CIMAT

Centro de Investigación en Matemáticas, A.C.

LOCAL DISCONTINUOUS GALERKIN METHODS FOR DETERMINISTIC AND STOCHASTIC PARABOLIC CONSERVATION LAWS

T E S I S

Que para obtener el grado de

Maestro en Ciencias

con Orientación en

Matemáticas Aplicadas

Presenta

Sebastián Gutiérrez Hernández

Director de Tesis:

Dra. Silvia Jerez Galiano

Guanajuato, Gto., 02 de Julio de 2021

CIMAT

Centro de Investigación en Matemáticas, A.C.

LOCAL DISCONTINUOUS GALERKIN METHODS FOR DETERMINISTIC AND STOCHASTIC PARABOLIC CONSERVATION LAWS

T E S I S

Que para obtener el grado de

Maestro en Ciencias

con Orientación en

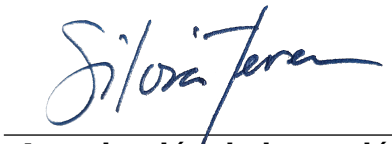
Matemáticas Aplicadas

Presenta

Sebastián Gutiérrez Hernández

Director de Tesis:

Dra. Silvia Jerez Galiano

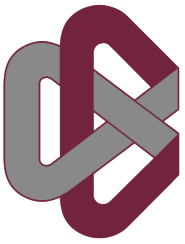


Autorización de la versión

Guanajuato, Gto., 02 de Julio de 2021

Centro de Investigación en Matemáticas

Matemáticas Aplicadas



CIMAT

Centro de Investigación en Matemáticas, A.C.

Local discontinuous Galerkin methods for deterministic and stochastic parabolic conservation laws

TESIS

QUE PARA OBTENER EL TÍTULO DE:

Maestro en Ciencias con Especialidad en Matemáticas Aplicadas

PRESENTA:

Sebastián Gutiérrez Hernández

ASESOR:

Dra. Silvia Jerez Galiano

Guanajuato, Guanajuato, México, 02 de Julio de 2020

‘

SEBASTIÁN GUTIÉRREZ HERNÁDEZ

Fecha

Firma

Agradecimientos

Agradezco al CONACYT por haberme proporcionado los medios para poder seguir preparándome como matemático. Al CIMAT por ser una institución que le preocupan sus estudiantes y que nos permite concéntranos en nuestro desarrollo académico. A mis profesores en CIMAT por la formación que me han brindado. En particular quiero agradecerle a mi asesora, la Dra. Silvia Jerez, por su apoyo durante el desarrollo de mi tesis. También quiero agradecer particularmente al Dr. Héctor Chang Lara, a la Dra. Eloisa Días Francés y al Dr. Miguel Nakamura Savoy por todo su apoyo durante la maestría.

Quiero agradecer también a mi familia, Norma Edith Hernández Abaraca, Eduardo Raúl Gutiérrez Albores y Eduardo Gutiérrrez Hernández por su apoyo incondicional y que aunque esté lejos sepan siempre los llevo conmigo. También agradecer a los excelentes amigos y compañeros de cubículo, que aunque nos hayamos distanciado por la pandemia, siempre los querré mucho. En particular quiero agradecer a Lili por su cariño y a Arturo y Yamil por su amistad. A viejos amigos que se han quedado conmigo, Mario, Fer, Axel, Lalo y Misha y que espero que siempre estemos cerca.

Abstract

The use of stochastic differential equations has been popularized in different areas such as Physics, Mathematical biology, and Finance. The analytic solution of this type of equations is in most cases extremely difficult or impossible and numerical methods must be used for their solution. In this work, we use the discontinuous Galerkin method (DG) for the solution of conservation laws with multiplicative white noise, and the local discontinuous Galerkin (LDG) for the solution of parabolic equations with multiplicative white noise. We first do a revision of classic, weak and entropy solutions for the conservation laws and parabolic PDEs. Once understood the deterministic problems we study the concept of Brownian motion, stochastic differential equations and add a multiplicative white noise in time to the conservation laws and parabolic problems. Then we proceed to study the discontinuous Galerkin method for deterministic conservation laws and the local discontinuous Galerkin for parabolic problems. In both cases, we include numerical examples to study their performance and order of convergence. In addition, we compare the performance of the LDG method with other methods for the convective-dominated convection-diffusion equations. As part of the numerical experimentation, we observe the oscillatory behavior that the DG method has for discontinuous solutions. To overcome this problem we study the use of a two step process called limiters. First, we detect elements of the domain that have oscillations and second we reconstruct the solutions in elements with oscillatory behavior. We work with the TVB and BDF limiters and propose a modified version of these limiters, the ATVB and the M-BDF limiters. A great disadvantage the DG and LDG methods have is the use of analytic integration in their weak formulations. We work with a nodal entropy DG that uses the Gauss-Lobatto quadrature to avoid analytical integration which ensures entropy stable solutions. Then we apply the DG for the stochastic convection equation and the stochastic Burgers equation, including the use of limiters and entropy stable solutions. The stochastic heat equation and linear convection-diffusion equations are solved with the LDG method. The inclusion of limiters and entropy stable solutions for stochastic conservation laws are a contribution made in the thesis, they had not been previously studied.

Contents

List of Figures

List of Tables

List of Algorithms

List of Symbols

The next list describes several symbols that will be later used within the body of the document

$(\Omega, \mathcal{F}, \mathbb{P})$ Probability space.

$[\cdot, \cdot]$ Bilinear form.

$\mathcal{A}$ Linear partial differential equation in divergence form.

$\mathcal{C}^p$ Set of continuously differentiable up to order p.

$\mathcal{D}'(\cdot)$ Space of Schwartz distributions.

$\mathbb{E}$ Expectation.

$\mathcal{F}$ Sigma field.

$\mathcal{D}(\cdot)$ Set of infinite differentiable functions.

μ_k $k-$th moment of a random variable

$\nabla \cdot u$ Divergence of u.

∇u Gradient of u.

Ω State-space.

$\mathcal{O}$ Bounded domain.

$\overline{A}$ Principal part of a second order partial differential equation.

∂^α Multi index partial derivative.

Φ_j Hat function.

$\Phi_X(t)$ Characteristic function.

$\mathbb{P}$ Probability distribution.

σ_k $k-$th central moment of a random variable.

$\mathcal{W}^{(m,p)}$ Sobolev space of orders (m, p).

$\{X(\tau), \tau \in T\}$ Stochastic process, with parameter set T.

A Stiffness matrix.

B Load vector.

$F_X(x)$ Distribution function.

H Hilbert Space.

L Elliptic differential operator.

$M_X(t)$ Moment generating function.

$N(\mu, \sigma^2)$ Normal density.

$supp\phi$ Support of ϕ.

T Distribution.

u_x Partial derivative of function u respect to variable x.

X Random variable.

Chapter 1

Introduction

1.1 Motivation

Conservation laws and convection-diffusion equations in their deterministic version have a wide range of applications. Conservation laws are equations that describe quantities such as mass, momentum, and energy or the distribution of objects like cars, individuals in population models or microscopic particles in space and time. The basic principle is that the change in time of the total amount of the object of study in a region, equals to the amount of the object that enters or exits through the boundary of the region and what it is produced or consumed inside the domain. This intuition is formalized and generalized by the so called *Transport theorem*, which is a generalized version of the differentiation under an integral sign. This theorem assumes that the domain in which the total amount of the object of study is computed is changing its position and boundary in time. It relates the change in time of the total amount of the object of study to the amount of quantity that enters and exits the domain in time due to the flux and the produced or consumed quantity inside the domain in time. In addition, these laws assume that the *Continuum hypothesis* holds. This hypothesis allows us to consider a continuous model for situations that are not continuous, for example, the modeling of a population. In the population example, we are interested in knowing how the number of individuals, a discrete quantity, behaves in space and time. The idea is that if we view the population from far away enough we lose the view of the individuals and the population seems to be continuous. Think of a wave in a football stadium, if we focus on an individual we do not see a wave, but if we watch it from far away (on the other side of the field or with a drone) we do not see individuals and only see the wave.

We will work with two examples, the convection equation and Burgers' equation. Both of them can be interpreted in terms of a traffic model. The first assumes that there is heavy traffic and that the cars move at a constant speed in one direction. The second one models the velocity profile of vehicles $u(x,t)$. At each time $\hat{t}$ the map $x \to u(x,\hat{t})$ describes the velocity of a vehicle at the spatial position x at the time $\hat{t}$. Denoting the trajectory of a vehicle by $x(t)$, this model assumes that the speed of each vehicle $\dot{x}(t)$ satisfies $\dot{x}(t) = u(x(t),t)$, with $u(x,t)$ the solution to the burgers equation evaluated at the trajectory of the vehicle $x(t)$. It can be shown that under those hypothesis, the acceleration of the vehicle $\ddot{x}(t)$ is zero. The vehicle density is allowed to vary on a range, the highway is full or empty. The importance of the Burgers' equation is that is the simplest model with a flux that is nonlinear

and it is helpful to understand the behaviours of these kind of equations. The Burgers equation has applications in interference dynamics, cosmology, statistical mechanics and hydrodynamical turbulence [10].

The second type of equations we consider are the convection-diffusion equations, they add diffusion to a conservation law. Which means that in addition to the transport of our quantity of interest (dynamics of the conservation law), we add to the macroscopic model the effects of what is occurring at a molecular level. We can understand intuitively the diffusive part by thinking in the following experiment. Consider a glass with water in which a droplet of colorant is put on the center of the glass of water. Assume that we have put the colorant with a lot of care so that there is no impulse by the effect of introducing the droplet. As time progresses the water will start to be mixed with the colorant. This mixing occurs due to the movement of the colorant molecules, which flow from highly colorant concentrated areas to less concentrated areas. The convection-diffusion equation is the mixture of these two phenomena, the transport of our quantity of interest plus the macroscopic effects of the molecular behavior.

An example where convection-diffusion equations can be used is in the modeling of the concentration of a pollutant in a river. The flow of the river transports the pollutant from one part to another of the river, the effect of convection. If we moved at the river's speed we would be able to see the effects of the diffusion of the contaminant in the water, going from regions with higher density to regions with lower density. The example of the concentration of a pollutant in a river helps us to introduce a type of problem we will focus on, namely the convective-dominated equations. These equations are dominated by the convective part of the equation. For example, the speed of the river is much greater than the speed at which the contaminant is diffusing in the river. Since the convective part is dominant, the solution to these equations is closer to the solution of conservation laws. We will study that the solutions of conservation laws can be problematic and have discontinuities, something unexpected for a differential model. Due to the dominance of the convective phenomena, most numerical schemes have problems while solving convective-dominated equations.

Although the previous models can help to study what is observed in nature they still miss an important fact: we do not always certainly know how to tune the parameters of the equations to match a particular event. For example, assume that there is a factory that is dropping pollutants on a river. If the factory is not willing to provide information on what exactly they are dropping into the river and how much pollutant they are dropping, how can the model be used? We could do a chemical analysis of the water and estimate the density of the pollutant (a key variable for the model) by measuring it at different times of the day. If these measurements vary randomly during the day we can not assign a simple function to the density. It might be better to model the variation with a random variable and assume each measurement we did is a realization of the random variable. With respect to how much pollutant the company is leaking, we could as well model it with a random variable. Note that although these two phenomena are related to the same process, each introduces a different type of stochastic process. In the first one, a random variable modeling the density, the quantity of interest itself becomes a stochastic process. Whereas the second case, a random variable modeling the amount of pollutant introduced, introduces a stochastic process to the model and therefore the solution becomes one.

The solution of these two types of equations changes a lot, even numerically. In this work, we focus on the second type of equation, where we model the introduction of the pollutant with a random variable. In particular, we will use a white noise which is the "derivative" of a Brownian motion. The reason to consider this model is that we are assuming that the pollutant is being introduced

continuously in time. This model is the easiest one to work with. More complicated models, for example, using a Levy process are out of the scope of this work. Note that we are assuming that the position of where the pollutant is being introduced is deterministic. We could model the position varying with a Brownian motion as well, but this would correspond to a white noise in space time, a much more complicated problem. The particular model we assume, is a *multiplicative* white noise. Which means that we add to the deterministic equation a function depending on the quantity of interest multiplied by the white noise. In case of the function being constant or independent to the quantity of interest, the equation has additive white noise. Thus the formulation of numerical solution for problems with multiplicative white noise is more general than a formulation for equations with additive noise.

1.2 Background

The study of numerical solutions to conservation laws and convection-diffusion equations is a well establish area of study, in particular in the computational fluid dynamics community [16], [25] , [27], [46], [53] and [72].

The numerical solution of the stochastic conservation laws and convection-diffusion equations with multiplicative white noise had been studied using finite differences techniques [6],[7] and [9]. The use of Discontinuos Galerkin (DG) method and Local Discontinuous Galerkin (LDG) method have been recently studied in [47] and [48].

Conservation laws might have discontinuous solutions even for continuous initial conditions. Most numerical methods suffer from the Gibbs phenomena in discontinuous solutions, oscillations appear near discontinuities. For the original formulation of the DG method, Cockburn in [16] borrowed a technique from the finite differences community called the slope limiter to control oscillations. As part of the development of the DG and LDG methods new limiting techniques have been developed. In [77] there is a panoramic perspective of the articles published up to 2013. More modern formulations are the use of wavelets [14], WENO reconstructions [59], artificial neural networks [60], convolutional neural networks [69] and the use of a multi-layer perception [74].

The DG and LDG methods use the weak formulation of the PDE to define the numerical method. It is known that weak solutions, these that satisfy the weak form of the equation, may not unique which is a disadvantage. To select the physically relevant solution, the concept of entropy solutions was developed, see for example [46]. To set up the DG and LDG numerical schemes some of the integrals in the weak formulation have to be performed analytically, which can be really difficult to do in some cases. In [13] and [66], the authors explore the use of numerical integration for the set up of DG and LDG ensuring entropy stable numerical solutions.

1.3 Objectives

Our main objective was to study the present framework of Discontinuous Galerkin methods to solve deterministic and stochastic convection-diffusion equations and extend the entropy stable formulation to the stochastic case. To achieve this objective, the following specific objectives were proposed:

i) Study the DG and LDG methods to gain a deep understanding of the methods, their uses, limitations and understand their differences.

ii) Explore limiting techniques known for DG methods and analyze their efficiency.

iii) With some points to improve, determine more suitable limiters for hyperbolic conservation laws.

iv) Understand the entropy stable numerical integration for the set up of DG and LDG and apply it to stochastic conservation laws.

v) Validation of the proposed modifications in the deterministic and stochastic setting.

1.4 Methodology

The methodological techniques used for the development of this work are described and structured in the following chapters:

- *Chapter 2* Introduce formally the conservation laws, convection-diffusion problems, their weak, and entropy solutions. Study what stochastic differential equations are and give examples of stochastic conservation laws and stochastic convection-diffusion equations with their corresponding solutions.
- *Chapter 3* Formal study of DG and LDG methods. In the introduction of the LDG methods, we include a brief discussion on why the DG method does not work for parabolic problems. For other formulation of the Galerkin methods, we explored the continuous Galerkin method, the Petrov-Galerkin method, and the Interior-Penalty methods. We compared them with the LDG method for convection-diffusion and convective-dominated equations and determined the best method is the LDG method.
- *Chapter 4* Study of limiters. A technique used to improve the DG solutions for discontinuous solutions of conservation laws. In particular, we work with the TVB and BDF limiters, which we modify to propose the ATVB and M-BDF limiters.
- *Chapter 5* Study of nodal entropy stable DG solutions. A technique used to avoid doing analytic integration in the DG and LDG methods employing the Gauss-Lobatto quadrature to ensure entropy solutions. We only worked with the DG case.
- *Chapter 6* Use of the DG method for stochastic conservation laws and of the LDG method for stochastic parabolic equations. The improvements done as compared to [47] are the numerical incorporation of limiters and nodal entropy stable solutions, they were not included in [47] and left as future work.

Chapter 2

Deterministic and Stochastic Conservation Laws

2.1 Introduction

In this chapter we first review hyperbolic equations or conservation laws, as examples we include the solution of the linear convection equation and the Burgers' equation. Second, we study parabolic equations, in particular the heat equation and the convection-diffusion equation.Thirdly, we give a brief introduction to stochastic differential equations, formulate the method of characteristics for stochastic conservation laws, solve the stochastic convection and Burgers equation, and state the stochastic version of the heat and convection-diffusion equation. For each type of equation we introduce the concept of weak solutions and entropy solutions.

2.2 Deterministic case

A scalar partial differential equation (PDE) of two variables has the general form:

$$F(x, y, u_x, u_y, u_{xx}, u_{xy}, \ldots) = 0 \tag{2.1}$$

where x and y are independent variables in a set $O \subset \mathbb{R}^2$ and $u(x, y)$ is the dependent variable. We denote by u_x the derivative of u respect to x, that is, $u_x = \frac{\partial u}{\partial x}$.

In particular we consider *linear PDEs*, which in the *divergence form* can be written as

$$\mathcal{A}u := -\nabla \cdot (a\nabla u) + b \cdot \nabla u + cu = f, \tag{2.2}$$

where the coefficients satisfy

$$a(\vec{x}) \geq a_0 > 0,\ c(\vec{x}) - \frac{1}{2}\nabla \cdot b(\vec{x}) \geq 0 \text{ for } \vec{x} \in \mathcal{O}. \tag{2.3}$$

These previous conditions are needed to establish important theorems regarding the solutions of PDEs [20]. The symbol ∇u denotes the gradient of the function u and $\nabla \cdot b$ the divergence of b.

For the problem to be well posed, equation (2.3) must be complemented with some initial conditions and possibly boundary conditions. The initial condition from now on, is denoted by $u_0(x)$ and unless stated differently the boundary conditions are assumed to be periodic.

The coefficients of the second order derivatives of u, can be arranged into a matrix $\overline{A}(\vec{x})$, called the *principal part* of the equation. The principal part defines a classification for PDEs.

Definition 2.2.1 (Classification of linear PDEs.)**.** A linear PDE might be classified as one of the following:

a) The equation (2.2) is said to be *elliptic* at the point $\vec{x}$ provided $\overline{A}(\vec{x})$ is positive definite.

b) The equation (2.2) is said to be *parabolic* at the point $\vec{x}$ provided $\overline{A}(\vec{x})$ has 0 as an eigenvalue and the remaining eigenvalue is positive.

c) The equation (2.2) is said to be *hyperbolic* at the point $\vec{x}$ provided $\overline{A}(\vec{x})$ has one negative and one positive eigenvalue.

An equation is called *elliptic, parabolic or hyperbolic* if the corresponding property is satisfied in all the points of its domain.

An example of such classification is the convection-diffusion equation:

Linear one dimensional convection-diffusion equation. Let α and D be non-negative constants, the one dimensional convection-diffusion equation is given by

$$u_y + \alpha u_x - D u_{xx} = 0. \tag{2.4}$$

The principal part of the equations is

$$\overline{A}(x) = \begin{bmatrix} -D & 0 \\ 0 & 0 \end{bmatrix}.$$

Thus, the equation is parabolic.

When the diffusive part u_{xx} is suppressed, $D = 0$, this equation becomes hyperbolic .

$$\begin{aligned} \alpha^2 u_{xx} &= \alpha(\alpha u_x)_x = \alpha(-u_y)_x \\ &= \alpha(-u_x)_y = -(\alpha u_x)_y \\ &= -(-u_y)_y = u_{yy} \\ \implies u_{yy} - \alpha^2 u_{xx} &= 0 \\ \implies \overline{A}(x) &= \begin{bmatrix} -\alpha^2 & 0 \\ 0 & 1 \end{bmatrix}, \end{aligned}$$

for $\alpha = 1$ this equation is the so called *wave equation.*

The relative values of α and D determine the dominant character in the convection-diffusion equation. Whenever $D << |\alpha|$ the problem is said to be convective-dominated. The wave equation is an example of a conservation law. The main type of equations we are going to deal with are conservation laws and convection-diffusion equations.

2.2.1 Hyperbolic problems

Let $u(x,t)$ be a quantity representing mass, momentum or energy with $x \in \mathcal{O} = [a_1, a_2]$ and $t \in [0, T]$. Such quantity, depending on the studied problem might be assumed to be conserved. Which means that the change of the total quantity in a sub-domain $\hat{\mathcal{O}} \subset \mathcal{O}$ in time is given by the amount of quantity

that flows in and out of $\hat{\mathcal{O}}$ and the amount of u generated in the sub-domain. The flow into $\hat{\mathcal{O}}$ is the amount of quantity that is transported through the boundaries of $\hat{\mathcal{O}}$. Assuming that the rate of flow or *flux* is given by a function $f \in C^\infty$ and that there is no generation in the sub-domain, the conservation law, by integration by parts, is given by

$$\frac{\partial}{\partial t}\int_{\hat{\mathcal{O}}} u dx + \int_{\hat{\mathcal{O}}} (f(u(x)))_x dx = 0. \tag{2.5}$$

Equation (2.5) is the *integral form* of the conservation law. Since $\hat{\mathcal{O}}$ is arbitrary, (2.5) is equivalent to

$$u_t + (f(u))_x = 0, \tag{2.6}$$

the *differential form* of the conservation law.

Classic applications of conservation laws are models for population, traffic, kinetic equations and the Euler equations [46]. Some of these models rely on two important theorems of the theory of hyperbolic equations, the Transport theorem and the Continuum hypothesis, for more information, we refer the reader to [53]. The particular examples we consider are when $f(u) = \alpha u$ and $f(u) = \dfrac{u^2}{2}$, called the convection equation and *Burgers' equation*, respectively,

$$u_t + (\alpha u)_x = 0 \text{ convection equation}, \tag{2.7}$$

$$u_t + \left(\frac{u^2}{2}\right)_x = 0 \text{ Burgers' equation }. \tag{2.8}$$

An important method for solving hyperbolic equations is the so called *method of characteristics*. The aim of the method is to find curves in the (x, t, z) space that reduce the PDE to a system of ordinary differential equations (ODE).

Since u is a solution of (2.6) it satisfies

$$(f_u(u), 1, 0) \cdot (u_x, u_t, -1) = 0.$$

The vector $(u_x, u_t, -1)$ is normal to the surface $S := \{(x, t, u(x, t))\}$ at each point $(x, t, u(x, t))$. Thus, the vector $(f_u(u), 1, 0)$ lies in the tangent plane to S. Consequently, to find a solution of (2.6), we need to find the surface S such that at each point (x, t, z) the vector $(f_u(u), 1, 0)$ is in its tangent plane. For doing so, we construct curves that lie in S parametrized as

$$(x(s), t(s), z(s)) = (x(s), t(s), u(x(s), t(s))),$$

$s \geq 0$. These curves must satisfy the following system of ODE

$$\frac{dx}{ds} = f_u(z(s)), \tag{2.9}$$

$$\frac{dt}{ds} = 1, \tag{2.10}$$

$$\frac{dz}{ds} = 0, \tag{2.11}$$

with the initial condition $(x(0) = x_0, t(0) = 0, z(0) = u_0(x_0))$ for $s = 0$, so that the vector $(f_u(z(s)), 1, 0)$ will be tangent to such curves.

The ODE system (2.9),(2.10) and (2.11) are called the *characteristic equations* of (2.6) and their solutions are the *characteristic curves*.

The equation for t and z have the solution

$$t(s) = s + c_2, \quad z(s) = c_3,$$

with initial conditions

$$t(s) = s, \quad z(s) = u_0(x_0).$$

Then

$$\begin{aligned} \frac{dx}{ds} &= f_u(z(s)), \\ \implies \frac{dx}{ds} &= f_u(u_0(x_0)), \\ \implies x(s) &= f_u(u_0(x_0))s + c_1, \end{aligned}$$

by the initial condition

$$x(s) = f_u(u_0(x_0))s + x_0.$$

Then

$$x_0 = x(s) - f_u(u_0(x_0)),$$

substituting x_0 in $z(s)$,

$$z(s) = u_0(x(s) - f_u(u_0(x_0))s).$$

We then conclude that the solution along the characteristic $(x(s), t(s)) = (f_u(u_0(x_0))s + x_0, s)$ is

$$u(x,t) = u_0(x_0 - f_u(u_0(x_0))t). \tag{2.12}$$

The solutions of the convection equation and the Burgers' equation given $x_0 \in \mathcal{O}$ are

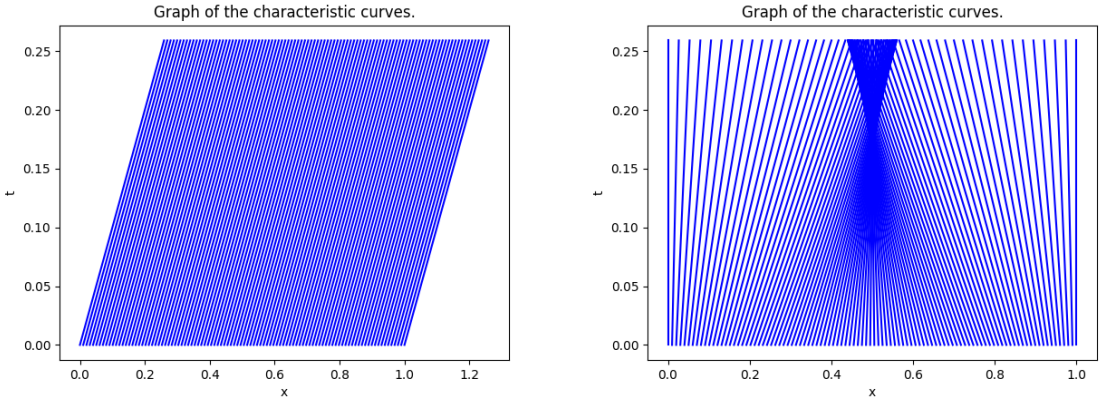
$$u(x,t) = u_0(x_0 - \alpha t), \quad u(x,t) = u_0(x_0 - u_0(x_0)t),$$

respectively, for x in the characteristic determined by x_0.

The solution of the convection equation is just a displacement of the initial condition u_0 with a speed α, which might move to the right if $\alpha > 0$ and to the left if $\alpha < 0$. The Burgers' equation solution is defined implicitly and for some initial conditions might be explicitly obtain.

For the more general equation

$$a(x,t)u_x + b(x,t)u_t = c(x,t),$$



(a) Characteristic curves convection equation, sine inertial condition

(b) Characteristic curves Burgers' equation, sine inertial condition

Figure 2.1: Characteristic curves for convection and Burgers' equations with sine initial condition

a similar procedure can be done and the characteristic equations are

$$\frac{dx}{ds} = a(x,t), \quad \frac{dt}{ds} = b(x,t), \quad , \frac{dz}{ds} = c(x,t),$$

for a more formal approach and details on the method of characteristics see [20].

Let us consider the domain $\mathcal{O} = [0,1]$, the initial condition $u_0(x) = \sin(2\pi x)$ and solve the convection equation and the Burgers' equation.

The characteristic curves for the convection equation and the Burgers' equation starting at a point $x_0 \in \mathcal{O}$ are

$$x(t) = \alpha s + x_0, \quad x(t) = \sin(x_0)s + x_0,$$

respectively. Both sets of curves are straight lines. For the convection equation they are parallel with slope α Figure 2.1a. Therefore they never intersect and the method of characteristics is well defined for all $t > 0$. For the Burgers' equation (Figure 2.1b) the slope of the curves depends on the initial point x_0. In fact, if $0 < x_0 < \frac{1}{2}$, then $\sin(2\pi x_0) > 0$ while for $x_1 = \frac{1}{2}$, $\sin(2\pi x_1) = 0$. Consequently, there is a coordinate $(\hat{x}, \hat{t})$ such that the characteristics starting at x_0 and x_1 intersect and the solution is multi-valued. In Figure 2.2 we show how the initial condition evolves in time and that it becomes multi-valued and two particular characteristics intersecting. How is $u(\hat{x}, \hat{t})$ then defined? This shows that conservation laws might not be well defined for all times. When the solution becomes multi-valued, as in the previous example, it is said that the equation has a *shock discontinuity*. It can be shown that for a smooth initial condition $u_0(x)$ for which $u_0'(x)$ is negative in an open interval. The solution to the Burgers' equation has its first shock at time

$$T_b = \frac{-1}{\min u_0'(x)}, \tag{2.13}$$

called the *breaking time*. The general result, states that for a convex f, if $u_0'(x)$ is negative in an open intervar, the breaking time is given by

$$T_b = -\frac{1}{\min(f''(u_0(x))u_0'(x)).} \tag{2.14}$$

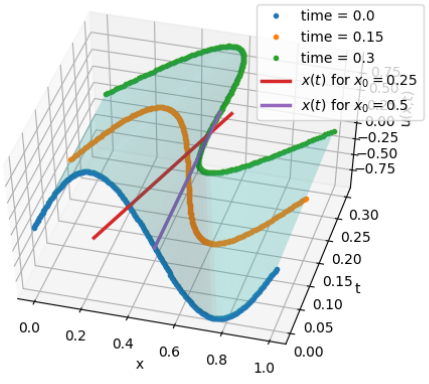


Figure 2.2: Multivalued solution for Burgers' equation

To show this, first note that for any $x_0 < x_1$, such that $f_u(u_0(x_0)) \neq f_u(u_0(x_1))$, the time at which the characteristics intersect is given by

$$t(x_0, x_1) = -\frac{1}{\frac{f_u(u_0(x_1)) - f_u(u_0(x_0))}{x_1 - x_0}}. \tag{2.15}$$

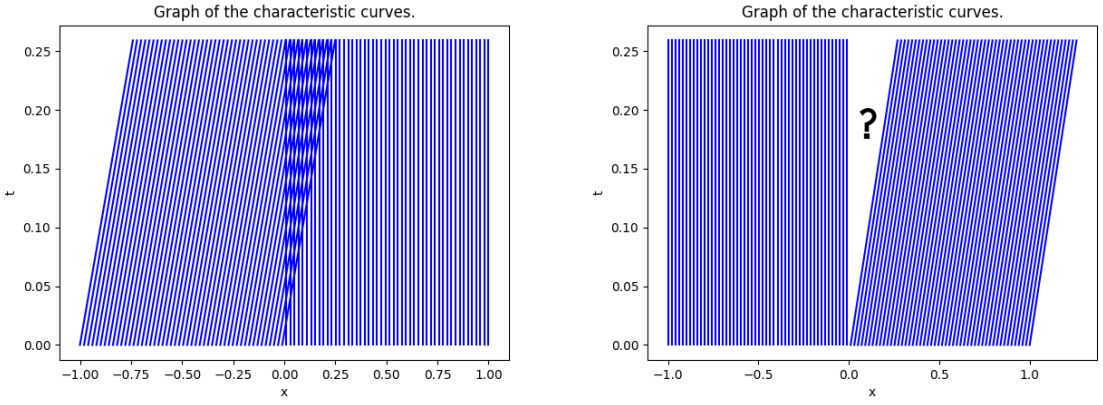
Therefore, since f is convex, the required condition for a shock to occur in a positive time is that $u_0(x_0) > u_0(x_1)$. Let us assume that for the point $\hat{x}$ there is a left neighborhood $B_{\hat{x}}$ in which $u_0(x)$ is decreasing. This hypothesis is valid since there is an open interval where $u_0'(x) < 0$. The characteristics that start in the neighborhood will intersect the characteristic of $\hat{x}$ in a positive time. Define $t_{\hat{x}}(x) := t(x, \hat{x})$ with $t(x, \hat{x})$ as in (2.15). For two points $x_0, x_1 \in B_{\hat{x}}$ with $x_0 < x_1$, $t(x_0)_{\hat{x}} > t(x_1)_{\hat{x}}$. Thus, $t_{\hat{x}}$ is a decreasing function. Therefore, the time at which the first intersection with the characteristic at $\hat{x}$, is obtained by taking the limit $x \to \hat{x}^-$, resulting in

$$T_b(\hat{x}) = -\frac{1}{f''(u_0(\hat{x}))u_0'(\hat{x})}. \tag{2.16}$$

Note that if characteristic starting a point outside of $B_{\hat{x}}$, intersects the characteristic of $\hat{x}$, the time of the intersection, by (2.15), is necessarily greater than $T_b(\hat{x})$ (2.16). Thus, the first positive time at which a characteristic starting at a point x intersects any characteristic (if it does it) is given by $T_b(x)$. Taking the minimum over the domain, we obtain (2.14).

At the breaking time the solution is set to be discontinuous instead of multivalued, to at least satisfy being a function. Consequently is not differentiable and cannot satisfy (2.6). Therefore, there is a need to have a better definition of what being a solution to the equation (2.6) means. For this, we introduce the concept of a *weak solution to the conservation law*.

Shock discontinuities are not the only type of discontinuities that conservation laws might have. The distinctive feature of shock discontinuities is that at least two characteristics intersect. As we know, the convection equation has parallel characteristics. If we consider a discontinuous initial condition for such equation, the discontinuity in the solution is not produced by the intersection of characteristics, they remain parallel at all times. This type of discontinuity is called *contact discontinuities*. The



(a) Characteristic curves convection equation, Riemann initial condition $u_r < u_l$.

(b) Characteristic curves Burgers' equation, Riemann initial condition $u_r > u_l$.

Figure 2.3: Characteristic curves for convection and Burgers' equations with Riemann initial condition

distinctive feature of these discontinuities is that the characteristic curves are parallel. For systems of conservation laws, contact discontinuities can appear in more examples. The formal requirement for these discontinuities is that the characteristic field is *linearly degenerate*, see [44]. In the setting we work on this thesis, this condition can only be met by the fluxes convection equation.

If we multiply $u_t + (f(u))_x = 0$ by a test function (see Appendix A) $\phi(x,t)$ with support in $\mathcal{O}$ and integrate in space and time, we obtain

$$\int_0^\infty \int_{\mathcal{O}} [\phi u_t + \phi (f(u))_x] dx dt = 0,$$

integrating by parts

$$\int_0^\infty \int_{\mathcal{O}} \phi u_t + \phi_x f(u) dx dt = -\int_{\mathcal{O}} \phi(x,0) u(x,0) dx. \tag{2.17}$$

Definition 2.2.2. The function $u(x,t)$ is called a *weak solution* of the conservation law if (2.17) holds for all function $\phi \in C_0^1(\mathcal{O} \times \mathbb{R}^+)$.

This definition allows the possibility of solutions in the Sobolev space H_0^1 (see Appendix A) to the conservation law and to consider discontinuous initial conditions. For understanding how to define a solution when a shock wave appears, we study the solutions for the Riemann initial condition (2.18). Let $\mathcal{O} = (-\infty, \infty)$, the *Riemann initial condition* is defined as

$$u_0(x) = \begin{cases} u_l & x < 0 \\ u_r & x > 0. \end{cases} \tag{2.18}$$

Note that if $u_r < u_l$ Figure 2.3a the solution instantly develops a shock. To define a solution for such equation, we add a viscous term ϵu_{xx} to it and study the behaviour of the solution $u_\epsilon(x,t)$, when $\epsilon \to 0$. The new equation is known as the *viscid Burgers' equation,*

$$u_t + \left(\frac{u^2}{2}\right)_x = \epsilon u_{xx}, \tag{2.19}$$

for a small $\epsilon > 0$.

Let us assume that $u(x,t) = v(\xi)$, with $\xi = x - ct$, where c is a constant to be determined. Using the chain rule, $u_x = v'$ and $u_t = -cv'$, obtaining the ODE

$$-cv' + vv' = \epsilon v'',$$

with boundary condition $v(-\infty) = u_l$ and $v(\infty) = u_r$ and $v'(\pm\infty) = 0$. Integrating

$$\frac{v^2}{2} - cv = \epsilon v' + C. \tag{2.20}$$

Then

$$-cu_l + \frac{u_l^2}{2} = C = -cu_r + \frac{u_r^2}{2}.$$

Consequently $c = \dfrac{u_l + u_r}{2}$, then $C = \frac{-u_l u_r}{2}$. Substituting C and c in (2.20), the equation can be solved by separation of variables, obtaining

$$v(\xi) = u_r + \frac{u_l - u_r}{e^A + 1},$$

with $A = \xi\dfrac{(u_l - u_r)}{2\epsilon} + C$. Using the trigonometric identity

$$1 - \frac{e^{\frac{x}{2}} - e^{\frac{x}{2}}}{e^{\frac{x}{2}} + e^{\frac{x}{2}}} = 1 - \tanh\left(\frac{x}{2}\right),$$

substituting $\xi = x - ct$,

$$u_\epsilon(x,t) = \frac{u_r + u_l}{2} - \frac{u_l - u_r}{2}\tanh\left(\frac{(x - ct)(u_l - u_r)}{4\epsilon}\right). \tag{2.21}$$

In Figure 2.4 we show the solution surface for different values of ϵ. This figure shows that the solution when $\epsilon \to 0$ tends to the initial condition moving at the velocity $c = \dfrac{u_l + u_r}{2}$. This motivates to propose the solution

$$u(x,t) = \begin{cases} u_r & \text{if } x < st \\ u_l & \text{if } x > st, \end{cases} \tag{2.22}$$

with $s = c$. It can be shown that this is in fact a weak solution of the Burgers' equation and that $u_\epsilon \to u$ when $\epsilon \to 0$. The solution (2.22) is the initial condition traveling at a speed $s = \dfrac{u_l + u_r}{2}$, thus s is known as the *shock speed,* or more generally the *Rankine-Hugoniot condition.* The solution obtained by adding the viscous term ϵu_{xx} and then studying the limit of the solutions when $\epsilon \to 0$ is generally know as a *vanishing viscosity solution,* and the solution for a fixed ϵ is called the *viscous profile* for the shock wave. For an introduction viscous solutions see [38].

The Rankine-Hugoniot condition can be obtained from the weak formulation of the conservation law (2.17) by integrating in a vicinity of a shock wave, for the details see [20] or [46]. The Rankine-Hugoniot condition for a general conservation law reads

$$s = \frac{f(u_l) - f(u_r)}{u_l - u_r}. \tag{2.23}$$

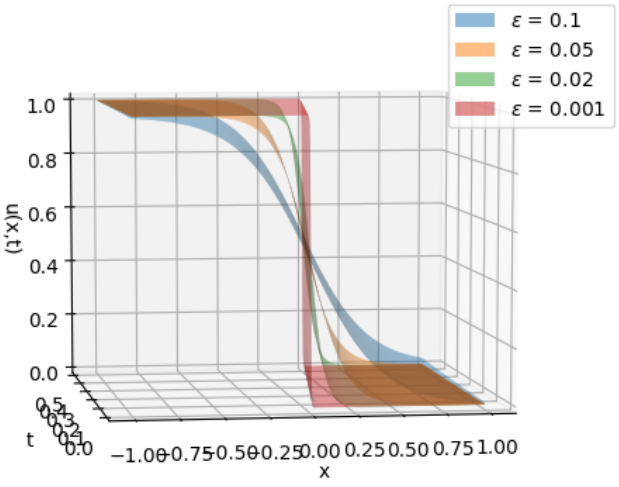


Figure 2.4: Viscous solutions for $\epsilon = 0.1, 0.05, 0.02, 0.001$

Assume that the conservation law develops a shock in the time interval $t_1 \leq t \leq t_2$, (where t_2 could equal infinity) and that there is a smooth curve in the (x,t) plane parametrized by $x = a(t)$, for $t_1 \leq t \leq t_2$ that describes the path of the shock in the (x,t) plane. Since u is discontinuous along the curve, for each $t \in [t_1, t_2]$, the one-sided limits

$$u_+(t) = \lim_{x \to a(t)^+} u(x,t), \quad u_-(t) = \lim_{x \to a(t)^-} u(x,t),$$

exist and are not equal. Due to the Rankine-Hugoniot condition, the curve satisfies

$$a'(t) = \frac{f(u_+) - f(u_-)}{u_+ - u_+}.$$

This ODE can be solved to obtain the curve that the shock follows, note that the initial condition is obtained at the point (x, t_1) where the shock is formed. Therefore, the Rankine-Hugoniot condition allows us to define a solution in the case of a shock wave, but as the following example shows, is not enough to obtain uniqueness.

Consider the Burgers' equation with the Riemann initial condition (2.18) in the case $u_l < u_r$. Note that the characteristics do not fill completely the (x,t) plane. Thus, it is not clear how to define a solution in the region without characteristics Figure 2.3b. There are at least two ways of filling the void space left by the characteristics.

