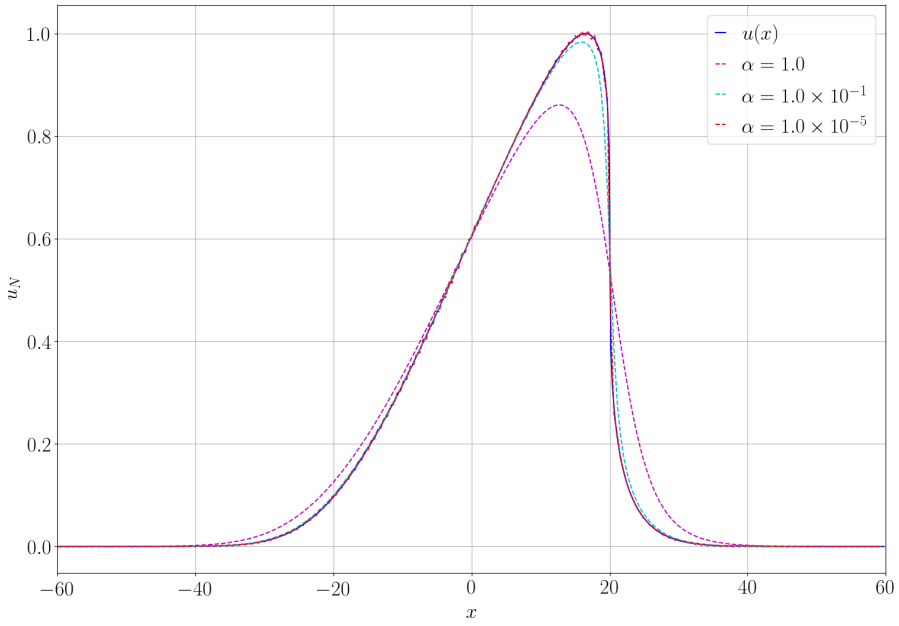


Figure 3.11: Exact solution for (2.57) and different approximations using (3.19) with $N = 256$, and $\Delta t = 1.0 \times 10^{-3}$.

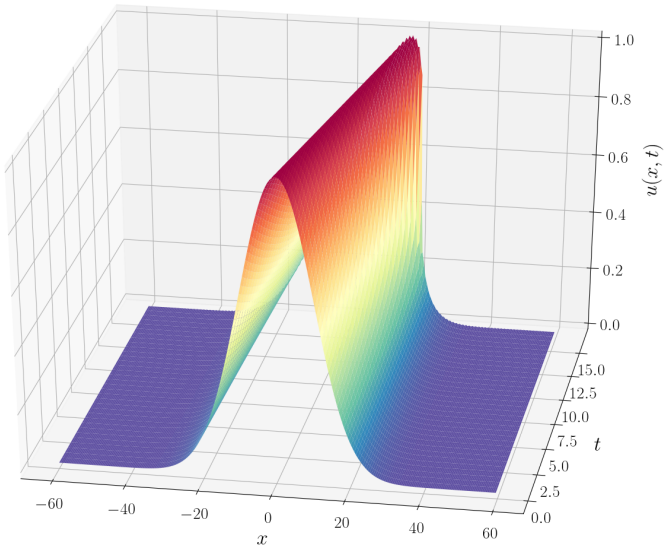


Figure 3.12: Numerical solution for (1.3) using (3.19) with $\alpha = 1.0 \times 10^{-5}$, $N = 256$, and $\Delta t = 1.0 \times 10^{-3}$.

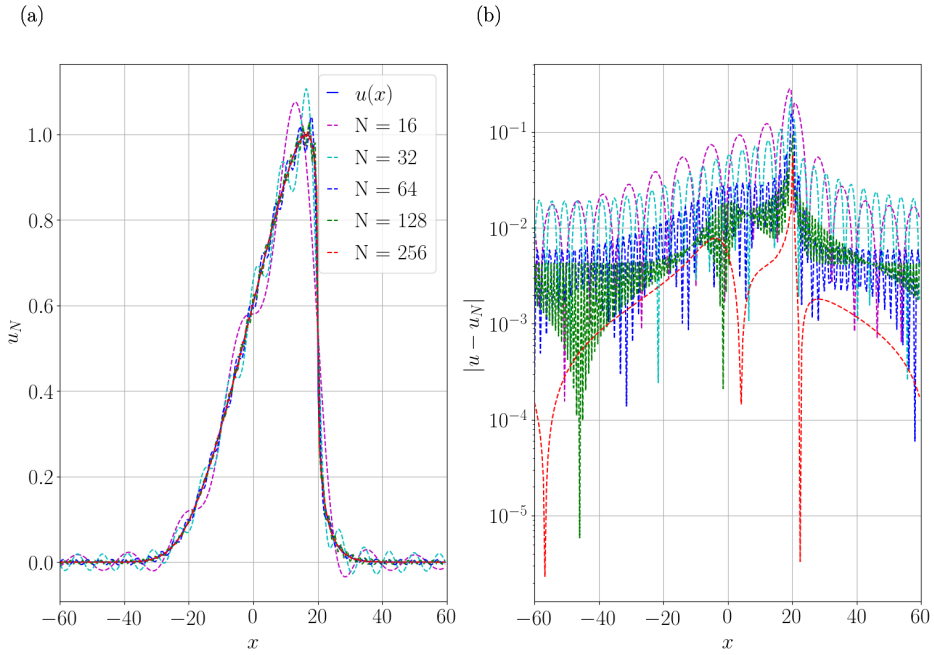


Figure 3.13: Numerical solution for (1.3) using (3.19) at the time T_c with $\alpha = 1.0 \times 10^{-5}$, and $\Delta t = 1.0 \times 10^{-3}$. (b) Point-wise error of approximation.