The first is called the *rarefaction wave*,

$$u(x,t) = \begin{cases} u_l & x < u_l t \\ \frac{x}{t} & u_l t \leq x \leq u_r t \\ u_r & x > u_r t, \end{cases} \tag{2.24}$$

this solution can be analytically found by realising that conservation laws have the property of self-similarity, i.e., the equation is invariant when replacing the variables x, t by $\lambda x, \lambda t$. Using the change of variables $u(x,t) = v\left(\frac{x}{t}\right) = v(\xi)$, substituting $v(\xi)$ in the conservation law (2.6) and using the chain rule, results after some algebra in

$$\left(f'(v(\xi) - \frac{x}{t}\right) v' = 0.$$

In the nontrivial case $v' \neq 0$, it must be satisfied

$$f'(v(\xi)) = \frac{x}{t}.$$

Assuming f to be strictly convex leads to the expression

$$v(x,t) = (f')^{-1}\left(\frac{x}{t}\right),$$

from which (2.24) is deduced.

The second solution is

$$u(x,t) = \begin{cases} u_l & x < \frac{u_l+u_r}{2}t \\ u_r & x > \frac{u_l+u_r}{2}t. \end{cases} \tag{2.25}$$

Both solutions can be shown to be a weak solution for the Burgers' equation, but is one of them a "better" solution? The first solution is clearly a continuous function, whereas the second one is a discontinuous solution that satisfies the Rankine-Hugoniot condition. Additionally, the characteristics behave differently for each case. In (2.24) they behave like a fan, fulfilling the void space without intersecting. Whereas in (2.25) the characteristics start from the shock.

In the previous example adding viscosity to the equation helped us to find a solution. A natural question to ask is if adding viscosity in this example and taking the limit of the viscous coefficient to zero will converge to any of these solutions. For the particular case of the viscous Burgers' equation the solution can be found using the Cole-Hopf transformation (we explain this method in the next section) but in general, adding viscosity to the conservation law will result in an equation difficult to solve. Although it can be shown that the vanishing viscosity solution for an arbitrary conservation law is a weak solution [64].

The concept of *entropy solutions* is used to avoid working with vanishing viscosity solutions. This concept is related to the behaviour of the characteristics with respect to a shock. Characteristics represent the flow of the initial condition information, in this example there is a void of information. As a heuristic, since information is taken from the initial data it should not be created at a shock. This means that the characteristic curves must flow into the shock and not start from them.

In our example when two characteristics intersect at a point on a shock the solution on the characteristic from the left of the shock is u_l and the solution on the right characteristic to the shock is u_r, with $u_l > u_r$. For the characteristics to intersect in the direction of the shock, the slope of the characteristic curve from the left must be greater than the speed of the shock and the speed of the shock must be greater than the slope of the characteristic of the curve from the right. Recall that the characteristic curves are

$$x(t) = f_u(u)t + x_0,$$

then the heuristic in terms of the slopes of the characteristics and the speed of the shock s, is

$$f_u(u_l) > s > f_u(u_r), \tag{2.26}$$

known as the *Lax entropy condition.* With this condition solution (2.25) is not admissible. In this example $f(u) = \frac{u^2}{2}$, which is a convex function and therefore if $u_l > u_r$, $f_u(u_l) > s > f_u(u_r)$. The Lax entropy condition (2.26) is not suited for an arbitrary conservation law, it requires that the flux f is a convex function [46].

For a more general entropy formulation, we again consider the viscous conservation problem.

$$u_t^\epsilon + (f(u^\epsilon))_x = \epsilon u_{xx}^\epsilon. \tag{2.27}$$

Let $\eta = \eta(u)$ be an arbitrary strictly convex function, which we call the *entropy function*, and define the function

$$\psi(u) = \int_0^u f'(s)\eta'(s)ds. \tag{2.28}$$

The functions ψ and η satisfy the relation

$$\psi' = \eta' f'. \tag{2.29}$$

Multiplying $\eta'(u)$ in both sides of (2.27), using the chain rule and the relation (2.29), we obtain

$$\eta(u^\epsilon)_t + \psi(u^\epsilon) = \epsilon\eta(u^\epsilon)_{xx} - \epsilon\eta''(u^\epsilon).$$

Integrating over the rectangle $[x_1, x_2] \times [t_1, t_2]$ gives

$$\int_{t_1}^{t_2}\int_{x_1}^{x_2} \eta(u^\epsilon)_t + \psi(u^\epsilon)_x dxdt = \epsilon \int_{t_1}^{t_2} \left[\eta'(u^\epsilon(x,t))u_x^\epsilon(x,t)\Big|_{x=x_1}^{x=x_2}\right] dt - \epsilon\int_{t_1}^{t_2}\int_{x_1}^{x_2} \eta''\cdot(u_x^\epsilon)^2 dxdt.$$

As $\epsilon \to 0$, the first term on the right side vanishes even for the discontinuous case. However, the second term includes the integration of $(u_x^\epsilon)^2$ over the rectangle $[x_1, x_2]\times[t_1, t_2]$, that for a discontinuous u will not vanish in the limit. By assumption η is convex thus the product $\epsilon\eta''\cdot(u_x^\epsilon)^2 > 0$. This implies that the weak solution satisfies

$$\int_{t_1}^{t_2}\int_{x_1}^{x_2} \eta(u)_t + \psi(u)_x dxdt \le 0, \tag{2.30}$$

for all x_1, x_2, t_1 and t_2. The inequality (2.30) is called the *entropy inequality,* the function ψ is the *entropy flux* and (η, ψ) an entropy pair [64]. The previous reasoning, can be made instead over an arbitrary rectangle, integrating over the hole space time domain with an arbitrary test function. The result is the same entropy inequality, but in the distribution sense. Using this formulation and formally integrating in space,

$$\frac{d}{dt}\int_{\mathcal{O}} \eta(u)dx \le 0, \tag{2.31}$$

meaning that the total entropy is non-decreasing with respect to time.

An alternative weak form of the entropy inequality can be formulated by integrating (2.30) against a smooth test function ϕ, obtaining the *weak form of the entropy inequality,*

$$\int_0^\infty \int_{\mathcal{O}} \phi_t \eta(u(x,t)) + \phi_x \psi(u(x,t)) dx dt \leq -\int_{\mathcal{O}} \phi(x,0) u_0(x) dx. \tag{2.32}$$

For the Burgers' equation with $f(u) = \dfrac{u^2}{2}$ and $\eta = u^2$, implies $\psi = \frac{2}{3}u^3$. Then,

$$\eta_u + \psi_x = (u^2)_t + \left(\frac{2}{3}u^3\right) x,$$

if u is smooth using the chain rule,

$$2u\left(u_t + \left(\frac{u^2}{2}\right)_x\right) = 0.$$

For discontinuous u, we integrate over an infinitesimal rectangle $[x_1, x_2] \times [t_1, t_2]$ so that the characteristic curve of the discontinuity approximately divides the rectangle in two triangles through the diagonal connecting (x_1, t_1) and (x_2, t_2). Over the triangle defined by the vertices (x_1, t_1), (x_1, t_2) and (x_2, t_2) we have $u = u_l$ and over the triangle (x_1, t_1), (x_2, t_1) and (x_2, t_2) we have $u = u_r$. Then, it can be shown [46] that

$$\int_{t_1}^{t_2} \int_{x_1}^{x_2} (u^2)_t + \left(\frac{2}{3}u^3\right)_x dx dt = -\frac{1}{6}(u_l - u_r)^3 (t_2 - t_1) + O((t_2 - t_1)^2),$$

for small $t_2 - t_1 > 0$ the $O((t_2 - t_1)^2)$ term will not affect the sign of the integral. Therefore (2.30) is satisfied if and only if (iff) $u_l > u_r$.

As a general rule if a convex flux f develops a discontinuity in its solution, with values u_l and u_r at the left and right of the discontinuity, respectively. The entropy solution can be determined as follows: If $u_l < u_r$ the solution is a rarefaction wave, otherwise, $u_r < u_l$ the solution is a shock wave. For non convex flux the entropy solution is determined by the *Oleink condition* [64] and it states

- $(u_l < u_r)$ the solution is a shock iff the graph of f restricted to the interval $[u_l, u_r]$, is situated above its chord.
- $(u_l > u_r)$ the solution is a shock iff the graph of f restricted to the interval $[u_r, u_l]$, is situated below its chord.

In conclusion, conservation laws might be solved analytically using the method of characteristics, but they present two major problems. First, the evolution of the initial condition in time can become multi-valued and the solution develops a shock. Second, the characteristic might not define a solution for all the points in the domain. The solution to the first problem is to construct a shock using the Rankine-Hugoniot condition and to the second one is to use a rarefaction wave. In both cases, the solutions must satisfy the entropy inequality (2.30).

Definition 2.2.3. We say that a function $u \in L^\infty(\mathcal{O} \times (0, \infty))$ is an *weak entropy* solution of the initial-value problem

$$\begin{cases} u_t + f(u)_x = 0 & \text{in } \mathcal{O} \times (0, \infty) \\ u(x, 0) = u_0(x) & \text{on } \mathcal{O} \times \{t = 0\}, \end{cases} \tag{2.33}$$

provided

$$\int_0^\infty \int_{\mathcal{O}} u\phi_t + f(u)\phi_x dxdt + \int_{\mathcal{O}} \phi(x,0)u_0(x)dx = 0$$

for all test function ϕ with compact support in $\mathcal{O}$, and for all entropy pairs (η, ψ)

$$\int_0^\infty \int_{\mathcal{O}} \phi_t \eta(u(x,t)) + \phi_x \psi(u(x,t))dxdt \leq -\int_{\mathcal{O}} \phi(x,0)u_0(x)dx.$$

There is a theorem due to Kružkov that ensures the existence and uniqueness of entropy solutions for the scalar conservation laws, [64].

Theorem 2.2.1. *For every bounded measurable initial condition u_0 on $\mathcal{O} = \mathbb{R}$ and every flux $f \in C^\infty$ flux, there exists one and only one entropy solution of (2.6) in $L^\infty(\mathcal{O} \times (0,T)) \bigcap C([0,T))$, so that it satisfies the maximum principle*

$$||u||_{L^\infty(\mathcal{O}\times(0,T))} = ||u_0||_{L^\infty(\mathcal{O})}. \tag{2.34}$$

The discontinuous Galerkin method requires a particular definition of weak solution [16]. We include this particular definition in this chapter. For it, we require a partition of the interval $[0,1]$, without loss of generality. Lets assume a uniform partition of width $h = \frac{1}{N}$, $N \in \mathbb{N}$ i.e. $x_j = \frac{j}{N}$, for $j = 0, \ldots, N$, denote by $I_j = (x_j, x_{j+1})$.

Definition 2.2.4. The function $u(x,t)$ is called a weak solution of the conservation law if for every interval I_j and every test function ϕ, the following equalities are satisfied

$$\int_{I_j} u_t \phi dx - \int_{I_j} f(u)\phi_x dx = -(f(u)\phi)\big|_{x=x_j}^{x=x_{j+1}} \tag{2.35}$$

$$\int_{I_j} u(x,0)\phi dx = \int_{I_j} u_0(x)\phi(x)dx. \tag{2.36}$$

In the following we solve analytically the benchmark problems that will be used for the validation of the methods presented in this work.

Example 1. Linear convection equation. We consider the two initial conditions, a sine wave

$$u_0(x) = \sin(2\pi x), \tag{2.37}$$

and an impulse

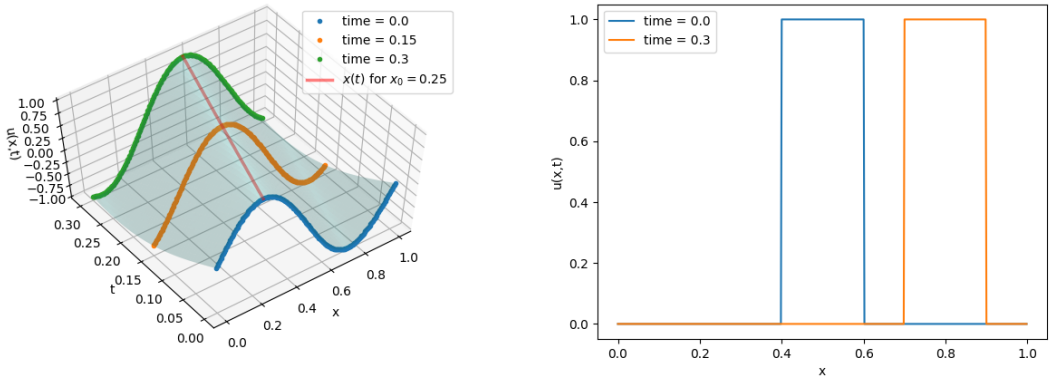
$$u_0(x) = \begin{cases} 0 & \text{if } x \in [0, 0.4] \cup [0.6, 1] \\ 1 & \text{if } x \in (0.4, 0.6), \end{cases} \tag{2.38}$$

in the domain $\mathcal{O} = [0,1]$.

The solution to the convection ($\alpha = 1$) with both initial conditions for a given $x_0 \in \mathcal{O}$ is given by

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

see Figures 2.5a and 2.5b. In the Figure 2.5a we include a characteristic to graphically show that the solution is constant along the characteristics.



(a) Solution surface convection equation sine initial condition.

(b) Solution convection equation impulse initial condition.

Figure 2.5: Solutions to convection equation.

Example 2. Burgers' equation. We as well solve this equation for the sine and impulse initial conditions

We have previously worked with the sine initial condition and we know that the method of characteristics is valid up to the time

$$T_b = \frac{-1}{\min u_0'(x)} = \frac{1}{2\pi}.$$

For $t > T_b$ we need to determine the characteristic curve for the shock. We know that this characteristic curve satisfies

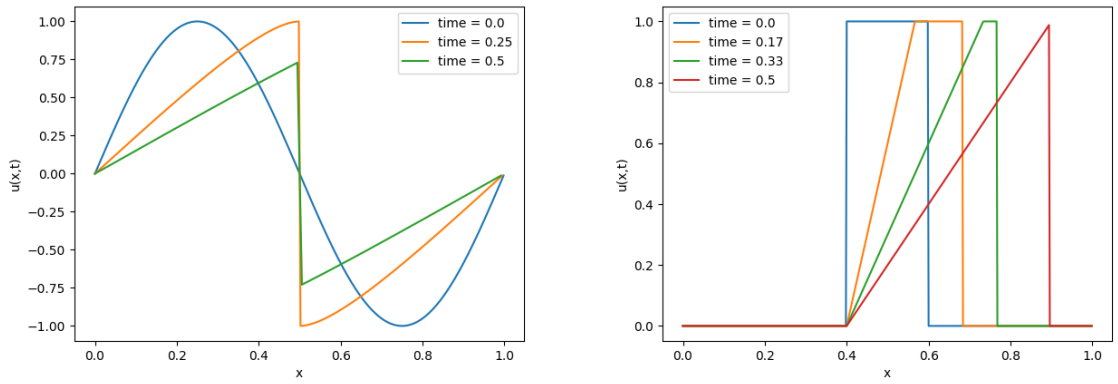
$$x'(t) = \frac{f(u_l) - f(u_r)}{u_l - u_r} = \frac{u_l + u_r}{2}.$$

From the Figure 2.1b it can be seen that the shock occurs at $x = \frac{1}{2}$. Note that the initial condition is symmetric with respect to $x = \frac{1}{2}$, then $u_l = -u_r$, thus

$$\begin{aligned} x'(t) &= 0 \\ \implies x(t) &= \frac{1}{2} \text{ for all } t \geq T_b, \end{aligned}$$

due to the initial condition $x\left(\frac{1}{2\pi}\right) = \frac{1}{2}$. This means that the shock stays constant at $x = \frac{1}{2}$. There is no explicit formula for the solution after the time T_b, but it can be constructed as follows. For $t > T_b$ find $0 < \hat{x} < \frac{1}{2}$ such that $\frac{1}{2} = \hat{x} + u_0(\hat{x})t$. The values at the discontinuity are $u_l = u_0(\hat{x})$ and $u_r = -u_0(\hat{x})$. The rest of the solution can be constructed using the characteristics with initial points in $[0, \hat{x}] \cup [1 - \hat{x}, 1]$, Figure 2.6a.

The impulse initial condition is the combination of the Riemann condition with $u_l < u_r$ at $x = 0.4$ and $u_l > u_r$ at $x = 0.6$. Therefore the entropy solution of the problem is to consider a rarefaction wave



(a) Solution Burgers' equation sine initial condition.

(b) Solution Burgers' equation impulse initial condition.

Figure 2.6: Solutions to Burgers' equation.

starting at $x = 0.4$ and a shock solution at $x = 0.6$ with speed $s = \frac{1}{2}$. The method of characteristics thus results in

$$u(x,t) = \begin{cases} 0 & 0 \leq x < 0.4 \\ \frac{x}{t} & 0.4 < x < t + 0.4 \\ 1 & 0.4 + t < x < \frac{t}{2} + 0.6 \\ 0 & \frac{t}{2} + 0.6 < x. \end{cases} \tag{2.39}$$

The characteristic $x(t) = t + 0.4$ defining the rarefaction wave and the characteristic $x(t) = \frac{t}{2} + 0.4$ defining the shock intersect at $(x,t) = (0.8, 0.4)$. For $t \geq 0.4$ the shock wave is parametrized by $x(\cdot)$, with $u_l(t) = \frac{x}{t}$ and $u_r(t) = 0$. Using the Rankine-Hugoniot condition,

$$\begin{aligned} x'(t) &= \frac{\frac{1}{2}\left(\frac{x(t)}{t}\right)^2}{\frac{x(t)}{t}} = \frac{x(t)}{2t} \\ \implies x(t) &= \left(\frac{8}{5}t\right)^{\frac{1}{2}}. \end{aligned}$$

Thus, for $t \geq 0.4$ the solution is

$$u(x,t) = \begin{cases} 0 & 0 < x < 0.4 \\ \frac{x}{t} & 0.4 < x < \left(\frac{8}{5}t\right)^{\frac{1}{2}} \\ 0 & x > \left(\frac{8}{5}t\right)^{\frac{1}{2}}, \end{cases} \tag{2.40}$$

Figure 2.6b.

2.2.2 Parabolic problems

In this section we briefly deal with the heat equation or diffusion equation and its properties. Later we focus on our main parabolic example, the convection-diffusion equation. This equation is the

combination of a conservation law and the diffusion equation.

The diffusion or heat equation is a mathematical expression that explains how some physical quantities, for example heat, move in a medium. Heat flows from hotter regions to colder regions, and the more drastic the temperature difference between two regions, the more rapid the flow of the heat. Then the substance or heat moves through a cross section at x at time t proportionally to its density gradient u_x. In other words, the flux f is proportional to the gradient u_x. Heat moving from hotter regions to colder regions, is expressed by saying that if $u_x > 0$, then $J < 0$ (the heat moves to the right), and if $u_x < 0$, then $J > 0$ (the substance moves to the right), i.e.,

$$f(x,t) = -Du_x(x,t).$$

This relationship is known as *Fick's Law* [51], the positive constant D is called the *diffusion constant.* Using the conservation law (2.6) we obtain

$$u_t = Du_{xx}. \tag{2.41}$$

The interpretation of the diffusion equation as a heat flow allows to give a probabilistic interpretation to its solution. A heat flow is governed by the random movement of its molecules and their collisions, which cause the energy to disperse from high-energy regions to low-energy regions [51]. We can assume that $u(x,t)$ represents a probability density function for the random variable $X = X(t)$ (see Appendix B.0.1) giving the location of a particle at time t. We assume that the movement of the particle occurs in one spatial dimension and at discrete time steps τ. At each time step the particle jumps a distance h to the left or right with equal probability. Then the probability of a particle being in an interval $[x, x+h]$ at the time t is $u(x,t)h$. For a position $(x, t+\tau)$, a particle could have only come from the intervals $[x-h, x]$ or $[x+h, x+2h]$, then the probability of a particle being in the interval $[x, x+h]$ can be written as

$$u(x,t+\tau) = \frac{1}{2}u(x-h,t) + \frac{1}{2}u(x-h,t),$$

known as the *master equation.* Expanding each term of the master equation with their Taylor series at (x,t) with the increments h in space and τ in time and simplifying,

$$u_t + \frac{\tau}{2}u_{tt} = \frac{h^2}{2\tau}u_{xx} + \dots.$$

The coefficients of the elements after u_{xx} are a positive power of h times $\dfrac{h^2}{2\tau}$. Then taking $h, \tau \to 0$ with the ratio $\dfrac{h^2}{2\tau} = D$ constant, we get the diffusion equation

$$u_t = Du_{xx}.$$

A more general result explaining the relationship of the diffusion equation and stochastic processes is the *Fokker-Plank equation* [3]. This argument is correct but it lacks some formality, we later use the *Laplace-De Moivre Theorem* to give a formal proof.

The solution of the diffusion equation (2.41) when $\mathcal{O} = (-\infty, \infty)$ for a continuous and bounded initial condition $u_0(x)$, is given by

$$u(x,t) = \int_{\mathbb{R}} K(x,y,t)u_0(u)dy,$$

with

$$K(x,y,t) = \frac{1}{(4\pi t)^{\frac{1}{d}}} \exp\left(-\frac{(x-y)^2}{4t}\right),$$

known as the *fundamental solution*. It can be shown that such $u(x,t)$ is of class C^∞ and that

$$\lim_{t\to\infty} u(x,t) = f(x).$$

Although the case $\mathcal{O} = (-\infty, \infty)$ is interesting, for studying numerical methods we require a bounded domain. Once again we consider $\mathcal{O} = [0,1]$, for $T > 0$ define,

$$\begin{aligned} \mathcal{O}_T &:= \mathcal{O} \times [0,T] \\ \partial^*\mathcal{O}_T &:= (\bar{\mathcal{O}} \times \{0\}) \cup (\partial\mathcal{O} \times \overline{(0,T)}), \end{aligned}$$

the set $\partial^*\mathcal{O}_T$ is called the reduced boundary of $\mathcal{O}_T$.

We say that u solves the heat equation with initial condition u_0 and periodic boundary if

$$\begin{aligned} u_t &= u_{xx} \text{ for } (x,t) \in (0,1) \times (0,T) \\ u(0,t) &= u(1,t) \text{ for } t \in (0,T) \\ u(0,x) &= u_0(x) \text{ for } x \in (0,1). \end{aligned} \tag{2.42}$$

The typical method for solving (2.42) is separation of variables see [36] and Fourier decomposition [31]. Let

$$\phi_k(x) = e^{2\pi i k x}, \text{ for } k \in \mathbb{Z},$$

then assume

$$u(x,t) = \gamma(t)\phi_k(x), \quad \gamma(0) = 1.$$

Substituting $u(x,t)$ in the diffusion equation an ODE for γ is obtained. Solving it and using the periodic condition, the solution $u(x,t)$ is

$$u(x,t) = e^{-4\pi^2 k^2 t}\phi_k(x).$$

We assume the initial condition $u_0 \in L_2[0,1]$, then it can be expressed as the Fourier series,

$$u_0(x) = \sum_{k\in\mathbb{Z}} \alpha_k \phi_x(x), \tag{2.43}$$

with

$$\alpha_k = \int_0^1 u_0(x)\phi_k(x)dx.$$

The *superposition principle* states that if u and v solve a PDE, their sum $u+v$ is a solution. Then using the superposition principle, we conclude that the solution to the problem (2.42) is given by,

$$u(x,t) = \sum_{k\in\mathbb{Z}} \alpha_k e^{-4\pi^2 k^2 t}\phi_k(x). \tag{2.44}$$

The following are main results from the theory of PDE on heat equations, see [36] and [20] for the proofs and more general formulations.

Theorem 2.2.2. *Maximum Principle. Let u be a solution of (2.42) with $\mathcal{O}$ open and bounded and*

$$u_{xx} - u_t \geq 0 \text{ in } \mathcal{O}_T. \tag{2.45}$$

We then have

$$\sup_{\overline{\mathcal{O}}_T} = \sup_{\partial^*\mathcal{O}_T} u, \tag{2.46}$$

if $T < \infty$, we can take max instead of sup.

The maximum principle theorem, states that the maximum of a the solution to the heat equation is attained in the boundary $\partial^*\mathcal{O}_T$. A corollary or this theorem is the uniqueness of solutions for the heat equation.

Corollary 2.2.2.1. *Let u, v be solutions of (2.42) with $u = v$ on $\partial^*\mathcal{O}_T$. Then $u = v$ on $\overline{\mathcal{O}}_T$.*

We now have an introductory perspective of what the diffusion equation is and what to expect from its solutions. Then we are prepared to define the convection-diffusion equation, simply stated as

$$u_t + (f(u))_x = Du_{xx}. \tag{2.47}$$

We now present some examples with known solution for linear and nonlinear parabolic conservation laws.

Example 3. Linear convection-diffusion equation. Consider $f(u) = \alpha u$, for $\alpha \in \mathbb{R}$,

$$\begin{aligned} u_t + \alpha u_x &= Du_{xx} \text{ for } (x,t) \in (0,1) \times (0,T) \\ u(0,t) &= u(1,t) \text{ for } t \in (0,T) \\ u(0,x) &= u_0(x) \text{ for } x \in (0,1). \end{aligned} \tag{2.48}$$

From the work done on the conservation law and the diffusion equation, is straightforward to verify that, expressing the initial condition with its Fourier series (2.43). The solution of the linear convection-diffusion equation is

$$u(x,t) = \sum_{k\in\mathbb{Z}} \alpha_k e^{-4\pi^2 k^2 Dt} \phi_k(x - \alpha t). \tag{2.49}$$

For the initial condition $u_0(x) = \sin(x)$ with $\mathcal{O} = [0, 2\pi]$ the solution is

$$u(x,t) = e^{-t}\sin(x), \tag{2.50}$$

see Figure 2.7.

The linear convection diffusion equation satisfies the maximum principle,

$$\min_{\mathcal{O}} u_0(x) \leq u(x,t) \leq \max_{\mathcal{O}} u_0(x).$$

Example 4. Nonlinear convection-diffusion equation. Consider $f(u) = \dfrac{u^2}{2}$,

$$\begin{aligned} u_t + \left(\frac{u^2}{2}\right)_x &= Du_{xx} \text{ for } (x,t) \in (0,1) \times (0,T) \\ u(0,t) &= u(1,t) \text{ for } t \in (0,T) \\ u(0,x) &= u_0(x) \text{ for } x \in \mathcal{O}. \end{aligned} \tag{2.51}$$

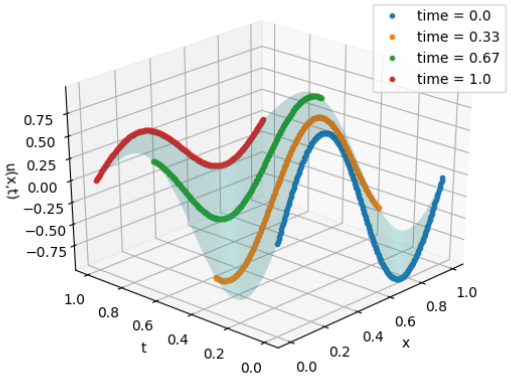


Figure 2.7: Solution of the linear convection-diffusion equation with sine initial condition, $\alpha = d = 1$

It should be noted that generally (2.51) is usually known as the Burgers' equation or viscid Burgers' equation and what we called the Burgers' equation is known as the inviscid Burgers' equation. There is a general technique called the Cole-Hopf transformation for solving (2.51) when $\mathcal{O} = (-\infty, \infty)$ [20].

The transformation is defined by

$$u = -2D\frac{\phi_x}{\phi},$$

then

$$u_t = \frac{2D(\phi_t\phi_x - \phi\phi_{xt})}{\phi^2}, \quad uu_x = \frac{4D^2\phi_x(\phi\phi_{xx} - \phi_x^2)}{\phi^3} \quad \text{and}$$
$$Du_{xx} = -\frac{2D^2(2\phi_x^3 - 3\phi\phi_{xx}\phi_x + \phi^2\phi_{xxx})}{\phi^3}.$$

Substituting into the nonlinear convection-diffusion equation,

$$\begin{aligned}\frac{2D(-\phi\phi_{xt} + \phi_x(\phi_t - D\phi_{xx} + D\phi\phi_{xxx}))}{\phi^2} = 0 \quad &\iff \quad -\phi\phi_{xt} + \phi_x(\phi_t - D\phi_{xx}) + D\phi\phi_{xxx} = 0 \\ &\iff \quad \phi_x(\phi_t - D\phi_{xx}) - \phi(\phi_{xt} - D\phi_{xxx}) = 0 \\ &\iff \quad \phi_x(\phi_t - D\phi_{xx}) - \phi(\phi_t - D\phi_{xx})_x = 0. \quad (2.52)\end{aligned}$$

Thus, if ϕ is solution to the heat equation,

$$\begin{aligned}\phi_t - D\phi_{xx} &= 0 \\ \phi_0(x) &= \exp\left(-\int_0^x \frac{u_0(y)}{2D}dy\right),\end{aligned}$$

the equality (2.52) is valid and u defined by

$$u = -2D(\log\phi)_x$$

is a solution to the nonlinear convection-diffusion equation. The solution ϕ can be "easily" be found using the fundamental solution. The solution using the Cole-Hopf transformation was followed from [43].

Note that for u to be defined $\phi(x) > 0$, for all $x \in \mathbb{R}$. The numerical methods developed in this thesis require periodic boundary conditions. Then for the numerical tests, we require a periodic initial condition that is everywhere positive. That is why we consider the initial condition $\phi_0(x) = 2+\sin(x)$. Using the fundamental solution, it can be shown that $\phi(x,t) = 2 + e^{-Dt}\sin(x)$. From the Cole-Hopf transformation we obtain

$$u(x,t) = -2D\frac{\cos(x)}{2e^{Dt} + \sin(x)}, \tag{2.53}$$

with initial condition

$$u_0(x) = -2D\frac{\cos(x)}{2 + \sin(x)},$$

which is periodic on $\mathcal{O} = [0, 2\pi]$.

If the magnitude of D on any of theses examples is small enough, the convective part of the equation will dominate and we say the problem is *convective-dominated*. In such case, it can be expected that the solution of the problem behaves like the solution of conservation laws. Although theoretically the solutions do not have discontinuities and are properly defined in their domains at all times, numerically this will not be the case. Therefore, we require to define weak solutions for the convection-diffusion equations. We give two different formulations, the first one is more classic and its an extension of Definition 2.2.2. The second one was specifically designed for the local discontinuous Galerkin [16].

Definition 2.2.5. (Weak solution version 1, Ref. [31]) We say u is a weak solution of the convection-diffusion equation (2.47) if for every test function ϕ with $\phi(0) = 0$ and use integration by parts it follows,

$$\int_0^1 u_t\phi - Du_x\phi_x + f(u)_x\phi dx = 0 \tag{2.54}$$

$$\int_0^1 u(x,0)\phi dx = \int_0^1 u_0(x)\phi dx. \tag{2.55}$$

This definition can be extended to be an entropy solution as follows [33]. Let (η, ψ) and entropy pair and define the entropy 3-tuple (η, ψ, r) with r the *diffusion entropy flux* satisfying

$$r' = \eta' D. \tag{2.56}$$

By adding artificial viscosity to the convection-diffusion equation, it can be derived the entropy inequality

$$\eta(u)_t + \psi(u)_x - r(u)_{xx} \geq 0, \tag{2.57}$$

satisfied in the distribution or weak sense. Entropy solutions for convection-diffusion equations become particularly important for degenerated equations. Where the diffusion coefficient is variable and becomes zero in some parts of the domain.

Definition 2.2.6. We say u is a *weak entropy solution* to the convection-diffusion equation, provided it satisfies (2.54) and (2.61) for all test functions ϕ and that satisfies (2.57) for all entropy 3-tuple (η, ψ, r).

For the definition of the second formulation we require a partition of the interval $[0,1]$, let us assume a uniform partition of width $h = \frac{1}{N}$, $N \in \mathbb{N}$ i.e. $x_j = \frac{j}{N}$, for $j = 0, \ldots, N$, denote by $I_j = (x_j, x_{j+1})$. Define the new variable $q = \sqrt{D}u_x$ and rewrite (2.51) as follows

$$\begin{aligned} u_t + (f(u) - \sqrt{D}q)_x &= 0 \text{ in } \mathcal{O}_T & (2.58)\\ q - g(u)_x &= 0 \text{ in } \mathcal{O}_T & (2.59)\\ u(x,0) &= u_0(x), \text{ on } (0,1), & (2.60) \end{aligned}$$

where $g(u) = \sqrt{D}u$.

Definition 2.2.7. (Weak solution version 2, Ref. [16]) We say u is a solution of (2.58)-(2.60) if for each interval I_j, $j = 1, \ldots, N$ and for every test functions ϕ_1, ϕ_2, ϕ_3 the following equalities hold

$$\begin{aligned} \int_{I_j} u_t\phi_1 dx - \int_{I_j} (f(u) - \sqrt{D}q)(\phi_1)_x dx + ((f(u) - \sqrt{D}q)(\phi_1))\big|_{x=x_j}^{x=x_{j+1}} &= 0 & (2.61)\\ \int_{I_j} q\phi_2 dx - \int_{I_j} g(u)(\phi_2)_x dx + (g(u)(\phi_2))\big|_{x=x_j}^{x=x_{j+1}} &= 0 & (2.62)\\ \int_{I_j} u(x,0)\phi_3 dx - \int_{I_j} u_0(x)\phi_3 dx &= 0. & (2.63) \end{aligned}$$

An analogous definition of weak entropy solution can be made with the weak solutions defined in Definition 2.2.7.

2.3 Stochastic case

In this work we focus on stochastic parabolic conservation laws (stochastic convection-diffusion equations). Recall that if the diffusion term vanishes we have hyperbolic conservation laws and on the other hand, the stochastic heat equation is obtained when the transport term is not considered. The stochastic counterparts are obtained by adding a Brownian motion in the time variable to the deterministic equations. The intuition of the stochastic PDEs we work with is that the white noise acts on the reaction term of the equations. In the deduction of the integral form of the conservation law equation (2.5), we assumed that there was no production or loss of the quantity of interest. If we assume such production by a function F, the integral form of the conservation law reads

$$\frac{d}{dt}\int_{\mathcal{O}} u(x,t)dx = -\int_{\partial\mathcal{O}} f(u)_x + \int_{\mathcal{O}} F(x,t,u(x,t))dx,$$

from which the differential equation

$$u_t + (f(u))_x = F(x,t,u(x,t)),$$

is obtained. The function F is called the reaction. In the following we will assume that F is modeled by a multiplicative white noise

$$F(x,t,u) = g(x,t,u)dW_t. \tag{2.64}$$

Thus, the problems that we are going to consider, can be presented by the general convection-diffusion equation perturbed by a *multiplicative white noise* of the form

$$\begin{aligned} du &= (-f(u)_x + Du_{xx})dt + g(\omega,t,u,u_x)dW_t \quad \text{in } \mathcal{O}\times(0,T]\times\Omega, & (2.65)\\ u(\omega,x,0) &= u_0(x) \quad \text{on } \mathcal{O}\times\Omega, \end{aligned}$$

The initial condition $u_0(x)$ is deterministic and $g(u)$ is a multiplicative noise, it is said to be multiplicative because $g(u)$ depends on the solution u . As expected from our study of the convection-diffusion equation, for $D > 0$, and sufficient regularity of f and g, the solutions of (2.65) exists and are unique. When $D = 0$, (the conservation law), the solutions will have shocks or rarefactions, so we require a definition of weak entropy solutions for (2.65).

A formal study of solutions to the stochastic hyperbolic case of equation (2.65) has been done by studying viscous solutions and taking the limit of the viscosity term to zero [54]. Additionally, solutions to this problem have been constructed by a splitting method [30]. Due to the white noise being multiplicative, it is not expected that the mean of a stochastic conservation law with multiplicative white noise is the solution of the deterministic equation. A formal approach to the stochastic parabolic equations of the form (2.65) can be found in [73]. For the nonlinear case, $f(u) = \frac{u^2}{2}$, the SPDE can be transformed into a stochastic heat equation through a procedure called the *Wick exponential substitution*, see [29].

In most of the literature the stochastic forcing is introduced by using the Brownian motion, in the article [5] a Poisson point field is used for the Burgers' equation. The author, reformulates the problem in terms of the associated Hamilton-Jacobi equation. With the Hamilton-Jacobi formulation of the Burgers' equation, the problem is solved by minimizing particle trajectories. The random forcing is introduced as an external potential forcing, defining the contribution of the potential in a path as the Poissonian points that the path passes through. Under such perspective, the forcing turns into a random filed, so "questions about Burgers equations with randomness become random media questions" [5]. Thus, for the Burgers' equation the random forcing can be selected to model certain random media. There are no general results of existence and uniqueness for arbitrary random forcing.

For the one spatial dimensional case there is a direct relationship between Hamilton-Jacobi equations and conservation laws [55]. Thus, the Hamilton-Jacobi formulation of stochastic conservation laws, allows to understand the random forcing as a random environment. Therefore, the solution of the stochastic conservation law can be understood as the minimization of paths in an environment modeled by the stochastic forcing.

The multiplicative white noise is not only more general than additive noise, the theoretical framework required for showing existence and uniqueness of solutions with white noise is more complicated. It can be shown that additive white noise can be transformed into an equation where the flux depends on the stochastic term and deterministic methods can be used to solve such problem. For the multiplicative white noise this procedure cannot be done [8].

It is important to point out that approach (2.64) is not the only form to introduce the environmental uncertainty into conservation laws. Another approach is to assume that $f(u)_x = \alpha u_x$, with α being a realization of a normal random variable $N(\mu, \sigma)$. With such model we do not certainly know the speed at which the quantity of interest is being transported, but that it has been observed to oscillate around μ with variance σ. In this case the solution of the equation is by definition a stochastic process. More abstract constructs include a Brownian motion in the flux function [49].

As it will be showed later, the Galerkin method, reduces the solution of these SPDE to the solution of systems of stochastic differential equations (SDE). We refer the reader to [22], [28], [39] and [76] for a formal approach to SPDEs. We do present a brief introduction to stochastic processes and in particular to the Brownian motion, stochastic integration and stochastic differential equations.

2.3.1 Theory of SDE

An *stochastic process* is a collection of *random variables* $\{X(\tau), \tau \in T\}$, the set T is called the *parameter set*. The random variables take values in the called *state-space,* Ω, of the stochastic process. The set T might be assumed to be discrete $T = \{0, 1, 2, \dots, \}$ or continuous $T = [0, \infty)$, the usual notation is $\{X_n, n \geq 0\}$ and $\{X(t), t \geq 0\}$, respectively. The state-space might as well be discrete or continuous. A particular, but unspecified outcome in Ω is denoted by ω. This section assumes the reader has an introductory knowledge of probability see Appendix B.0.1 for a review of basic concepts. In the Appendix B.0.2 there are relevant definitions and theorems related to the Brownian motion and SDEs.

The objective of this section is to understand how to formulate an ordinary differential equations of the form

$$du(t) = F(\omega, t, u(t))dt + G(\omega, t, u(t))dW_t, \quad u(0) = u_0 \tag{2.66}$$

with W_t a stochastic process called the *Brownian motion.* This expression is a *stochastic differential equation.* The function F is the *drift term* and G the *diffusive term* We first study the Brownian motion, then define the integral with respect a Brownian motion and finally give the existence and uniqueness theorem of the solution of (2.66). The content and the missing proofs of this section can be found in [21] . A more formal approach can be found in [3], [78] and [57].

In the section of parabolic problems, we made a deduction for the heat equation by its interpretation of particles moving randomly in space to the left or right and the density function of such particles. We can describe the same phenomena by means of the trajectory of a single particle. Recall that the time step is τ and the length of the steps is h. The following computations were followed from [21]. Let $X(t)$ denote the position of a particle at time $t = n\tau$ $(n = 0, 1, \dots)$. This stochastic process is known as a *random walk.* Define $S_n := \sum_{i=1}^{n} X_i$, where the X_i are independent random variables such that

$$\begin{aligned} \mathbb{P}(X_i = 0) &= \frac{1}{2} \\ \mathbb{P}(X_i = 1) &= \frac{1}{2}, \end{aligned}$$

for $i = 1, \dots,$ representing that the particle moved to the left $X_i = 0$ or to the right $X_i = 1$ at the time $i\tau$. Without loss of generality it is assumed that the particle starts at $x = 0$. The variables X_i satisfy $\mathbb{V}(X_i) = \frac{1}{4}$. The random variable S_n is the number of steps to the right of the particle by time $t = n\tau$. Therefore, summing the distance advanced to the right $S_n h$, plus the distance advanced to the left $(n - S_n)h$ we obtain the position of the particle at time $t = n\tau$

$$X(t) = S_n h + (n - S_n)h = (2S_n - n)h.$$

Then

$$\begin{aligned} \mathbb{V}(X(t)) &= (h)^2 \mathbb{V}(2S_n - n) \\ &= (h)^2 n \\ &= \frac{h^2}{\tau} t. \end{aligned}$$

Again assume $\dfrac{h^2}{\tau} = D$,

$$X(t) = (2S_n - n)h = \left(\frac{S_n - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \right) \sqrt{n} h = \left(\frac{S_n - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \right) \sqrt{tD}.$$

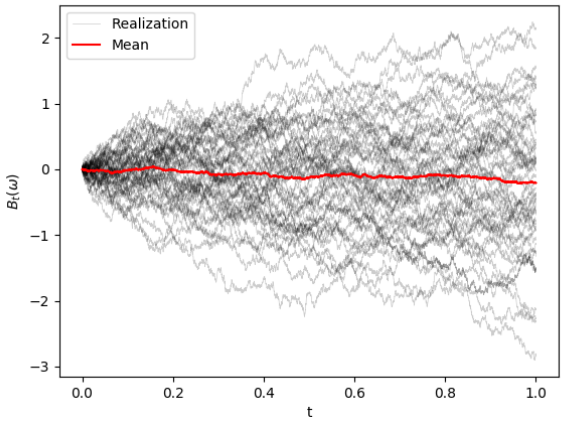


Figure 2.8: Brownian motion, 50 realizations and mean

Then the Laplace-De Moivre Theorem (see Appendix B.0.1) implies

$$\begin{aligned}\lim_{n\to\infty}\mathbb{P}(a\le X(t)\le b) &= \lim_{n\to\infty}\left(\frac{a}{\sqrt{tD}}\le\frac{S_n-\frac{n}{2}}{\sqrt{\frac{n}{4}}}\le\frac{b}{\sqrt{tD}}\right)\\ &= \frac{1}{\sqrt{2\pi}}\int_{\frac{a}{\sqrt{tD}}}^{\frac{b}{\sqrt{tD}}}\exp^{-\left(\frac{x^2}{2}\right)}\\ &= \frac{1}{\sqrt{2\pi Dt}}\int_a^b\exp-\left(\frac{x^2}{2Dt}\right)dx.\end{aligned}$$

Thus, we obtain that the position of a particle at a time t has a $N(0,Dt)$ distribution.

Definition 2.3.1. A real valued stochastic process $W(t)=W_t$, $t\ge 0$ is called a *Brownian motion* or *Winner process* if

i) $W(0)=0$ almost surely (a.s.).

ii) W_{t-s} is $N(0,t-s)$ for all $t\ge s\ge 0$, i.e., it has stationary increments.

iii) For all times $0<t_1<t_2<\cdots<t_n$, the random variables $W(t_1), W(t_2)-W(t_1),\ldots,W(t_n-t_{n-1})$, are independent, i.e., it has independent increments.

We use the notation W_t or $W(t)$ indistinctly, Figure 2.8. It can be shown that for each sample path w the function $t\to W_t(w)$ is continuous almost everywhere (a.e.) but it fails to be differentiable.

Now that we know what the Brownian motion is, we want to understand how to solve the equation (2.66), intuitively we could say that the solution is given by

$$u(t)=u_0+\int_0^t F dt+\int_0^t G dW_t.$$

But, what exactly does

$$\int_0^t G dW_t$$

mean? Let us work with the case $G = W(t)$ and define a left Riemann-Stieltjes sum approximation.

Definition 2.3.2. For the interval $[0, T]$ define a *partition* P of $[0, T]$ as finite collection of ordered points in $[0, T]$:

$$P := \{0 = t_0 < t_1 < \ldots < t_n = T\}.$$

The *mesh size* of P is $|P| := \max_{0 \leq k \leq n-1} |t_{k+1} - t_k|$. Then the left Riemann-Stieltjes sum approximation of $\int_0^T W dW$ is defined as

$$R = R(P) := \sum_{k=0}^{n-1} W(t_k)(W(t_{k+1}) - W(t_k)).$$

Lemma 2.3.1. *If P^n denotes a sequence of partitions such that $|P^n| \to 0$ as $n \to \infty$, the sequence*

$$R_n := \sum_{k=0}^{n-1} W(t_k^n)(W(t_{k+1}^n) - W(t_k^n)),$$

satisfies the limit

$$\lim_{n \to \infty} \mathbb{E}\left(\left(R_n - \frac{W(T)^2}{2} - \frac{T}{2}\right)^2\right) = 0,$$

which is,

$$\lim_{n \to \infty} R_n = \frac{W(T)^2}{2} - \frac{T}{2},$$

in $L^2(\Omega)$.

Therefore, we conclude that

$$\int_0^T W_s dW_s = \frac{W(T)^2}{2} - \frac{T}{2}. \tag{2.67}$$

An important observation to make is that we used left Riemann-Stieltjes sums to obtain (2.67). In a more general formulation, instead of left Riemann-Stieltjes sums, the intervals $[t_k, t_{k+1}]$ are parametrized by a straight line and the Riemann-Stieltjes sums are defined by

$$R_n(\lambda) := \sum_{k=0}^{n-1} W((1-\lambda)t_k^n + \lambda t_{k+1})(W(t_{k+1}^n) - W(t_k^n)), \quad 0 \leq \lambda \leq 1.$$

It can be shown that the integral $\int_0^T W_s dW_s$ depends on the value λ. Left Riemann-Stieltjes sums corresponds to $\lambda = 0$ and we will define it as the *Itô integral*. The Itô integral satisfies several properties that make it more advantageous. In the literature the object dW_t is usually called *white noise*.

Let $W(t)$ be a 1-dimensional Brownian motion defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$

Definition 2.3.3. We denote by $\mathbb{L}^p(0, T)$, $p = 1, 2$ the space of all real valued, progressively measurable stochastic processes $G(t)$ such that

$$\mathbb{E}\left(\int_0^T |G|^p dt\right) < \infty.$$

The construction of the Itô stochastic integral for functions in $\mathbb{L}^2$ resembles the construction of the Lebesgue integral. First it is defined for some process called "step processes", which are constant on finite disjoint intervals of $[0,T]$. Then it is shown that any stochastic process in $\mathbb{L}^2$ can be approximated by a sequence of step processes in $\mathbb{L}^2(\Omega)$. The Itô stochastic integral of an arbitrary $\mathbb{L}^2$ stochastic process is then defined as the limit of the stochastic integrals of the step processes that approximate the stochastic process.

Definition 2.3.4. Suppose that $X(t)$ is a real valued stochastic process satisfying

$$X(r) = X(s) + \int_s^r F dt + \int_s^r G dW_t,$$

for some $F \in \mathbb{L}^1(0,T)$, $G \in \mathbb{L}^2(0,T)$ and all times $0 \leq s \leq r \leq T$. We say that $X(t)$ has the *stochastic differential*

$$dX_t = F dt + G dW_t,$$

for $0 \leq t \leq T$.

Definition 2.3.5. (*Itô Formula*) Suppose that $X(t)$ has a stochastic differential

$$dX_t = F dt + G dW_t,$$

for $F \in \mathbb{L}^1(0,T)$, $G \in \mathbb{L}^2(0,T)$. Assume $u : \mathbb{R} \times [0,T] \to \mathbb{R}$ is continuous and that u_t, u_x, u_{xx} exist and are continuous. Define

$$Y(t) := u(X(t), t).$$

Then, $Y(t)$ has the stochastic differential

$$dY_t = \left(u_t + u_x F + \frac{1}{2} u_{xx} G^2 \right) dt + u_x G dW_t, \tag{2.68}$$

equivalently, $Y(t)$ satisfies

$$Y(r) - Y(s) = \int_s^r u_t(X,t) + u_x(X,t)F + \frac{1}{2} u_{xx}(X,t) G^2 dt + \int_s^r u_x(X_t,t) G dW_t.$$

Equality (2.68) is known as the Itô formula or thr Itô chain rule.

The following example not only illustrates an application of the Itô formula, but as well is the first example of an *stochastic differential equation.*

Take $X(t) = W_t$, $u(x,t) = \exp(\lambda x - \frac{\lambda^2 t}{2})$, $\lambda \in \mathbb{R}$, $F = 0$ and $G = 1$. Then for $Y(t) = u(X(t), t)$ the Itô formula implies

$$d\left(\exp(\lambda X_t - \frac{\lambda^2 t}{2}) \right) = \left(-\frac{\lambda^2}{2} \exp(\lambda X_t - \frac{\lambda^2 t}{2}) + \frac{\lambda^2}{2} \exp(\lambda X_t - \frac{\lambda^2 t}{2}) \right) dt + \lambda \exp(\lambda X_t - \frac{\lambda^2 t}{2}) dW$$

Then

$$dY_t = \lambda Y_t dW_t,$$

with initial condition $Y(0) = 1$ is solved by $Y(t) = u(X,t)$, Figure 2.9.

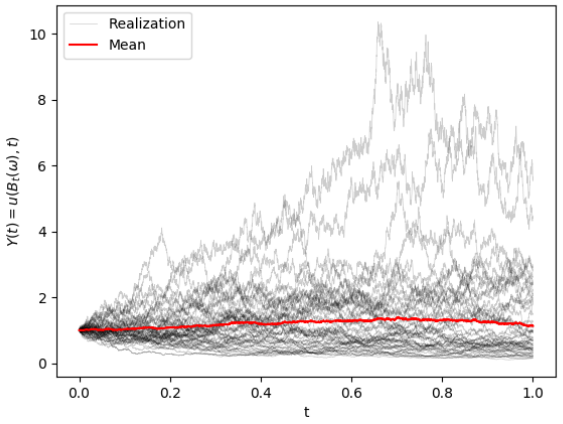


Figure 2.9: Stochastic process $Y(t)$, 50 realizations and mean, $\lambda = 1$.

Definition 2.3.6. We say that the stochastic process $X(t)$ is an *Itô process solution* of the *Itô stochastic differential equation*

$$dX_t = F(X_t, t)dt + G(X_t, t)dW_t \tag{2.69}$$
$$X(0) = X_0, \tag{2.70}$$

for $0 \leq t \leq T$, provided

i) $X(t)$ is progressively measurable with respect to $U(\cdot)$,

ii) $F \in \mathbb{L}^1(0,T)$, $G \in \mathbb{L}^2(0,T)$,

iii) $X(t) = X_0 + \int_0^T F(X(s), s)ds + \int_0^t G(X(s), s)dW_s$ a.s. for all $0 \leq t \leq T$.

With the following example, we show how to use the Itô formula to proof a stochastic process is a solution of a stochastic differential equation.

Consider the stochastic differential equation

$$dX_t = fX_t dt + gX_t dW_t,$$

with initial condition $X(0) = 1$.

Consider the stochastic process

$$Y(t) = (f - \frac{1}{2}g^2)dt + gdW_t,$$

and define $X(t) = e^{Y(t)}$. Then, using the Itô formula

$$\begin{aligned} dX(t) &= e^{Y(t)}\left(\left(\left(f - \frac{1}{2}g^2\right) + \frac{1}{2}g^2\right)dt + gdW_t\right) \\ &= e^{Y(t)}\left(fdt + gdW_t\right) \\ &= fX_t dt + gX_t dW_t. \end{aligned}$$

Theorem 2.3.2. *(Existence and uniqueness theorem) Suppose that* $F : \mathbb{R} \times [0,T] \to \mathbb{R}$ *and* $G : \mathbb{R} \to [0,T] \to \mathbb{R}$, *are continuous and satisfy the following conditions*

i) (Uniformly Lipschitz in space)

$$\begin{aligned} |F(x,t) - F(\hat{x},t)| &\leq L|x-\hat{x}| \\ |G(x,t) - G(\hat{x},t)| &\leq L|x-\hat{x}|, \end{aligned}$$

for all $0 \leq t \leq T$, $x, \hat{x} \in \mathbb{R}$.

ii)

$$\begin{aligned} |F(x,t)| &\leq L(1+|x|) \\ |G(x,t)| &\leq L(1+|x|), \end{aligned}$$

for some constant L.

iii) Let X_0 *be a random variable in* $\mathbb{R}$ *such that*

$$\mathbb{E}(|X_0|^2) < \infty$$

and

$$X_0 \text{ is independent of } \mathcal{W}^+(0),$$

where $W(t)$ *is a given Brownian motion.*

Then there exists a unique solution $X_t \in \mathbb{L}^2(0,T)$ *of the stochastic differential equation (SDE):*

$$dX_t = F(X_t,t)dt + G(X_t,t)dW_t \quad (0 \leq t \leq T) \tag{2.71}$$

$$X(0) = X_0. \tag{2.72}$$

Note that uniqueness is understood as follows: If $X_t, \hat{X}_t \in \mathbb{L}^2(0,T)$ have continuous sample paths almost surely, and both solve the same SDE, then

$$\mathbb{P}(X(t) = \hat{X}(t) \text{ for all } 0 \leq t \leq T) = 1.$$

2.3.2 Entropy weak solution

We first state and define weak entropy solutions for a general SPDE [19]. We then focus on particular examples of stochastic conservation laws [47] and parabolic problems [48].

Definition 2.3.7. The function u is a *stochastic weak solution* of (2.65) if almost surely,

$$\int_0^T \int_{\mathcal{O}} \phi_t(x,t)u(x,t) + (f(u) - Du_x)\phi_x dxdt = \int_0^T \int_{\mathcal{O}} g(\omega,x,t,u)\phi dxdtdW_t, \tag{2.73}$$

for all test functions $\phi \in \mathcal{D}(\mathcal{O} \times [0,T])$

To give the definition of a *stochastic entropy solution*, we require some notation. Let $\mathcal{D}^+(A)$ be the non-negative test functions of $\mathcal{D}(A)$, $N^2_\omega(0,T,L^2(\mathbb{R}))$ be the space of the predictable valued processes $L^2((0,T)\times\Omega, L^2(\mathbb{R}))$ for the product measure $dt \times dP$ on the $\sigma-$field generated by the sets $\{0\} \times \mathcal{F}_0$ and the rectangles $(s,t] \times A$ for any $A \in \mathcal{F}_s$. Let Σ be the set of non-negative even convex functions in $C^2(\mathbb{R})$ that approximate the absolute-valued function, such that $\eta(0) = 0$ and there exists a $\delta > 0$ with $\eta'(x) = 1$ (resp. -1) if $x > \delta$ (resp- $x < \delta$).

Definition 2.3.8. The function u in $N^2_\omega(0,T,L^2(\mathbb{R})) \in L^\infty(0,T;L^2(\Omega\times L^2(\mathbb{R})))$ is a *stochastic entropy solution* of (2.65) with initial condition $u_0 \in L^2(\mathbb{R})$, if almost surely for any 3-tuple $(\eta,\psi,r) \in \Sigma\times\mathbb{R}\times \mathcal{D}^+([0,T)\times\mathbb{R})$

$$\begin{aligned}\int_{\mathcal{O}_t} (\eta(u-k)\phi_t - F^\eta(u,k)\phi_x + K^\eta(u,k)\phi_{xx})dxdt \\ + \int_{\mathcal{O}_t} \eta'(u-k)g(u)\phi dxdtdW_t + \frac{1}{2}\int_{\mathcal{O}_t}\eta''(u-k)h^2(u)\phi dxdt \\ \geq -\int_{\mathbb{R}}\eta(u_0-k)\phi(0)dx, \end{aligned} \tag{2.74}$$

where F^η is the entropy flux for the function f, $F^\eta(a,b) = \int_a^b \eta'(\sigma-b)f'(\sigma)d\sigma$, and $K^\eta(a,b) = D\eta(a-b)$ the flux for D.

An entropy solution in this sense satisfies being a stochastic weak solution of (2.65). For a more complete description of stochastic entropy solutions see [9] and references therein.

As for the deterministic case, we give a definition of weak solutions for the stochastic conservation laws [47] and stochastic parabolic equations [48] in terms of an arbitrary partition.

Definition 2.3.9. The function $u(\omega,x,t)$ is a weak solution for (2.65) with $D=0$ if for any $(\omega,t) \in \Omega\times[0,T]$ and any test function ϕ

$$\int_{I_j} du_h dx = \left(\int_{I_j} f(u)\phi_x dx - (f\phi)\big|_{x=x_j}^{x=x_{j+1}}\right)dt + \left(\int_{I_j} g\phi dx\right)dW_t \tag{2.75}$$

The stochastic parabolic equation requires to be rewritten as the first order system

$$\begin{aligned} du &= f(u) - w_x + gdW_t \text{ in } \mathcal{O}_T && (2.76)\\ v &= u_x \text{ in } \Omega\times\mathcal{O}_T && (2.77)\\ w &= Dv \text{ in } \mathcal{O}_T && (2.78)\\ u(x,0) &= u_0(x), \text{ on } \mathcal{O}, && (2.79)\end{aligned}$$

where $g(u) = \sqrt{D}u$.

Definition 2.3.10. We say u is a solution of (2.76)-(2.79) if for each interval I_j, $j=1,\dots,N$ and for every test functions $\phi_1,\phi_2,\phi_3,\phi_4$ the following equalities hold

$$\begin{aligned}\int_{I_j}\phi_1 du(\omega,x,t)dx &= \left(\int_{I_j} -w(\phi_1)_x + f\phi dx + (w\phi_1)\big|_{x=x_j}^{x=x_{j+1}}\right)dt + \int_{I_j} g\phi dx dW_t, && (2.80)\\ \int_{I_j} v(\omega,x,t)\phi_2 dx &= -\int_{I_j} u(\phi_2)_x dx + (u\phi_2)\big|_{x=x_j}^{x=x_{j+1}}, && (2.81)\\ \int_{I_j} w(\omega,x,t)\phi_3 dx &= D\int_{I_j} v\phi_3 dx && (2.82)\\ \int_{I_j} u(\omega,x,0)\phi_4 dx &= \int_{I_j} u_0(x)\phi_4 dx && (2.83)\end{aligned}$$

Analogously to the previous sections we present benchmark problems where the solution is known which will be used for the validation process. But let us first formulate the characteristic method for a stochastic conservation law of the form

$$du = -(f(u))_x + g(x,t,u)dW_t.$$

The system of characteristic equations (2.9)-(2.11), is modified as follows,

$$\begin{aligned} \frac{dx}{ds} &= f_u(z(s)), && (2.84)\\ \frac{dt}{ds} &= 1, && (2.85)\\ dz &= g dW_t, && (2.86) \end{aligned}$$

note that the characteristic for z is now an stochastic differential equation. The solution of (2.85) and (2.86) are

$$t(s) = s \quad \text{and} \quad z(s) = u_0(x_0) + \int_0^t g dW_t,$$

thus

$$x(s) = x_0 + \int_0^s f_u \left(u_0(x_0) + \int_0^\xi g dW_s d\xi \right) ds.$$

We can then conclude that over the characteristic curve

$$(x(t),t) = \left(x_0 + \int_0^t f_u \left(u_0(x_0) + \int_0^\xi g dW_s d\xi \right) ds, t \right), \tag{2.87}$$

the solution is

$$u(x(t),t) = u_0(x_0) + \int_0^t g dW_t. \tag{2.88}$$

Example 1. Constant-coefficient linear stochastic equation with $\mathcal{O} = [0, 2\pi]$

$$\begin{aligned} du &= -u_x dt + b u dW_t \text{ in } \mathcal{O} \times (0,T] \times \Omega, && (2.89)\\ u(\omega, x, 0) &= u_0(x) \quad \text{on } \mathcal{O} \times \Omega. \end{aligned}$$

The characteristic curves are (2.87)

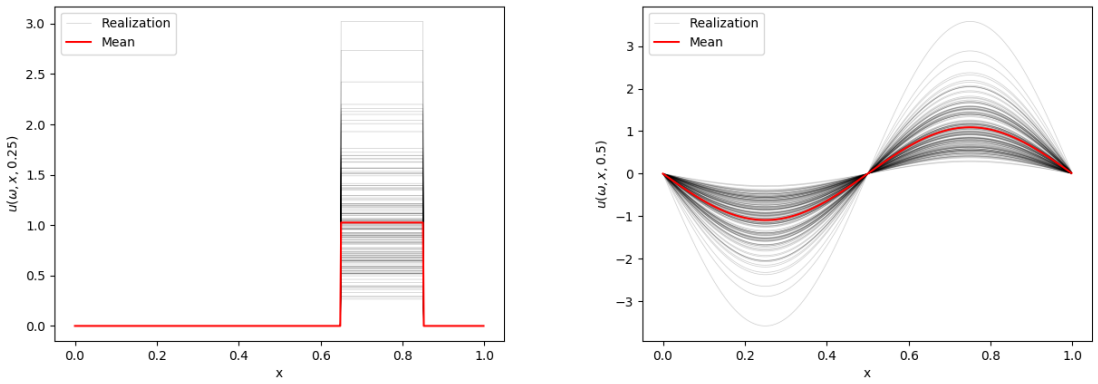
$$(x(t),t) = (x_0 + t, t),$$

and the solution is

$$\begin{aligned} du(x(t),t) &= b u dW_t \\ \implies u(x(t),t) &= u_0(x_0) \exp\left(bW_t - \frac{1}{2} b^2 t \right). \end{aligned}$$

Substituting $x_0 = x - t$, the exact solution of (2.89) is

$$u(\omega, x, t) = u_0(x - t) \exp\left(bW_t(\omega) - \frac{1}{2} b^2 t \right)$$



(a) Solution stochastic convection equation impulse initial condition, 50 realizations and mean.

(b) Solution stochastic convection equation sine initial condition, 50 realizations and mean.

Figure 2.10: Solutions to stochastic conservation law $b = 0.5$.

with the impulse initial condition

$$u_0(x) = \begin{cases} 1 & \text{if } x \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right], \\ 0 & \text{if } x \in \left[0, \frac{\pi}{2}\right) \cup \left(\frac{3\pi}{2}, 2\pi\right], \end{cases}$$

Figure 2.10a and a sine initial condition

$$u_0(x) = \sin(x),$$

Figure 2.10b. In Figures 2.11a and 2.11b the lateral and upward view of a realization of the solution to (2.89) up to time $T = 0.3$ are shown. It can be seen that the solutions are not differentiable in time and that the initial condition travels to the right as for the deterministic case.

Example 2 Stochastic Burgers equation,

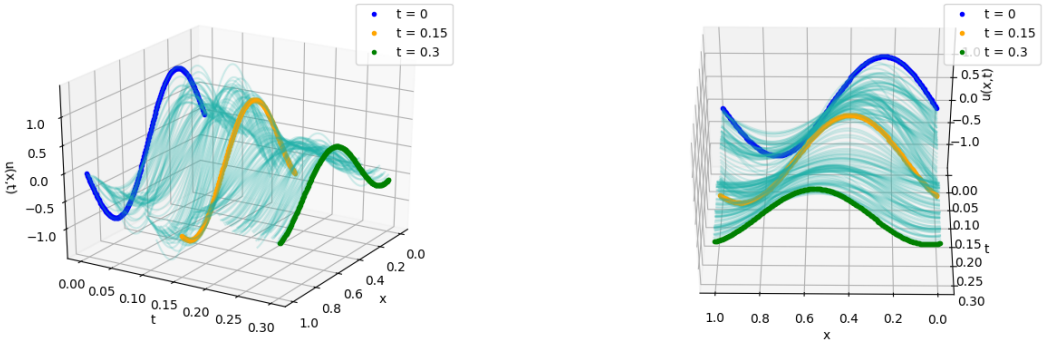
$$\begin{aligned} du &= -\left(\frac{1}{2}u^2\right)_x dt + b dW_t \text{ in } \mathcal{O} \times (0,T] \times \Omega, \\ u(\omega, x, 0) &= u_0(x) \quad \text{on } \mathcal{O} \times \Omega. \end{aligned} \tag{2.90}$$

The characteristic curve $x(t)$ (2.87) is

$$\begin{aligned} x(t) &= x_0 + u_0(x_0)t + b\int_0^t W_s ds \\ \implies x(t) - b\int_0^t W_s ds &= x_0 + u_0(x_0)t. \end{aligned}$$

If we consider a deterministic Burgers' equation with the same initial condition $u_0(x)$,

$$\begin{aligned} v_t + \left(\frac{1}{2}u^2\right)_x &= 0 \\ v(x,0) &= u_0(x), \end{aligned}$$



(a) Lateral view. (b) Upward view.

Figure 2.11: Realization of surface solution for stochastic convection equation $b = 2$.

its solution over the deterministic characteristic $x(t) = x_0 + u_0(x_0)t$ is

$$v(x(t), t) = u_0(x_0).$$

Then, substituting

$$x(t) - b\int_0^t W_s ds = x_0 + u_0(x_0)t$$

in $v(x(t), t)$,

$$v\left(x(t) - b\int_0^t W_s ds, t\right) = u_0(x_0).$$

Then the solution, over the characteristic $(x(t), t)$ is

$$\begin{aligned} u(x(t), t) &= u_0(x_0) + bW_t && (2.91)\\ &= v\left(x(t) - b\int_0^t W_s ds, t\right) + bW_t. && (2.92)\end{aligned}$$

See Appendix B for the details on the computation of the random variable $\int_0^t W_s ds$.

Example 3. Stochastic heat equation.

$$\begin{aligned} du &= u_{xx}dt + budW_t \text{ in } \mathcal{O} \times (0, T] \times \Omega, \\ u(\omega, x, 0) &= \sin(x) \quad \text{on } \mathcal{O} \times \Omega. \end{aligned} \tag{2.93}$$

The exact solution of (2.93) is

$$u(\omega, x, t) = \sin(x) \exp\left(bW_t(\omega) - \frac{1}{2}b^2 t - t\right),$$

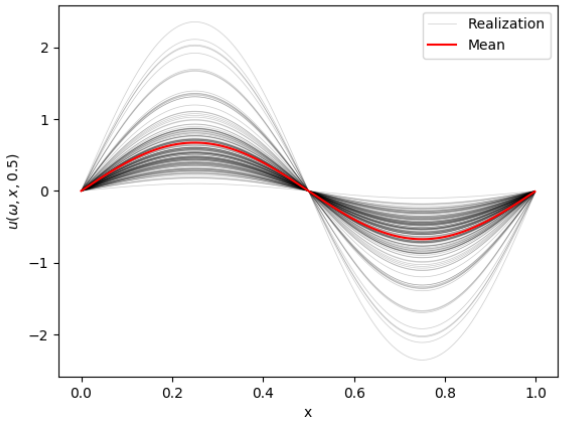


Figure 2.12: Solution to the stochastic linear equation, 50 realizations and mean $b = 0.5$.

Example 4. Stochastic nonlinear convection-diffusion equation.

$$\begin{aligned} du &= \left[\frac{\sigma^2}{2}u_{xx} - \left(\frac{1}{2}u^2\right)_x\right]dt + (\sigma u_x + b)dW_t \text{ in } \mathcal{O}\times(0,T]\times\Omega, \\ u(\omega,x,0) &= \sin(x) \quad \text{on } \mathcal{O}\times\Omega. \end{aligned} \tag{2.94}$$

The exact solution of (2.94) is

$$u(\omega,x,t) = v(x - b\int_0^t W_s ds + \sigma W_t, t) + bW_t,$$

with v the solution to the Burgers' equation, with initial condition $v_0(x)\sin(x)$.

It is important to note that the function $v(x,t)$ for both (2.90) and (2.94) develop a shock at a time T_b. In both cases the solutions are made up of the solution v and an additive Brownian motion. Therefore, the shock in both solutions will still be present for the stochastic case at the time T_b.

Chapter 3

Discontinuous Galerkin Methods

3.1 Introduction

In this chapter we formulate the discontinuous Galerkin (DG) method for hyperbolic conservation laws and the local discontinuous Galerkin (LDG) method for the parabolic case. We apply them to the deterministic examples studied in the previous chapter and show how it outperforms other Galerkin methods and other classic methods for the convective dominated problems. Furthermore, we provide as well the pseudoalgorithms for the implementations of the DG and LDG. The theory of this chapter can be found in [16], [25] and [71].

The discontinuous Galerkin method was originally introduced by Reed and Hill in 1973 [61]. This method was later adapted to solve elliptic equations and generalized to the local discontinuous Galerkin for studying convection-dominated problems [16]. Simultaneously and independently the *interior penalty* (IP) methods were developed [62]. The IP methods have a great resemblance to the DG. They are constructed using the weak formulation, but they weakly impose inter-element continuity. Due the resemblance of the methods some authors do not distinguish the IP methods and simply call them discontinuous Galerkin methods. This can be confusing and misleading since the classic DG and the IP are at first glance different and produce different results. In [2] two formulations of the DG methods were developed, the flux formulation (generalization of the original DG method) and the primal formulation (generalization of the IP method). This development allowed to unify the theory of discontinuous Galerkin methods in a single framework and to understand the relationships and the differences between the IP methods and DG methods. It was shown that the IP methods were a particular case of the flux formulation by a careful selection of the numerical fluxes. In [2], it is shown that 9 different formulations of DG methods can be deduced from a proper choice of flux in the flux formulation, including DG, LDG and IP. The author of this work has the opinion that the IP method should not be simply called a discontinuous Galerkin method, it must be properly distinguished when used. The IP method uses parameters to define weak inter-element continuity and its results drastically depend on the selection of the parameters.

The DG can be developed in two ways, the nodal formulation and the modal formulation. Both formulations use the basic idea of expressing the solution of the EDP as a linear combination of a basis. For the nodal formulation, points of the spatial domain are selected and the corresponding

Lagrange polynomials are used as a base. With such selection, the method only allows to know the solution of the problem at the selected spatial points at each time step. For the modal formulation, the basis is usually the Legendre polynomials, but this is not necessary. With this formulation the method finds the weights of the solution in terms of the spatial basis at each time step. Thus, the solution at an arbitrary point in the space dimension can be computed at each time step. When the Legendre polynomials are used in the modal formulation, the nodal and modal formulations can be related, see [25]. In this section we work with the modal formulation. The nodal formulation is going to be used later on in the "Entropy stable discontinuous Galerkin method" chapter.

3.2 Spatial discretization

Prior to the study of the DG and LDG methods, we introduce some concepts and useful notation. We assume without loss of generality $\mathcal{O} = [0, 1]$. Also let us assume a uniform partition of width $h = \frac{1}{N}$, $N \in \mathbb{N}$ i.e. $x_j = \frac{j}{N}$, for $j = 0, \ldots, N$, denote by $I_j = (x_j, x_{j+1})$ and define $x_{j+\frac{1}{2}} = \frac{x_j + x_{j+1}}{2}$. The intervals I_j are generally known as *elements* and the partition elements x_j are called *inter-elements*. Define the set of piecewise polynomials of degree at most K by

$$V_h^K = \{\phi \in L^1(0,1) : \phi|_{I_j} \in P^K(I_j), j = 1, \ldots, N-1\}.$$

A polynomial basis commonly used to generate V_h^k are the Legendre polynomial, we use P_j to denote the Legendre polynomial of order j. The first four Legendre polynomial are

$$\begin{aligned} P_0 &= 1, \qquad P_2 = \frac{3x^2 - 1}{2}, \\ P_1 &= x, \qquad P_3 = \frac{5x^3 - 3x}{2}. \end{aligned}$$

Properties of the Legendre polynomials that we are going to used are

$$\int_{-1}^{1} P_m(x) P_l(x) dx = \frac{2}{2l+1}\delta_{ml} \tag{3.1}$$

$$\int_{-1}^{1} P_n(x) = 0 \text{ for } n \geq 1 \tag{3.2}$$

$$\frac{d}{dx} P_{n+1}(x) = (2n+1)P_n(x) + (2(n-2)+1)P_{n-2}(x) + (2(n-4)+1)P_{n-4} + \ldots, \tag{3.3}$$

$$P_n(1) = 1, \quad P_n(-1) = (-1)^n. \tag{3.4}$$

where δ_{ml} is the Kronecker delta.

3.3 Time discretization.

The DG and LDG reduce the solution of PDEs to solve a system of ODEs. For the solution of the system of ODEs the TVD-Runge-Kutta (TVDRK) method is used in both cases. We first explain the time integrator, so that later we can freely use it in the DG and LDG methods. The order of the TVDRK method is related to the polynomial order used for generating V_h^K, the corresponding TVDRK order is $K + 1$.

Assume we have a system of differential equations for the approximate solution u_h of the form

$$\frac{d}{dt}u_h(t) = L_h(u_h, t), \text{ for all } t \in (0, T) \tag{3.5}$$
$$u_h(0) = u_0^h. \tag{3.6}$$

Then for the time partition $t_n = n\tau$, the TVD Runge-Kutta methods of order two, three and four are

i) Second order

$$\begin{aligned} w^1 &= u_h^n + \tau L_h(u_h^n), \\ u_h^{n+1} &= \frac{1}{2}u_h^n + \frac{1}{2}(w^1 + \tau L_h(w^1)) \end{aligned}$$

ii) Third order

$$\begin{aligned} w^1 &= u_h^n + \tau L_h(u_h^n), \\ w^2 &= \frac{3}{4}u_h^n + \frac{1}{4}(w^1 + \tau L_h(w^1)) \\ u_h^{n+1} &= \frac{1}{3}u_h^n + \frac{2}{3}(w^2 + \tau L_h(w^2)) \end{aligned}$$

iii) Fourth order

$$\begin{aligned} L_h^0 &= L_h(u_h^n) \\ w^1 &= u_h^n + \tau L_h^0 \\ L_h^1 &= L_h(w^1) \\ w^2 &= \frac{1}{2}(u_h^n + w^1) + \frac{\tau}{4}(-L_h^0 + 2L_h^1) \\ L_h^2 &= L_h(w^2) \\ w^3 &= \frac{1}{9}(u_h^n + 2w^1 + 6w^2) + \frac{\tau}{9}(-L_h^0 - 3L_h^1 + L_h^2) \\ L_h^3 &= L_h(w^3) \\ u_h^{n+1} &= \frac{1}{3}(w^1 + w^2 + w^3) + \frac{\tau}{6}(L_h^1 + L_h^3). \end{aligned}$$

The prefix TVD means that the method has the *total variation diminishing property.* Meaning that the discrete total variation TV,

$$TV(u^n) = \sum_j |u_{j+1}^n - u_j^n|,$$

where $u_j^n = u(x_j, t_n)$ satisfies

$$TV(u^{n+1}) \leq TV(u^n),$$

see [46] and also see Algorithms 1 and 2 for the implementation of the TVD Runge-Kutta method.

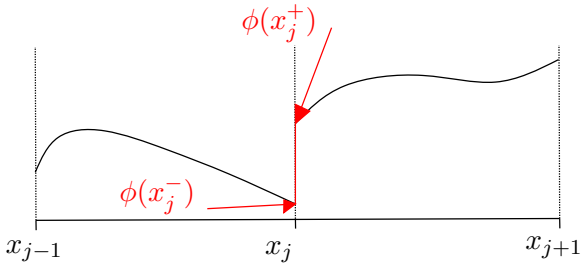


Figure 3.1: Example test function discontinuous at x_j.

3.4 Discontinuous Galerkin method

Recall that the conservation law is

$$u_t + (f(u))_x = 0.$$

Let us rewrite such equation using the weak formulation from Definition 2.2.4 with a test function $\phi \in V_h^K$ and consider an initial condition, then

$$\begin{aligned} \int_{I_j} u_t \phi dx &= \int_{I_j} f(u)\phi' dx + f(u(x_j^+,t))\phi(x_j^+) - f(u(x_{j+1}^-,t))\phi(x_{j+1}^-), \\ \int_{I_j} u(x,0)\phi(x) &= \int_{I_j} u_0(x)\phi(x)dx. \end{aligned}$$

The notation x_j^+ denotes the limit from the right and x_{j+1}^- the limit from the left. These limits are required because the polynomial ϕ is only continuous in each interval I_j, that is, it is no required that $\phi(x_j^-) \neq \phi(x_j^+)$, see Figure 3.1.

We look for an approximation $u_h \in V_h^K$ to u at each time $t \in (0,T)$. Then there is no reason for u_h to be continuous at the inter-element points x_j, in such case there is an ambiguity on how to define $f(u_h(x_j^\pm))$. Thus, the flux $f(u_h(x_j^\pm))$ is substituted by a *numerical flux* $\hat{f}$ that depends on the left and right values of $u_h(x_j^\pm)$,

$$\hat{f}_j(u_h,t) := \hat{f}(u_h(x_j^-,t), u_h(x_j^+,t)).$$

Then the approximation solution u_h for each time $t \in (0,T)$ is defined as the function in V_h^K that satisfies the weak formulation

$$\int_{I_j} (u_h)_t \phi dx = \int_{I_j} f(u_h)\phi' dx + \hat{f}_j(u_h,t)\phi(x_j^+) - \hat{f}_{j+1}(u_h,t)\phi(x_{j+1}^-), \tag{3.7}$$

$$\int_{I_j} u_h(x,0)\phi(x) = \int_{I_j} u_0(x)\phi(x)dx, \tag{3.8}$$

for all $\phi \in V_h^k$ and all I_j, $j = 0,1,\ldots,N-1$.

The numerical flux is chosen so that

i) (*Locally Lipschitz*) For each point x_0 of the domain there is a neighborhood where the function is Lipschitz.

ii) (*Consistent*) $\hat{f}(u,u) = f(u)$.

iii) (*Monotone*) Nondecreasing of its first argument and nonincreasing of its second argument.

Classic examples of numerical fluxes are

i) The Lax-Friedrichs flux:

$$\begin{aligned} \hat{f}^{LF}(a,b) &= \frac{1}{2}[f(a)+f(b)-C(b-1)], \\ C &= \max_{\inf u_0(x)\le s\le \sup u_0(x)} |f'(s)|. \end{aligned}$$

ii) The local Lax-Friedrichs flux:

$$\begin{aligned} \hat{f}^{LLF}(a,b) &= \frac{1}{2}[f(a)+f(b)-C(b-1)], \\ C &= \max_{\min(a,b)\le s\le \max(a,b)} |f'(s)|. \end{aligned}$$

iii) The Godunov flux:

$$\hat{f}^{G}(a,b) = \begin{cases} \min_{a\le u\le b} f(u) & a \le b \\ \max_{b\le u\le a} f(u) & b < a. \end{cases}$$

There is not a definite rule for choosing a numerical flux. In most of the literature, [16] ,[25], [59] and [71] the preferred limiter, due to its simplicity, is the LLF limiter. In [16] the authors commented that in their experience the flux selection becomes irrelevant as K and N are increased. In most of the numerical experiments this was the case, we point out the examples where the G flux is better. In [25], the authors show that the LF is more dissipative than the LLF, the authors mention that even though for some problems other flux limiters might give better results, the simplicity of the LLF makes it a more suitable choice.

We choose the Legendre polynomials $(P_l)_{l=0}^K$ to be a basis of the polynomials at each interval I_j, to do so, we consider the local change of variables

$$\xi_j(x) = 2\frac{x - x_{j+\frac{1}{2}}}{h}.$$

Since for each time $t \in (0,T)$ we assume that $u_h \in V_h^k$, then for each $j = 0,\dots,N-1$ there are K coefficients u_j^l dependent on time, such that

$$u_h(x,t)|_{I_j} = \sum_{l=0}^{K} u_j^l(t) P_l(\xi_j(x)). \tag{3.9}$$

Substituting (3.9) in (3.7) and (3.8) and using (3.1) and (3.4), the weak form simplifies to

$$\frac{d}{dt}u_j^l(t) = \frac{2l+1}{h}\left(\int_{-1}^{1} f(u_h(\xi,t))P_l'(\xi)d\xi + (-1)^l \hat{f}_j^+ - \hat{f}_{j+1}^-\right) \tag{3.10}$$

$$u_j^l(0) = \frac{2l+1}{2}\int_{-1}^{1} u_0(\xi)P_l(\xi)d\xi, \tag{3.11}$$

for all $j = 0, \ldots, N-1$ and all $l = 0, \ldots, K$. Equation (3.10) is usually called the *semidiscrete form* of the problem. See Algorithm 4 for the implementation.

Note that at each x_j, $j = 1, \ldots, N-1$, there are two values of u_h, one from the approximation of u_h in the interval I_{j-1} and one from the approximation in the interval I_j. These two values are the ones evaluated in the flux limiter $\hat{f}$.

The solution of the system of ODEs (3.10) can be represented in a vector of length $(K+1)N$,

$$u_h = \begin{bmatrix} u_0^0 & \ldots & u_0^K, & u_1^0 & \ldots & u_1^K & \ldots & u_{N-1}^0 & \ldots & u_{N-1}^K \end{bmatrix}.$$

With this notation, the ODE system can be solved using the TVD Runge-Kutta method (3.5), with L_h represented by

$$[L_h(u_h, t)]_i = \frac{2l+1}{h}\left(\int_{-1}^{1} f(u_h(\xi, t))P_l'(\xi)d\xi + (-1)^l \hat{f}_j^+ - \hat{f}_{j+1}^-\right),$$

for $j \in \{0, \ldots, N-1\}$ and $l \in \{0, \ldots, K\}$ such that $i = 2(K+1)j + l$.