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 (2.57) and the approximation for (1.3) with $\alpha = 1.0 \times 10^{-5}$.

Chapter 4

Numerical Solution to Stochastic Burgers' equation

In this chapter, we will study a spectral method to solve the Burgers' stochastic equation given by (4.19). In the previous chapters we focus on spectral methods based on trigonometric polynomials to build a base on space, now in this method that we will present below, it will be constructed from another family of polynomials called Hermite polynomials. First, in the first two sections, we will study the theoretical basis of the method and its implementation. In the third section, the Burgers' stochastic equation method will be implemented. Finally, in the final section, we will present some important results of the numerical analysis of the method.

4.1 Elemental Theory to Numerical Method

In statistical mechanics, the Fokker-Planck-Kolmogorov (FPK) equation is a partial differential equation that describes the temporal evolution of the probability density function of the velocity of a particle under the influence of drag forces and random forces, as in Brownian motion. The equation applies to systems that can be modified by a small number of "macrovariables", where other parameters can be quickly modified over time that can be treated as "noise" or a disturbance. The main idea is to associate the equation (FPK) with one (SPDE) through a stochastic differential equation. To do this, let's define the following.

Set a separable infinite-dimensional Hilbert space $\mathcal{H}$ with inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$. We define a Gaussian measure μ with mean zero and nuclear covariance operator Λ with $Tr(\Lambda) < +\infty$. Setting the following stochastic differential equation in $\mathcal{H}$

$$dX_t = AX_t dt + B(X_t)dt + \sqrt{Q}dW_t \tag{4.1}$$

where the operator $A : D(A) \subset \mathcal{H} \to \mathcal{H}$ is the infinitesimal generator of a strongly continuous semigroup $e^{tA} \in \mathcal{H}$, Q is a bounded operator from another Hilbert space U to $\mathcal{H}$ and $B : D(B) \subset \mathcal{H} \to \mathcal{H}$ is a nonlinear mapping.

We define the next function

$$u(x,t) = \mathbb{E}\left[u_0(X_t^x)\right] \tag{4.2}$$

where $u_0 : \mathcal{H} \to \mathbb{R}$ and X_t^x is the solution to (4.1) with initial conditions $X_0 = x \in \mathcal{H}$. Then the equation (4.1) can be associated with the Kolmogorov equation as following

$$\frac{\partial u}{\partial t} = \frac{1}{2} Tr(QD^2 u) + \langle Ax, Du \rangle_{\mathcal{H}} + \langle B(x), Du \rangle_{\mathcal{H}}, \ x \in D(A) \tag{4.3}$$

where u defined in (4.2) satisfies (4.3).

Suppose that the operators $-A$ and Q have the same eigenfunctions e_k with eigenvalues λ_k and ρ_k respectively. We define the operator $\mathcal{L}$ as

$$\mathcal{L}u = \frac{1}{2} Tr(QD^2 u) + \langle Ax, Du \rangle_{\mathcal{H}}, \ \ x \in \mathcal{H} \tag{4.4}$$

the operator $\mathcal{L}$ given above satisfies the following result.

Lemma 4.1. *Let $H_n(h)$ be a Hermite polynomial functional given by (B.5). Then the following holds: First, define $\mathcal{J}$ as follows*

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

hence

$$\mathcal{L} H_n(h) = -\lambda_n H_n(h) \tag{4.5}$$

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

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

Using Lemmas B.1 and 4.1, we have that $\{H_n\}$ forms a complete orthonormal system at $\mathbb{H}$. Therefore, if $u \in S(\mathbb{H})$ then

$$u(x,t) = \sum_{n \in J} u_n(t) H_n(x), \ \ x \in \mathcal{H}, \ \ t \in [0,T], \tag{4.6}$$

where $u_n : [0,T] \to \mathbb{R}$ and $H_n(x)$ are the Hermite functionals.
Substituting (4.6) into the equation (4.3) in the left side we get

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

the two first terms of the right side are obtained as

$$\begin{aligned} \mathcal{L}u &= \mathcal{L}\left(\sum_{n \in J} u_n(t) H_n(x) \right) \\ &= \sum_{n \in J} u_n(t) \mathcal{L} H_n(x), \end{aligned}$$

and by Lemma 4.1 we have

$$\mathcal{L}u = -\sum_{n\in J} u_n(t)\lambda_n H_n(x). \tag{4.8}$$

The last term of the equation is obtained as following

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

$$= \sum_{n\in J} u_n(t)\left(B(x), D_x H_n(x)\right)_{\mathcal{H}} \tag{4.9}$$

where D_x denoting the Frechet Derivative. Therefore, by (4.7-4.9) the Kolmogorov equation becomes

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

Multiplying the above equation by $H_m(x)$, $m \in J$ and integrating over $\mathcal{H}$ w.r.t. $\mu(dx)$ we have

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

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

$$\dot{u}_m(t) = -u_m(t)\lambda_m + \sum_{n\in\mathcal{J}} u_n(t)C_{n,m}, \quad n, m \in \mathcal{J} \tag{4.10}$$

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

$$C_{n,m} = \int_{\mathcal{H}} H_m(x)\left(B(x), D_x H_n(x)\right)_{\mathcal{H}}\mu(dx) \tag{4.11}$$

Existence and uniqueness to above equation has been proven in [1].

4.2 Numerical Approximation and Its Description

Define the set of finite multi-index $J^{M,N}$ as

$$J^{M,N} = \{\gamma = (\gamma_i,\ 1 \geq \gamma_i \geq M) \mid \gamma_i \in \{0, 1, \cdots, N\}\} \tag{4.12}$$

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

The Approximation of the solution to equation (4.3) by (4.6) is as following

$$\hat{u}_N(x,t) = \sum_{n \in J^{M,N}} u_n(t) H_n(x), \quad x \in \mathcal{H}, \quad , t \in [0,T]. \tag{4.13}$$

Consider the same value M as in $J^{M,N}$ and $m_1, m_2, \cdots, m_M \in J^{M,N}$. We define the finite system of equations, which is truncated the infinite system given by (4.10).

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

rewrite the above system in vectorial form as

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

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

and the matrix A is given by

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

where $\lambda_i = \lambda_{mi}$ and $C_{i,j} = C_{n_i,m_j}$ para $1 \leq i,j \leq M$. Notice that, given the expression (4.11), in general the matrix A is not symmetric. We can now write the system (4.14) as a matrix differential equation:

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

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 of the eigenvalues and eigenvectors, or at least one of them, take values in the complex field we can still have real solutions. Indeed, suppose that we have the case with one complex eigenvalue and eigenvector then it is known that we will have $M-2$ real eigenvalues but we can obtain two real solutions from the complex eigenvalue.

Let us write one of the complex eigenvalues and eigenvectors as

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

then we can write two real solutions as follows:

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

4.2.1 Initial Conditions

In contrast to several types of differential equations, whether ordinary or partial, deterministic or stochastic, for FPK equations there is no standard way to determine the initial conditions. This is because in this type of equations we must choose a functional that acts on the initial condition, this implies that depending on the functional chosen we must adapt the method.

Here we present the method for two examples of functionals, but only one be devoloped. We will consider two cases:

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

$$u_0(g) := \int_0^1 g(z)dz.$$

For the first functional, define the set points into the set $[0,1]$ as

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

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 (4.2) we obtain

$$u(0,x) = \mathbb{E}[u_0^{z_i}(X_0^x)] = X^x(0,z_i) = x(z_i) \tag{4.17}$$

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 (B.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

$$u_m^{z_i}(0) \approx \sum_{k=1}^{M} \frac{e_k}{\lambda_k} \int_{\mathbb{R}} P_{m_k}(\xi_k)\xi_k(z_i)\mu(d\xi_k) \tag{4.18}$$

Setting the equation (4.18) 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

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

and the constants c_j are calculated as

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

4.3 Numerical Approximation to Stochastic Burgers' Equation

Let $\mathcal{H} = L^2(0,1)$. We consider the problem given by (1.14) over the interval $[0,1]$,

$$dX(\xi,t) = \left[\alpha\partial_\xi^2 X(\xi,t) + \frac{1}{2}\partial_\xi\left(X^2(\xi,t)\right)\right]dt + dW_t(\xi,t), \quad \xi \in [0,1] \tag{4.19}$$

The boundary condition and its initial condition are respectively

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

where W is a cylindrical Wiener process on $\mathcal{H}$ as was given by B.0.10, 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.

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 (4.1), with $Q = Id$, the equation (4.19) can be rewritten as

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

As we know the operator A is a infinitesimal generator of a strongly continuous semigroup e^{At} in $\mathcal{H}$, also is self-adjoint that has a complete orthonormal system of 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 $\Lambda^{-1/2}e_k = \sqrt{2\alpha}\pi|k|e_k$.

If we define $u(x,t) \in S(\mathbb{H})$ as was defined by (4.2), then u satisfies Kolmogorov equation given by (4.3). So, the descomposition for $u \in S(\mathbb{H})$ given by (4.6) gives

$$u(t,x) = \sum_{n\in\mathcal{J}} u_n(t)H_n(x), \ x \in \mathcal{H}, \ t \in [0,T], \tag{4.20}$$

to get the following system as was given by (4.10)

$$\dot{u}_m(t) = -u_m(t)\lambda_m + \sum_{n\in\mathcal{J}} u_n(t)C_{n,m}, \ n,m \in \mathcal{J} \tag{4.21}$$

We need to calculate the value of the constants $C_{n,m}$, then we need to calculate expressions such as $B(x)$, $D_xH_n(x)$. 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 $D_xH_n(x)$ we have

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

Therefore, $C_{n,m}$ given by (4.11) gives

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

For $N_1 \in \mathbb{N}$ define as before 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

$$\begin{aligned}\bar{C}_{n,m} =& \frac{1}{2}\sum_{j=1}^{\infty}\sqrt{2\alpha}\pi|j|\int_{\mathbb{R}^{\mathcal{M}}} P_{m_j}(\xi_j)P'_{n_j}(\xi_j)\mu(d\xi_j) \\ & \cdot \prod_{i=1, i\neq j}^{M} P_{m_i}(\xi_i)P_{n_i}(\xi_i)\mu(d\xi_i)\sum_{l=1}^{M}\sum_{k=1}^{M}\beta_l\beta_k(e_l e'_k + e'_l e_k)\end{aligned}$$

So, the finite system of ordinary differential equations has the following form

$$\dot{u}_m(t) = -u_m(t)\lambda_m + \sum_{n\in S_N} u_n(t)\bar{C}_{n,m},\ n, m \in S_N \tag{4.22}$$

Therefore (4.22) approximates to the infinite system of ordinary differential equations (4.10) when $N, M \to \infty$.

4.4 Something about Numerical Analysis

In this section, we will study some results of the numerical analysis of the method described above that have been studied in [35]. Something that we are going to observe is that the analysis is very similar to the deterministic case we saw in the previous chapter.