Notice that the method requires that the integrals

$$\int_{-1}^{1} f(u_h(\xi, t))P_l'(\xi)d\xi \quad \text{and} \quad \int_{-1}^{1} u_0(\xi)P_l(\xi)d\xi,$$

are performed analytically, which can be difficult, time consuming and even a bit tedious. The first integral is usually done analytically, numerical quadratures can seriously affect the accuracy of the method or even produce non-entropy solutions, specially for non-polynomial f. In Chapter 5, we work on a numerical quadrature using Gauss-Lobatto points that produces entropy solutions. In case of f being a polynomial, the properties (3.2) and (3.3) might be quite handy. The second integral can be performed numerically without a great loss in accuracy.

Remark If instead of the Legendre polynomials (P_l) other polynomials were chosen as a basis, the system of equations (3.10) would not be diagonal. As part of the work of the thesis, we did the computations and algorithms for two other set of polynomial basis, Radau polynomials $(R_l(x))$ and Bernstein polynomials $(B_l(x))$ for the linear convection equation and Burgers' equation. Surprisingly, the numerical results obtained with the Radau and Bernstein polynomials were exactly the same to the ones obtained by using the Legendre polynomials. For the Radau polynomials this is reasonable, since the l-th Radau polynomial is defined by

$$R_l(x) = \frac{(-1)^l}{2}(P_l - P_{l-1}).$$

So implicitly the base is still generated by the Legendre polynomials. For the Bernstein basis it was more intriguing that the numerical results were exactly the same. First the basis generates continuous functions in the interval $[0, 1]$, instead of $[-1, 1]$, and second that the Bernstein polynomials are not defined by means of the Legendre polynomials, so it is reasonable to expect different results. As an example of the ODE system obtained with these polynomials we show the systems for $K = 2$ for an

arbitrary interval I_j

$$\begin{bmatrix} 2 & 1 & 0 \\ 1 & \frac{2}{3} & \frac{1}{6} \\ 0 & \frac{1}{6} & \frac{4}{15} \end{bmatrix} \begin{bmatrix} \frac{d}{dt}u_j^0 \\ \frac{d}{dt}u_j^1 \\ \frac{d}{dt}u_j^2 \end{bmatrix} = L_h(u_j) \quad \text{Radaua polynomials,}$$

$$\begin{bmatrix} 9 & -9 & 3 \\ -9 & 21 & -9 \\ 3 & -9 & 9 \end{bmatrix} \begin{bmatrix} \frac{d}{dt}u_j^0 \\ \frac{d}{dt}u_j^1 \\ \frac{d}{dt}u_j^2 \end{bmatrix} = L_h(u_j) \quad \text{Bernstenin polynomials,}$$

the left side is not explicitly computed since it depends on the flux f. Since these basis did not had a change in the results and were slower, we did not pursued continuing exploring their performance for other examples. We do not include the results of the error in this work.

The DG method solution u_h satisfies for the square entropy $\eta(u) = \frac{u^2}{2}$ a cell entropy inequality in each interval I_j for the entropy function $\psi' = uf'$,

$$\int_{I_j} \left(\frac{u_h^2(x,t)}{2} \right)_t dx + \psi_{j+1}(t) - \psi_j(t) \leq 0. \tag{3.12}$$

This inequality leads to L^2 stability of the scheme,

$$\frac{d}{dt} \int_{\mathcal{O}} \left(\frac{u_h^2(x,t)}{2} \right) dx \leq 0.$$

In particular for a convex flux $f(u)$ and in conjunction with the RKTVD method and the TVB limiter (next section), the solution is TVB and will converge towards the unique entropy solution [34].

When polynomials of degree K are used, the corresponding order of the Runge-Kutta must at least $(k+1)$ so that the order of the method is $O(h^{k+1})$ in the L^∞ norm away from discontinuities. The theoretical study of error bounds, stability and more theoretical properties of the RKDG (Runge-Kutta DG) is out of the scope of this work, we refer the interested reader to [16].

In the examples of this thesis, we choose $N = [10, 20, 40, 80, 160, 320]$, so the numerical order $\hat{K}$ can be computed by

$$\hat{K} = \frac{\log\left(\frac{u_h^n}{u_h^{n+1}} \right)}{\log(2)}.$$

The notation $u_h^{(K)}$ is used in the tables related to all the examples to denote the $K-$th order polynomial approximation, $K = 1, 2, 3$.

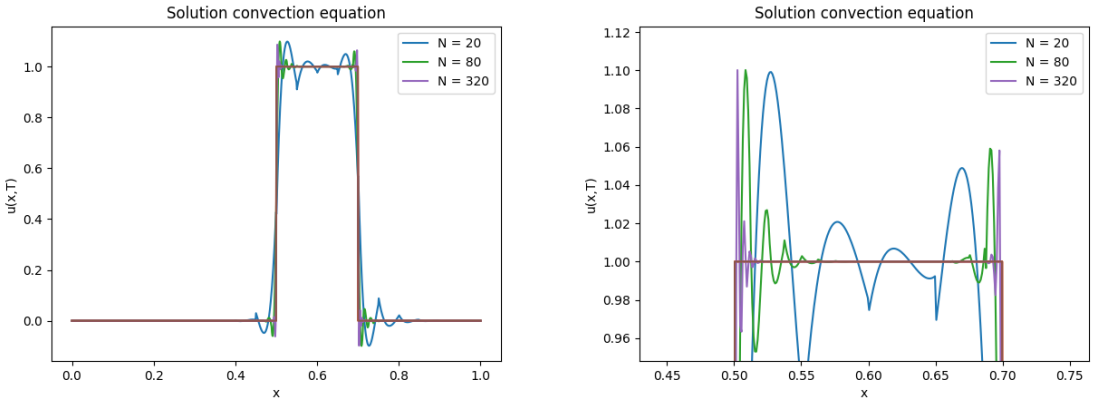
3.4.1 Examples

Example 1. Linear convection equation,

$$u_t + u_x = 0.$$

In the previous chapter we worked with two initial conditions a sine and an impulse function (2.37) and (2.38). For the impulse initial condition Figure 3.2a and Figure 3.2b it can be seen that the DG presents the Gibbs effect and oscillates around discontinuities. In the Table 3.1 the error is shown for

the different number of elements and different polynomial orders of approximation. This table shows that the error is indeed decreasing as a function of the elements but is not order $O(h^{K+1})$. In addition the error almost does not decreases when the polynomial order is increased. Which happens because oscillations near discontinuities do not disappears when K increases. In the case of the sine initial condition, see Table 3.2 and Figure 3.3 it can be seen that the order satisfies to be $O(h^{K+1})$ and that the DG approximations are indistinguishable from the exact solution for $K = 3$. The CFL condition was set to 0.01 for $K = 1$, 0.001 for $K = 2$ and 0.0001 for $K = 3$.



(a) General view

(b) Zoom on oscillatory behaviour.

Figure 3.2: Comparison of DG solutions for the linear convection equation $\alpha = 1$, $K = 3$, $T = 0.1$ impulse initial condition.

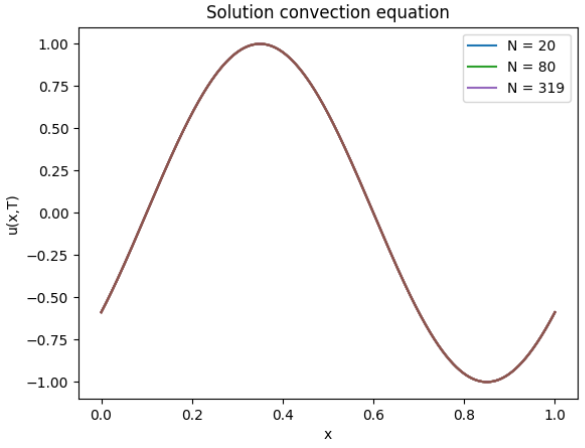


Figure 3.3: Comparison of DG solutions of the linear convection equation $\alpha = 1$, $K = 3$, $T = 0.1$ sine initial condition.

Example 2. Burgers' equation

$$u_t + \left(\frac{u^2}{2}\right)_x = 0.$$

Both of the initial conditions (2.37) and (2.38) develop a shock and (2.38) develops a rarefaction wave.

N	$\|\|u_h^{(1)} - u\|\|_{L^\infty}$	Ord. $K=1$	$\|\|u_h^{(2)} - u\|\|_{L^\infty}$	Ord. $K=2$	$\|\|u_h^{(3)} - u\|\|_{L^\infty}$	Ord. $K=3$
10	7.6×10^{-2}	-	5.1×10^{-2}	-	4.1×10^{-2}	-
20	4.7×10^{-2}	0.67	3×10^{-2}	0.23	2.31×10^{-2}	0.82
40	2.8×10^{-2}	0.72	1.7×10^{-2}	0.20	1.3×10^{-2}	0.82
80	1.7×10^{-2}	0.76	1×10^{-2}	0.18	7.2×10^{-3}	0.85
160	1×10^{-2}	0.73	5.8×10^{-3}	0.16	4×10^{-3}	0.85
320	6.2×10^{-3}	0.72	3.2×10^{-3}	0.15	2.2×10^{-3}	0.85
Time	$6.43s$	-	$8.23s$	-	$39.33s$	-

Table 3.1: DG solutions linear convection equation $\alpha = 1$ impulse initial condition.

N	$\|\|u_h^{(1)} - u\|\|_{L^\infty}$	Ord. $K=1$	$\|\|u_h^{(2)} - u\|\|_{L^\infty}$	Ord. $K=2$	$\|\|u_h^{(3)} - u\|\|_{L^\infty}$	Ord. $K=3$
10	1.24×10^{-2}	-	6.11×10^{-4}	-	2.27×10^{-5}	-
20	3.13×10^{-3}	2	8.72×10^{-5}	2.97	1.44×10^{-6}	3.97
40	7.8×10^{-4}	2	9.77×10^{-6}	3	9.24×10^{-8}	3.96
80	2×10^{-4}	2	1.23×10^{-6}	2.99	5.74×10^{-9}	4
160	4.9×10^{-5}	2	1.51×10^{-7}	3.03	3.63×10^{-10}	3.98
320	2.2×10^{-5}	1.98	1.9×10^{-8}	2.98	2.3×10^{-11}	3.97
Time	$6.43s$	-	$8.23s$	-	$39.33s$	-

Table 3.2: DG solutions linear convection equation $\alpha = 1$ sine initial condition.

For the sine initial condition, prior to the shock time $T_b = \frac{1}{2\pi}$, the solution has the expected order and for $K = 3$ and it can not be distinguished from the analytical solution, see Table 3.3 and Figure 3.4. After the shock time, the DG losses its order and performance. Both, the order and performance, decay and the solutions do not greatly improve by increasing neither the elements nor the polynomial order, see Table 3.4 Figure 3.5a and Figure 3.5b. For the impulse initial condition, the DG has no problem in the rarefaction wave, but once again it has troubles in the discontinuity, see Table 3.4, Figure 3.5a and Figure 3.5b. This example is the only one so far to improve by changing to the G flux, see Table 3.5 and Figure 3.6. Sadly this behaviour depends on the parity of N, as it seen in Figure 3.7a and Figure 3.7b.

The previous examples show that the DG method is not exempt of the Gibbs phenomena. In the next chapter we work with *limiters*, a technique that helps to control this oscillatory behaviour.

N	$\|\|u_h^{(1)} - u\|\|_{L^\infty}$	Ord. $K=1$	$\|\|u_h^{(2)} - u\|\|_{L^\infty}$	Ord. $K=2$	$\|\|u_h^{(3)} - u\|\|_{L^\infty}$	Ord. $K=3$
10	3.37×10^{-2}	-	8.28×10^{-3}	-	9.11×10^{-4}	-
20	9.08×10^{-3}	1.89	1.17×10^{-3}	2.81	6.69×10^{-5}	3.76
40	2.26×10^{-3}	2	1.45×10^{-4}	3.02	8.84×10^{-6}	2.92
80	5.73×10^{-4}	1.98	2.86×10^{-5}	2.34	1.01×10^{-6}	3.11
160	1.57×10^{-4}	1.87	4.61×10^{-6}	2.63	1.78×10^{-7}	2.51
320	3.9×10^{-5}	2.01	6.92×10^{-7}	2.73	6.93×10^{-8}	1.36
Time	$7.35s$	-	$9.98s$	-	$50.76s$	-

Table 3.3: DG solutions Burgers' sine initial condition $T = T_b - 0.05$.

N	$\|\|u_h^{(1)} - u\|\|_{L^\infty}$	Ord. $K=1$	$\|\|u_h^{(2)} - u\|\|_{L^\infty}$	Ord. $K=2$	$\|\|u_h^{(3)} - u\|\|_{L^\infty}$	Ord. $K=3$
10	6.86×10^{-2}	-	8.28×10^{-2}	-	4.38×10^{-4}	-
20	3.72×10^{-2}	0.89	5.33×10^{-2}	0.81	2.43×10^{-2}	0.84
40	2.05×10^{-2}	0.85	3.06×10^{-2}	0.89	1.29×10^{-2}	0.91
80	1.17×10^{-2}	0.81	1.64×10^{-3}	0.81	8.43×10^{-3}	0.61
160	5.52×10^{-3}	1.08	4.11×10^{-3}	1.18	5.59×10^{-3}	0.59
320	4.76×10^{-3}	0.21	3.54×10^{-3}	0.21	5.18×10^{-3}	0.11
Time	$9.76s$	-	$12.83s$	-	$111.34s$	-

Table 3.4: DG solutions Burgers' sine initial condition $T = T_b + 0.1$.

N	$\|\|u_h^{(1)} - u\|\|_{L^\infty}$	Ord. $K=1$	$\|\|u_h^{(2)} - u\|\|_{L^\infty}$	Ord. $K=2$	$\|\|u_h^{(3)} - u\|\|_{L^\infty}$	Ord. $K=3$
10	7.14×10^{-3}	-	1.36×10^{-3}	-	7.38×10^{-4}	-
20	1.22×10^{-3}	2.70	1.51×10^{-4}	3.17	1.44×10^{-5}	5.68
40	2.23×10^{-4}	2.45	2.38×10^{-5}	4.26	5.24×10^{-7}	4.78
80	9.99×10^{-5}	0.90	8.01×10^{-6}	1.57	7.06×10^{-8}	2.89
160	3.71×10^{-5}	1.43	1.37×10^{-6}	2.55	5.80×10^{-8}	0.28
320	8.17×10^{-6}	2.18	3.48×10^{-7}	1.98	2.74×10^{-8}	1.08
Time	$9.76s$	-	$12.83s$	-	$111.34s$	-

Table 3.5: DG solutions Burgers' sine initial condition $T = T_b + 0.1$ G flux.

3.5 Local discontinuous Galerkin method

Recall the convection-diffusion equation

$$u_t + (f(u))_x = Du_{xx}.$$

The DG method cannot be applied directly to it. In [75] the authors show that if this is done, the obtained solution is stable but inconsistent.

The following construction of the LDG method can be found in [16]. First transform the second order PDE into the system of first order PDEs,

$$\begin{aligned} u_t + (f(u) - \sqrt{D}q)_x &= 0, \\ q - g(u)_x &= 0, \end{aligned}$$

with $g(x) = \sqrt{D}u$. To define the LDG solution to the system we use the second version of weak formulation (2.61)-(2.63).

Define the flux $h = (h_u, h_q)^t$ by

$$h(u, q) = (f(u) - \sqrt{D}q, -g(u))^t.$$

We seek for an approximate solution $w_h = (u_h, q_h)^t$, with $u_h, q_h \in V_h^k$ of $w = (u, q)$. To construct it, we need to find $w_h = (u_h, q_h)^t$ such that for any test functions $\phi_1, \phi_2, \phi_3 \in V_h^K$ the weak formulation

N	$\|\|u_h^{(1)} - u\|\|_{L^\infty}$	Ord. $K=1$	$\|\|u_h^{(2)} - u\|\|_{L^\infty}$	Ord. $K=2$	$\|\|u_h^{(3)} - u\|\|_{L^\infty}$	Ord. $K=3$
10	6.86×10^{-2}	-	3.46×10^{-2}	-	2.72×10^{-2}	-
20	3.72×10^{-2}	0.89	1.68×10^{-2}	1.03	1.50×10^{-2}	0.85
40	2.05×10^{-2}	0.85	8.91×10^{-3}	0.92	6.38×10^{-3}	1.23
80	1.17×10^{-2}	0.81	4.92×10^{-3}	0.85	4.12×10^{-3}	0.62
160	5.52×10^{-3}	1.08	2.74×10^{-3}	0.84	2.16×10^{-3}	0.93
320	4.76×10^{-3}	0.21	1.47×10^{-3}	0.89	1.23×10^{-3}	0.80
Time	$9.76s$	-	$12.83s$	-	$111.34s$	-

Table 3.6: DG solutions Burgers' impulse initial condition $T = \frac{1}{2\pi} + 0.1$.

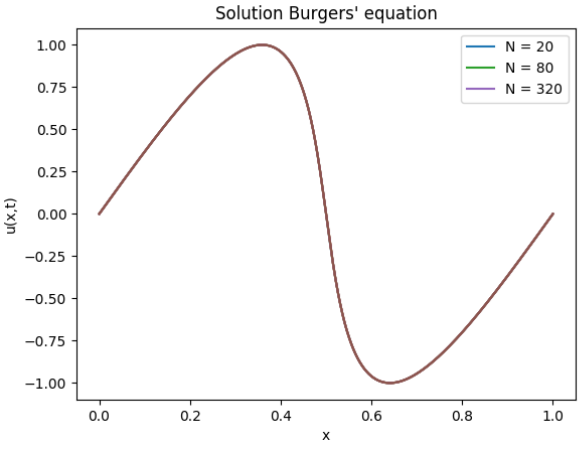


Figure 3.4: Comparison of DG solutions for the Burgers' $K = 3$, $T = T_b - 0.05$ sine initial condition.

(2.61)-(2.63)

$$\int_{I_j} (u_h)_t \phi_1 dx - \int_{I_j} h_u(w_h)(\phi_1)_x dx = \hat{h}_u(w_h(x))\phi_1(x)\big|_{x=x_j}^{x=x_{j+1}}, \tag{3.13}$$

$$\int_{I_j} q_h \phi_2 dx - \int_{I_j} h_q(w_h)(\phi_2)_x dx = \hat{h}_q(w_h(x))\phi_2(x)\big|_{x=x_j}^{x=x_{j+1}}, \tag{3.14}$$

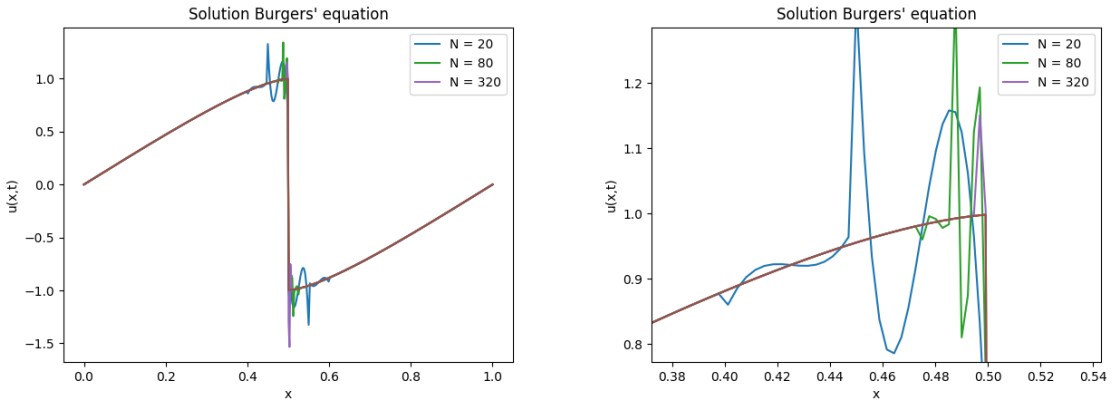
$$\int_{I_j} u_h(x,0)\phi_3 dx = \int_{I_j} u_0(x)\phi_3 dx, \tag{3.15}$$

is satisfied. Once again, neither u_h nor q_h are required to be continuous in the inter-elements x_j so fluxes h_u and h_q are replaced by numerical fluxes, as previously done for the DG method. To define the numerical flux $\hat{h}$ is useful to consider the following notation

$$[u] = u^+ - u^-, \quad \overline{u} = \frac{1}{2}(u^+ + u^-), \quad u_j^\pm = u_j^\pm = u(x_j^\pm).$$

The numerical flux is written as the sum of a diffusive flux and a convective flux:

$$\hat{h}(w^-, w^+) = \hat{h}_{diff}(w^-, w^+) + \hat{h}_{conv}(w^-, w^+).$$



(a) General view

(b) Zoom on oscillatory behaviour.

Figure 3.5: Comparison of DG solutions for the Burgers' $K = 3$, $T = T_b + 0.1$ sine initial condition.

The diffusive flux is defined by

$$\hat{h}_{diff}(w^-, w^+) = \left(-\frac{[g(u)]}{[u]}\overline{q}, -\overline{g(u)}\right)^t - C_{diff} \times [w],$$

with

$$C_{diff} = \begin{bmatrix} 0 & c_{12} \\ -c_{12} & 0 \end{bmatrix}$$

where $c_{12} = c_{12}(u^-, u^+)$ is locally Lipschitz. The convective flux is given by

$$\hat{h}_{conv}(w^-, w^+) = (\hat{f}(u^-, u^+), 0)^t,$$

where $\hat{f}$ is any locally Lispchitz function consistent with the nonlinearity of f.

Note that once u_h has been determined at a time step, equation (3.14) determines q_h by using the flux

$$\hat{h}_q = -\overline{g(u)} - c_{12}[u],$$

without requiring the solution of any ODE.

Once the convective and diffusive fluxes are defined, we can rewrite the numerical flux $\hat{h}$ as follows

$$\hat{h}(w^-, w^+) = \left(\frac{[\psi(u)]}{[u]} - \frac{[g(u)]}{[u]}\overline{q}, -\overline{g(u)}\right)^t - C \times [w], \tag{3.16}$$

where

$$C = \begin{bmatrix} c_{11} & c_{12} \\ -c_{12} & 0 \end{bmatrix}, \quad c_{11} = \frac{1}{[u]}\left(\frac{[\psi(u)]}{[u]} - \hat{f}(u^-, u^+)\right)^t. \tag{3.17}$$

Substituting (3.17) into (3.16)

$$\hat{h}(w^-, w^+) = (\hat{f}(u^-, u^+) - \frac{[g(u)]}{[u]}\overline{q} + c_{12}[q], -\overline{g} - c_{12}[u]).$$

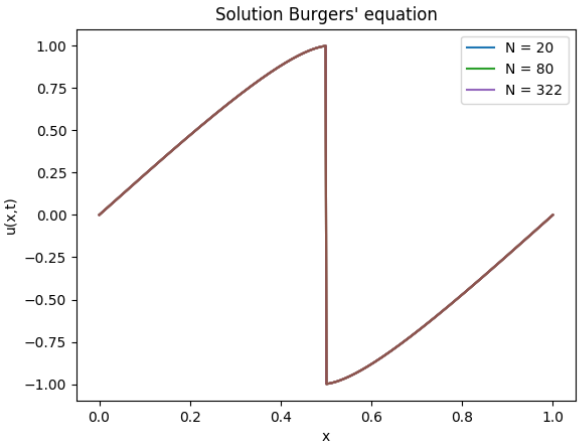


Figure 3.6: Comparison of DG solutions for the Burgers' $K = 3$, $T = T_b + 0.1$ sine initial condition G flux.

Considering c_{12} as the standard three point central difference

$$c_{12} = -\frac{\sqrt{D}}{2},$$

and recalling $g(u) = \int^u \sqrt{D} ds = \sqrt{D}u$, it is concluded that

$$\hat{h}(w^-, w^+) = (\hat{h}_u, \hat{h}_q) = (\hat{f}(u^+, u^-) - \sqrt{D}q^-, -\sqrt{a}u^+)^t. \tag{3.18}$$

Note that (3.18) does not depend on f, we can use it either with $f(u) = \alpha u$ or $f(u) = \dfrac{u^2}{2}$.

Once again considering the polynomial basis generated by the Legendre polynomial P_l the approximations of u and q at each I_j, for $j = 0, \ldots, N-1$ are

$$u_h(x,t)\big|_{I_j} = \sum_{l=0}^{K} u_j^l(t) P_l(\xi_j(x)) \quad \text{and} \quad q_h(x,t)\big|_{I_j} = \sum_{l=0}^{K} q_j^l(t) P_l(\xi_j(x)). \tag{3.19}$$

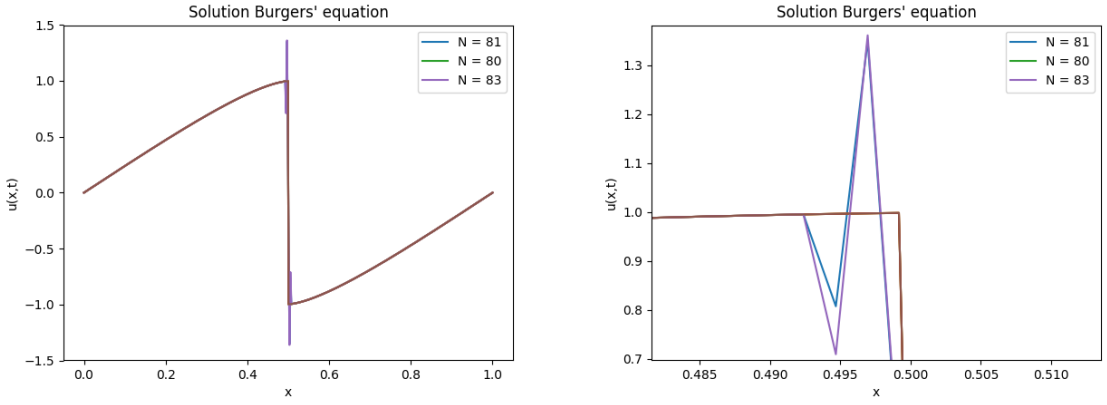
Thus, substituting in (3.13) and (3.14) and using Legendre polynomials as test functions the system of ODEs for the coefficients u_j^l is

$$\frac{d}{dt}u_j^l = \frac{2l+1}{h}\left[\int_{-1}^{1} (f(u_h(\xi,t)) - \sqrt{D}q_h(\xi,t))P_l'(\xi)d\xi + (-1)^l(\hat{f}_j - \sqrt{D}q_j^-) - (\hat{f}_{j+1} - \sqrt{D}q_{j+1}^-)\right], \tag{3.20}$$

and q_h is given by

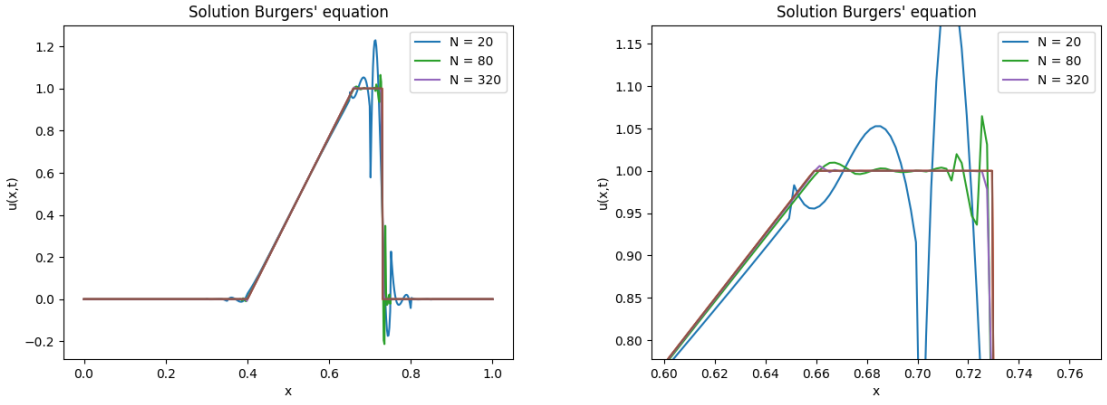
$$q_j^l = \frac{2l+1}{h}\left[\int_{-1}^{1} \sqrt{D}u_h(\xi,t)P_l'(\xi) + \sqrt{D}u_{j+1}^+ - (-1)^{l+1}u_j^+\right], \tag{3.21}$$

see Algorithm 5 for the implementation. We can now use the TVDRK method to solve (3.20). If a polynomial order K is used for the approximation, recall that it is required that the order of the RK method is $K+1$. If so, the order of the method is in the L^∞ norm is $O(h^k)$ [16].



(a) General view (b) Zoom on oscillatory behaviour.

Figure 3.7: Dependence on parity of N of the G flux.



(a) General view (b) Zoom on oscillatory behaviour.

Figure 3.8: Comparison of DG solutions for the Burgers' $K = 3$, $T = \frac{1}{2\pi} + 0.1$ impulse initial condition.

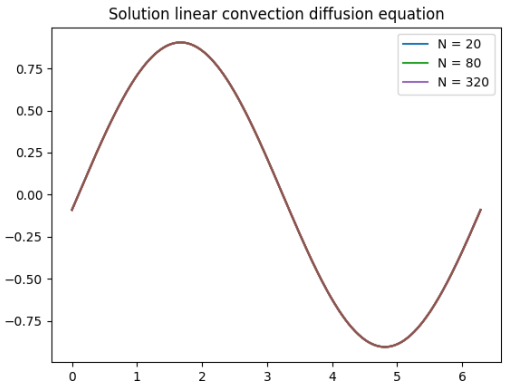
3.5.1 Examples

In the following examples we used $CFL = 0.001$ for $K = 1$, $CFL = 0.0001$ for $K = 2$, and $CFL = 0.00001$ for $K = 3$.

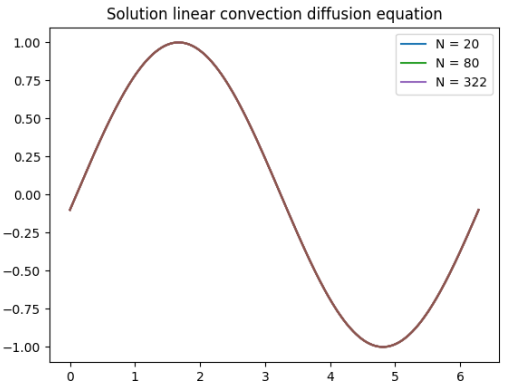
Example 1. Linear convection-diffusion equation

$$u_t + \alpha u_x = D u_{xx}.$$

In Tables 3.7 and 3.8 the L^∞ norm of the error and the order of the solutions for the polynomial orders $K = 1, 2, 3$ are shown. With the choice of CFLs for $D = 1$ the order in general was $O(h^{k+1})$, except for $N = 80, 160, 320$ where the error did not decreased as expected. In the case $D = 10^{-5}$ the order was not consistently $O(h^{k+1})$, but it was at least $O(h^k)$ for most of the cases. In the Figure 3.9a and Figure 3.9b we can see a comparison of different LDG solutions. When both figures are compared we can see that the decrement on the diffusion coefficient affects the amplitude of the sine wave. For $D = 1$ the amplitude decreases faster than for $D = 10^{-5}$.



(a) $D = 1$.



(b) $D = 10^{-5}$.

Figure 3.9: Comparison of LDG solutions for the linear convection-diffusion equation $K = 3$, $T = 0.1$.

N	$\|\|u_h^{(1)} - u\|\|_{L^\infty}$	Ord. $K = 1$	$\|\|u_h^{(2)} - u\|\|_{L^\infty}$	Ord. $K = 2$	$\|\|u_h^{(3)} - u\|\|_{L^\infty}$	Ord. $K = 3$
10	1.03×10^{-2}	-	2.88×10^{-4}	-	2.13×10^{-5}	-
20	2.67×10^{-3}	1.93	5.42×10^{-5}	2.88	1.38×10^{-6}	3.99
40	6.93×10^{-4}	1.95	7.32×10^{-5}	2.41	1.9×10^{-7}	2.81
80	1.89×10^{-4}	1.86	1.37×10^{-6}	2.77	2.94×10^{-8}	2.69
160	4.99×10^{-5}	1.92	2.01×10^{-7}	3.17	2.91×10^{-8}	0.01
320	1.11×10^{-5}	2.17	6.38×10^{-8}	1.79	1.32×10^{-8}	1.13
Time	$8.01s$	-	$10.83s$	-	$73.77s$	-

Table 3.7: LDG solutions for the linear convection–diffusion equation $\alpha = 1$, $D = 1$, $T = 0.1$.

Example 2. Nonlinear convection-diffusion equation

$$u_t + \left(\frac{u^2}{2}\right)_x = Du_{xx}.$$

In Tables 3.9 and 3.10 the L^∞ norm of the error and the order of the solutions for the polynomial orders $K = 1, 2, 3$ are shown. In this example for both $D = 1$ and $D = 10^{-5}$ the numerical orders are in general $O(h^{K+1})$. Note that the error for $D = 10^{-5}$ is much lower than the one for $D = 1$. The reason for this can be seen in the Figures 3.10a and 3.10b and in the equation 2.53, the solution is bounded at all times by $2D$.

3.5.2 Convective-dominated problems

In the two previous examples we have focused on what happens when D is small, i.e., the convective dominated case. The reason of the interest is that convective dominated problems are problematic for numerical solvers. As part of the thesis work, we programmed the Petrov-Galerkin method (PG) [31], the continuous Galerkin method (CG) [35] and the interior penalty method (IP) with first order polynomial approximations [62]. These methods as well transform the PDE into the solution of a system of ODEs. We used the backward Euler (BE) and Crank-Nicolson (CN) methods for the solution of the ODE systems. The IP methods depend on three parameters $(\sigma_0, \sigma_1, \epsilon)$ for weakly imposing continuity.

N	$\|\|u_h^{(1)} - u\|\|_{L^\infty}$	Ord. $K=1$	$\|\|u_h^{(2)} - u\|\|_{L^\infty}$	Ord. $K=2$	$\|\|u_h^{(3)} - u\|\|_{L^\infty}$	Ord. $K=3$
10	4.98×10^{-2}	-	6.14×10^{-4}	-	1.36×10^{-4}	-
20	1.45×10^{-2}	1.78	5.09×10^{-4}	2.08	7.26×10^{-6}	4.23
40	4.30×10^{-3}	1.75	8.67×10^{-5}	2.55	9.08×10^{-7}	3.00
80	1.04×10^{-3}	2.05	1.60×10^{-5}	2.43	4.61×10^{-8}	4.30
160	2.52×10^{-4}	2.05	1.82×10^{-6}	3.14	3.59×10^{-8}	0.36
320	5.88×10^{-5}	2.10	3.33×10^{-7}	2.45	1.62×10^{-8}	1.14
Time	$8.01s$	-	$10.83s$	-	$73.77s$	-

Table 3.8: LDG solutions for the linear convection-diffusion equation $\alpha = 1$, $D = 10^{-5}$, $T = 0.1$.

N	$\|\|u_h^{(1)} - u\|\|_{L^\infty}$	Ord. $K=1$	$\|\|u_h^{(2)} - u\|\|_{L^\infty}$	Ord. $K=2$	$\|\|u_h^{(3)} - u\|\|_{L^\infty}$	Ord. $K=3$
10	1.19×10^{-1}	-	2.08×10^{-2}	-	1.59×10^{-3}	-
20	3.05×10^{-2}	1.96	2.47×10^{-3}	3.07	1.25×10^{-4}	3.67
40	7.68×10^{-3}	1.99	3.00×10^{-4}	3.04	6.37×10^{-6}	4.29
80	1.78×10^{-3}	2.11	4.39×10^{-5}	2.77	5.25×10^{-7}	3.60
160	4.66×10^{-4}	1.93	5.54×10^{-6}	2.99	7.12×10^{-8}	2.88
320	1.18×10^{-4}	1.98	1.28×10^{-6}	2.11	2.32×10^{-8}	1.62
Time	$8.01s$	-	$10.83s$	-	$113.68ss$	-

Table 3.9: LDG solutions for the non linear convection-diffusion equation $D = 1$, $T = 0.1$.

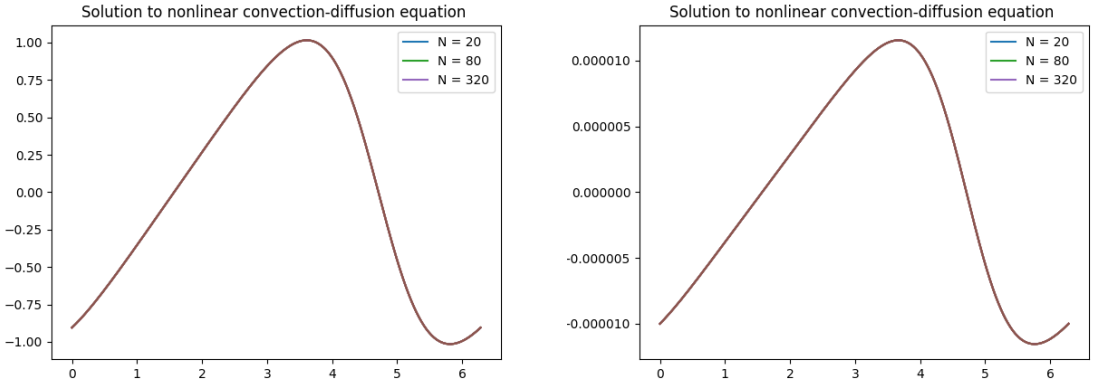
This feature makes the use of the IP methods complicated. There are values for which the method is well characterized and there are result about their stability and convergence, for example: $(1, 0, 1)$ nonsymmetric interior penalty Galerkin (NPIG), $(0, 0, 1)$ NPIG0, $\epsilon = 0$ incomplete interior penalty Galerkin (IIPG) and $(\sigma_0, 0, -1)$ with σ_0 bounded below by a large enough constant symmetric interior penalty Galerkin (SIPG) [62]. For other values of $(\sigma_0, \sigma_1, \epsilon)$ at least to the best knowledge of the author, there are no general results that ensure convergence.

We compare in Table 3.11 the CG, PG, NPIG0 and LDG methods using order 1 polynomials for $\alpha = D = 1$. The LDG has the best performance. For the convective-dominate case Table 3.12 the PG method did not converge, the other methods did converge but the LDG still performed better. We had to use the parameters $(\sigma_0, \sigma_1, \epsilon) = (0, 0, 10^{-4})$ for the IP method to converge in the convective dominated case, but we do not know any results of its stability or order of convergence. Finding the parameters so that the IP method converges is not an easy task and requires a lot of numerical experimentation. Thus, we can conclude that among different formulations of the Galerkin method, the best suited for convection-diffusion equations is the LDG method. A similar type of problems are the degenerated problems where the parameter of the diffusion is not constant and becomes zero in some regions of the domain. The IP methods do work for these kinds of problems [58]. We tried to adapt the ideas developed in [58] to convective-dominated problems, but the method requires a nonempty set where the diffusion vanishes which is not the case. Since the LDG outperformed the other methods, we did not developed them for further examples and we do not include their construction. The interested reader is referred to the references [31], [35] and [58] for more information.

The difficulties for solving convective-dominated equations do not only appear when using Galerkin methods in [23], the authors compare the LDG method with the numerical solutions obtained by fully the implicit finite difference method (IFDM) and a mixed finite difference and boundary element

N	$\|\|u_h^{(1)} - u\|\|_{L^\infty}$	Ord. $K=1$	$\|\|u_h^{(2)} - u\|\|_{L^\infty}$	Ord. $K=2$	$\|\|u_h^{(3)} - u\|\|_{L^\infty}$	Ord. $K=3$
10	7.79×10^{-7}	-	1.46×10^{-7}	-	1.05×10^{-8}	-
20	1.99×10^{-7}	1.97	1.61×10^{-8}	3.18	8.74×10^{-10}	3.58
40	4.89×10^{-8}	2.03	1.94×10^{-9}	3.05	5.03×10^{-11}	4.12
80	1.20×10^{-8}	2.03	2.94×10^{-10}	2.73	3.66×10^{-12}	3.78
160	3.17×10^{-9}	1.92	3.70×10^{-11}	2.99	2.15×10^{-13}	4.09
320	7.62×10^{-10}	2.05	4.29×10^{-12}	3.11	2.03×10^{-14}	3.40
Time	$8.01s$	-	$10.83s$	-	$113.68ss$	-

Table 3.10: LDG solutions for the non linear convection–diffusion equation $D = 10^{-5}$, $T = 0.1$.



(a) $D = 1$. (b) $D = 10^{-5}$.