Consider the following linear stochastic equation

$$du_t = Au_t dt + dW_t,\ \ u_0 = h \in \mathcal{H},$$

where $A : \mathcal{D}(A) \subset \mathcal{H} \to \mathcal{H}$ is the infinitesimal generator of a strongly continuous semi-group e^{At} in $\mathcal{H}$ and W_t is a Q-Wiener process in $\mathcal{H}$.

The solution of the above equation is a time-homogeneous Markov process with transition operator P_t defined for $\Phi \in \mathbb{H}$ given by

$$(P_t\Phi)(h) = \int_{\mathcal{H}} \Phi(v)\mu_t^h(dv) = \mathbb{E}\left[\Phi(u_t^h)\right].$$

Let $\Phi \in S(\mathbb{H})$ be a smooth simple functional. By setting $\varphi_k = e_k$ in (B.5), it takes the form $\Phi(h) = \phi(l_1(h), \cdots, l_n(h))$, where $l_k(h) = (h, \Lambda^{-1/2}e_k)$. Define a differential operator A_0 on $S(\mathbb{H})$ by

$$A_0\Phi(v) = \frac{1}{2}Tr[RD^2\Phi(v)] + \langle Av, D\Phi(v)\rangle,\quad v \in \mathcal{H}$$

which is well defined, since $D\Phi \in D(A)$ and $\langle Av, D\Phi(v)\rangle_{\mathcal{H}} = \langle v, AD\Phi(v)\rangle_{\mathcal{H}}$.

Lemma 4.2. *Let P_t be the transition operator as defined above. Then the following properties hold:*

1. $P_t : S(\mathbb{H}) \to S(\mathbb{H})$ *for* $t \geq 0$.
2. $\{P_t, t \geq 0\}$ *is a strongly continuous semigroup on* $S(\mathbb{H})$ *so that, for any* $\Phi \in S(\mathbb{H})$, *we have* $P_0 = I$, $P_{t+s}\Phi = P_t P_s \Phi$, *for all* $t, s \geq 0$, *and* $\lim_{t \downarrow 0} P_t \Phi = \Phi$.
3. A_0 *is the infinitesimal generator of* P_t *so that, for each* $\Phi \in S(\mathbb{H})$,

$$\lim_{t \downarrow 0} \frac{1}{t}(P_t - I)\Phi = A_0 \Phi.$$

Lemma 4.3. *Let $H_n(h)$ be a Hermite polynomial functional given by (B.5). Then the following hold:*

$$\begin{aligned} A_0 H_n(h) &= -\lambda_n H_n(h), \\ P_t H_n(h) &= \exp(-\lambda_n t) H_n(h), \end{aligned}$$

for any $n \in J$ *and* $h \in \mathcal{H}$, *where* $\lambda_n = \sum_{i=1}^{\infty} n_i \lambda_i$.

By above Lemma and B.1, for $\Phi \in \mathbb{H}$, it can be represented as

$$\Phi(v) = \sum_{n=0}^{\infty} \phi_n H_n(v),$$

where $n = |n|$ and $n \in J$. Notice that we can think in n as a vector of r dimension, i.e., $n = (n_1, \ldots, n_r)$. Let $\alpha_n = \alpha_{n_1}, \ldots, \alpha_{n_r}$ be a sequence of positive numbers with $\alpha_n > 0$, such that $\alpha_n \to \infty$ as $n \to \infty$. Define

$$|||\Phi|||_{k,\alpha} = \left[\sum_n (1 + \alpha_n)^k |\phi_n|^2 \right]^{1/2}$$

$$|||\Phi|||_{0,\alpha} = |||\Phi||| = \left[\sum_n |\phi_n|^2 \right]^{1/2}$$

which is $L^2(\mu)$-norm of Φ. For the given sequence $\alpha = \{\alpha_n\}$, let $\mathbb{H}_{k,\alpha}$ denote the completion of $S(\mathbb{H})$ with respect to the norm $|||\cdot|||_{k,\alpha}$.Then $\mathbb{H}_{k,\alpha}$ is called a Gauss-Sobolev space of order k with parameter α. The dual space of $\mathbb{H}_{k,\alpha}$ is $\mathbb{H}_{k,-\alpha}$. From now on, we will fix the sequence $\alpha_n = \lambda_n$, where λ_n is given in Lemma 4.3. We shall simply denote $\mathbb{H}_{k,\alpha}$ by $\mathbb{H}_k$ and $|||\Phi|||_{k,\alpha}$ by $|||\Phi|||_k$.

Consider the following Kolmogorov equation,

$$\frac{\partial}{\partial t}\psi(v,t) = A\psi(v,t) + \langle B(v), D\psi(v,t)\rangle_{\mathcal{H}}, \text{ a.e. } v \in \mathbb{H}_2, \tag{4.23}$$

$$\psi(v,0) = \phi(v), \tag{4.24}$$

where $A : \mathbb{H}_2 \to \mathbb{H}$ is given as in (4.4)

$$A\Phi = \frac{1}{2}Tr[RD^2\Phi(v)] + \langle Av, D\Phi(v)\rangle$$

We will study a weak solution of the above equation. Let $\lambda > 0$ be a parameter. By changing ψ to $e^{\lambda t}\psi$ in above equation we get the following equation:

$$\frac{\partial}{\partial t}\psi(v,t) = A_\lambda\psi(v,t) + \langle B(v), D\psi(v,t)\rangle_{\mathcal{H}}, \text{ a.e. } v \in \mathbb{H}_2,$$

$$\psi(v,0) = \phi(v),$$

where $A_\lambda = A - \lambda I$, with I the identity operator in $\mathbb{H}$.

Denote by P_t the semigroup with infinitesimal generator A_λ. Then, we can rewrite the last equation in an integral form by using the semigroup P_t

$$\psi(v,t) = e^{-\lambda t}(P_t\phi)(v) + \int_0^t e^{-\lambda(t-s)}\left[P_{t-s}(B, D\Psi_s)\right](v)ds,$$

where $\phi = \phi(\cdot)$ and $\Psi_s = \Psi(\cdot, s)$.

Let $\mathbb{X}_T$ denote the Banach space $C([0,T];\mathbb{H})$ with the norm

$$|||\Psi|||_T := \sup_{0\le t\le T} |||\Psi|||.$$

***Theorem* 4.1.** *Suppose that $B : \mathcal{H} \to \mathcal{H}_0$ satisfies $(B, D\Phi) \in L^2((0,T);\mathbb{H})$ for any $\Phi \in \mathbb{H}$ and*

$$\sup_{v\in\mathcal{H}} \|\Lambda^{-1/2}B(v)\|_{\mathcal{H}} < +\infty.$$

Then, the unique weak solution $\Phi \in C((0,T);\mathbb{H})$ for (4.23) depends continuously on the initial conditions, i.e.,

$$|||\psi_t^\varphi - \psi_t^\varphi||| \le \exp(Ct)|||\varphi - \psi|||.$$

Existence and uniqueness is ensure provided that operator $\mathbb{Q}$ in $\mathbb{X}_T$ defined as

$$\mathbb{Q}\Psi = e^{-\lambda t}P_t\Phi + \int_0^t e^{-\lambda(t-s)}P_{t-s}(B, D\Psi_s)ds,$$

for any $\Psi \in \mathbb{X}_T$, is a contraction in $\mathbb{X}_T$ under the same assumptions as the previous theorem, i.e., for $\Psi, \Psi' \in \mathbb{X}_T$ and for small T

$$|||\mathbb{Q}\Psi - \mathbb{Q}\Psi'|||_T \le C\sqrt{T}|||\Psi_s - \Psi'|||_T.$$

which is obtained as a consequence of

$$|||(B(v), D\Phi(v))|||^2_{-1} \leq C|||\Phi(v)|||^2$$

for any $\Phi \in \mathbb{H}$, $v \in \mathbb{H}_2$, and for some $C > 0$.

Recall that the solution $u(x,t)$ for (4.3) is given by

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

where $\varphi : \mathcal{H} \to \mathbb{R}$ and X_t^x is the solution to (4.1) with initial conditions $X_0 = x \in \mathcal{H}$. Then by Lemma 4.3 we can written the solution ψ_t^x as follows

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

The following result show the continuity with respect to the initial conditions for a numerical approximation of the Kolmogorov equation associated with an SPDE. Here we understand that a numerical scheme is stable respect to initial conditions if this method reproduces the same behavior when the continuous problem satisfies continuity respect initial conditions.

***Theorem* 4.2.** *Assume that the eigenvalues of Λ, satisfies that for every $k \in \mathbb{N}$, $\lambda_k < \lambda_{k+1} \to \infty$. Assume that the functional φ is Lipschitz. Then, the numeric approximation Ψ_t^φ to the solution of the Kolmogorov equation $\psi \in C((0,T),\mathbb{H})$ depends continuously on the initial conditions, i.e., for two different initial values $x, y \in \mathcal{H}$*

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

for some $C > 0$ and $f(t)$ is given by

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

Equivalently, if $\|x-y\|_{\mathcal{H}} \leq \delta$ then

$$\|\Psi_t^x - \Psi_t^y\| \leq G(t)\delta.$$

For ilustrate the above Theorem, numerical experiments were performed using as initial condition $x(\xi)$ for equation (4.19) and its truncated Chebyshev expansion with polynomials of the first kind given by

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

with a discretization of 2048 points in spatial variable ξ over $[0,1]$, 1024 points in time variable t over $[0,10]$, and with parameters $\alpha = 0.01$, $N = 5$, $M = 11$ for using approximation given by (4.22).

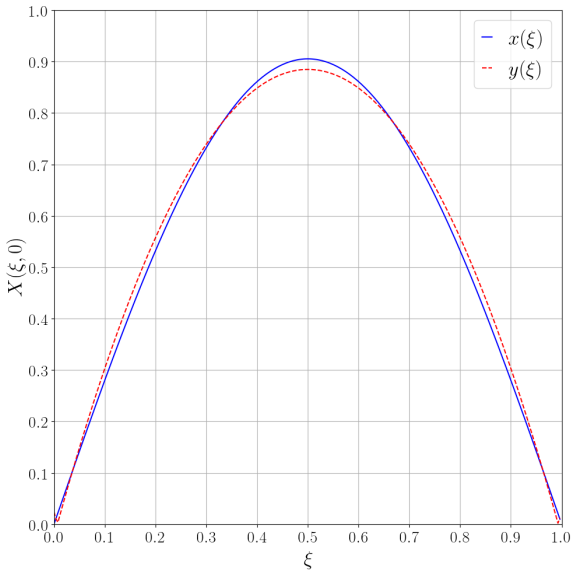


Figure 4.1: Initial condition for (4.19) and its approximation.