Figure 3.10: Comparison of LDG solutions for the nolinear convection-diffusion equation $K = 3$, $T = 0.1$.

method (BEM). They numerically show that the most efficient method for the nonlinear convection-diffusion equations with the diffusion coefficients $D = 1, 0.1, 0.01, 0.001$ and a sine initial condition is the LDG method, see Figure 3.11.

N	CG		NPIG 0		PG		LDG $K=1$
-	BE	CN	BE	CN	BE	CN	RK
100	9.09×10^{-2}	8.95×10^{-2}	5.56×10^{-2}	5.7×10^{-2}	7.56×10^{-2}	7.4×10^{-2}	4.68×10^{-4}
200	2.53×10^{-2}	2.49×10^{-2}	2.22×10^{-2}	2.3×10^{-2}	1.58×10^{-2}	1.51×10^{-2}	9.00×10^{-5}
300	1.31×10^{-2}	1.35×10^{-2}	1.32×10^{-2}	1.41×10^{-2}	5.34×10^{-3}	5.42×10^{-3}	5.66×10^{-5}

Table 3.11: Comparison norm L_1 of error for different methods, linear convection-diffusion $D = 1$

3.6 Benefits of the DG and LDG methods

The DG and LDG methods outperform other numerical schemes in the solution of conservation laws and convection-diffusion equations, which is not the only advantage that they have. They are well suited for complex geometries, they have high-order accuracy, *hp*-adaptivity is easily implemented, they have

N	CG		$(\sigma_0, \sigma_1, \epsilon) = (0, 0, 10^{-4})$		LDG $K = 1$
-	BE	CN	BE	CN	RK
100	7.73×10^{-1}	7.65×10^{-1}	3.96×10^{-2}	4.02×10^{-2}	6.07×10^{-4}
200	1.55×10^{-2}	1.51×10^{-2}	1.95×10^{-2}	1.93×10^{-2}	2.37×10^{-4}
300	5.78×10^{-3}	5.72×10^{-3}	1.42×10^{-2}	1.36×10^{-2}	1.21×10^{-4}

Table 3.12: Comparison norm L_1 of error for different methods, linear convection-diffusion $D = 10^{-2}$

Method	t	$x = 0.1$	$x = 0.3$	$x = 0.5$	$x = 0.7$	$x = 0.9$
IFDM [5]	0.5	0.12182	0.36206	0.59079	0.79416	0.93322
BEM [2]		0.12079	0.36113	0.59559	0.81257	0.97184
LDG ($k = 1$)		0.12110	0.36016	0.58861	0.79363	0.93867
LDG ($k = 2$)		0.12113	0.36022	0.58880	0.79323	0.93767
Exact		0.12114	0.36027	0.58870	0.79349	0.93811
IFDM [5]	2	0.04367	0.13095	0.21800	0.30466	0.38024
BEM [2]		0.04300	0.12877	0.21468	0.30075	0.37452
LDG ($k = 1$)		0.04296	0.12883	0.21453	0.29997	0.37325
LDG ($k = 2$)		0.04296	0.12883	0.21455	0.29998	0.37325
Exact		0.04296	0.12884	0.21456	0.30000	0.37328
IFDM [5]	4	0.02364	0.07092	0.11817	0.16499	0.17226
BEM [2]		0.02324	0.06935	0.11550	0.16125	0.16515
LDG ($k = 1$)		0.02310	0.06930	0.11549	0.16120	0.16604
LDG ($k = 2$)		0.02310	0.06931	0.11549	0.16121	0.16605
Exact		0.02310	0.06931	0.11549	0.16121	0.16606

Figure 3.11: Comparison of methods for the nonlinear convection diffusion equation $D = 10^{-2}$, extracted from [23]

an explicit semidiscrete form and they are parallelizable [25]. Note that they are parallelizable because the computations in each interval I_j are independent to the computations in other intervals I_k. We used the package NUMBA [42] from python to parallelize the codes of this thesis. In each table shown we include the time taken to obtain the sequential solution for $N = 10, 20, 40, 80, 160, 320$ on a Lenovo Legion Y720 CORI i7 7th Gen with NVIDIA GEFORCE GTX 1060.

3.7 Algorithms

The general implementation of the DG and LDG algorithms is Algorithm 2. To initialize it we obtain an approximation u_h^0 of the initial condition using Equation 3.8, select the order of polynomial approximation K, the number of elements N, the CFL condition, the flux f, the derivative of the flux f', the numerical flux $\hat{f}$ and the left side of the Runge-Kutta method L_h (use Algorithm 4 for DG and Algorithm 5 for the LDG method).

Algorithm 1 TVD Runge-Kutta

procedure RUNGE_KUTTA(K, u_0, τ, T, $f(u)$, $f'(u)$, $\hat{f}$, N, h, L_h)
  if The polynomial order K is 1 **then**
    $w^1 = u_0 + \tau L_h(u_0, f(u), f'(u), \hat{f}, N, h)$
    $u_1 = \frac{1}{2}u_0 + \frac{1}{2}(w^1 + \tau L_h(w^1, f(u), f'(u), \hat{f}, N, h))$
  if The polynomial order K is 2 **then**
    $w^1 = u_0 + \tau L_h(u_0, f(u), f'(u), \hat{f}, N, h)$
    $w^2 = \frac{3}{4}u_0 + \frac{1}{4}(w^1 + \tau L_h(w^1), f(u), f'(u), \hat{f}, N, h)$
    $u_1 = \frac{1}{3}u_0 + \frac{2}{3}(w^2 + \tau L_h(w^2), f(u), f'(u), \hat{f}, N, h)$
  if The polynomial order K is 3 **then**
    $L_h^0 = L_h(u_0, f(u), f'(u), \hat{f}, N, h)$
    $w^1 = u_0 + \tau L_h^0$
    $L_h^1 = L_h(w^1, f(u), f'(u), \hat{f}, N, h)$
    $w^2 = \frac{1}{2}(u_0 + w^1) + \frac{\tau}{4}(-L_h^0 + 2L_h^1)$
    $L_h^2 = L_h(w^2, f(u), f'(u), \hat{f}, N, h)$
    $w^3 = \frac{1}{9}(u_0 + 2w^1 + 6w^2) + \frac{\tau}{9}(-L_h^0 - 3L_h^1 + L_h^2)$
    $L_h^3 = L_h(w^3, f(u), f'(u), \hat{f}, N, h)$
    $u_1 = \frac{1}{3}(w^1 + w^2 + w^3) + \frac{\tau}{6}(L_h^1 + L_h^3)$
return u^1

Algorithm 2 Solution of system of ODEs

procedure SOLVER_ODE(K, u_0, N, CFL T, $f(u)$, $f'(u)$, $\hat{f}$, N, h, L_h)
  $\tau = \frac{CFL}{N}$.
  $dt = \tau$
  $n = 1$
  while $dt = n\tau < T$ **do**
    Compute u_n using Algorithm 1

$$u_n = \text{ runge\_kutta}(K, u_{n-1}, \tau, T, f(u), f'(u), \hat{f}, N, h, L_h)$$

    $n+ = 1$
return Solution of the system of ODEs u_n

Algorithm 3 local Lax-Friedrichs flux

procedure LAX_FRIEDRICHS(a, b, f, f')
  Define $C = \alpha$ for $f(u) = \alpha u$ linear or $C = \max(a, b)$ for $f(u) = \frac{u^2}{2}$
  Compute $\hat{f}(a, b) = \frac{f(a)+f(b)-C(b-a)}{2}$
return Numerical flux $\hat{f}(a, b)$

Algorithm 4 DG Conservation law

procedure L_H($u_h, f(u), f'(u), \hat{f}, N, h, L_h$)
  Define the solution vector u_1.
  for the element $j = 0, \ldots, N-1$ **do**
    Compute by (3.9) $u_j^- = u_h(x_j^-, t), u_j^+ = u_h(x_j^+, t), u_{j+1}^- = u_h(x_{j+1}^-, t)$ and $u_{j+1}^+ = u_h(x_{j+1}^+, t)$.
    for each order polynomial $l = 0, \ldots, K$ **do**
      Compute the numerical flux difference

$$\hat{f}_j^{\pm} = (-1)^l \hat{f}_j^+ - \hat{f}_{j+1}^-$$

      Compute and save the entry $2j(K+1)+l$ of the solution u_1

$$u_1[2j(K+1)+l] = \frac{2l+1}{h}\left(\int_{-1}^{1} f(u_h(\xi,t))P_l'(\xi)d\xi + \hat{f}_j^{\pm}\right).$$

  return Solution $u_1 = L_h(u_0)$

Algorithm 5 LDG convection-diffusion equations.

procedure L_H($u_h, f(u), f'(u), \hat{f}, D, N, h, L_h$)
  Define the solution vectors u_1, q.
  for the element $j = 0, \ldots, N-1$ **do**
    Compute by (3.19) $u_j^+ = u_h(x_j^+, t)$ and $u_{j+1}^+ = u_h(x_{j+1}^+, t)$.
    for each order polynomial $l = 0, \ldots, K$ **do**
      Compute and save the entry $2j(K+1)+l$ of the solution q

$$q[2j(K+1)+l] = \frac{2l+1}{h}\left[\int_{-1}^{1} \sqrt{D}u_h(\xi,t)P_l'(\xi) + \sqrt{D}u_{j+1}^+ - (-1)^{l+1}u_j^+\right].$$

  for the element $j = 0, \ldots, N-1$ **do**
    Compute by (3.19) $q_j^- = q_h(x_j^-, t), u_j^- = u_h(x_j^-, t), q_{j+1}^- = q_h(x_{j+1}^-)$ and $u_{j+1}^- = u_n(x_{j+1}^-, t)$.
    for each order polynomial $l = 0, \ldots, K$ **do**
      Compute and save the entry $2j(K+1)+l$ of the solution u_1

$$\begin{aligned} u_1[2j(K+1)+l] &= \frac{2l+1}{h}\left[\int_{-1}^{1}(f(u_h(\xi,t)) - \sqrt{D}q_h(\xi,t))P_l'(\xi)d\xi\right] + \\ &+ \frac{2l+1}{h}\left[(-1)^l(\hat{f}_j - \sqrt{D}q_j^-) - (\hat{f}_{j+1} - \sqrt{D}q_{j+1}^-)\right] \end{aligned}$$

  return Solution $u_1 = L_h(u_0)$

Chapter 4

Limiting Techniques

4.1 Introduction

Solutions by the discontinuous Galerkin method present the Gibbs phenomenon and oscillations appear near discontinuities. To address this problem, "limiting techniques" can be used. Such techniques consist in a two step process; first identify “troubled cells" where the solution has a highly oscillatory behavior. Second, modify the coefficients of the DG polynomial approximation in the troubled cells. The development of troubled cells detectors and reconstruction methods is an active area of research. Some classic and modern slope limiters are: a multi-wavelet type [14], the minmod-type TVB [16], the moment limiter [40], WENO reconstruction [59], artificial neural networks as a troubled-cell indicator [60], a convolution neural network shock detector [69], and the TVB method with a multi-layer perceptron [74]. For a panoramic perspective on limiters developed up to 2013 and their performance, see [77].

In this chapter we include an introduction to limiters, we show the performance of the minmod-based TVB limiter, the moment limiter and our proposals for limiters, without any formal results.

4.2 Two step process: identify and reconstruct

As we have mentioned earlier, the identifying and reconstruction processes are still an active area of research. The material in this section explains the motivation and ideas of the slope limiter [16] and moment limiter [40].

The limiting process was motivated from the slope limiter methods for finite difference methods for conservation laws [46]. Just as the solution of the DG method, finite difference schemes have oscillations near discontinuities and this method was created to obtain nonoscillatory discontinuities for high order schemes. Finite difference methods compute the solution to PDEs at certain predetermined space-time grid points (x_j, t_n). A natural interpretation of the numerical solution is that it approximates the exact pointwise values of the true solution. A second interpretation is that it approximates the average value

$\overline{u}_j^n$ of the true solution at a time t_n in an interval of width h that has x_j as its midpoint, that is,

$$\overline{u}_j^n = \frac{1}{h}\int_{I_j} u(x,t_n)dx,$$

with $I_j = [x_j - \frac{h}{2}, x_j + \frac{h}{2}]$. The value $\overline{u}_j^n$ is called the *cell average* of $u(x,t_n)$. Note that this idea is reasonable by the integral form of the conservation law.

Let $u_h(x,t)\big|_{I_j} = \sum_{l=0}^{l} u_j^l(t)P_l(\xi_j(x))$ be the DG approximation for a conservation law and consider as in the previous section $I_j = [x_j, x_{j+1}]$, $j = 0, \ldots, N-1$ to be a partition of the domain $\mathcal{O}$. We are looking for a way to determine when $u_h(x,t)$ has a highly oscillatory behaviour, as the ones shown in Figure 3.2b, Figure 3.5b and Figure 3.8b. Clearly at the intervals I_j where a function oscillates, there is an abrupt change in the spatial derivatives. The idea of the identifying process of the slope and moment limiters is to approximate the spatial derivatives using two different methods and compare the different approximations to determine abrupt changes in the sign and magnitudes of the derivatives.

We first explain the slope limiter, since it is more intuitive. At each interval I_j the approximate solution u_h is truncated as a first order polynomial approximation $\hat{u}_h(x,t) = u_j^0(t) + u_j^1(t)P_1(\xi_j(x))$. Consider the cell averages $\overline{u}_{j-1}, \overline{u}_j$ and $\overline{u}_{j+1}$, we dropped the time index since for this analysis is not relevant. The first spatial derivative approximation of $u(x_j,t)$, uses the neighboring elements cell averages solutions to compute the *external* forward and backward differences

$$\Delta_- u_j = \overline{u}_j - \overline{u}_{j-1} \quad \text{and} \quad \Delta_+ u_j = \overline{u}_{j+1} - \overline{u}_j.$$

The second approximation, uses the values $u_j^+ = u_h(x_j^+,t)$ and $u_{j+1}^- = u_h(x_{j+1}^-,t)$ to compute the *internal* forward and backward approximations

$$\hat{u}_j = \overline{u}_j - u_j^+ \quad \text{and} \quad \check{u}_j = u_{j+1}^- - \overline{u}_j.$$

The terms "external" and "internal" forward and backward differences are introduced in this thesis to help the explanation, they are not a formal definition from any of the references. What do theses quantities actually mean? We use their sign and magnitude to understand them. The external backward difference tells us if in average the values of u_h increased or decreased (its sign) and by how much (its magnitude) in I_j with respect to the average of the values of u_h in I_{j-1}. The external forward difference has an analogous interpretation with the intervals I_j and I_{j+1}. The internal backward difference tells us if in average the function u_h increases or decreases (its sign) and by how much (its magnitude) in I_j with respect to the value of u_h in x_j. The internal forward difference has an analogous interpretation. Thus, these quantities express if u_h is increasing or decreasing in I_j with respect to the solution in the neighboring intervals and with respect to its own values in the boundaries of I_j.

What happens when there is an oscillation in an interval I_j? How can oscillations be characterized? Looking at Figure 3.2b, Figure 3.5b and Figure 3.8b we see that oscillations have inflection points such as minimum and maximum points and the functions rapidly increase and decrease in small regions. Thus we can expect from an oscillatory function that in average there is a change of sign in the derivative or that the solution increases or decreases in average faster in I_j as to the average increase or decrease of the function with respect to either I_{j-1} or I_{j+1}. To determine a change in the sign derivative we can verify if $sign(\hat{u}_j) = sign(\Delta_- u_j) = sign(\Delta_+ u_j)$ and if $sign(\check{u}_j) = sign(\Delta_- u_j) = sign(\Delta_+ u_j)$. The first equality verifies a change of sign with respect to the behaviour of u_h in the first half of I_j, the second equality has an analogous interpretation. The solution does not grow faster in I_j if

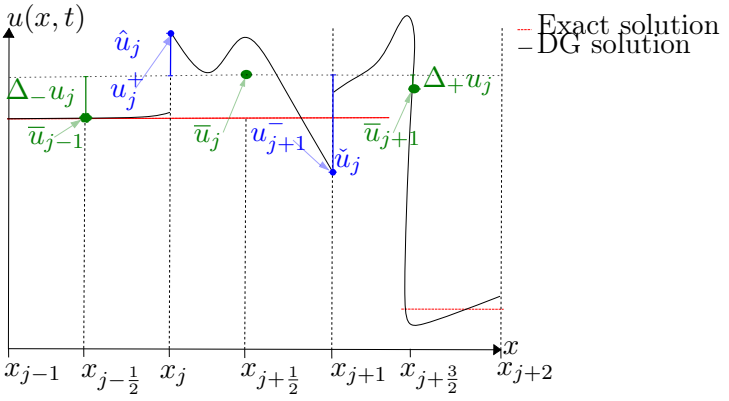


Figure 4.1: Example of oscillatory solution.

$|\hat{u}_j| = \min(|\hat{u}_j|, |\Delta_- u_j|, |\Delta_+ u_j|)$ and if $|\check{u}_j| = \min(|\check{u}_j|, |\Delta_- u_j|, |\Delta_+ u_j|)$. See Figure 4.1 for an example where an oscillation makes the previous conditions fail. On x_j there is an abrupt change on the sign of the derivative and on x_{j+1} the magnitude of $\check{u}_j$ is greater.

Recall that we are considering an DG approximation truncated to the first polynomial order. Thus, integrating

$$\overline{u}_j^n = \frac{1}{h}\int_{I_j} \hat{u}_h(x, t_j)dx = u_j^0.$$

Then

$$\hat{u}_j = \overline{u}_j - u_j^+ = u_j^0 - (u_j^0 - u_j^1) = u_j^1 = u_j^0 + u_j^1 - u_j^0 = u_{j+1}^- - \overline{u}_j = \check{u}_j.$$

Thus, the internal forward and backward approximations are equal $\hat{u}_j = u_j^1 = \check{u}_j$. From the previous discussion we define an interval to be a troubled cell if the following conditions do not hold simultaneously

i) $sign(u_j^1) = sign(\Delta_+ u_j) = sign(\Delta_- u_j)$

ii) $u_j^1 = \min(u_j^1, \Delta_+ u_j, \Delta_- u_j)$.

These conditions can be summarized using the minmod function m and defining a troubled cell with the minmod limiter (or slope limiter) as any interval I_j such that $u_0^1 \neq m(u_0^1, \Delta_+ u_j, \Delta_- u_j)$, with

$$m(a_1, a_2, \ldots, a_j) = \begin{cases} s \cdot \min_{1\leq j\leq n} |a_j|, & \text{if } sign(a_1) = sign(a_2) = \ldots = sign(a_j) = s, \\ 0, & \text{otherwise.} \end{cases} \tag{4.1}$$

Great disadvantages the minmod limiter has are that it can identify an interval where the solution actually has an extrema, it truncates the approximate solution to a first polynomial order in the intervals where oscillations are detected, and it does not take into account the information given by the higher order spatial derivatives. The truncation to the first order is required for the solution with this limiter to have total variation diminishing property, which only first order methods have, and thus ensure the stability of the limited solution. To over come the detection in regions with extrema, we introduce the moment limiter [40]. Note that with the first order truncation $\overline{u}_j = u_j^1$ for $j = 0, \ldots, N-1$. Thus,

$$\Delta_- u_j = u_j^0 - u_{j-1}^0 \quad \text{and} \quad \Delta_+ u_j = u_{j+1}^0 - u_j^0.$$

Which shows that the coefficients of the order zero polynomial can be used to compute the external forward and backward first differences. In [40] the author shows through Taylor expansions and backward and froward differences of higher order that the coefficient u_j^l is an internal approximation of the l-th derivative of $u_h(x_j,t)$, the $l-$th external forward difference is

$$\Delta_+^l = \frac{u_{j+1}^{l-1} - u_j^{l-1}}{2l-1},$$

and the $l-$th external backward difference is

$$\Delta_-^l = \frac{u_j^{l-1} - u_{j-1}^{l-1}}{2l-1}.$$

The moment limiter identifies the interval I_j as a troubled cell if for $l = K$ the equality

$$u_j^l = m\left(u_j^l, \Delta_+^l, \Delta_-^l\right), \tag{4.2}$$

does not hold.

For the reconstruction process we explain two approaches. The first one is known as the MUSCL reconstruction procedure [25], where for a troubled cell I_j we define

$$u_h\big|_{I_j}(x,t) = u_j^0 + m(u_j^1, u_{j+1}^0 - u_j^0, u_j^0 - u_{j-1}^0).$$

This procedure can be used with either of the slope and moment identifiers, although it is the natural reconstruction for the slope limiter. Its disadvantage, is that in the identified cells, the approximation decreases its order of approximation from K to 1.

The second reconstruction process is the counter part of the moment limiter. Thus, we name it the moment reconstruction process. For each troubled cell I_j apply sequentially from $l = K$ up to $l = 1$ the following verification process. Define

$$\tilde{u}_j^l = m\left(u_j^l, \Delta_+^l, \Delta_-^l\right),$$

if $u_j^l \neq \tilde{u}_j^l$ redefine u_j^l by the assignation $u_j^l = \tilde{u}_j^l$. If $u_j^l = \tilde{u}_j^l$ stop the process and continue to the next troubled cell. The author of the article mentions that this limiter can have an additional parameter α_l, with the range

$$\frac{1}{2(2l-1)} \leq \alpha_l \leq 1.$$

The modified limiter reads

$$\tilde{u}_j^l = m\left(u_j^l, \alpha_l(u_{j+1}^{l-1} - u_j^{l-1}), \alpha_l(u_j^{l-1} - u_{j-1}^{l-1})\right).$$

The reason to introduce such modification is because with the strict forward and backward differences the limiter introduces diffusion, the parameter helps to control the diffusivity, $\alpha_l = 1$ is the least diffusive parameter. For the numerical experiments we use $\alpha_l = 1$.

Note that this reconstruction process does not diminish the order of the approximation and it is a natural generalization of MUSCL, if $K = 1$ they coincide.

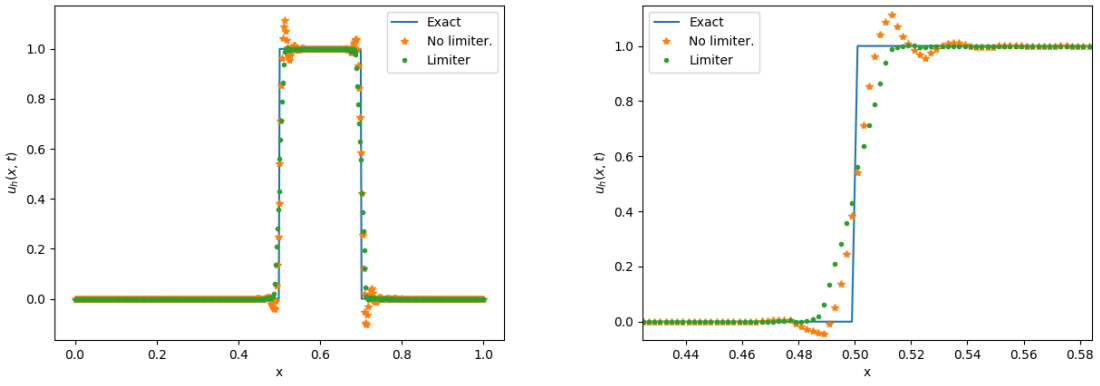
The references we gave can be divided in the following categories

Indicators	Reconstruction	Both
[60] and [69]	[14] and [59]	[16],[40] and [77]

4.3 Efficiency

We establish here the limiters that we are going to use for numerical experimentation and to propose a limiter. We define each limiter and explain the effects of their particular parameters.

Prior to make the description of the limiters it is important to make the following observation. In Figure 4.2a and Figure 4.2b we notice the DG solutions to the convection equation with the impulse initial condition with a limiter and without limiter. The error with the $||\cdot||_{L^\infty}$ norm is 0.47 for the solution with the limiter and 0.50 for the solution without the limiter. It is clear that the solution with the limiter is better, why does not the error show it? In Figure 4.2b we show a zoom on the left shock and show the points at which the solutions are evaluated. Due to the approximate solutions being continuous at each I_j, they do not have the discontinuity and a smooth change takes place. At a small neighborhood of the discontinuity at least a point of the approximated solution is compared with the exact solution and at such points due to the smooth transition, the error will be approximately the mid-value of the left and right values of the discontinuities see Figure 4.2b. Motivated by this, instead of measuring the error in the hole domain, we measure it outside of a neighborhood of the shocks. For example, in this case, assume that the shocks have the characteristics $x_1(t)$ and $x_2(t)$. The error is measured in the interval $[0, x_1(t) - \epsilon) \cup (x_1(t) + \epsilon, x_2(t) - \epsilon) \cup (x_2(t) + \epsilon, 1]$, in the examples we take $\epsilon = 0.02$. This might seem like a trick to make look the results better, but, the objective of the limiters is to control oscillations. With this error, we can in fact view the effects of the oscillations instead of a natural error caused by the discontinuity. In addition, in [16] the same idea is used to show that the method has the theoretical order of convergence away from discontinuities, they use $\epsilon = 0.1$. In [59], from the figures, it seems like the authors avoid placing points on a neighborhood of the discontinuity, avoiding this problem. In the other referenced papers this method for measuring the error is not done. For problems with shocks they explain the results with pictures, here we give quantitative results. Thus, measuring the error outside small neighborhoods of discontinuities is justified.



(a) General view.

(b) Zoom on discontinuity.

Figure 4.2: Comparison of DG solutions to the convection equation, impulse initial condition with and without limiter.

4.3.1 Standard limiters

i) **Minmod-based TVB limiter (TVB) [16].** Denote

$$u_{j+1}^{-} = u_j^0 + \tilde{u}_j, \quad u_j^{+} = u_j^0 - \tilde{\tilde{u}}_j, \tag{4.3}$$

with

$$\tilde{u}_j = \sum_{l=1}^{K} u_j^l P_l(\xi_j(x_{j+1})), \quad \tilde{\tilde{u}}_j = -\sum_{l=1}^{K} u_j^l P_l(\xi_j(x_j)). \tag{4.4}$$

Consider the modified minmod function $\hat{m}$

$$\hat{m}(a_1, a_2, \ldots, a_j, M, h) = \begin{cases} a_1, & \text{if } |a_1| \leq Mh^2, \\ m(a_1, a_2, \ldots, a_j), & \text{otherwise,} \end{cases} \tag{4.5}$$

h is the width of the cell, M is a positive constant dependent on the problem and m is the minmod function. The modified minmod function is used to define

$$\tilde{u}_j^{mod} = \tilde{m}(\tilde{u}_j, u_{j+1}^0 - u_j^0, u_j^0 - u_{j-1}^0, M, h), \tag{4.6}$$

$$\tilde{\tilde{u}}_j^{mod} = \tilde{m}(\tilde{\tilde{u}}_j, u_{j+1}^0 - u_j^0, u_l^0 - u_{j-1}^0, M, h). \tag{4.7}$$

If either $\tilde{u}_j \neq \tilde{u}_j^{mod}$ or $\tilde{\tilde{u}}_j \neq \tilde{\tilde{u}}_j^{mod}$ the cell is marked as a troubled cell. The reconstruction process is the first order approximation MUSCL. Thus, in the troubled cells the polynomial approximation decays from degree $K+1$ to order 1. This prefix TVB means that the solution has a bounded total variation. In [65] the author shows a general method to construct TVB methods form TVD methods, increasing the order of accuracy but loosing the TVD property. The importance of TVB methods is that it has been shown that the solutions converge to the unique entropy solution [17]. See Algorithm 7 for the TVB implementation. For the DG solution with the TVB limiter see Algorithm 9 selecting Algorithm 7 as the limiter procedure and defining $Z = [M]$.

ii) **Moment limiter of Biswas, Devine, and Flaherty (BDF limiter) [40]** with Krivodonova's modification. The modified coefficients of the DG expansion are defined as follows:

$$u_j^{l,(mod)} = m(u_j^l, \alpha_j(u_{j+1}^{l-1} - u_j^{l-1}), \alpha_j(u_j^{l-1} - u_{j-1}^{l-1})), \tag{4.8}$$

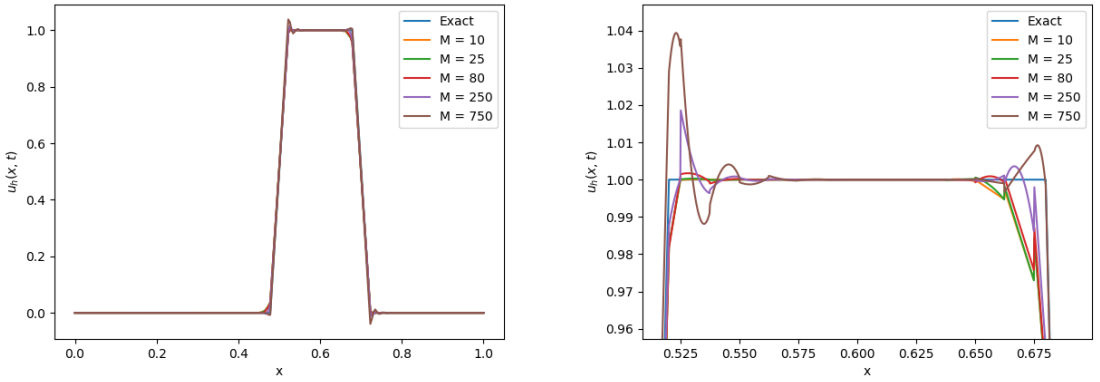
with α_j in the following range

$$\frac{1}{2(2n-1)} \leq \alpha_j \leq 1.$$

The limiter starts at the top level $l = N$, if $u_j^l = u_j^{l,(mod)}$ the cell I_j is not marked as troubled. If $u_j^l \neq u_j^{l,(mod)}$ the cell is marked as troubled. The reconstruction process substitutes u_j^l with $u_j^{l,(mod)}$ and continues to compare u_j^{l-1} with $u_j^{l-1,(mod)}$. If they are equal the process stops, if they differ, u_j^{l-1} is changed with $u_j^{l-1,(mod)}$ and the process is continued decreasing in l. The reconstruction stops when the first $u_j^l = u_j^{l,(mod)}$ or $l = 1$. In other words, for a troubled cell the reconstruction starts from the coefficient of the highest order polynomial to the coefficient of the first order polynomial. It is stopped at the first polynomial order l that satisfies $u_j^l = u_j^{l,(mod)}$. See Algorithm 8 with $M = 0$ for the implementation. For the DG solution with the BDF limiter, see Algorithm 9 selecting Algorithm 8 as the limiter procedure and defining $Z = [M] = [0]$.

4.3.2 Disadvantages of standard limiters

i) **TVB limiter.** The constant M depends on the initial condition. It introduces dissipation or oscillations near discontinuities, depending on the value of M. In Figure 4.3a and Figure 4.3b we show the solution to the linear convection equation up to time $T = 2$, $N = 80$. Small values of M diffuse the solution whereas larger values of M makes the solution oscillatory. Selecting M requires experimental fitting, but once it is found, the solution improves significantly. For scalar conservation laws with a smooth initial condition, M is proportional to the curvature of the initial condition near smooth extrema. For problems in a higher dimension or non smooth initial data such as the Riemann problems there is no rule for choosing M.

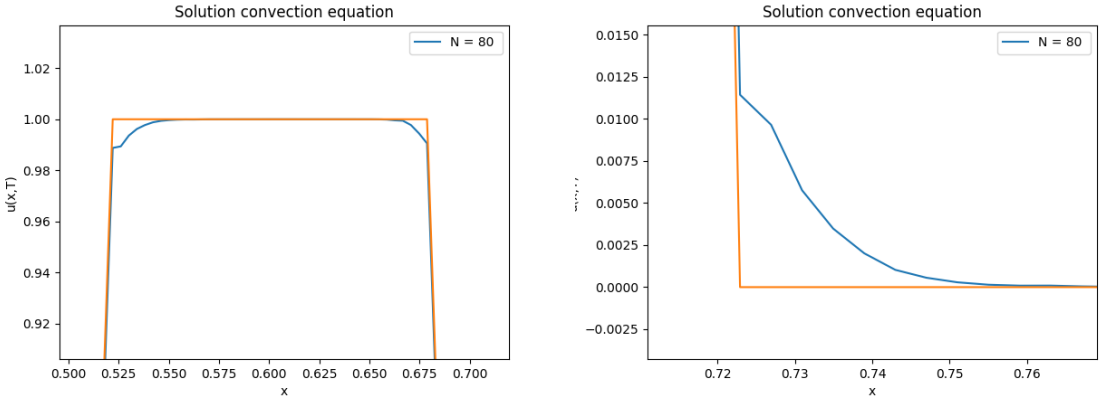


(a) General view.

(b) Zoom on diffusive behaviour.

Figure 4.3: Comparison of different values of M for the TVB limiter.

ii) **BDF limiter.** This limiter can produce diffusive solutions, see Figures 4.4a and 4.4b.



(a) Zoom on top part of discontinuity.

(b) Zoom on bottom part of discontinuity.

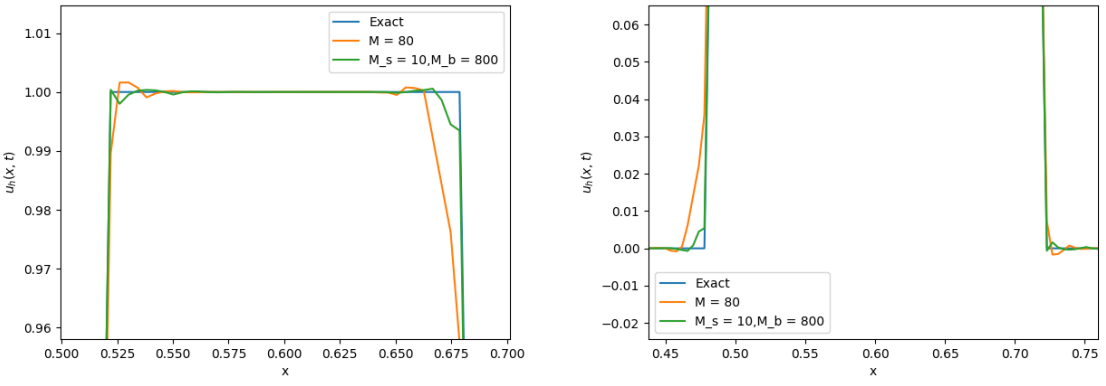
Figure 4.4: Diffusive behaviour of BDF limiter.

4.3.3 Modified limiters.

i) **Adaptive TVB limiter (ATVB).** As we notice earlier, the parameter M in the modified minmod function can introduce dissipation if M is small or oscillations near the discontinuity if M is large enough. The ATVM limiter tries to exploit this feature to avoid searching for a parameter for each initial condition. The idea is to use the TVB limiter, but choose two values for M, one small M_s and one big M_b. We fix the values $M_s = 10$ and $M_b = 800$. For a certain number of iterations $\hat{\tau}$ in the time discretization use the TVB with M_b allowing that oscillations appear in the solution. In the following $\check{\tau}$ steps of the time discretization, use TVB with M_s, so that the introduced oscillations diffuse. Thus, we want to create a balance between the diffusion and the oscillations, so that the solution is not oscillatory nor diffusive.

Instead of explicitly introducing the two parameters $\hat{\tau}$ and $\check{\tau}$, at each time iteration we compute the percentages of cells flagged as troubled. We introduce a threshold θ such that if the percentage of flagged cells is less or equal than θ the M used in the modified minmod function is M_b, otherwise use M_s. Although the threshold θ has to be determined experimentally for each conservation law, it seems to not depend on the initial condition. If its chosen poorly, the solution will be similar to the TVB with $M = M_b$ or $M = M_s$. As a first example of the improvement of the ATVB method, in Figures 4.5a and 4.5b TVB and ATVB methods are compared, $M = 150$ for TVB. The error for the TVB method is 4.25×10^{-2}, whereas the error for the ATVB is 6.54×10^{-3}. Showing a clear improvement in the solution.

The use of the percentage of troubled cells has been used in [59] for comparing their performance, the more a limiter marks cells as troubled, the more computationally expensive it is. For the DG solution using the ATVB limiter, see Algorithm 10 selecting Algorithm 7 as the limiter procedure and defining $Z = [M_s, M_b]$.



(a) Zoom on top part of discontinuity.

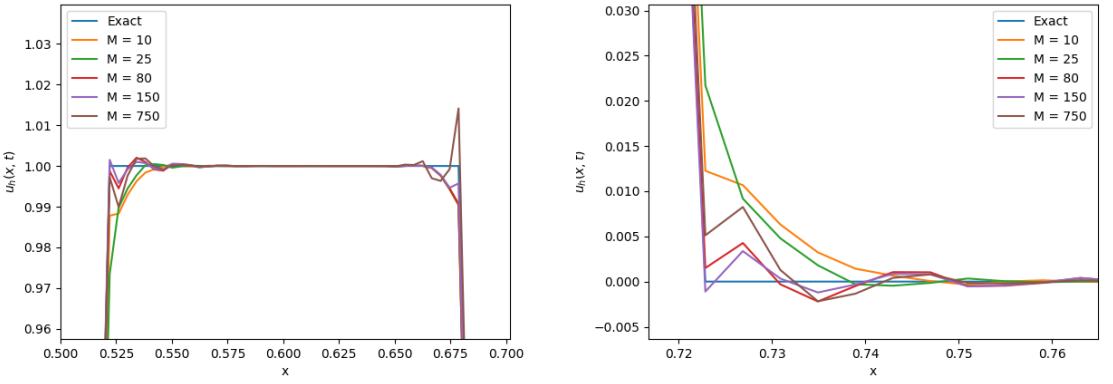
(b) Zoom on bottom part of discontinuity.

Figure 4.5: Comparative example of ATVB improvements.

ii) **Modified BDF limiter (M-BDF)** To the best of the author's knowledge, this naive modification has not been proposed. We apply the BDF limiter but change the minmod function in (4.8) with the modified minmod function, that is,

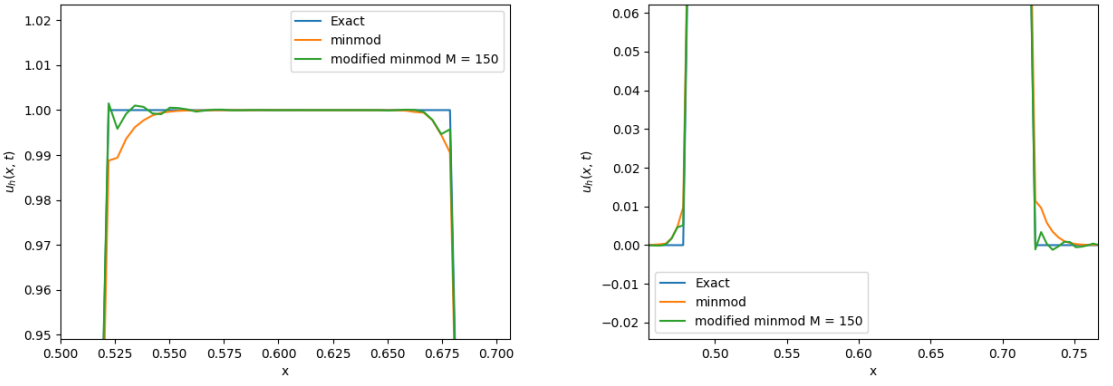
$$u_j^{l,(mod)} = \hat{m}(u_j^l, \alpha_j(u_{j+1}^{l-1} - u_j^{l-1}), \alpha_j(u_1^{l-1} - u_{j-1}^{l-1}), M, h).$$

As it could be expected, introducing the parameter M in the limiter gives a degree of freedom to diffuse the solutions or makes them more oscillatory, see Figure 4.6a and Figure 4.6b for a comparison of different values of M for the solution of the convection equation with an impulse initial condition $T = 2$, $N = 80$. In Figures 4.7a and 4.7b we compare the BDF with the M-BDF limiter choosing $M = 150$. The errors are BDF 1.14×10^{-2} and MBDF 5.35×10^{-2}. See Algorithm 8 for the implementation of MBDF. For the DG solution with the MBDF limiter, see Algorithm 9 selecting Algorithm 8 as the limiter procedure and defining $Z = [M]$.



(a) Zoom on top part of discontinuity. (b) Zoom on bottom part of discontinuity.

Figure 4.6: Diffusive behaviour of BDF limiter.



(a) Zoom on top part of discontinuity. (b) Zoom on bottom part of discontinuity.

Figure 4.7: Comparative example of MBDF improvements.

iii) **Remark** An intuitive next step is to apply the adaptive technique to the M-BDF limiter and to avoid choosing M. Although the numerical implementation might seem straight forward, the tuning of the threshold θ is more complicated. We leave this fact as an open question for future research.