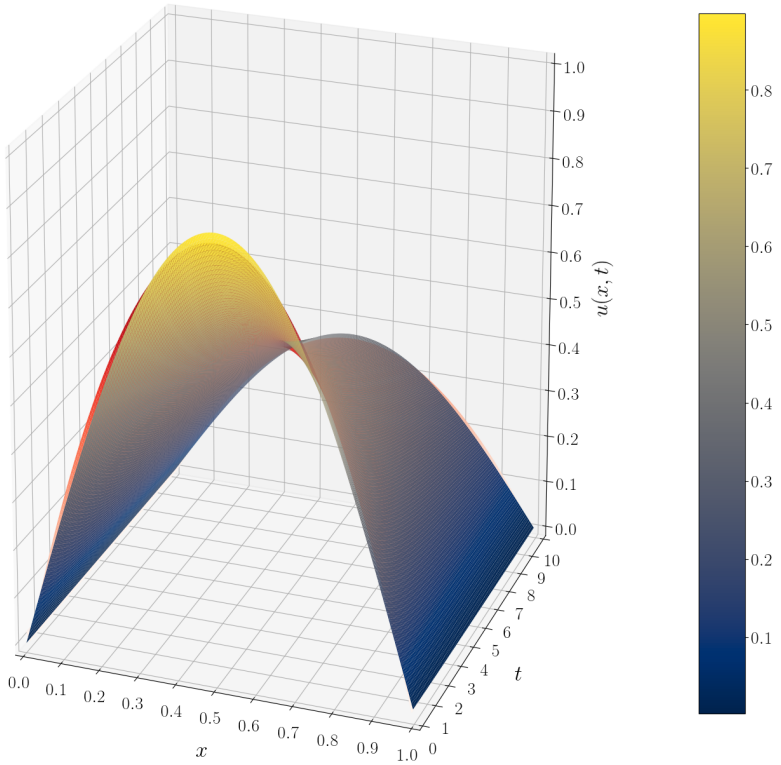


Figure 4.2: Numerical solutions for (4.19) with initial conditions $x(\xi)$ and $y(\xi)$.

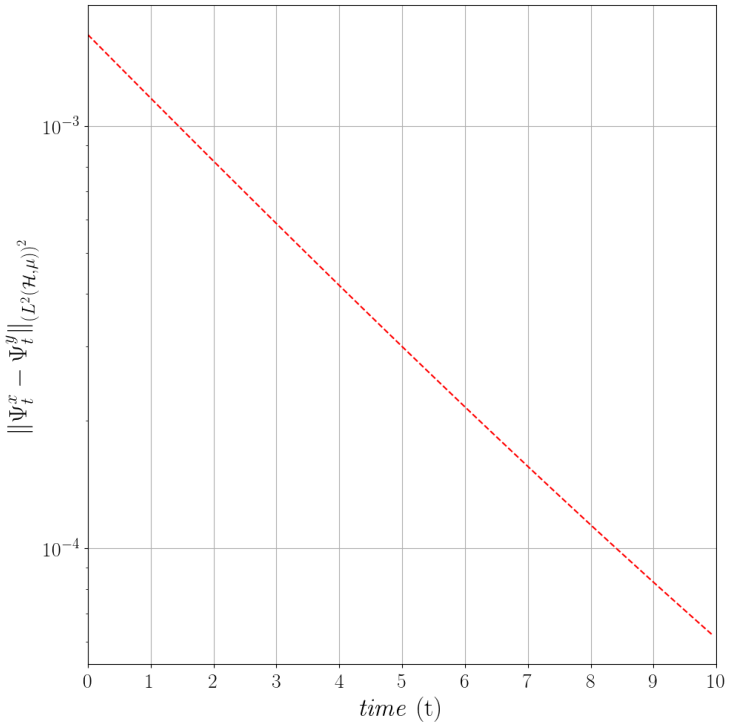


Figure 4.3: Distance between the numerical solutions for equation (4.19) with initial conditions $x(\xi)$, and $y(\xi)$.

The stability version presented above says merely that method preserve the continuity respect to initial conditions, i.e., if a given problem satisfies certain regularity conditions, then two of its solution remain closed if its initial function conditions are close. So, we desire that a numerical method reproduce this behavior and if it is the case, we say that an underlying method is stable in this context. However, the stability theory for spectral methods is still under construction and is an active research area.

This kind of stability, combining with the weak approximation approach, would save computation time. Since the scheme asks specific conditions to obtain a weak numerical solution of an underlying SPDE, so convert the stochastic problem into a deterministic ODE for the first moment. This procedure overcome Montecarlo type simulations to approximate moments or distributionssimulate many realization of the numerical stochastic process to approximate distributions or moments.

Chapter 5

Discussion and Conclusions

Based on the study and analysis performed in Chapters 2 and 3, we can ensure that spectral methods are an excellent choice to solve problems where solutions are well behaved or are smooth, such as periodic problems or rather solutions that vanishes at bounds.

In addition, it was possible to understand some of the advantages and disadvantages that may arise in practice, such as in the cases for small viscosity coefficients studied at the end of Chapter 3 we could see that the order of convergence was lower but the solution approached the case of zero viscosity which if it has an exact solution and that could be used as a good approximation.

However, it was possible to note that the problem for these cases was due to a discontinuity or excessive changes in function, which produces strong oscillations around them and is also known as the Gibbs phenomenon, which can be reduced by choosing a adequate initial condition or consider longer intervals in the spatial variable. Although for this there are some techniques that can help improve the accuracy of these methods, one of them is to use non-uniform discretizations in the spatial variable such as Chebyshev nodes and using the known Chebyshev polynomials as bases.

In the same way with respect to the stochastic version of Burgers' equation, we can say that the methods are good choice to approximate their solutions, since apparently in the Chapter 4 they are very similar in theory and implementation in comparison with the deterministic version.

But nevertheless in general, the great advantage of these methods, whether they are implemented to the deterministic or stochastic version, is that they are easy to implement and develop very efficient computational codes in order to make numerical analysis studies or rather predictions of some phenomenon.

It would be very interesting to be able to extend these same ideas for more complex models, either for non-linear problems where their solutions are not smooth and discontinuities occur such as the case of Burgers' equation without viscosity, or even more, studying under this approach the famous problem of Navier-Stokes, whether in his classic deterministic version or his stochastic version, which could be an excellent work in the future to study.

Appendix A

Some results of Hilbert's space theory

A.1 Important Inequalities in Hilbert Spaces and Some about Semi-Group Theory

A.1.1 The Cauchy-Schwarz inequality

Let X be a Hilbert space, endowed with the inner product $\langle u, v\rangle$ and the associated norm $\|u\|$. The Cauchy-Schwarz inequality states that

$$|\langle u, v\rangle| \leq \|u\|\|v\|, \text{ for all } u, v \in X$$

A.1.2 The Poincaré Inequality

Let v be a function of $H^1(a, b)$. We know that v is continuous on $[a, b]$. Assume that at a point $x_0[a, b]$, $v_0(x_0) = 0$. The Poincaré inequality states that there exists a constant C (depending upon the interval length) such that

$$\|v\|_{L^2[a,b]} \leq C\|v'\|_{L^2[a,b]}$$

i.e., the L^2-norm of the function is bounded by the L^2-norm of the derivative. The Poincaré inequality applies to functions belonging to $H_1^0(a, b)$, where

$$H_0^1(a, b) = \{v \in H^1(a, b) : v(a) = v(b) = 0\},$$

for which $x_0 = a$ or b. and also to functions of $H^1(a, b)$ that have zero average on (a, b), since necessarily such functions change sign in the domain.

Definition A.1.1. A collection $\{T(t)\}$, where $t \in [0, \infty)$, of bounded linear operators in X is called strongly continuous if

(1) $T(s + t) = T(s)T(t)$ for all s, $t > 0$

(2) $T(0) = I$ (identity operator)

(3) For each $x \in X$, $T(t)x$ is continuous in t on $[0, \infty)$.

Definition A.1.2. Let $h > 0$. Then, A is called an infinitesimal generator of the semigroup $\{T(t)\}$ if

$$Ax = \lim_{h\to 0} \frac{T(h)x - x}{h} \tag{A.1}$$

Definition A.1.3. Let $\{T(t)\}$ be a C_0 semigroup on a Banach Space X with infinitesimal generator A. Then, $\{T(t)\}$ is said to be an analytic semigroup if:

(1) For some $\phi \in (0, \pi/2)$, $T(t)$ can be extended to Δ_ϕ, where:

$$\Delta_\phi = \{0\} \cup \{t \in \mathbb{C} : |\arg(t)| < \phi\}$$

(2) For all $t \in \Delta_\phi - 0$, we have that $T(t)$ is analytic in t in the uniform operator topology.

In a less formal manner, analytic semigroups are C_0 semigroups in which each $T(t)$ has an analytic continuation to the sector Δ_ϕ, in which the local power series representation of $T(t)$ converges in norm. As we shall see, this type of semigroup has a natural association to the Abstract Cauchy Problem

There are some elementary definitions of the exponential function, we will use the following definition.

$$e^{At} = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!} \tag{A.2}$$

***Theorem* A.1.** *Let $A : X \to X$ be a bounded linear operator. Then,*

$$T = \left\{ T(t) = e^{At} = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!} \right\}$$

is a uniformly continuous semigroup.

Proof. Firstly, $\|A\| \leq \infty$ since our operator is bounded. We first show that $T(s)T(t) = T(s+t)$.

$$T(s)T(t) = e^{As}e^{At} = \sum_{n=0}^{\infty} \frac{(sA)^n}{n!} \sum_{n=0}^{\infty} \frac{(s+t)A)^n}{n!} = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!}$$

This of course holds by the properties of the exponential. Also, setting $t = 0$, it is obvious that the only term in our summation is I,the identity operator.

Finally, to show this is a uniformly continuous semigroup, we need to show that $T(t) \to I$ as $t \to 0$ in norm. We see:

$$\|T(t) - I\| = \|\sum_{n=1}^{\infty} \frac{(tA)^n}{n!}\| \leq \sum_{n=1}^{\infty} \frac{(t\|A\|)^n}{n!} = e^{t\|A\|} - 1.$$

Letting $t \to 0$, we see that the norm tends to 0. □

Lemma A.1. *Suppose that above conditions are satisfied. Then $-A$ is the infinitesimal generator of an analytic semigroup in X, denoted by $\left\{e^{-tA}\right\}_{t \geqq 0}$. For a proof of this lemma, see Kato [30].*

The next lemma is the version of Gronwall's inequality we are going to use.

Lemma A.2. *Let T, α, β, ν be positive constants. $0 < \nu < 1$. Then for any continuous functions $f : [0, T] \to [0, \infty)$ satisfying*

$$f(t) \leqq \alpha + \beta \int_0^t (t-s)^{-\nu} f(s) ds, \quad 0 \leqq t \leqq T,$$

we have

$$f(t) \leqq C\alpha \; exp\big\{C\beta^{1/(1-\nu)}\big\}, \quad 0 \leqq t \leqq T,$$

with a positive constant C that depends only on ν.