4.3.4 Numerical simulations

For a direct comparison of the methods in Tables 4.1, 4.2, 4.3 and 4.4 we compare the solution of the convection equation with the impulse initial condition, the convection equation with the sine initial condition, the solution of the Burgers' equation with the sine initial condition and the solution of the Burgers' equation with the impulse initial condition. For these examples we show that the ATVB θ parameter only depends on the conservation law and the advantages the M-BDF method over the BDF method. The polynomial order is fixed to $K = 2$. This decision was made on the fact that the improvement of the solution due to limiters varies a lot from example to example. The characteristic behavior of each limiter is really similar for the different polynomial orders. We do not want to overwhelm the reader with too many tables and distract our attention from the comparison of the methods within an example and among the different examples. For all the examples, in their respective tables the error with the $||\cdot||_{L^\infty}$ is shown for the different limiters with $N = 10, 20, 40, 80, 160, 320$ and $K = 2$. All figures are computed with $K = 2$, $N = 80$, the time is specified at each figure. The M_s and M_b chosen for all the alternate method are $M_s = 10$ and $M_b = 800$, for the convection equation $\theta = 3$ and for the Burgers' equation $\theta = 23$.

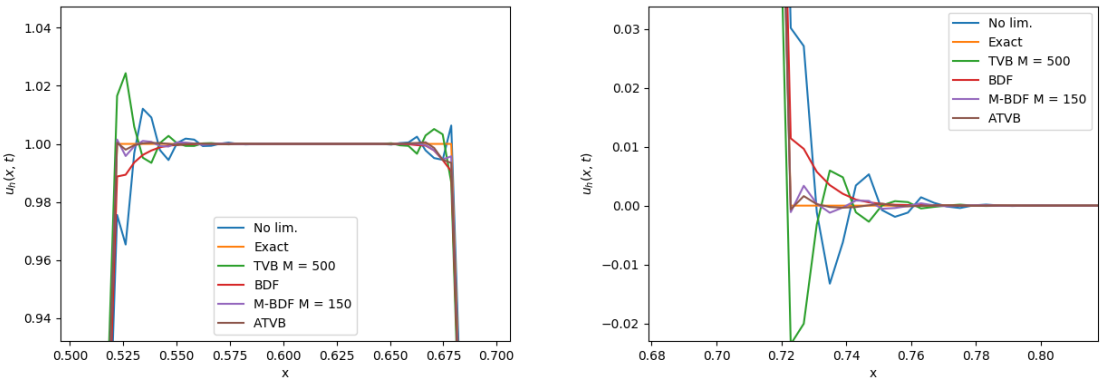
Example 1. Convection equation with impulse initial condition. See Table 4.1 for the norm of the errors and Figures 4.8a and 4.8b for a graphical comparison of the methods. The table shows that the inclusion of the limiter helps the error to decrease. Comparing TVB and ATVB we notice that the error of the ATVB method decreases faster than the error of the TVB. The M-BDF method performed better than the BDF method at each N. All the methods perform almost similarly for $N = 320$. From Figures 4.8a and 4.8b we notice that the proposed methods achieve their goal of being less oscillatory and less diffusive than their original counterparts.

N	No lim.	TVB($M = 500$)	ATVB(3%)	BDF	M-BDF($M = 150$)
10	2.19×10^{-1}	2.19×10^{-1}	1.75×10^{-1}	3.24×10^{-1}	1.75×10^{-1}
20	9.50×10^{-2}	9.50×10^{-2}	9.76×10^{-2}	2.56×10^{-1}	1.06×10^{-1}
40	8.07×10^{-2}	6.92×10^{-2}	1.61×10^{-2}	1.01×10^{-1}	3.63×10^{-2}
80	3.28×10^{-2}	2.45×10^{-2}	6.54×10^{-3}	3.29×10^{-2}	5.35×10^{-3}
160	1.06×10^{-2}	2.44×10^{-3}	1.47×10^{-3}	4.54×10^{-3}	8.44×10^{-4}
320	7.63×10^{-4}	2.12×10^{-5}	3.53×10^{-5}	3.39×10^{-5}	3.14×10^{-5}
Time	$4.70s$	$4.44s$	$4.44s$	$4.55s$	$4.38s$

Table 4.1: Error comparison for linear convection equation $T = 0.1$ $K = 2$

Example 2. Convection equation with sine initial condition. In Table 4.2 we show the errors for the different limiters. The main objective of including this example is to show that without changing the parameter θ the ATVB still converges. If we had used $M = 40$ in the TVB method in the previous example, the results would have been much more diffusive. We add that the BDF method and that the M-BDF method seem to converge faster than the other limiters. In addition, numerical experimentation showed that for the M-BDF the change in M did not produce diffusion in the extrema of the solution.

Example 3. Burgers' equation with sine initial condition. See Table 4.3 for the norm of the errors . The table shows that for this example, the use of the limiters only improves the solution for small N and the order of convergence. Recall that the DG is $O(h^{k+1})$ far away from the discontinuities [16]. Thus, it could be expected that for a large enough N the DG solution would have a high degree of



(a) Zoom on top part of the discontinuity.

(b) Zoom on bottom part of the discontinuity.

Figure 4.8: Comparison of limiters for the convection diffusion equation impulse initial condition $T = 0.1$, $K = 2$, $N = 80$.

N	No lim.	TVB($M = 40$)	ATVB(3%)	BDF	M-BDF($M = 150$)
10	2.21×10^{-3}	2.69×10^{-2}	2.21×10^{-3}	5.49×10^{-3}	2.21×10^{-3}
20	1.07×10^{-3}	1.03×10^{-2}	1.11×10^{-3}	2.81×10^{-3}	1.07×10^{-3}
40	1.86×10^{-4}	1.10×10^{-3}	1.07×10^{-3}	3.15×10^{-4}	1.86×10^{-4}
80	8.44×10^{-5}	5.94×10^{-4}	5.94×10^{-4}	1.05×10^{-4}	8.44×10^{-5}
160	8.15×10^{-7}	7.20×10^{-6}	7.20×10^{-6}	4.04×10^{-6}	8.15×10^{-7}
320	1.03×10^{-7}	1.02×10^{-7}	1.02×10^{-7}	6.30×10^{-7}	1.03×10^{-7}
Time	$4.70s$	$4.44s$	$4.44s$	$4.55s$	$4.61s$

Table 4.2: Error comparison for linear convection equation sine initial condition $T = 0.1$ $K = 2$

accuracy with the method we are using for computing the error. In this example the TVB outperforms the ATVB although their results are really similar. In the case of BDF and M-BDF the results almost match. Although recall that in M-BDF we could have chosen $M = 0$ or $M = 0.1$ and exactly match the BDF. We chose $M = 10$ to show that this method has flexibility in the parameter M as compared with the TVB. If we would have chosen $M = 10$ in TVB the results would be similar to the ones obtained with ATVB.

Example 4. Burgers' equations with impulse initial condition. See Table 4.4 for the norm of the errors and Figures 4.9a and 4.9b for a graphical comparison of the methods. For this example, surprisingly, the best method is the DG without any limiter. Let us understand why. Note that at $t = 0$ the initial condition has two discontinuities and limiters act in both. As time develops, the left discontinuity turns into a rarefaction wave, a continuous function. Explaining why the limiters diffuse this part of the solution. Note in Figure 4.9a that the ATVB oscillates a bit more than the other solution, showing that through the change in M is trying to eliminate the wrongly introduced diffusion. In the same figure we can see that the M-BDF is not as diffusive as the BDF. In general the error for the M-BDF method is lower than the error of the BDF method.

This example has the difficulty of a rarefaction wave colliding with a shock. Thus, limiters on the one hand have to diffuse the oscillations while not diffusing the rarefaction in an space interval that

N	No lim.	TVB($M = 0.01$)	ATVB(23%)	BDF	M-BDF ($M = 10$)
10	4.01×10^{-1}	3.81×10^{-1}	4.17×10^{-1}	3.62×10^{-1}	3.13×10^{-1}
20	2.81×10^{-1}	2.42×10^{-1}	4.90×10^{-2}	8.95×10^{-2}	8.50×10^{-2}
40	3.68×10^{-1}	1.80×10^{-2}	2.16×10^{-3}	1.42×10^{-4}	7.67×10^{-3}
80	3.54×10^{-2}	7.60×10^{-3}	5.32×10^{-4}	2.39×10^{-5}	2.39×10^{-5}
160	2.03×10^{-6}	5.31×10^{-5}	7.47×10^{-5}	3.64×10^{-6}	3.64×10^{-6}
320	2.00×10^{-6}	9.28×10^{-6}	1.31×10^{-5}	5.33×10^{-7}	5.33×10^{-7}
Time	$8.74s$	$9.5s$	$9.47s$	$9.22s$	$13.37s$

Table 4.3: Error comparison for Burgers' equation sine initial condition $T = T_b + 0.1$, $K = 2$

decreases as time progresses. It would be interesting to create a plot showing which cells were marked at each time step to properly understand what is happening in this example.

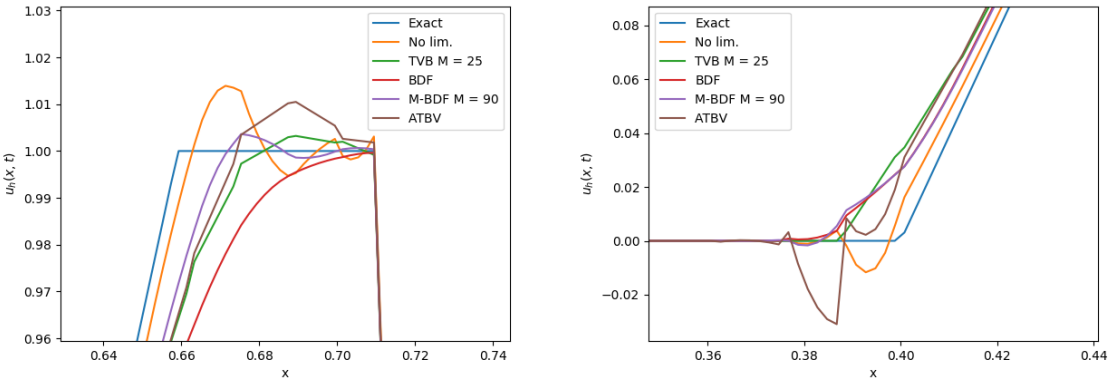
In [40],[60] and [59] the authors include examples with rarefaction waves, from the figures in those articles, the solutions seem to generate diffusion in the rarefaction wave. The authors consider that making further research on the effect of limiters on solutions with rarefaction waves might be interesting to consider.

N	No lim.	TVB($M = 25$)	ATVB(23%)	BDF	M-BDF($M = 1000$)
10	5.54×10^{-1}	4.30×10^{-1}	5.54×10^{-1}	3.62×10^{-1}	2.98×10^{-1}
20	2.27×10^{-1}	1.93×10^{-1}	2.38×10^{-1}	2.10×10^{-1}	9.49×10^{-2}
40	5.15×10^{-1}	6.26×10^{-2}	6.82×10^{-2}	7.36×10^{-2}	5.73×10^{-2}
80	2.73×10^{-2}	3.56×10^{-2}	3.47×10^{-2}	4.56×10^{-2}	2.81×10^{-2}
160	1.39×10^{-2}	2.41×10^{-2}	2.48×10^{-2}	3.25×10^{-2}	2.50×10^{-2}
320	9.76×10^{-3}	1.36×10^{-2}	1.42×10^{-2}	1.51×10^{-2}	1.54×10^{-2}
Time	$9.12s$	$9.23s$	$9.58s$	$9.26s$	$9.08s$

Table 4.4: Error comparison for Burgers' equation impulse initial condition $T = T_b + 0.1$, $K = 2$

Conclusions. We have seen four examples where the performance of the limiters varies a lot. In the first one they improve the solution, in the second and third one the limiters only improve the solution of small enough N and a fourth example where the limiters make the solution worst. This wide panorama allows us to understand why the study of identifying troubled cells and reconstructing solutions is still an active area of research.

In the examples showed here we could achieve the objectives proposed for the ATVB and M-BDF. For the ATVB method we were able to find a θ that worked for all the initial conditions, clearly this does not ensure that is going to work for all initial conditions. It is at least a first step in this direction. In addition, it shows that the concept of using and adaptive M leads to better results. It is still unclear if the proposed method for changing M is the best one, for future research we could investigate more techniques for automatic switching. The importance of automatic switching is that for higher dimensions and non smooth initial condition, M for the TVB method has to be determined experimentally. This idea might help to avoid the numerical experimentation. The search of avoiding numerical experimentation to determine M is an active area of research, in 2021 the article [74] was published. In it the authors implement a multi-layer perceptron for the estimation of the parameter M.



(a) Zoom on top part of the discontinuity.

(b) Zoom on bottom part of the discontinuity.

Figure 4.9: Comparison of limiters for the Burgers' equation impulse initial condition $T = T_b + 0.1$, $K = 2$, $N = 200$.

The M-BDF method showed in general better results, the advantage is that for $M = 0$ the BDF method is recovered and upon necessary, better solutions can be looked for by increasing M.

4.4 Algorithms.

An important part of the implementation of limiters is the position where the limiter is applied. For example, it could be applied each time $L_h(u_h)$ is computed in the TVDRK method, we found that this introduces too much diffusion to the solution. Here we applied the limiter each time a complete iteration of the TVDRK method is made. The modified solver for the system of ODEs is shown in Algorithm 9 for non adaptive limiters and Algorithm 10 for adaptive algorithms. In most of the algorithms, the forward and backward differences are required, recall that the boundary conditions are considered to be periodic. Then for computing the backward difference in the interval I_0 use the solution in I_N and for the forward difference in the interval I_N use the solution in I_0.

Algorithm 6 Modified minmod function.

1: **procedure** MINMOD(a,b,c,delta,M)
2: **if** absolute value of a is less or equal than $M \cdot delta^2$ **then**
return a
3: **else**
4: **if** the sign of a, b and c is the same **then**
return the minimum of a, b and c with the sign of a.
5: **else**
return 0

Algorithm 7 TVB limiter

procedure TVB(Pol. order, number of intervals, vector of coefficients u, delta, M)
  Define limited vector u_l
  for each interval I_j **do**
    Compute $\tilde{u}_j$ and $\tilde{\tilde{u}}_j$

$$\tilde{u}_j = \sum_{l=1}^{K} u_j^l P_l(\xi_j(x_{j+1})), \quad \tilde{\tilde{u}}_j = -\sum_{l=1}^{K} u_j^l P_l(\xi_j(x_j)).$$

    Compute forward and backward differences $(u_{j+1}^0 - u_j^0), (u_j^0 - u_{j-1}^0)$.
    Compute $\tilde{u}_j^{(mod)}$ and $\tilde{\tilde{u}}_j^{(mod)}$ using Algorithm 6

$$\begin{aligned} \tilde{u}_j^{mod} &= \text{MINMOD}(\tilde{u}_j^{mod}, (u_{j+1}^0 - u_j^0), (u_j^0 - u_{j-1}^0), M, h), \\ \tilde{\tilde{u}}_j^{mod} &= \text{MINMOD}(\tilde{\tilde{u}}_j^{mod}, (u_{j+1}^0 - u_j^0), (u_j^0 - u_{j-1}^0), M, h). \end{aligned}$$

    if either $\tilde{u}^l \neq \tilde{u}^{l,(mod)}$ or $\tilde{\tilde{u}}^l \neq \tilde{\tilde{u}}^{l,(mod)}$ **then**
      Mark the cell as troubled.
      Save in the limted vector u_l the zero and first order polynomial coefficients of u for the interval I_j
      Set to 0 the rest coefficients of the limited solution u_l at I_j.
    else
      The solution does not need to be limited, continue with the next interval I_j.
  return the limited solution u_l and the proportion of troubled cells.

Algorithm 8 Moment TVB-modified limiter.

procedure M-BDF(Pol. order N, number of intervals, vector of coefficients u, delta, M, α)

  for each interval I_j **do**
    for the polynomial coefficient $l = N$ with step -1 up to $l = 1$ **do**
      Compute forward and backward differences $(u_{j-1}^{l+1} - u_{j-1}^l), (u_{j-1}^l - u_{j-1}^{l-1})$.
      Define $\alpha_j = \alpha[l]$
      Compute $u_j^{l,(mod)}$ using Algorithm 6

$$u_j^{l,(mod)} = \text{MINMOD}(u_j^l, \alpha_j(u_{j+1}^{l-1} - u_j^{l-1}), \alpha_j(u_j^{l-1} - u_{j-1}^{l-1}), M, h).$$

      if $u_j^l = u_j^{l,(mod)}$ **then**
        Stop the limiting procedure on the interval I_j and continue to the interval I_{k+1}.
      else
        Update u_j^l with $u_j^{l,(mod)}$.

  return Limited vector of polynomial coefficients.

Algorithm 9 Solution of system of ODEs using limiters

procedure SOLVER_ODE_LIM(K, u_0, N, CFL, T, $f(u)$, $f'(u)$, $\hat{f}$, N, h, L_h, limiter, Z vector of limiter parameters)
    $\tau = \frac{CFL}{N}$.
    $dt = \tau$
    $n = 1$
    while $dt = n\tau < T$ **do**
        Compute u_n using Algorithm 1

$$u_n = \text{ runge\_kutta}(K, u_{n-1}, \tau, T, f(u), f'(u), \hat{f}, N, h, L_h)$$

        Apply the chosen limiter procedure with its corresponding parameters Z to u_n.
        $n+ = 1$
return Solution of the system of ODEs u_n

Algorithm 10 Solution of system of ODEs using adaptive limiters

procedure SOLVER_ODE_ALIM(K, u_0, N, CFL, T, $f(u)$, $f'(u)$, $\hat{f}$, N, h, L_h, limiter, Z vector of limiter parameters,θ)
    $\tau = \frac{CFL}{N}$.
    $dt = \tau$
    $n = 1$
    Define steps_for_change $= [\hat{\tau}, \hat{\tau}+1, \ldots, \hat{\tau}+\check{\tau}]$.
    while $dt = n\tau < T$ **do**
        Compute u_n using Algorithm 1

$$u_n, \text{proportion of troubled cells} = \text{ runge\_kutta}(K, u_{n-1}, \tau, T, f(u), f'(u), \hat{f}, N, h, L_h)$$

        if $\theta <$ the proportion of troubled cells **then**
            Apply the chosen limiter procedure with its corresponding parameters Z and $M = M_s$ to u_n.
        else
            Apply the chosen limiter procedure with its corresponding parameters Z and $M = M_b$ to u_n.
        $n+ = 1$
return Solution of the system of ODEs u_n

Chapter 5

Entropy Stable Discontinuous Galerkin Method

As mentioned in the second chapter, a difficulty the DG and LDG methods have is that for each flux f several integrals have to be performed for defining the system of ODEs (3.7) and (3.20). An alternative is to compute such integrals using numerical integration. For this, it is necessary to verify that the solution obtained using numerical integration converge and are entropy solutions.

The DG satisfy the entropy inequality for the square entropy function (3.12) and we saw that this inequality leads to L^2 stability and thus implies, with the use of the TVDRK, convergence to entropy solutions. The proof of L^2 stability for the DG method, assumes that the integrals in the weak formulation are done analytically. An important step of the proof uses integration by parts, a result that in general is not used when numerical integration is done. In this section we study how to do numerical integration in the weak formulation of the DG by using Gauss-Lobatto quadrature points and obtain entropy solutions. The importance of using the Gauss-Lobatto quadrature is to apply a discrete integration by parts formula, called "summation by parts". The entropy inequalities are rephrased discretely using the concept of "entropy stable solutions" following the definitions introduced by Tadmor in [70]. The L^2 stability is generalized in the nonlinear case by entropy stability. We followed [13] for most of the content of this chapter. The vector notation for the DG method (5.11) and the decomposition of the entropy stable flux in an entropy conservative flux and a dissipation function comes from [66].

5.1 Gauss-Lobatto integration

Recall that the weak formulation of the DG method (3.7) for a hyperbolic conservation law is

$$\int_{I_j} (u_h)_t \phi dx = \int_{I_j} f(u_h)\phi' dx + \hat{f}_j(u_h,t)\phi(x_j^+) - \hat{f}_{j+1}(u_h,t)\phi(x_{j+1}^-).$$

Previously, we performed the integrals analytically to obtain the semidiscrete form. If the flux function f is not a polynomial, it can be difficult or impossible to compute the second integral. Here we study how to use the Gauss-Lobatto quadrature for doing these integration numerically.

The Gauss-Lobatto quadrature of a function $f(x)$ using K points for the integration on the

interval $[-1,1]$ is of the form:

$$\int_{-1}^{1} f(x)dx = \frac{2}{K(K-1)}[f(1)+f(-1)] + \sum_{i=1}^{K-1} w_i f(x_i) + R_K,$$

the point x_i is the $i-1$ zero of the $P'_{K-1}(x)$ Legendre polynomial. The weights are defined as

$$w_i = \frac{2}{K(K-1)[P_{K-1}(x_i)]^2}, \quad x_i \neq \pm 1,$$

and the reminder is

$$R_K = \frac{-K(K-1)^3 2^{2K-1}[(K-2)!]^4}{(2K-1)[(2K-2)!]^3} f^{(2K-2)}(\xi), \quad -1 < \xi < 1.$$

From now on, when we refer to the K point Gauss-Lobatto quadrature, we refer to the points $x_0 = -1 < x_1 < \ldots < x_{K-1} < x_K = 1$, with x_i, $i = 1, \ldots, K-1$ the ordered zeros of the P'_{K-1} Legendre polynomial.

With the change of variables $\xi_j(x) = 2\dfrac{x - x_{j+\frac{1}{2}}}{h}$ $j = 0, \ldots, N-1$ in the weak form (3.7), we obtain

$$\frac{h}{2}\int_I \frac{\partial}{\partial t} u_h \phi d\xi = \int_I f(u_h)\phi' d\xi + \hat{f}_j(u_h,t)\phi(x_j^+) - \hat{f}_{j+1}(u_h,t)\phi(x_{j+1}^-). \tag{5.1}$$

Now, consider the set of Lagrange basis polynomials generated by the Gauss-Lobatto points $(x_i)_{i=0}^K$,

$$L_l(\xi) = \prod_{m=0,\ m\neq l}^{K} \frac{\xi - x_m}{x_l - x_m},$$

for the V_h^K space. Then the DG expansion is

$$u_h(x,t)\big|_{I_j} = \sum_{l=0}^{K} u_j^l L_l(\xi_j(x)), \tag{5.2}$$

substituting in the weak formulation (5.1) with $\phi \in V_h^K$ arbitrary, we obtain

$$\frac{h}{2}\sum_{l=0}^{K} \frac{d}{dt} u_j^l(t) \int_I L_l(\xi)\phi(\xi)d\xi = \int_I f(u_h)\phi(\xi)d\xi + \hat{f}_j(u_h,t)\phi(-1) - \hat{f}_{j+1}(u_h,t)\phi(1). \tag{5.3}$$

We now compute the integrals in equation (5.3) with the Gauss-Lobatto quadrature and use the following vector notation for the evaluation on the Gauss-Lobatto points x_i. The function u_h in the interval I_j evaluated on the points x_i is represented by

$$\overrightarrow{u_j} = \begin{bmatrix} u_h(\xi_j(x_0)) & \cdots & u_h(\xi_j(x_K)) \end{bmatrix}^T = \begin{bmatrix} u_j^0 & \cdots & u_j^K \end{bmatrix}^T,$$

analogously for the functions ϕ and f, we define

$$\overrightarrow{\phi_j} = \begin{bmatrix} \phi(\xi_j(x_0)) & \cdots & \phi(\xi_j(x_K)) \end{bmatrix}^T, \quad \overrightarrow{f_j} = \begin{bmatrix} f(u_j^0) & \cdots & f(u_j^K) \end{bmatrix}^T.$$

The numerical flux has the vector notation

$$\overrightarrow{f_j^*} = \begin{bmatrix} \hat{f}_j & 0 & \cdots & 0 & \hat{f}_{j+1} \end{bmatrix}.$$

The first quadrature integral involves the evaluation of the Lagrange polynomials at the points that define them, due to their construction $L_i(x_j) = \delta_{ij}$ with δ_{ij} the Kronenker delta. Define the mass matrix M, the difference matrix D and the boundary matrix B by

$$[M]_{ij} = w_i \delta_{ij}, \quad [D]_{ij} = L_i'(x_j), \quad B = \text{diag}\{-1, 0, \cdots, 0, 1\}.$$

Then the Gauss-Lobatto quadrature applied to (5.3) with the notation just defined gives

$$\frac{h}{2} \overrightarrow{\phi_j}^T M \frac{d}{dt} \overrightarrow{u_j} = (D\overrightarrow{\phi_j})^T M \overrightarrow{f_j} - \overrightarrow{\phi_j}^T B \overrightarrow{f_j^*}.$$

Note that the test function ϕ is arbitrary, we then can conclude

$$\frac{h}{2} M \frac{d}{dt} \overrightarrow{u_j} = (D)^T M \overrightarrow{f_j} - B \overrightarrow{f_j^*}. \tag{5.4}$$

This last equality is a particular case of the *nodal DG formulation* [25]. The general formulation of the nodal DG does not requires to use neither the Gauss-Lobbato points nor the numerical integration. It only considers a sequence of point $(x_i)_{i=0}^K$ in each I_j and defines the Lagrange basis with such $(x_i)_{i=0}^K$. Then it uses delta functions in each x_i as test function and it gets a similar equation to (5.4).

Define the continuous $\langle \cdot, \cdot \rangle$ and discrete $\langle \cdot, \cdot \rangle_w$ inner products as

$$\langle u, v \rangle = \int_{-1}^{1} uv dx, \quad \langle u, v \rangle_w = (\overrightarrow{u^k})^T M \overrightarrow{v^k},$$

with M the mass matrix. Then the stiffness matrix S is defined as

$$[S]_{ij} = \langle L_i, L_j' \rangle = \langle L_i, L_j' \rangle_w.$$

The last equality holds due the Gauss-Lobatto property of being exact for polynomial of orders up to degree $2K - 3$ for K integration points. It can be shown that the stiffness, mass, difference and boundary matrices satisfy the following discrete analog of integration by parts *summation-by-parts* property (SBP)

$$S = MD, \quad MD + D^T M = S + S^T = B. \tag{5.5}$$

Using the SBP property (5.5) the nodal DG (5.4) has the equivalent formulation

$$\begin{aligned} \frac{h}{2} M \frac{d}{dt} \overrightarrow{u_j} + S \overrightarrow{f_j} &= B(\overrightarrow{f_j} - \overrightarrow{f_j^*}) \\ \frac{h}{2} \frac{d}{dt} \overrightarrow{u_j} + D \overrightarrow{f_j} &= M^{-1} B(\overrightarrow{f_j} - \overrightarrow{f_j^*}). \end{aligned} \tag{5.6}$$

Note that the prior discussion can be rewritten assuming different number of Gauss-Lobatto points at each I_j. If this approach is taken, then there would be a mass, stiffness, difference and boundary matrices for each I_j with their dimensions depending on the number of Gauss-Lobatto points, we refer the interested reader to [66]. In addition, the authors of such article do a more profound development on the SBP property and the *simulations approximation term* (SAT's), the left side of (5.6). In the same article, there is a wider explanation of the relation between the modal an nodal formulation. We do recommend the reader to prior read [13] and then for further explanations see [66]. The second reference is much more abstract and due to its generality can be confusing for a first approach.

5.2 Entropy stable formulation

This nodal DG form (5.6) does not satisfy any entropy condition [66], its total (discrete) entropy (2.31) cannot be proven to diminish in time. To be entropy stable we modify (5.6) using the entropy conservative fluxes and entropy stable fluxes introduced by Tadmor [70]. We proceed to their definition.

Let (η, ψ) an entropy pair. Recall that η is a strictly convex function and ψ satisfies for a flux function f

$$\psi' = \eta' f'.$$

By convexity of η we can define the entropy variable $v = \eta'(u)$ which is invertible and thus the mapping $v \to u$, is well defined. The potential flux Ψ is defined by

$$\Psi(v) = vf(u(v)) - \psi(u(v)).$$

Definition 5.2.1. A numerical flux function $f_S(u_-, u_+)$ is called *entropy conservative* with respect to some entropy η if it is consistent, symmetric and satisfies the following equality:

$$(v_+ - v_-) f_S(u^-, u_+) = \Psi_+ - \Psi_-, \tag{5.7}$$

(η, ψ) are an entropy pair. The notation u_- and u_+ are the left and right limits of u at a point x.

The scalar case has a unique entropy conservative numerical flux

$$f_S(u_-, u_+) = \begin{cases} \frac{\Psi_+ - \Psi_-}{v_+ - v_-} & \text{if } u_- \neq u_+ \\ f(u_-) & \text{otherwise } u_- = u_+. \end{cases} \tag{5.8}$$

Example Consider the Burgers' equation

$$u_t + \left(\frac{u^2}{2}\right)_x = 0,$$

with the square entropy $\eta(u) = \frac{u^2}{2}$ results in the entropy variable $v(u) = u$, the flux function ψ, and the potential flux Ψ

$$\psi = \frac{1}{3}u^3, \quad \Psi(v) = \frac{1}{6}v^3.$$

Then the entropy conservative flux is

$$f_S(u_-, u_+) = \frac{1}{6}(u_-^2 + u_- u_+ + u_+^2).$$

If we consider $\eta(u) = e^u$, then the entropy variable is $v(u) = e^u$, the flux function ψ, and the potential flux Ψ are

$$\psi(u) = e^u(u-1), \quad \Psi(u) = e^u\left(\frac{u^2}{2} - u + 1\right).$$

The entropy conservative flux f_S does not simplifies as in the previous example.

Definition 5.2.2. A numerical flux function $\hat{f}(u_-, u_+)$ is called *entropy stable with respect to some entropy* η if it is consistent, single-valued and satisfies the following inequality:

$$(v_+ - v_-)\hat{f}(u_-, u_+) \leq \Psi_+ - \Psi_-. \tag{5.9}$$

An entropy stable numerical flux can be obtained by adding some entropy dissipation to the entropy conservative flux as follows:

$$\hat{f}(u_-, u_+) = f_S(u_-, u_+) + \epsilon \hat{d}(u_-, u_+),$$

where the entropy dissipation function $\hat{d}(u_-, u_+)$ satisfies the following conditions:

1. Consistency: $\hat{d}(u, u) = 0$.
2. Single-valuedness: $\hat{d}(u_+, u_-) = -\hat{d}(u_-, u_+)$.
3. Entropy dissipation: $(v_+ - v_-)\hat{d}(u_-, u_+) \leq 0$.

In this work we use the local Lax-Friedrichs dissipation function

$$\hat{d}(u_-, u_+) = -\lambda(u_+ - u_-),$$

where $\lambda \geq 0$ is the maximum of $|f'(u)|$. The constant $0 < \epsilon \leq 1$ is used to control the added dissipation.

We modify (5.6) on an fixed but arbitrary element I_j, thus we omit the subscript j. The modified nodal DG is

$$\frac{h}{2}\frac{d}{dt}u^i + 2\sum_{l=0}^{K} D_{il} f_S(u^i, u^l) = \frac{\delta^i}{w_i}(f^i - f^{*,i}), \tag{5.10}$$

the superscript represents the corresponding entry of the vectors $\overrightarrow{u}$, $\overrightarrow{f}$, $\overrightarrow{f^*}$, D is the difference matrix, w_i are the Gauss-Lobatto weights and

$$\delta^i = \begin{cases} -1 & i = 0 \\ 1 & i = k \\ 0 & 1 \leq i \leq K-1. \end{cases}$$

The inclusion of the entropy stable flux $\hat{f}$ occurs in the numerical flux vector $\overrightarrow{f^*}$. Recall that the entries of this vector are zero for $i = 1, \ldots, K-1$, and for $i = 0, K$

$$\overrightarrow{f^{*,0}} = \hat{f}_j(u_h) \quad \text{and} \quad \overrightarrow{f^{*,K}} = \hat{f}_{j+1}(u_h).$$

Theorem 5.2.1. *[13] If $f_S(u^i, u^l)$ is consistent and symmetric, then (5.10) is conservative and high order accurate. If we further assume that $f_S(u^i, u^l)$ is entropy conservative in the sense of (5.7), then (5.10) is also entropy conservative within a single element.*

We now have that the nodal DG is conservative in a single element, we extend this result for all the elements and assume that the numerical flux $\hat{f}$ is entropy stable.

Theorem 5.2.2. *[13] If the numerical flux $\hat{f}$ at the element interface is entropy stable, then the scheme (5.10) is entropy stable. Meaning that the discrete total entropy is non-decreasing with respect to time (2.31). Recalling that w_j are the Gauss-Lobatto weights, the discrete total entropy reads,*

$$\frac{d}{dt}\sum_{j=0}^{N-1}\left(\sum_{l=0}^{K}\frac{h}{2}w_l\eta_l\right) \leq 0,$$

for every entropy function η.

Equation (5.10) can be rewritten in the following vector form

$$\frac{d}{dt}\overrightarrow{u_j} = \frac{2}{h}\left(-2D \circ F_s(\overrightarrow{u_j}, \overrightarrow{u_j})\overrightarrow{1}) - M^{-1}B(\overrightarrow{f_j} - \overrightarrow{f_j^*})\right), \tag{5.11}$$

where $\circ$ is the Hadamard (pointwise) product and F_S is the matrix of pairwise combinations of entropy conservative fluxes

$$F_S(\overrightarrow{u_j}, \overrightarrow{u_j}) = \begin{bmatrix} f_S(u_j^0, u_j^0) & \dots & f_S(u_j^0, u_j^K) \\ \vdots & \ddots & \vdots \\ f_S(u_j^K, u_j^0) & \dots & f_S(u_j^K, u_j^K) \end{bmatrix}.$$

5.3 Numerical simulations

In this section we include the solution of the convection diffusion equation and Burgers' equation with a sine initial condition and an impulse initial condition. For the numerical integration we consider $K = 3, 4, 5$. The entropy stable DG method is more computationally expensive than the usual DG method. Thus, we only consider $N = 10, 20, 40, 80$. The solution of the ODE systems is as well parallelized. The numerical solution is compared with the exact solution at a uniform partition of the domain using (5.2). In this case the relation between the increment in time τ and the spatial discretization is $\tau = CFL * h^{\frac{K+1}{3}}$. The order for the nodal entropy DG is $O(h^K)$. This convergence order is again assuming a proper choice of the TVDRK method order. We only have it up order 4, this might be the cause for not achieving the theoretical order of convergence in the examples. In both examples, we use the square entropy $\eta(u) = \frac{u^2}{2}$.

Example 1. Linear convection equation,

$$u_t + u_x = 0.$$

Since $f(u) = u$ and $\eta(u) = \frac{u^2}{2}$ then $v = u$ and

$$\Psi(u) = \frac{u^2}{2}.$$

For the sine initial condition (2.37) in the Table 5.1, we observe that the error decreases as a function of N and of K, as expected.

N	$\|\|u_h^{(3)} - u\|\|_{L^\infty}$	Order $K = 3$	$\|\|u_h^{(4)} - u\|\|_{L^\infty}$	Order$K = 4$	$\|\|u_h^{(5)} - u\|\|_{L^\infty}$	Order $K = 5$
10	8.49×10^{-3}	-	1.65×10^{-3}	-	1.75×10^{-5}	-
20	1.03×10^{-3}	3.05	1.94×10^{-4}	3.10	4.35×10^{-7}	5.33
40	1.08×10^{-4}	3.25	4.26×10^{-6}	5.51	1.32×10^{-8}	5.04
Time	$23.74s$		$46.02s$		$123.75s$	

Table 5.1: Error of linear convection equation with the sine initial condition $T = 0.1$

For the impulse initial condition (2.38) the solutions have oscillations near discontinuities. The nodal formulation allows the use of limiters, we leave this as future work. To show the accuracy of the nodal entropy DG, we could compute the error in the whole domain, remove a neighborhood of the discontinuities or graph the numerical and exact solutions, we choose the later. For this problem the

value of ϵ determines how much diffusion is introduced to the solution, in this case we choose $\epsilon = 1$. For the behaviour of the solutions with different values of ϵ, Figure (5.1a) has solutions with $\epsilon = 0.15$, Figure (5.1c) with $\epsilon = 0.45$ and Figure 5.1e for $\epsilon = 1$. The effects of a higher ϵ is that the solutions are less oscillatory. In Figures 5.1e, 5.2a and 5.2b we show the solutions for $K = 3, 4, 5$, respectively. The improvement for $N = 10$ is clear as K increases. For $N = 40, 80$ the improvements are more subtle. In both cases, the solution curves oscillate closer to the discontinuity and the oscillation region gets smaller as K increases.

Example 2. Consider Burgers' equation

$$u_t + \left(\frac{u^2}{2}\right)_x = 0.$$

Since $f(u) = \frac{u^2}{2}$ and $\eta(u) = \frac{u^2}{2}$ then $v = u$ and

$$\Psi(u) = \frac{u^3}{6}.$$

In Table 5.2 we show the L^∞ norm of the errors for $T = 0.1$, prior to the shock. Although the numerical order is below the optimal, the results are still good. In [13] they obtained as well that the order with the L^1, L^2 and L^∞ norm is below optimal and specially the L^∞.

N	$\|\|u_h^{(3)} - u\|\|_{L^\infty}$	Order $K = 3$	$\|\|u_h^{(4)} - u\|\|_{L^\infty}$	Order$K = 4$	$\|\|u_h^{(5)} - u\|\|_{L^\infty}$	Order $K = 5$
10	6.72×10^{-2}	-	1.28×10^{-2}	-	4.46×10^{-3}	-
20	2.23×10^{-2}	1.58	2.16×10^{-3}	2.55	7.29×10^{-4}	2.61
40	5.66×10^{-3}	1.98	4.90×10^{-4}	2.14	6.38×10^{-5}	3.51
80	1.12×10^{-3}	2.34	1.02×10^{-4}	2.26	9.69×10^{-6}	2.71
Time	$60.91s$		$208.02s$		$794.s$	

Table 5.2: Error Burgers' equation sine initial condition $T = 0.1$

In Figures 5.3a, 5.3b and 5.4 we show the solutions for the sine initial condition (2.37) at time $T = T_b + 0.1$. We could obtain solutions that are not highly oscillatory by using $\epsilon = 0.3$. The improvements as K grows are the decrease of the oscillations and the reduction of the oscillation region.

In Figures 5.5a, 5.5b and 5.6 we show the solution of the Burgers' equation with the impulse initial condition (2.38) for $K = 3, 4, 5$, respectively. On the shock part of the solution, they have a similar behaviour to the solution of the linear convection equation with the same initial condition. On the left side of the rarefaction wave, there are as well oscillations, but on the right side of the rarefaction wave the solution is not as diffusive as with the modal DG. We as well choose $\epsilon = 1$ for the incorporation of the diffusion in the entropy stable flux.

Conclusions. Numerical simulations show that by considering a proper nodal DG formulation entropy solutions can be obtained. Although these algorithms were much more slower than their modal counter part. They are advantageous for fluxes f that are not polynomial, where the integrals in the weak formulation might not even be possible to make. This not only affects the preparation of the algorithm, avoiding to explicitly computing L_h for each f. It makes it easier to make a general code for an arbitrary f. Additionally, this formulation makes it easier to consider high order polynomial K. The integration of the parameter ϵ is not part of the standard implementation, we included it here to

aid the control of oscillations without the use of limiters. Including limiters would complement the action of the parameter ϵ. Note that they are applied at different stages of the the ODE solver. The entropy stable flux is used to define the operator L_h of the TVDRK method. Whereas the limiters are applied after a complete step of the TVDRK solver is made. The value of ϵ should not change if the techniques are used simultaneously. The best choice of ϵ will give to the limiter a more diffused solution and so the limiter will detect less troubled cells and have a lower computational cost.

The entropy stable DG method is much more expensive than the original DG. In the Table 5.3 the computation for the solution of the Burgers' equation with $K = 3$, $CFL = 0.01$ and $T = 0.1$ for different N. In such table, we can see that the computational cost for the entropy stable DG grows much more faster than for the DG method.

N	10	20	40	80
DG	$3.86s$	$3.87s$	$3.98s$	$3.94s$
Entropy stable DG	$13.74s$	$15.13s$	$20.34s$	$49.07s$

Table 5.3: Comparison of times for different N.

There is plenty of work to continue in this direction. In [13] the authors develop the corresponding nodal formulation for the LDG to solve parabolic problems. Other area to explore is the use of other quadratures, in [66] the authors explicitly include the Legendre-Gauss-Lobatto quadrature and a general framework for other quadratures.

5.4 Algorithms

For the solution of the ODE system (5.11) we use the Algorithm 2 with the third order TVDRK, Algorithm 1. The L_h argument of the Runge-Kutta method is Algorithm 13. The $\hat{d}$ argument in Algorithm 13 is a modification of Algorithm 3, instead of returning $\frac{f(a)+f(b)-C(b-a)}{2}$, only return $C(b-a)$. Algorithm 13 requires the functions f, f', Ψ, v as an argument and no further integration is required. The matrix M is constructed with the weights from Algorithm 11. The difference matrix D Algorithm 12, uses the nodes of the Gauss-Lobatto quadrature obtained in Algorithm 11. We do not include an algorithm for the boundary matrix B due to its simplicity. These matrices are only required to be computed once. Therefore their computation is done in the Algorithm 2 prior to the first call of the TVDRK method. If the number of quadrature points K_j varied in each element I_j there would be a M, D, B for each different K_j quadrature points. In such case, Algorithm 13 needs to be modified to use the proper M_j, D_j, B_j at each I_j. Once again, we recall that in the elements I_0 and I_{N-1} the periodic boundary condition is used for the computation of "u_{-1}^K" and "u_N^0", respectively.

Algorithm 11 Gauss-Lobatto points and weights

1: **procedure** GAUSS_LOB(K)
2: In the vector *nodes* save the Gauss-Lobatto points $x_0, x_1, \ldots, x_K, x_{K+1}$, $x_0 = -1$, x_{K+1}
3: In the vector *weights* save the corresponding weights $w_0, \ldots, w_{K+1}$.
4: **return** *nodes*, *weights*

Algorithm 12 Differentiation matrix D.

1: **procedure** DIFF_MAT($nodes$, K)
2: Define a matrix D of size $K \times K$
3: **for** $i = 0, 1, \ldots, K$ **do**
4: **for** each node $x_i \in nodes$ **do**
5: $[D_{i,j}] = L_i'(x_j)$
return D

Algorithm 13 Nodal DG

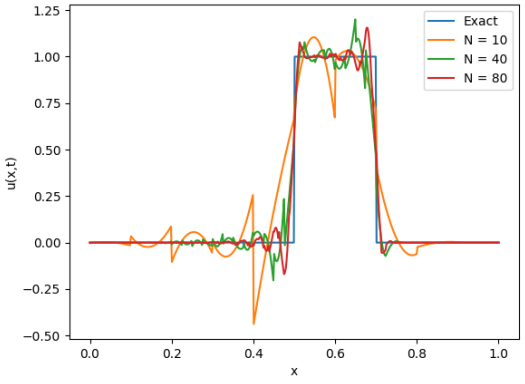
1: **procedure** NODAL_DG(u, D,M, B, K, f, f', N, $\hat{d}$, Ψ, v, h, ϵ)
2: Define solution vector u^1
3: **for** each element I_j $j = 0 \ldots, N-1$ **do**
4: Compute $\overrightarrow{f_j^*} = \begin{bmatrix} \hat{f}_j & 0 & \ldots & 0 & \hat{f}_{j+1} \end{bmatrix}^T$,

$$\begin{aligned} \hat{f}_j &= f_S(u_{j-1}^K, u_j^0) + \epsilon \hat{d}(u_{j-1}^K, u_j^0) \\ \hat{f}_{j+1} &= f_S(u_j^K, u_{j+1}^0) + \epsilon \hat{d}(u_j^K, u_{j+1}^0) \end{aligned}$$

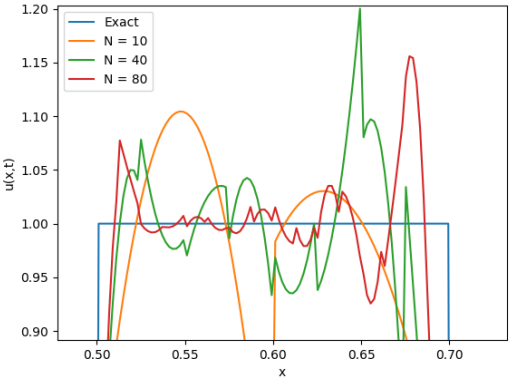
5: Update the corresponding entries in the solution vector u^1 by (5.11),

$$u_j^1 = \frac{2}{h}\left(-2D \circ F_s(\overrightarrow{u_j}, \overrightarrow{u_j})\,\overrightarrow{1}) - M^{-1}B(\overrightarrow{f(u_j)} - \overrightarrow{f_j^*})\right),$$

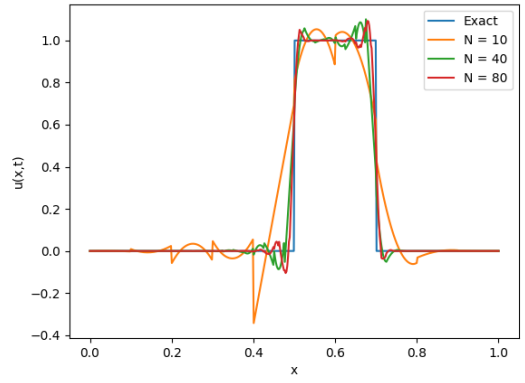
return



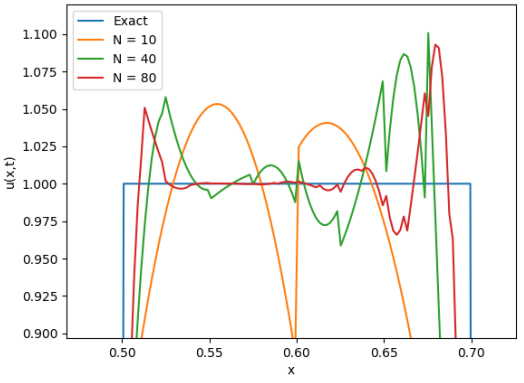
(a) $\epsilon = 0.15$.



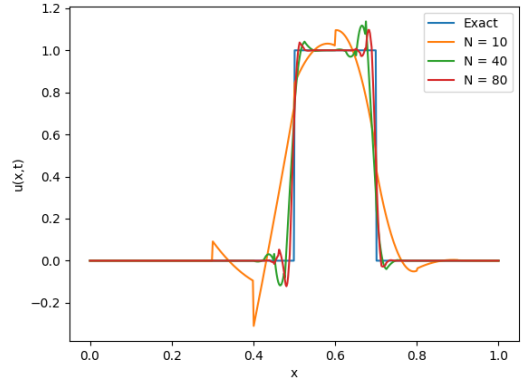
(b) $\epsilon = 0.15$ zoom.



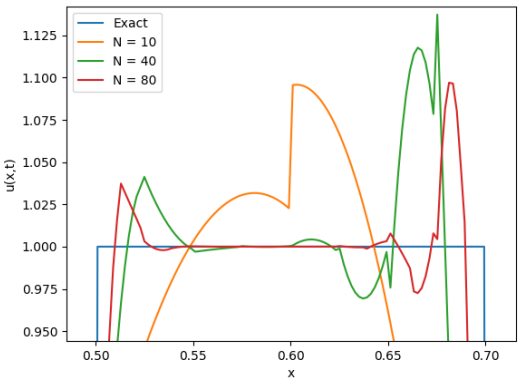
(c) $\epsilon = 0.45$.



(d) $\epsilon = 0.45$ zoom.

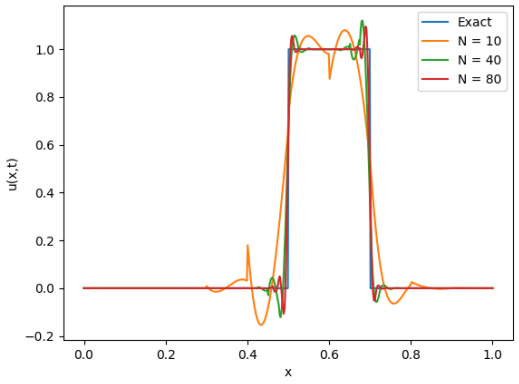


(e) $\epsilon = 1$

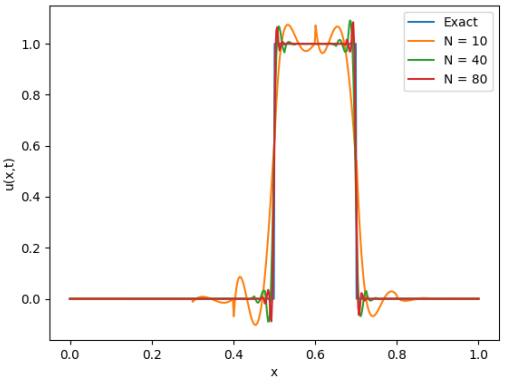


(f) $\epsilon = 1$ zoom.

Figure 5.1: Solution comparison of the convection equation with impulse initial condition for different values of ϵ.

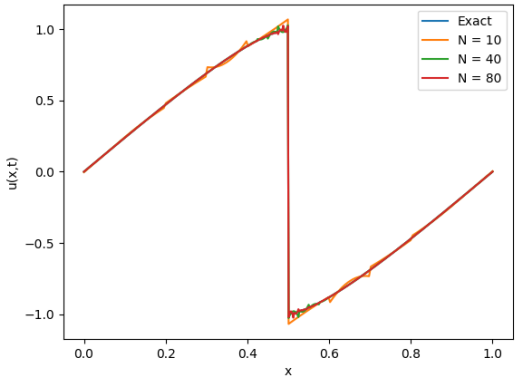


(a) Solution with $K = 4$.

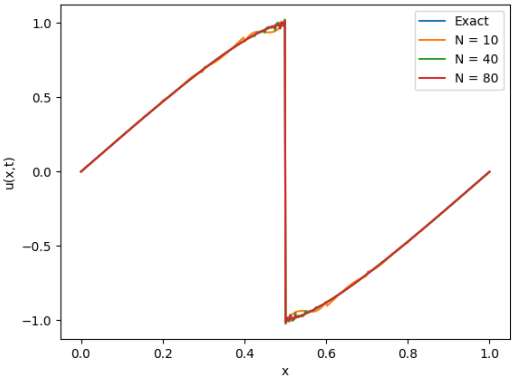


(b) Solution with $K = 5$.

Figure 5.2: Solutions of the convection equation with impulse initial condition for $K = 4, 5$.



(a) Solution with $K = 3$.



(b) Solution with $K = 4$.

Figure 5.3: Solutions to the Burgers equation with sine initial condition $K = 3, 4$, $T = T_b + 0.1$

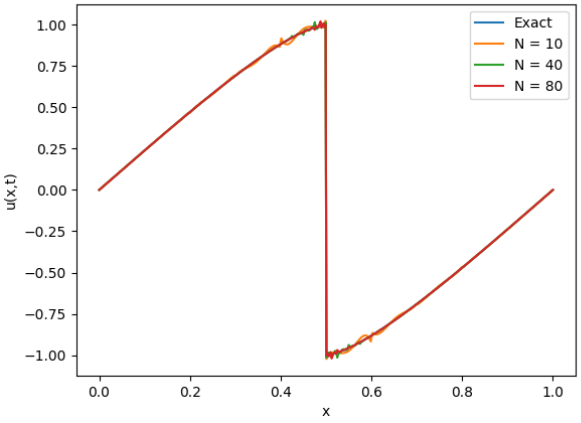
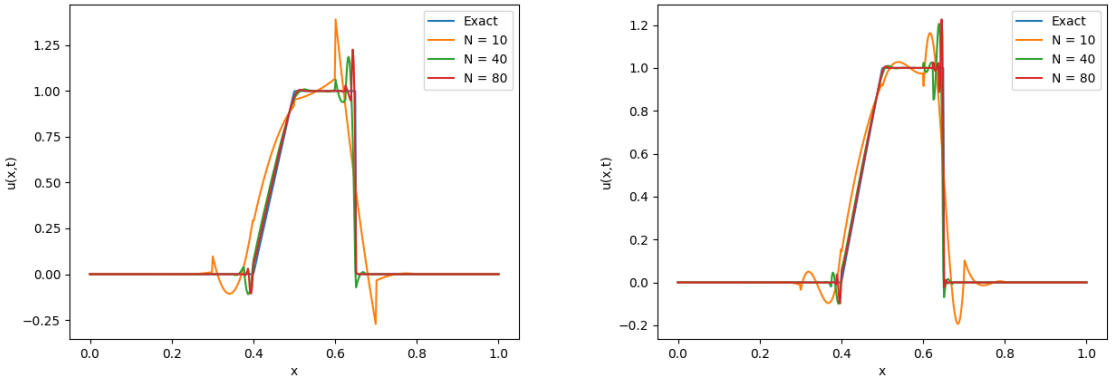


Figure 5.4: Solutions to the Burgers equation with sine initial condition $K = 5$, $T = T_b + 0.1$



(a) Solution with $K = 3$.

(b) Solution with $K = 4$.

Figure 5.5: Solutions to the Burgers' equation with impulse initial condition $K = 3, 4$.

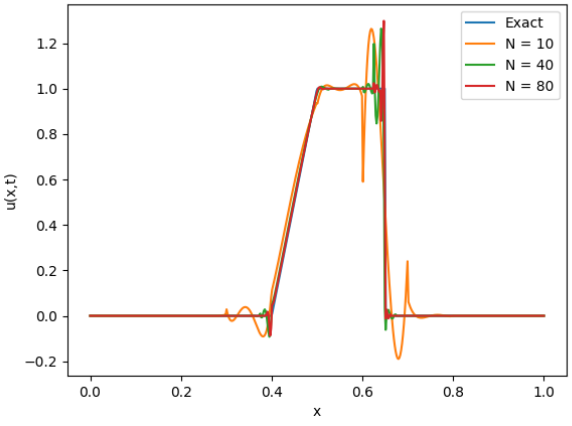


Figure 5.6: Solutions to the Burgers' equation with impulse initial condition $K = 5$.

Chapter 6

Discontinuous Galerkin Methods for Stochastic Equations

The motivation of introducing a stochastic source in a conservation law equation is to reduce the uncertainty of a variable which is not known. For example, the modelling of the concentration of a pollutant in a river. It might be the case that we a priory not know how the rate at which the pollutant is poured, but we can model it using a white noise. Solving this stochastic conservation law will help us to determine what is the mean behaviour of the concentration of the pollutant in the river and, if required, estimate the variance of this behaviour.

The solution for a general f in either the conservation law or convection-diffusion equation is not a trivial matter nor in the deterministic case neither in the stochastic case. Finite volume schemes have previously been used for the solution of such equations [6], [9] and [32]. From the other formulations of the discontinuous Galerkin, the IP methods have been developed for elliptic SPDE using the Karhunen-Loève expansion, see [4], [50] and the references therein. The SPDEs in [4] and [50] assume an stochastic flux. In terms of the example of the concentration of pollution in a river, they assume that the density of the pollutant is modelled with a random variable.

Recall that for the deterministic case, we defined a spatial discretization I_j and the space V_h^K of picewise continuous polynomials of degree at most K. We then constructed an approximate solution u_h such that for each $t \in [0,T]$, $u_h(\cdot,t) \in V_h^K$. In the stochastic case, we use the same spatial discretization but we require to the approximate solution u_h that for each $(\omega,t) \in \Omega \times [0,T]$, $u_h(\omega,\cdot,t) \in V_h^K$. In the deterministic case, this assumption lead to a system of ODEs for finding the coefficients of the expansion of the approximated solution in terms of the basis generated by the Legendre polynomials. Considering the same assumption in the stochastic case, leads to a system of SODEs for the coefficients of the expansion in terms of the Legendre polynomial basis.

Since the stochastic case leads to SODEs, we need to substitute the TVDRK method with a suitable solver. We first review the numerical solution of stochastic differential equations using the Itô-Taylor expansion. Then we work on the discontinuous Galerkin method. In the deterministic framework, the DG method gave inconsistent solutions for parabolic problems. The same behavior happens in the stochastic case. Thus, we use the DG for stochastic conservation laws and apply the

LDG for the stochastic convection-diffusion equations.

6.1 Time discretization

As in the deterministic case, we can use the Taylor expansion to come up with numerical methods. Although, for the stochastic case we need to use the *Itô-Taylor* expansion. The material of this section can be found in [37]. Let X_t be the process defined by

$$X_t = X_0 + \int_{t_0}^{t} F(X_s)ds + \int_{t_0}^{t} G(X_s)dW_s.$$

Then for a twice differentiable function f, by the Itô formula

$$f(X_t) = f(X_{t_0}) + \int_{t_0}^{t} L^0 f(X_s)ds + \int_{t_0}^{t} L^1 f(X_s)dW_s,$$

with

$$L^0 = F\frac{\partial}{\partial x} + \frac{1}{2}G^2\frac{\partial^2}{\partial x^2}, \quad L^1 = G\frac{\partial}{\partial x}.$$

Applying the Itô formula to $f = F$ and $f = G$ in the definition of X_t,

$$\begin{aligned} X_t &= X_{t_0} + \int_{t_0}^{t}\left(F(X_0) + \int_{t_0}^{s} L^0F(X_z)dz + \int_{t_0}^{s} L^1F(X_z)dz\right)ds + \\ &+ \int_{t_0}^{t}\left(G(X_0) + \int_{t_0}^{s} L^0G(X_z)dz + \int_{t_0}^{s} L^1G(X_z)dz\right)dW_s \\ &= X_{t_0} + F(X_{t_0})\int_{t_0}^{t} ds + G(X_{t_0})\int_{t_0}^{t} dW_s + R, \end{aligned} \tag{6.1}$$

with remainder

$$\begin{aligned} R &= \int_{t_0}^{t}\int_{t_0}^{s} L^0F(X_s)dzds + \int_{t_0}^{t}\int_{t_0}^{s} L^1F(X_z)dW_zds \\ &+ \int_{t_0}^{t}\int_{t_0}^{s} L^0G(X_s)dzdW_s + \int_{t_0}^{t}\int_{t_0}^{s} L^1G(X_z)dW_zdW_s. \end{aligned}$$

Applying the Itô formula to the term L^1G of the reminder we obtain the next order Itô-Taylor approximation,

$$X_t = X_{t_0} + F(X_{t_0})\int_{t_0}^{t} ds + G(X_{t_0})\int_{t_0}^{t} dW_s + L^1G(X_{t_0})\int_{t_0}^{t}\int_{t_0}^{s} dW_zdW_s + \hat{R}. \tag{6.2}$$

Using the Itô formula it can be showed that

$$\int_{t_0}^{t}\int_{t_0}^{z} dW_zdW_s = \frac{1}{2}[W(t) - W(t_0)]^2 - \frac{t - t_0}{2}.$$

With (6.2) Itô-Taylor expansion we can construct the *Euler-Maruyama method* and the *Milstein method*. For the interval $[t_i, t_{i+1}]$, assuming $X(t_i)$ is known, define $\tau = t_{i+1} - t_i$, and $\Delta W_i = W(t_{i+1}) - W(t_i)$, substituting in equation (6.2) yields

$$X(t_{i+1}) = X(t_i) + F(X(t_i))\tau + G(X(t_i))\Delta W_i + \frac{1}{2}G[X(t_i)]G'[X(t_i)][(\Delta W_i)^2 - \tau] + \hat{R}.$$

The first three terms define the *Euler-Maruyama method*

$$X(t_{i+1}) = X(t_i) + F(X(t_i))\tau + G(X(t_i))\Delta W_i,$$

the whole expression defines the *Milstein method*

$$X(t_{i+1}) = X(t_i) + F(X(t_i))\tau + G(X(t_i))\Delta W_i + \frac{1}{2}G[X(t_i)]G'[X(t_i)][(\Delta W_i)^2 - \tau].$$

The Milstein method requires the evaluation of the derivative of the diffusion term G. Higher order methods will require as well the evaluation of derivatives of the drift and diffusion functions. To avoid evaluating the derivatives they are approximated by finite differences, we call this schemes the *derivative free schemes*,

$$\begin{aligned} X(t_{i+1}) &= X(t_i) + F(X(t_i))\tau + G(X(t_i))\Delta W_i \quad \text{Euler-Maurayama} \\ X(t_{i+1}) &= X(t_i) + F(X(t_i))\tau + G(X(t_i))\Delta W_i + \frac{1}{4\sqrt{\tau}}(G(\gamma_+) - G(\gamma_-))((\Delta W_i)^2 - \tau) \quad \text{Milstein,} \end{aligned}$$

where

$$\gamma_\pm := X(t_i) + F(X(t_i))\tau \pm G(X(t_i))\sqrt{\tau}.$$

The next order numerical approximation has the form

$$\begin{aligned} X(t_{i+1}) &= X(t_i) + F(X(t_i))\tau + G(X(t_i))\Delta W_i \\ &+ \frac{1}{4\sqrt{\tau}}(G(\gamma_+) - G(\gamma_-))((\Delta W_i)^2 - \tau) \\ &+ \frac{1}{4}(F(\gamma_+) - 2F(X(t_i)) + F(\gamma_-))\tau \\ &+ \frac{1}{2\tau}(F(\gamma_+) - F(\gamma_-))\Delta Z_i \\ &+ \frac{1}{2\tau}(G(\gamma_+) - 2G(X(t_i)) + G(\gamma_-))(\Delta W_i\tau - \Delta Z_i) \\ &+ \frac{1}{4\tau}(G(\phi_+) - G(\phi_-) - G(\gamma_+) + G(\gamma_-))\left(\frac{1}{3}(\Delta W_i)^3 - \tau\right)\Delta W_i, \end{aligned} \tag{6.3}$$

where $\phi_\pm = \gamma_+ \pm G(\gamma_+)\sqrt{\tau}$ and ΔZ_i is a normally distributed with the following mean, variance and correlation

$$\mathbb{E}(\Delta Z_i) = 0, \quad \mathbb{E}((\Delta Z_i)^2) = \frac{1}{3}\tau^3, \quad \mathbb{E}(\Delta W_i \Delta Z_i) = \frac{1}{2}\tau^2.$$

For the simulation of the previous schemes we use *strong simulation* in which the discrete increments ΔW_i are directly sampled from independent Gaussian distributions in each computation interval. Then the error is computed as the expected value of the error. Denoting the exact solution by y_t and the numerical solution by $\hat{y}_t$, the expected value of the error is

$$\mathbb{E}|y_t - \hat{y}_t|_{\mathbb{L}^2}.$$

Definition 6.1.1. A method is said to be of strong order p if

$$\mathbb{E}|y_t - \hat{y}_t|_{\mathbb{L}^2} = \mathcal{O}^p.$$

It can be showed that that the Euler-Maruyama method is of strong order 0.5, the Milstein method has strong order 1, and (6.3) has strong order 1.5 from now on we refer to it as the *order* 1.5 *method* [52]. There is a definition of *weak simulation*, the main difference with the strong simulation is that the error of the weak simulation is the difference of the expected values of the true solution and the numerical solutions

$$|\mathbb{E}(y_t) - \mathbb{E}(\hat{y}_t)|.$$

For more information regarding the consistency, convergence and stability of the methods see [37] chapter 9.

The methods will be required to be used for system of SDE. To define the Brownian motion in more dimensions simply consider independent Brownian motions in each dimension. The theorems and methods of this section can be generalized to higher dimensions [37].

6.2 Discontinuous Galerkin method

The main reference for this subsection is [47]. Recall that the stochastic conservation law with a multiplicative stochastic perturbation is

$$\begin{aligned} du + f(u)_x dt &= g(\omega, x, t, u) dW_t \quad \text{in } \Omega \times \mathcal{O} \times (0, T), && (6.4)\\ u(\omega, x, 0) &= u_0(x) \quad \omega \in \Omega, x \in \mathcal{O}, && (6.5) \end{aligned}$$

and the equation is assumed to have a periodic boundary condition. Where $\{W_t, 0 \leq t \leq T\}$ is a standard one-dimensional Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\{\mathcal{F}_t, 0 \leq t \leq T\}$ the natural filtration. We make the following assumptions of the functions $f, g, u_0(x)$

i) The function g is $\mathcal{F} \times \mathcal{B}(\mathcal{O}) \times \mathcal{B}([0, T]) \times \mathcal{B}(\mathbb{R})$- measurable.

ii) The initial condition $u_0 \in L^2(\mathcal{O})$.

iii) The functions f and g are locally Lipschitz continuous. For and $M \in \mathbb{N}$ there is a positive constant $L(M)$ such that for all $(\omega, x, t) \in \Omega \times \mathcal{O} \times [0, T]$ and all $u, u' \in \mathbb{R}$ with $\max |u|, |u'| \leq M$,

$$\max\{|f(u) - f(u')|, |g(\omega, x, t, u) - g(\omega, x, t, u')|\} \leq L(M)|u - u'|.$$

iv) There exists a constant $C > 0$ such tat for any $(\omega, x, t, u) \in \Omega \times \mathcal{O} \times [0, T] \times \mathbb{R}$

$$\max\{|f(u)|, |g(\omega, x, t, u)|\} \leq C(1 + |u|).$$

Functions with this property are said to be *at most linear growing.*

As in the previous DG analysis we are looking for an approximation in the weak form of the SPDE, such that for any $(\omega, t) \in \Omega \times [0, T]$, $u_h(\omega, \cdot, t) \in V_h^K$ and for an arbitrary test function $\phi \in V_h^K$ the following is satisfied

$$\int_{I_j} \phi(x) du_h(\omega, \xi_j(x), t) dx = \left(\int_{I_j} f(u_h)\phi' dx - \hat{f}_{j+1}\phi_{j+1}^- + \hat{f}_j^+ \phi_j^+ \right) dt + \left(\int_{I_j} g(\omega, x, t, u_h)\phi dx \right) dW_t, \tag{6.6}$$

where $\hat{f}$ is a monotone numerical flux. The DG approximation with the Legendre polynomials P_l, at the interval I_j is

$$u_h(\omega, x, t)\big|_{I_j} = \sum_{l=0}^{K} u_j^l(\omega, t) P_l. \tag{6.7}$$

We can expect that when we substitute the DG approximation into the weak formulation of the SPDE we will obtain a system of stochastic differential equations, just as the ones we studied in the previous section.

Substituting (6.7) in (6.6), using the change of variables $\xi_j(x) = 2\frac{x - x_{j+\frac{1}{2}}}{h}$ and the properties of the Legendre polynomial (3.1) and (3.2), we obtain for each $l = 0, \ldots, K$

$$\begin{aligned} du_j^l(\omega, t) &= \frac{2l+1}{h}\left(\int_{-1}^{1} f(u_h) P_l'(\xi) d\xi - \hat{f}_{j+1} + (-1)^l \hat{f}_j\right) dt \\ &+ \frac{2l+1}{2}\left(\int_{-1}^{1} g(\omega, \xi, t, u_h) P_l(\xi) d\xi\right) dW_t \end{aligned}$$

This equation can be summarized as the vector stochastic differential equation

$$du(w, t) = F(u(\omega, t))dt + G(\omega, t, u(\omega, t))dW_t, \tag{6.8}$$

the vector u has length $(k+1)N$ and is represented by

$$u = \left[u_0^0 \quad \cdots u_0^K, \quad u_1^K \quad \cdots \quad u_1^K \cdots u_{N-1}^0 \quad \cdots \quad u_{N-1}^K\right],$$

the entries of the vectors F and G are

$$[F]_i = \frac{2l+1}{h}\left(\int_{-1}^{1} f(u_h) P_l'(\xi) d\xi - \hat{f}_{j+1} + (-1)^l \hat{f}_j\right) dt$$

and

$$[G]_i = \frac{2l+1}{2}\left(\int_{-1}^{1} g(\omega, \xi, t, u_h) P_l(\xi) d\xi\right) dW_t,$$

for $j \in \{0, \ldots, N-1\}$ and $l \in \{0, \ldots, K\}$ such that $i = 2(K+1)j + l$.

We now work with the modal formulation for the modal version of the stochastic conservation law. Consider the vectors $\overrightarrow{u_j}, \overrightarrow{f_j}, \overrightarrow{\phi_j}, \overrightarrow{f_j^*}$ as defined in Chapter 5. Then using the Gauss-Lobatto quadrature in (6.6) we obtain

$$\frac{h}{2}\overrightarrow{\phi_j} M d\overrightarrow{u_j} dx = (-S^T \overrightarrow{f_j} - B\overrightarrow{f_j^*}) + \frac{h}{2} M \overrightarrow{g_j} dW_t.$$

Using the SBP property we can deduce another equivalent relationship

$$du_j = \frac{2}{h}\left(-D\overrightarrow{f_j} + M^{-1}B\left(\overrightarrow{f_j} - \overrightarrow{f_j^*}\right)\right) dt + \overrightarrow{g_j} dW_t.$$

Similarly to the deterministic case, we modify this nodal DG on each element by first considering an entropy flux f_S defined in (5.7) obtaining, for each $l \in \{0, \ldots, K\}$

$$du^l(\omega, t) = \frac{2}{h}\left(-2\sum_{i=0}^{K} D_{li} f_S(u^l, u^i) + \frac{\delta_l}{w_l}(f^l - f^{*,l})\right) dt + g(\omega, t, u^i) dW_t, \tag{6.9}$$

where w_l are Gauss-Lobatto weights and δ_l as defined in Chapter 5. Thus, we can define a SODE in the form of (6.8). Just as in the deterministic case, the inclusion of the entropy stable flux $\hat{f}$ (5.9) occurs in the numerical flux vector $\overrightarrow{f^*}$. Recall that the entries of this vector are zero for $i = 1, \ldots, K-1$, and for $i = 0, K$

$$\overrightarrow{f^{*,0}} = \hat{f}_j(u_h) \quad \text{and} \quad \overrightarrow{f^{*,K}} = \hat{f}_{j+1}(u_h).$$

6.2.1 Numerical simulations

In this section we work with the constant-coefficient linear stochastic equation with the sine initial condition and the impulse initial condition. The first initial condition allows us to corroborate the order of convergence and the second to numerically validate the use of limiters in the stochastic case. In the original work [13], the authors do not include limiters to control oscillations. For the solutions with discontinuities, we measure the error outside a neighborhood of the discontinuities. In addition, we include the nodal entropy stable DG solution with $\epsilon = 1$, to numerically show that the its solutions are entropy solutions. In addition, we consider the solution of the stochastic Burgers' equation with a sine initial condition prior to the shock time and after the shock time, to show that limiters work in the nonlinear case. The relationship between the time discretization and the spatial discretization is $\tau = CFL(h)^{\frac{3}{2}}$.

In the following $\hat{M}$ is the number of realizations of the solutions. The error estimate is given using a Monte Carlo approximation of the $L^2(\Omega \times \mathcal{O} \times [0,T])$, given by

$$\mathbb{E}[||u_h(\cdot,\cdot,T) - u(\cdot,\cdot,T)||^2_{L^2(\mathcal{O})}] \approx e_2^2 \pm \nu,$$

with

$$e_2 := \left(\frac{1}{\hat{M}} \sum_{i=1}^{\hat{M}} z_i \right)^{\frac{1}{2}}, \quad \nu := \frac{2}{\sqrt{\hat{M}}} \left[\frac{1}{\hat{M}} \sum_{i=1}^{\hat{M}} z_i^2 - e_2^4 \right]^{\frac{1}{2}},$$

where $z_i := ||u_h(\omega_i,\cdot,T) - u(\omega_i,\cdot,T)||^2_{L^2(\mathcal{O})}$, $u_h(\omega_i,\cdot,T)$ and $u(\omega_i,\cdot,T)$ are the simulation with the path ω_i from the $\hat{M}$ paths. The order of convergence for the linear cases has been proved to be $O(h^{k+1})$ but there is no proof for the order of convergence in the nonlinear case. In the stochastic conservation laws and stochastic parabolic equations numerical experiments we use the strong order 1.5 method (6.3) for solving the SODEs.

Example 1 Constant-coefficient stochastic convection equation, $\mathcal{O} = [0,1]$

$$\begin{aligned} du + u_x dt &= budW_t \\ u(\omega, x, 0) &= \sin(x). \end{aligned}$$

For the numerical solution we used $(\hat{M}, CFL) = (5000, 0.01)$ for $K = 1$ and $(\hat{M}, CFL) = (2500, 0.001)$ for $K = 2$. As we can observe the order of convergence is close to the optimal order of convergence in both cases.

Example 2 Constant-coefficient stochastic convection equation with impulse initial condition. We show the results with the TVB 3% and BDF limiters in Table 6.2. The DG solutions for the stochastic equation resemble the results for the non limiter and limited solutions from the deterministic case. The TVB 3% had a smaller error for most of the N although the order of convergence of the BDF is greater. In general, the limiters proposed in this thesis did not have a great improvement in the

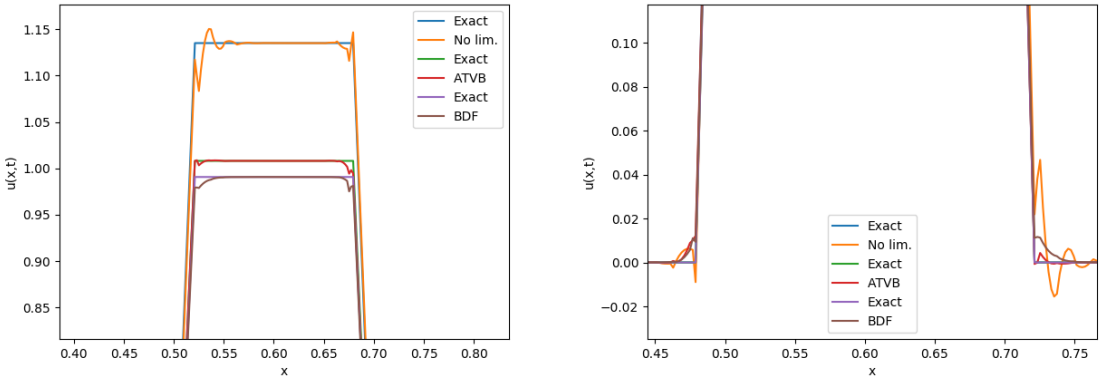
	$K=1$			$K=2$		
N	e_2	order	ν	e_2	order	ν
10	1.33×10^{-2}	-	1.62×10^{-6}	1.03×10^{-3}	-	1.55×10^{-8}
20	4.01×10^{-3}	1.73	1.45×10^{-7}	8.96×10^{-5}	3.53	1.16×10^{-10}
40	1.10×10^{-3}	1.87	1.10×10^{-8}	1.91×10^{-5}	2.23	5.19×10^{-12}
80	2.73×10^{-4}	2.01	6.76×10^{-10}	2.44×10^{-6}	2.96	8.62×10^{-14}
Time	$82.76s$			$273.40s$		

Table 6.1: Error stochastic convection equation sine initial condition with $T=0.1$.

solution, they have similar results to the original limiters. In Figures 6.1a and 6.1b the top and bottom parts of the discontinuity are shown for the DG solution without limiter and the DG solutions with the ATVB and BDF limiters.

	TVB 3%			BDF		
N	e_2	order	ν	e_2	order	ν
10	4.54×10^{-2}	-	2.69×10^{-5}	1.29×10^{-1}	-	2.15×10^{-4}
20	2.11×10^{-2}	1.11	5.80×10^{-6}	4.58×10^{-2}	1.50	2.79×10^{-5}
40	9.46×10^{-3}	1.15	2.20×10^{-6}	1.11×10^{-2}	2.04	1.66×10^{-6}
80	2.37×10^{-3}	2.00	1.92×10^{-7}	2.10×10^{-3}	2.41	5.70×10^{-8}
Time	$653.24s$			$654.04s$		

Table 6.2: Error stochastic convection equation square initial condition with $K=2$ and $T=0.1$.



(a) Top part of discontinuity

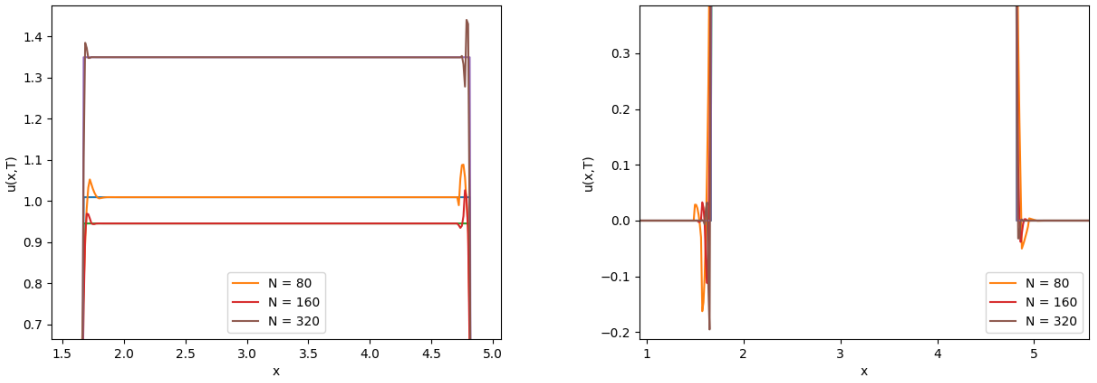
(b) Bottom part of discontinuity

Figure 6.1: Realization of the stochastic conservation law $K=2$, $N=160$, $T=0.1$

Example 3. Constant coefficient stochastic convection equation using Gauss-Lobatto quadrature. This algorithm is much more computationally expensive. For all the examples $b=1$. In the Table 6.3 for $K=1,2$ the order of convergence is below the optimal order of convergence, but the results are still acceptable. For $K=3$ we used $\hat{M}=100$ and for $K=4$, $\hat{M}=10$. In Figures 6.2a and 6.2b the solutions for $N=80,160,320$ using three points for the integration.

	$K=3$			$K=4$		
N	e_2	order	ν	e_2	order	ν
10	2.53×10^{-3}	-	3.61×10^{-7}	3.20×10^{-4}	-	2.20×10^{-8}
20	1.29×10^{-3}	0.97	1.07×10^{-7}	8.60×10^{-5}	1.90	1.59×10^{-9}
40	6.13×10^{-4}	1.07	2.19×10^{-8}	2.45×10^{-5}	1.81	1.22×10^{-10}
80	7.80×10^{-5}	2.97	3.92×10^{-10}	5.72×10^{-6}	2.10	5.59×10^{-12}
Time	$384.5s$			$721.51s$		

Table 6.3: Error stochastic convection equation, Guass-Lobatto quadrature $T=0.1$.



(a) Top part of discontinuity

(b) Bottom part of discontinuity

Figure 6.2: Realizations of the stochastic conservation law, Guass-Lobatto quadrature $K=3$, $T=0.1$

Example 4 Stochastic Burgers equation with additive noise

$$du = -\left(\frac{u^2}{2}\right)_x dt + bdW_t,$$

with the sine initial condition (2.37). In Table 6.4 the error and orders are shown time $T=0.1$, prior to the shock. Note that for both K the order of convergence is almost $K+1$. In the literature there is still no error estimate for the nonlinear case [47]. In Figures 6.3a and 6.3c we can compare the effectiveness of the BDF and MBDF limiters for a realization of the solution. In this case is more clear the improvement of the MBDF limiter over the BDF limiter, the diffusive behavior of the BDF limiter is overcome.