A.2 Distributions (or generalized functions) Theory

Definition A.2.1. The support of a function $f(x)$ is the lock of the set of all points x such that $f(x) = 0$. We will denote the support of a f function for $sop(f)$. If $sop(f)$ is a bounded set, then it is said that f has compact support.

Let us denote D_m as the complex function space $\varphi(x)$ defined in $\mathbb{R}^n$ with continuous partial derivatives up to order m and compact support. This view is also taken as the basis for the definition of an arbitrary generalized function. Accordingly, consider the space D consisting of functions at real values $\varphi(x) = \varphi(x_1, x_2, \ldots, x_n)$, such that the following is satisfied:

- $\varphi(x)$ is an infinitely differentiable function defined at every point of $\mathbb{R}^n$. This means that D_α exists for all multi-indexes. Such a function is also called a C^∞ function.
- There is a A number such that $\varphi(x)$ is canceled for $r > A$. This means that $\varphi(x)$ has compact support.

***Theorem* A.2.** *Let $U \subset R^n$ open and $f \in L^1_{loc}(U)$. Then f can be identified with the distribution in U $f : D(U) \to C$ by using the formula*

$$\langle f, \varphi \rangle = \int_{R^n} f(x)\varphi(x)dx.$$

In addition, a distribution defined by this formula is called regular. If a distribution is not regular, it is called singular.

Regular distributions can also be defined by partial differential operators. If $f(x) \in L^1_{loc}$, we can define a distribution as

$$\langle f, \varphi \rangle = \int_{R^n} f(x) D^\alpha \varphi(x) dx, \;\; \varphi \in D.$$

Let f be an integrable function especially compact, except for those that contain certain singular points: to simplify, we will assume a single singular point, in x_0. The functional f such that

$$\langle f, \varphi \rangle = \int_{R^n} f(x)\varphi(x)dx, \;\; \varphi \in D.$$

it is not a distribution over $\mathbb{R}^n$, the right side is only defined, in general, if the support excludes x_0; In other words, the functional f is only a mostly open distribution that excludes x_0.

The problem of regularization is as follows: find a distribution over $\mathbb{R}^n$ that is reduced to the f distribution over the open ones that exclude x_0. It will be said that this distribution (which extends, in the environment of x_0, the definition of the f distribution outside this environment), is a regularization of the f function.

Appendix B

Elements of Probability

Definition B.0.1. 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 B.0.2. 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.,

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

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 B.0.3. 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 B.0.4. 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 B.0.5. 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 $\mathbb{E}\left[|X - \mathbb{E}[X]|\right]$, $p = 1, 2, \ldots,$. When $p = 2$, the centered moment is also called the variance.

Definition B.0.6. 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 B.0.7. 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 B.0.8. 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.

Definition B.0.9. 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

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

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

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

Definition B.0.10. 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

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

The Hermite polynomial of grade n evaluated in $\mathbb{R}$ is given as follows

$$P_k(x) = \frac{(-1)^k}{(k!)^{1/2}} e^{\frac{x^2}{2}} \frac{d^k}{dx^k} e^{-\frac{x^2}{2}} \tag{B.4}$$

with $P_0 = 1$. Is well known that $\{P_k(\dot{)}\}_{k\in\mathbb{N}}$ is a complete orthonormal system for $L_2(\mathbb{R}, \mu_1(dx))$ with $\mu_1(dx) = \frac{1}{\sqrt{2\ pi}} e^{-\frac{x^2}{2}} dx$.

Define Λ as the nuclear covariance operator and Hermite functional given by as follows

$$H_n(h) = \prod_{i=1}^{\infty} P_{n_i}(l_i(h)), h \in \mathcal{H}_0, n \in J \tag{B.5}$$

where

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

and $\mathcal{J}$ is given by

$$\mathcal{J} = \{\alpha = (\alpha_i, i \geq 1) | \alpha_i \in \mathbb{N} \cup 0, |\alpha| := \sum_{i=0}^{\infty} \alpha_i < \infty\} \tag{B.6}$$

Then we have the following result (see [24])

Lemma B.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 $\{H_n : |n| = j\}$.

Hence we can the followin decomposition given by

$$u_N(t, x) = \sum_{n \in \mathcal{J}^{M,N}} u_n(t) H_n(x),\ x \in \mathcal{H},\ t \in [0, T] \tag{B.7}$$

which is known as the deterministic Wiener-Chaos descomposition.

We define the inner product as follows

$$\langle g, h \rangle_0 = \langle \Lambda^{-1/2} g, \Lambda^{-1/2} h \rangle_{\mathcal{H}}, \text{ for } g, h \in \Lambda^{-1/2}\mathcal{H}.$$

Set $\mathcal{H}_0$ denote the Hilbert subspace of $\mathcal{H}$, which is the completion of $\Lambda^{-1/2}\mathcal{H}$ with respect to the norm $\|g\|_0 = (g, g)_0^{1/2}$. Then $\mathcal{H}_0$ is dense in $\mathcal{H}$ and the inclusion map $i : \mathcal{H}_0 \longrightarrow \mathcal{H}$ is compact. The triple $(i, \mathcal{H}_0, \mathcal{H})$ forms an abstract Wiener space.

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

$$\langle \Phi, \Psi \rangle_{\mathbb{H}} = \int_{\mathcal{H}} \Phi(v)\Psi(v)\mu(dx), \quad \Phi, \Psi \in \mathbb{H}, \tag{B.8}$$

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 $\varphi_k \in \mathcal{H}$.

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 $h \in \mathcal{H}$

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

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 conservation 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 viscosity 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 NUMERICAL APPROXIMATION FOR KOLMOGOROV EQUATIONS *. 10.13140/RG.2.2.21593.47203.