6.3 Local discontinuous Galerkin method

The general form of the stochastic convection-diffusion partial differential equation with multiplicative noise is

$$du = \{Du_{xx} - f(u)_x\}dt + g(\cdot, x, t, u, u_x)dW_t, \quad (x,t) \in \mathcal{O}\times(0,T) \tag{6.10}$$

	$K=1$			$K=2$		
N	e_2	order	ν	e_2	order	ν
10	1.39×10^{-2}	-	1.27×10^{-6}	6.97×10^{-4}	-	8.05×10^{-9}
20	4.32×10^{-3}	1.69	1.82×10^{-7}	1.04×10^{-4}	2.75	2.20×10^{-10}
40	9.95×10^{-4}	2.12	1.65×10^{-8}	1.45×10^{-5}	2.84	3.60×10^{-12}
80	2.87×10^{-4}	1.80	2.13×10^{-9}	1.97×10^{-6}	2.88	2.39×10^{-13}
Time	$41.67s$			$378.19s$		

Table 6.4: Error stochastic Burgers' equation sine initial condition with $T=0.1$.

we assume the same hypothesis on f and g as in the previous section. As for the deterministic case we have to rewrite the previous equation as a first order system and then apply the DG method.

$$\begin{aligned} du &= [w_x-(f(u))_x]dt+g(\cdot,x,t,u,q)dW_t, \quad (x,t)\in\mathcal{O}\times(0,T] \\ q &= u_x, \qquad\qquad \text{in } \Omega\times\mathcal{O}\times(0,T] \\ w &= Dq \qquad\qquad (x,t)\in\mathcal{O}\times(0,T]. \end{aligned}$$

We look for approximations $u_h, v_h, w_h \in V_h^K$ for each $(\omega,t)\in\Omega\times(0,T]$ of u,v,w, in the weak sense. Recall that at the inter-elements the solution is not properly defined. Thus, we consider the numerical fluxes $\hat{u},\hat{q}$ and $\hat{f}$, for $\hat{u}$ and $\hat{w}$ the midpoint is used

$$\hat{u}_j=\frac{u_h(x_j^-)+u_h(x_j^+)}{2}$$

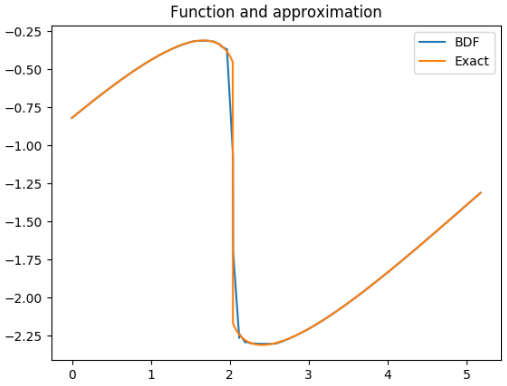
$$\begin{aligned} \int_{I_j}\phi_1 du_h(\omega,x,t) &= \left\{\int_{I_j}(-w_h+f(u_h))\phi_1'dx+\hat{w}_{j+1}\phi_{1,j+1}^- -\hat{w}_j\phi_{1,j}^+-\hat{f}_{j+1}\phi_{1j+1}^-+\hat{f}_j\phi_{1,j}^+)\right\}dt \\ &+ \int_{I_j} g(\omega,x,t,u_h,q_h)\phi_1(x)dxdW_t \qquad (6.11) \\ \int_{I_j} q_h(\omega,x,t)\phi_2 dx &= -\int_{I_j} u_h\phi_2'dx+\hat{u}_{j+1}\phi_{2,j+1}^- -\hat{u}_j\phi_{2,j}^+ \qquad (6.12) \\ \int_{I_j} w_h(\omega,x,t)\phi_3 dx &= \int_{I_j} Dq_h\phi_3 dx \qquad (6.13) \end{aligned}$$

where the notation $u_{h,j}^+$ is understood as $u_{h,j}^+=u_h(x_j^+)$ for $u_n, w_h, q_h, \phi_1, \phi_2$ and ϕ_3. Once again we consider that the basis at each I_j in V_h^K are the Legendre polynomials P_l, the LDG approximations u_h, q_h and w_h are then

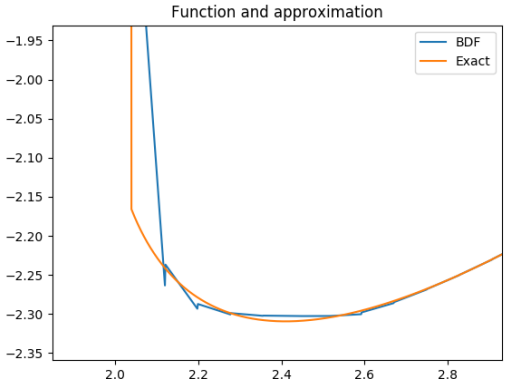
$$u_h(\omega,x,t)|_{I_j}=\sum_{l=0}^{K}u_j^l(\omega,t)P_l(\xi_j(x)), \quad q_h(\omega,x,t)|_{I_j}=\sum_{l=0}^{K}q_j^l(\omega,t)P_l(\xi_j(x))$$

and

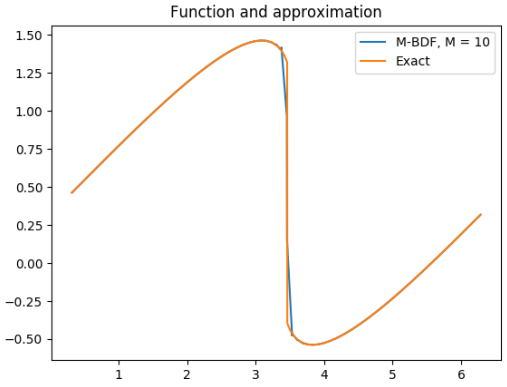
$$w_h(\omega,x,t)|_{I_j}=\sum_{l=0}^{K}w_j^l(\omega,t)P_l(\xi_j(x)).$$



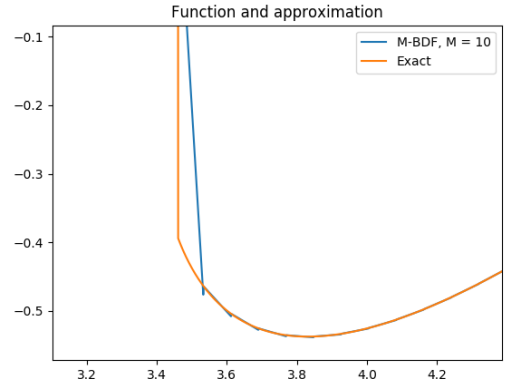
(a) Shock solution with BDF limiter



(b) Bottom part of discontinuity



(c) Shock solution with MBDF limiter, $M = 10$



(d) Bottom part of discontinuity

Figure 6.3: Reduction of oscillatory behaviour with BDF and MBDF limiters, realization of solutions.

Substituting LDG approximations in the system (6.11)-(6.13) and using the properties of the Legendre polynomials (3.1) and (3.2), we obtain for each $l \in \{0, 1, \ldots, K\}$

$$
\begin{aligned}
du_j^l &= \frac{2l+1}{h}\left[\int_{-1}^{1}(f(u_h)) - w_h)P_l'(\xi)d\xi + (-1)^l(\hat{f}_j - \hat{w}_j) - (\hat{f}_{j+1} - \hat{w}_{j+1})\right] \\
&+ \frac{2l+1}{2}\int_{-1}^{1} g(\omega, x, t, u_h, q_h)P_l d\xi dW_t \\
q_j^l &= \frac{2l+1}{h}\left[-\int_{-1}^{1} u_h(\xi, t)P_l'(\xi)d\xi + \hat{u}_{j+1} + (-1)^l\hat{u}_j\right] \\
w_j^l &= Dq_j^l
\end{aligned}
$$

these set of equations can be reduced to

$$
\begin{aligned}
du_j^l &= \frac{2l+1}{h}\left[\int_{-1}^{1}(f(u_h))-Dq_h)P_l'(\xi)d\xi+(-1)^l(\hat{f}_j-D\hat{q}_j)-(\hat{f}_{j+1}-D\hat{q}_{j+1})\right] \\
&+ \frac{2l+1}{2}\int_{-1}^{1} g(\omega,x,t,u_h,q_h)P_l d\xi dW_t \qquad (6.14)\\
q_j^l &= \frac{2l+1}{h}\left[-\int_{-1}^{1} u_h(\xi,t)P_l'(\xi)d\xi+\hat{u}_{j+1}+(-1)^l\hat{u}_j\right], \qquad (6.15)
\end{aligned}
$$

where $\hat{q}$ is the midpoint numerical flux. Equation (6.14) can be written as a SODE of the form (6.8) once u_h is determined q_h is determined as well by (6.15).

6.3.1 Numerical simulations

In each of the tables we show the L^2 error and order of convergence as explained in the previous numerical experiments. The theoretical order of convergence es $O(h^{k+1})$.

Example 1 Stochastic heat equation.

$$
\begin{aligned}
du &= u_{xx}dt+budW_t, \quad (\omega,x,t)\in\Omega\times\mathcal{O}\times[0,T]\\
u(x,0) &= \sin(x).
\end{aligned}
$$

In Tables 6.5 and 6.6 the errors and orders are shown for different values of the stochastic parameter b. For the first order approximations $K=1$, Table 6.5, the order of convergence is consistently below the theoretical order of convergence. For $K=2$, Table 6.6, the order behaves better and oscillates around $O(h^3)$. For $K=1$ we used $(\hat{M},CFL)=(500,0.01)$ and for $K=2$, $(\hat{M},CFL)=(250,0.001)$.

	$b=0.5$			$b=1$		
N	e_2	order	ν	e_2	order	ν
10	1.76×10^{-2}	-	8.87×10^{-6}	1.82×10^{-2}	-	1.94×10^{-5}
20	8.30×10^{-3}	1.09	2.03×10^{-6}	8.52×10^{-3}	1.10	5.18×10^{-6}
40	4.05×10^{-3}	1.04	4.63×10^{-7}	4.39×10^{-3}	0.96	1.28×10^{-6}
80	1.95×10^{-3}	1.05	1.06×10^{-7}	2.03×10^{-3}	1.11	2.53×10^{-7}
Time	$99.86s$			$47.66s$		

Table 6.5: Error of stochastic heat equation with sine initial condition $T=0.1$ and $D=1$

	$b=0.5$			$b=1$		
N	e_2	order	ν	e_2	order	ν
10	5.43×10^{-4}	-	2.27×10^{-8}	5.69×10^{-4}	-	2.25×10^{-8}
20	6.66×10^{-5}	3.03	1.27×10^{-10}	6.99×10^{-5}	3.03	2.88×10^{-10}
40	1.70×10^{-5}	1.97	8.15×10^{-12}	2.22×10^{-5}	1.65	2.93×10^{-11}
80	9.56×10^{-7}	4.15	2.77×10^{-14}	9.92×10^{-7}	4.49	6.88×10^{-14}
Time	$497.51s$			$490.42s$		

Table 6.6: Error stochastic heat equation with sine initial condition $T=0.1$, and $K=2$

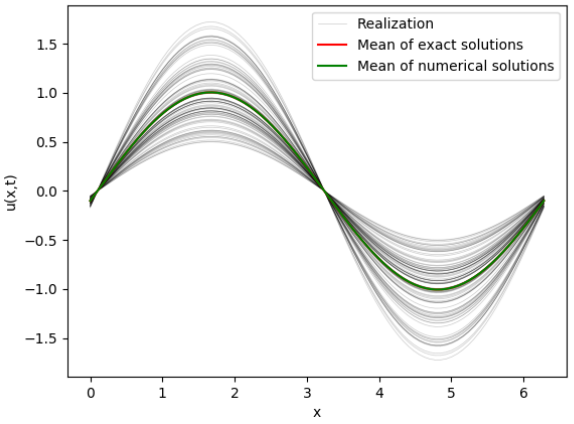


Figure 6.4: Realizations and mean of numerical and exact solutions of the stochastic linear convective dominated equation $b = 1$, $T = 0.1$.

Example 2 Stochastic linear convection-diffusion.

$$\begin{aligned} du &= (Du_{xx} - u_x)dt + budW_t, \quad (\omega, x, t) \in \Omega \times \mathcal{O} \times [0, T] \\ u(x, 0) &= \sin(x), \end{aligned}$$

the solution of this equation is

$$u(\omega, x, t) = \sin(x - t) \exp\left(bW_t - \frac{b^2 t}{2} - Dt\right).$$

In Tables 6.7 and 6.8 the errors and orders are shown for $b = 1$ and different values of the diffusion parameter, $D = 1$ and $D = 0.001$. In both cases the order behaves almost like the theoretical order of convergence, $O(h^3)$. For $K = 1$ we used $(\hat{M}, CFL) = (500, 0.01)$ and for $K = 2$, $(\hat{M}, CFL) = (250, 0.001)$. In Figure 6.4 we can see realizations and the numerical and theoretical mean for the solution of this example. Clearly, the means are indistinguishable, graphically showing that the solution converges as well in the sense of weak simulation.

	$K = 1$			$K = 2$		
N	e_2	order	ν	e_2	order	ν
10	1.49×10^{-2}	-	1.34×10^{-5}	5.54×10^{-4}	-	2.36×10^{-8}
20	5.08×10^{-3}	1.55	1.42×10^{-6}	7.35×10^{-5}	2.91	4.82×10^{-10}
40	1.54×10^{-3}	1.72	1.44×10^{-7}	2.69×10^{-5}	1.45	6.94×10^{-11}
80	3.93×10^{-4}	1.97	1.07×10^{-8}	1.00×10^{-6}	4.75	8.03×10^{-14}
Time	$452.71s$			$266.46s$		

Table 6.7: Error stochastic linear convection-diffusion equation sine initial condition $T = 0.1$, $D = 1$

Conclusions. In this section, we formulate the DG method and the LDG method for stochastic conservation and stochastic parabolic problems, respectively. For stochastic conservation laws, we solved linear and nonlinear problems using the DG method and constructed the stable entropy DG

	$K=1$			$K=2$		
N	e_2	order	ν	e_2	order	ν
10	1.52×10^{-2}	-	1.41×10^{-5}	7.76×10^{-4}	-	5.41×10^{-8}
20	4.33×10^{-3}	1.82	1.20×10^{-6}	1.24×10^{-4}	2.64	1.56×10^{-9}
40	1.12×10^{-3}	1.95	7.76×10^{-8}	2.25×10^{-5}	2.47	5.07×10^{-11}
80	2.63×10^{-4}	2.09	4.18×10^{-9}	1.58×10^{-6}	3.83	2.66×10^{-13}
Time	$461.51s$			$270.60s$		

Table 6.8: Error stochastic convective dominated equation sine initial condition $T=0.1$, $D=0.001$

formulation. In these examples, we were able to explore the order of convergence, the application of limiters to control oscillations, and the use of stable entropy solutions. The inclusion of limiters for the linear convection equation with the initial impulse condition did not show much improvement between the standard limiters and the ATVB and M-BDF. For Burgers' stochastic equation with $N=80$, the M-BDF showed better results than the BDF results. The numerical results obtained are promising to continue with the theoretical study of limiters and stable solutions of entropy. For limiters, we would first need to extend the definition of TVB and other properties to the stochastic case. The theoretical properties of the stable entropy solutions would require an extension of the conservative entropy flows (5.7) and the stable entropy flows (5.9) for the stochastic conservation laws. It may be the case that it is sufficient to ensure deterministic properties almost certainly. The stochastic parabolic solved are the stochastic heat equation and the linear stochastic diffusion-convection equation, including the convection-dominated case. For these examples, we study the convergence for different magnitudes of the parameter b and the effects of dominated convection. In both cases, the order of convergence is as expected. For future work we are interested in considering more examples and more initial conditions. In particular, the computation of the solution for non-linear equations becomes difficult, specially after shocks or non-smooth initial conditions due to the evaluation of the integral of the Brownian motion.

6.4 Algorithms

For the solution of the stochastic problems we use $\hat{M}$ times Algorithm 2 to compute the Monte Carlo error. In each iteration, instead of computing u_n with the RK algorithm, we use Algorithm 14 for the solution of the SODEs. The functions f, f' and $\hat{f}$ are the flux, its derivative and the numerical flux. The functions F and G vary from problem to problem. For conservation laws define F as Algorithm 4 for the nodal DG and as Algorithm 13 for the modal DG . For the LDG define F as Algorithm 5. For each Algorithm 4, 5 and 13 it is necessary to include their particular arguments. The function G is a function returning a vector with the $i-$th entry

$$[G]_i = \frac{2l+1}{2}\left(\int_{-1}^{1} g(\omega,\xi,t,u_h)P_l(\xi)d\xi\right),$$

with g defined for each particular problem. Due to its simplicity we do not include an algorithm for it. The relationship between τ and h is $\tau = CFL\delta^{\frac{3}{2}}$. The limiter option should only be included for the DG method. We do not include a particular algorithm for the adaptive case, the modification is done as in Algorithm 10.

Algorithm 14 SODE system solver order 1.5

procedure SODE_SOLVER(K, u_0, τ, h, T, f, f', $\hat{f}$, N, F, G, limiter, type of limiter, Z)
  Define $dt := \tau$ and vector solution u.
  $n = 1$
  while $dt = n\tau < T$ **do**
    Let N_1 and N_2 be a realization of i.i.d. standard normal.
    Define

$$\Delta Z = \left(N_1 + \frac{N_2}{\sqrt{3}}\right)\frac{\tau^{\frac{3}{2}}}{2}.$$

    Define $F_{eval} = F(u_{n-1}, f, fp, \hat{f}, N, h)$ and $G_{eval} = G(u_{n-1}, b)$.
    Compute

$$\begin{aligned}
\gamma_{\pm} &= u_0 + F_{eval}\tau \pm G_{eval}\sqrt{\tau} \\
G_{\pm} &= G(\gamma_{\pm}, b) \\
F_{\pm} &= F(\gamma_{\pm}, f, fp, \hat{f}, N, h) \\
\phi_{\pm} &= \gamma_{+} \pm G(\gamma_{+})\sqrt{\tau} \\
G(\phi_{\pm}) &= G(\phi_{\pm}, b)
\end{aligned}$$

    Update u_0 using (6.3)

$$\begin{aligned}
u_n &= u_{n-1} + F_{eval}\tau + G_{eval}N_1 \\
&+ \frac{1}{4\sqrt{\tau}}(G_{+} - G_{-})(N_1^2 - \tau) \\
&+ \frac{1}{4}(F_{+} - 2F_{eval} + F_{-})\tau \\
&+ \frac{1}{2\tau}(G_{+} - 2G_{eval} + G_{-})(N_1\tau - \Delta Z) \\
&+ \frac{1}{4\tau}(G(\phi_{+}) - G(\phi_{-}) - G_{+} - G_{-})\left(\frac{1}{3}N_1^3 - \tau\right)N_1.
\end{aligned}$$

    if limiter == True **then**
      Apply to u_n the type of limiter selected with its corresponding parameters Z.
    Update time step $dt += \tau$
  return Solution u_0.

Conclusions and future work

Our original objective, was to obtain an efficient discontinuous Galerkin formulation for stochastic parabolic conservation laws with dominant convection or degenerate diffusion. For that, we studied the discontinuous and local discontinuous Galerkin methods for hyperbolic and parabolic conservation laws as well as numerical techniques like limiters and entropy stable discretizations which have been recently proposed in the literature in the deterministic case. In such direction, we made two contributions to the development of the DG method. First, numerically implement the use of limiters and the nodal entropy stable discretization for stochastic conservation laws. Second, we propose modified limiters which we named as the ATVB and M-BDF limiters. The improvements of the new limiters were clear for the deterministic case, the error decreased faster and the solutions were less oscillatory and less diffusive when compared to the original limiters.

The contributions can be summarized as follows:

- Construct two limiters: The ATVB method serves as a starting point to develop future limiters that do not require the constant M of the TVB limiter, thus eliminating a heuristic calculation. The M-BDF limiter considers the modified minmod function to alleviate the diffusive solutions of the BDF limiter.
- Use the entropy stable formulation for the stochastic case. A great advantage that this formulation provides, its simplicity for using high order polynomial, in the deterministic examples we included up to $K = 5$ and for the stochastic case up to $K = 4$. In addition, the inclusion of the parameter ϵ in the entropy stable flux helps to decrease oscillations without the use of limiters.
- Validate the use of limiters and entropy stable solutions for stochastic conservation laws.

The future work can be synthesized in the following list:

- Study more techniques for the automatic switching between M_s and M_b, to avoid the estimation of M in the TVB limiter. Theoretically studying how to better choose M_s and M_b.
- Implement a more general formulation of the nodal entropy DG with limiters.
- Implement the entropy stable LDG.
- Study theoretically the properties of the proposed limiters and nodal entropy stable DG for the stochastic case.

- Implement the algorithms in C++ for better performance and parallelization.
- Work on higher-dimensional problems.
- Work on degenerated problems, which are convection-diffusion problems where the diffusion coefficient becomes zero in some regions.

Appendix A

Functional Analysis preliminaries

In this appendix we include the necessary background in functional analysis, in particular distribution theory, Hilbert and Sobolev spaces. The information in this appendix can be found in [15], for a wider panorama in distribution theory the reader is refereed to [18] and [68], further details in Hilbert and Sobolev spaces [1].

Definition A.0.1 (Norm,Normed space, Banach space)**.** Consider a vector space E, a map $||\cdot|| : E \to \mathbb{R}_+$, is called a *norm* if and only if

$$\begin{aligned} ||x|| &= 0 \iff x = 0 \\ ||\lambda x|| &= |\lambda|||x|| \text{ for all } \lambda \in \mathbb{R} \text{ , } x \in E \\ ||x+y|| &\leq ||x|| + ||y|| \text{ for all } x, y \in E. \end{aligned}$$

Then E is said to be a *normed space* and its norm is denoted by $||\cdot||_E$. Furthermore if E satisfies being complete respect to the following convergence

$$x_n \to x \text{ in E} \iff ||x_n - x||_E \to 0,$$

it is called a *Banach space.*

Definition A.0.2 (Bilinear form)**.** A bilinear form on a vector space V is a map $[\cdot,\cdot] : V \times V \to K$, where K is the field of scalars that for all $u, v, w, z \in V$ and for all $\lambda \in K$ verifies:

$$[\lambda(u+v), w] = \lambda([u, w] + [v, w])$$

$$[u, \lambda(v+w)] = \lambda([u, v] + [u, w])$$

Definition A.0.3 (Continuous bilinear form)**.** Let H be a Hilbert space. A bilinear form $[\cdot,\cdot] : H \times H \to \mathbb{R}$ is called *continuous* provided there exists $c > 0$ such that

$$[(u, v)] \leq C||u||_H||v||_H \text{ for all } u, v \in H.$$

Definition A.0.4 (Elliptic bilinear form)**.** A bilinear form B is called $H-$*elliptic,* or for short *elliptic* or *coercive*, provided form some $\alpha > 0$,

$$B(v, v) \geq \alpha||v||_H^2 \text{ for all } v \in H. \tag{A.1}$$

Every coercive bilinear form B, induces a norm via

$$||v||_B := \sqrt{B(v,v)} \tag{A.2}$$

This norm is called the *energy norm.*

Definition A.0.5 (Scalar product, Hilbert space)**.** Let H to be a lineal real space, a map $(\cdot,\cdot)_H : H \times H \to \mathbb{R}$ is a *scalar product* if

$$\begin{aligned} (x,y)_H &> 0 \iff x \neq 0 \\ (x,y)_H &= (y,x)_H \text{ for all } x,y \in H \\ (\lambda x + \mu y, z)_H &= \lambda(x,z)_H + \mu(y,z)_H \text{ for all } \lambda,\mu \in \mathbb{R},\ x,y,z \in H. \end{aligned}$$

If H is a Banach space respect to the norm associated to its scalar product

$$||x||_H = (x,x)_H^{\frac{1}{2}},$$

then H is a *Hilbert space.*

Definition A.0.6 (L^p spaces)**.** Let $p \in \mathbb{R}$ with $1 \leq p < \infty$, $O \subset \mathbb{R}^n$ and open set. The $L^p(O)$ spaces are defined as

$$L^p(O) = \{f | f : O \to \mathbb{R},\ f \text{ is measurable and } \int_O |f|^p dx < \infty.\}$$

Lemma A.0.1. *If $1 \leq p < \infty$, the set $L^p(O)$ is a Banach space with the norm*

$$||f||_{L^p} = \left[\int_O |f|^p dx\right]^{\frac{1}{p}}.$$

For $p = 2$, $L^2(O)$ we use the notation $||f|| := ||f||_{L^2(O)}$. In addition, this space is a Hilbert space with the scalar product

$$(f,g)_{L^2(O)} = \int_O fg dx.$$

We consider $\mathcal{D}(O)$ as the set of infinite differentiable functions in the set O. For $n \in \mathbb{N}$, $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, is called the multi index, we define

$$|\alpha| = \alpha_1 + \cdots + \alpha_n$$

and

$$\partial^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}},$$

where for $|\alpha| = 0$, $\partial^\alpha = 0$ is the identity.

Definition A.0.7 (Support)**.** Let $\phi : O \to \mathbb{R}$ a function, the *support* of ϕ denoted by $supp\phi$ is defined as follows

$$supp\phi = \overline{\{x \in O, \phi \neq 0\}} \cap O.$$

Definition A.0.8. A sequence $(\phi_n)_{n\in\mathbb{N}}$ in $\mathcal{D}(O)$ is said to be convergent to an element ϕ, iff (if and only if)

a) There is a compact set $K \subset O$ such that for all $n \in \mathbb{N}$, $supp\phi_n \subset O$.

b) For all $\alpha \in \mathbb{N}^n$, $\partial^\alpha \phi_n$ uniformly converges to $\partial^\alpha \phi$ in K.

Definition A.0.9 (Distribution)**.** A map $T : \mathcal{D}(O) \to \mathbb{R}$ is said to be a *distribution* in O iff,

a) T is lineal, i.e., for all $\lambda, \mu \in \mathbb{R}$, $\phi_1, \phi_2 \in \mathcal{D}(O)$,

$$T(\lambda\phi_1 + \mu\phi_2) = \lambda T(\phi_1) + \mu T(\phi_2)$$

b) If ϕ_n tends to ϕ in $\mathcal{D}(\mathcal{O})$, then $T(\phi_n)$ tends to $T(\phi)$.

We denote by $\mathcal{D}'(O)$ the space of *(Schwartz) distributions* in O. The usual notation for the application of distribution to a function is

$$T(\phi) = \langle T, \phi \rangle_{\mathcal{D}'(O),\mathcal{D}(O)}.$$

Definition A.0.10 (Derivative of a distribution)**.** Let $T \in \mathcal{D}(O)$, and $i = 1, \ldots, N$. The derivative $\frac{\partial T}{\partial x_i}$ is defined as

$$\langle \frac{\partial T}{\partial x_i}, \phi \rangle_{\mathcal{D}'(O),\mathcal{D}(O)} = -\langle T, \frac{\partial \phi}{\partial x_i} \rangle_{\mathcal{D}'(O),\mathcal{D}(O)}, \text{ for all } \phi \in D(O).$$

Definition A.0.11 (Weak derivative)**.** Let $u \in L^1_{loc}(O)$, with $L^1_{loc}(O)$ the set of locally integrable functions. A function $v_\alpha \in L^1_{loc}(O)$ is said to be the *weak derivative* of u if, for all $\phi \in \mathcal{D}(O)$,

$$\int_O u \partial^\alpha \phi dx = (-1)^{|\alpha|} \int_O v_\alpha \phi dx.$$

Definition A.0.12 (Sobolev spaces)**.** Let $1 \leq p < \infty$ and $m \in \mathbb{N}$. The *Sobolev space* $\mathcal{W}^{m,p}(O)$, is

$$\mathcal{W}^{(m,p)}(O) = \{u \in L^p(O) | \partial u^\alpha \in L^p(O) \ , \ 0 \leq |\alpha| \leq m\},$$

with ∂^α the weak derivative of u.

Definition A.0.13. The set $C_0^\infty(O)$ of the infinitely differentiable functions with compact support, is clearly a subset of $\mathcal{W}^{m,p}$ for any m and p. The closure of C_0^∞ as a subspace of $\mathcal{W}^{m,p}$ is denoted by $\mathcal{W}_0^{m,p}$.

For the case $m \in \mathbb{N}$, $p = 2$, denoted by $\mathcal{H}^m(O)$, we consider the following norm and seminorm

$$||v||_m = ||v||_{\mathcal{H}^m(O)} = (\sum_{|\alpha| \leq m} ||D^\alpha v||^2)^{\frac{1}{2}},$$
$$|v|_m = |v|_{\mathcal{H}^m(O)} = (\sum_{|\alpha| = m} ||D^\alpha v||^2)^{\frac{1}{2}},$$

Lemma A.0.2. *The space $\mathcal{H}^1(O)$ is a Hilbert space with the following scalar product,*

$$(v, w)_{\mathcal{H}^1(O)} = (v, w)_{L^2(O)} + \sum_{i=1}^{n} \left(\frac{\partial v}{\partial x_i}, \frac{\partial w}{\partial x_i} \right)_{L^2(O)}, \text{ for all } v, w \in \mathcal{H}^1(O).$$

We use the corresponding notation $\mathcal{H}_0^1(O)$.

Definition A.0.14 (Lipschitz boundary)**.** The boundary of O, ∂O is said to be *Lipschitz* if there is a constant $c > 0$ and a finite number M of local coordinate systems $(\hat{x}^m, x_N^m)$ and local maps Φ_m, defined in the sets

$$R_m = \{\tilde{x}^m \in \mathbb{R}^{N-1}, \tilde{x}^m = (x_1^m, \ldots, x_{N-1}^m), |x_i^m| \leq c\},$$

that are Lipschitz in their domain for $i = 1, \ldots, M$ and that

$$\partial\Omega = \cup_{m=1}^{M} \Gamma_m,$$

where

$$\Gamma_m = \{\hat{x}^m, x_N^m | x_N^m = \Phi_m(\hat{x}^m), \hat{x}^m \in R_m\}.$$

Intuitively this definitions says that a boundary is Lipschitz if it can be seen as the finite union of images of Lipschitz functions.

Theorem A.0.3 (Green Theorem)**.** *Let u, v be scalar functions and $w = (w_1, \ldots, w_n)$ a vector-valued function of $x \in R^n$. Then relation*

$$\int_O w \cdot \nabla v dx = \int_{\partial O} w \cdot n v ds - \int_O \nabla \cdot w v dx, \tag{A.3}$$

holds. Where n is the outward unit normal to ∂O.

When $w = \nabla u$ the formula becomes

$$\int_O \nabla u \cdot \nabla v dx = \int_{\partial O} \frac{\partial u}{\partial n} v ds - \int_O \Delta v dx, \tag{A.4}$$

where $\frac{\partial u}{\partial n} = \nabla u \cdot n$ is the exterior normal derivative of u on ∂O.

Theorem A.0.4. *(Sobolev inequality) Let O be a bounded domain in $\mathbb{R}$. Then $\mathcal{H}^1 \subset C(\overline{O})$, and there exists a constant $C = C(O)$ such that*

$$||v||_{C(O)} \leq C||v||_1, \text{ for all } v \in \mathcal{H}^1(O). \tag{A.5}$$

Theorem A.0.5 (The Lax-Milgram Theorem)**.** *Let V be a closed convex set in a Hilbert space H, and let $B : H \times H \to \mathbb{R}$ be an elliptic bilinear form. Then, for every linear functional $l \in H'$, the variational problem*

$$J(v) := \frac{1}{2} B(v, v) - \langle l, v \rangle, \tag{A.6}$$

has a unique minimum $v \in V$.

Appendix B

Probability

B.0.1 Probability review

Definition B.0.1 (Event)**.** An *event* is a subset of the sample space Ω. A relevant collection of events are the *sigma fields* $\mathcal{F}$, this collections satisfy the following properties:

a) The empty set, $\emptyset$ is in $\mathcal{F}$.

b) If A is in $\mathcal{F}$, its complement A^c is in $\mathcal{F}$.

c) If A_n is in $\mathcal{F}$ for $n = 1, 2, 3, \ldots,$, then $\bigcup_{n=1}^{\infty} A_n$ is in $\mathcal{F}$.

For a proper definition of a stochastic process, a *probability space* is required. A probability space is a triad $(\Omega, \mathcal{F}, \mathbb{P})$, with Ω a state-space, $\mathcal{F}$ a sigma algebra over Ω and $\mathbb{P}$ a *probability distribution (or measure)*

Definition B.0.2 (Probability distribution)**.** A function $\mathbb{P} : \mathcal{F} \to [0, 1]$, is said to be a *probability* if it satisfies:

a) $\mathbb{P}(\Omega) = 1$.

b) $0 \leq \mathbb{P}(A) \leq 1$, for any event A.

c) If $(A_n)_{n\in I}$ a finite or countably infinite pairwise disjoint collection, i.e., $A_j \bigcap A_k = \emptyset$ for $j \neq k$, then

$$\mathbb{P}(\bigcup_{n\in I} A_n) = \sum_{n\in I} \mathbb{P}(A_n).$$

Now we can properly define a random variable.

Definition B.0.3 (Random variable)**.** Given Ω and a probability distribution, a *random variable* is a function $X(\omega)$, defined on Ω taking values in $\mathbb{R}$.

The importance of random variables is that the "their probability " can be measured.

Definition B.0.4 (Distribution function)**.** The *distribution function* $F_X(x)$ of the random variable X is denoted by

$$F_X(x) = \mathbb{P}(B_x),$$

where

$$B_x = \{\omega : X(w) \leq x\}.$$

All random variables have a distribution function, there are two principle types of random variables: the discrete and the continuous. In this appendix we focus on continuous random variables, in this case, there may be a non-negative function $f_X(s)$ called the *density function* of X. If such function exists the distribution function $F(x)$ can be written in the form

$$F_X(x) = \int_{-\infty}^{x} f(x)dx. \tag{B.1}$$

An important example of such random variables is the *Normal density.*

Definition B.0.5 (Normal density)**.** For any constants $\mu \in \mathbb{R}$ and $\sigma^2 > 0$,

$$f(x) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathbb{1}(x)_{(\infty,\infty)} \tag{B.2}$$

We refer to this as $N(\mu, \sigma^2)$. The special case $\mu = 0$ and $\sigma^2 = 1$ is the *standard Normal Density*

Definition B.0.6 (Joint density function)**.** The pair of random variables (X, Y) is said to be *(jointly) continuous* with density $f(x, y)$ if, for all a, b, c, d

$$\mathbb{P}(a < X \leq b \cap c < Y \leq d) = \int_a^b \int_c^d f(x, y)dydx.$$

In particular this supplies the *joint distribution function* of X and Y, denotes by $F(x, y)$,

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{b} f(u, v)dvdu.$$

Definition B.0.7 (Marginal density)**.** The *marginal density* of X, is defined to be

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy.$$

The marginal of Y is defined analogously.

Definition B.0.8 (Independence of random variables)**.** The pair X and Y are *independent*, if for all x and y,

$$\mathbb{P}(X \leq x, Y \leq y) = F(x, y) = F_X(x)F_y(y) = \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y).$$

In particular for X and Y are independent random variables, if for all x and y,

$$f_{X,Y}(x, y) = f_X(x)f_Y(y).$$

Definition B.0.9 (Expectation)**.** Let X be a continuous random variable with density $f(x)$. The *expected value or mean* of X is denoted by $\mathbb{E}(X)$ and defined by

$$\mathbb{E}(X) = \int_{\mathbb{R}} xf(x)dx,$$

provided that $\int_{\mathbb{R}} |x|f(x)dx$ is finite.

Definition B.0.10. Let X be a continuous random variable and g real valued function , if $Y = g(X)$, then

$$\mathbb{E}(Y) = \int_{\mathbb{R}} g(x) f_X(x) dx$$

Definition B.0.11 (Moments)**.** The *moments and central moments* of a continuous random variable are defined as:

a) The $k-$th moment of X is $\mu_k = \mathbb{E}(X^k)$.

b) The $k-$th central moment of X is $\sigma_k = \mathbb{E}((X - \mathbb{E}(X))^k)$.

Definition B.0.12 (Covariance and correlation)**.** The *covariance* of X and Y is

$$cov(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))].$$

Their *correlation coefficient (or correlation)* is

$$\rho(X, Y) = \frac{cov(X, Y)}{\sqrt{var(X)var(Y)}},$$

where $var(X) = \sigma_2^2$.

Definition B.0.13 (Conditional density and expectation.)**.** If X and Y are jointly continuous then the *conditional density* of X given Y is defined by

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}, \text{ when } f_Y(y) > 0.$$

The *conditional expectation* of X given $Y = y$ is

$$\mathbb{E}(X|Y = y) = \int_{\mathbb{R}} x f_{X|Y}(x|y) dx,$$

provided the integral converges absolutely.

Definition B.0.14 (Moment generating function)**.** The *moment generation function* (mgf) of the random variable X is given by

$$M_X(t) = \mathbb{E}(e^{tX}),$$

for all real t where the expected value exists.

An important example of a mgf is for a normal distribution $N(\mu, \sigma^2)$. For $t \in \mathbb{R}$,

$$M_x(t) = \exp(\mu t + \frac{\sigma^2 t^2}{2}).$$

Definition B.0.15 (Characteristic function)**.** The function

$$\phi_X(t) = \mathbb{E}(\exp(itX)), \text{ for } t \in \mathbb{R}, \text{ where } i^2 = -1,$$

is called the *characteristic function,* or cf, of the random variable X.

Theorem B.0.1. *(Laplace-DeMovire) Let $X_1, \ldots, X_n$ be independent and identically distributed random variables, with*

$$\begin{aligned} \mathbb{P}(X_i = 1) &= p \\ \mathbb{P}(X_i = 0) &= q, \end{aligned}$$

for $p, q \geq 0$, $p + q = 1$. Define the sums

$$S_n := \sum_{i=1}^{n} X_i.$$

Then for all $-\infty < a < b < \infty$

$$\lim_{n\to\infty} \mathbb{P}\left(a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b \exp\left(-\frac{x^2}{2}\right).$$

B.0.2 Brownian motion and SDEs

Definition B.0.16. The $\sigma-$ algebra $\mathcal{W}(t) := \mathcal{F}(W(s)|0 \leq s \leq t)$ is called the *history* of the Brownian motion up to (and including) time t. The $\sigma-$ algebra $\mathcal{W}^+(t) := \mathcal{F}(W(s) - W(t)|s \geq t)$ is the *future* of the Brownian motion beyond time t.

Definition B.0.17. A family $\mathcal{U}(\cdot)$ of $\sigma-$algebras $\subset \mathcal{F}$ is called *nonanticipating* (with respect to $W(t)$) if

i) $\mathcal{U}(s) \subset \mathcal{U}(t)$ for all $t \geq s \geq 0$.

ii) $\mathcal{W}(t) \subset \mathcal{U}(t)$ for all $t \geq 0$.

iii) $\mathcal{U}(t)$ is independent of $\mathcal{W}^+(t)$ for all $t \geq 0$.

We refer to $\mathcal{U}(\cdot)$ as a *filtration*.

Definition B.0.18. A real-valued stochastic process G(t) is called *nonanticipating* (with respect to $\mathcal{U}(t)$) if for each time $t \geq 0$, $G(t)$ is $\mathcal{U}(t)$ measurable.

A stochastic process $G(t)$ being nonanticipating intuitively means that the values of the function G at a time t depend only on the values $W(u)$ for $0 \leq u \leq t$.

Definition B.0.19. A stochastic process $X(t)$ is said to be *progressively measurable* if, for every time t, the map $[0, t] \times \Omega \to \mathbb{R}$ defined by $(s, \omega) \to X_s(\omega)$ is $\mathcal{B}([0, t]) \times \mathcal{F}_t$ measurable. [57]

Theorem B.0.2. ***(Properties of Itô Integral)**. For all constants $a, b \in \mathbb{R}$ and for all $G, H \in \mathbb{L}^2(0, T)$, we have*

i) $\int_0^T aG + bH dW_t = a\int_0^T G dW_t + b\int_0^T H dW_t$.

ii) $\mathbb{E}\left(\int_0^T G dW_t\right) = 0$

iii) (Itô isometry) $\mathbb{E}\left(\left(\int_0^T G dW_t\right)^2\right) = \mathbb{E}\left(\int_0^T G^2 dt\right)$.

iv) $\mathbb{E}\left(\int_0^T G dW_t \int_0^T H dW_t\right) = \mathbb{E}\left(\int_0^T GH dt\right)$

Definition B.0.20. For $G \in \mathcal{L}^2(0,T)$, set

$$I(t) := \int_0^t G dW_t \quad (0 \le t \le T),$$

the *indefinite integral of* $G(\cdot)$.

Theorem B.0.3. *Assume X_t and Y_t are two Itô processes, then*

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y \rangle_t,$$

where $\langle X, Y \rangle$ is the quadratic covariation process of X and Y. If Y has bounded total variation

$$\langle X, Y \rangle = 0.$$

B.1 Approximation to the stochastic Wiener integral

Computation of the integral $\int_0^t W_s ds$, [47]. Define the process

$$v(s) = \frac{W_{t_n+\tau s} - W_{t_n}}{\sqrt{\tau}},$$

where $t_n = n\tau$, and W_t is a standard Brownian motion. Clearly the stochastic process $\{v(s), 0 \le s \le 1\}$ is a standard Brownian motion. In addition, define the random variables

$$\Delta W_n = \tau^{\frac{1}{2}} v(1), \quad \Delta Z_n = \tau^{\frac{3}{2}} \int_0^1 v(s) ds.$$

Then, if the variables $v(1)$ and $\int_0^1 v(s)ds$ are computed, the random variables ΔW_n and ΔZ_n can be obtained. These variables are the solution of the system

$$\begin{aligned} dx &= dv(s), \quad x(0) = 0, \\ dy &= x ds, \quad y(0) = 0. \end{aligned}$$

at time $t = 1$. We use a method of order 1.5 to compute the solution of the system. Denote by $x_k = x(x_k)$ and $y_k = y(s_k)$, for a partition $0 = s_0 < s_1 < \ldots < s_N = 1$, where

$$N = \lceil \tau^{-\frac{1}{3}} \rceil$$

and Then the system has the numerical solution

$$\begin{aligned} x_{k+1} &= x_k - (v(s_{k+1}) - v(s_k)) \\ y_{k+1} &= y_k + x_k \delta + \int_{s_k}^{s_{k+1}} (v(\theta) - v(s_k)) d\theta. \end{aligned}$$

The random variables $(v(s_{k+1}) - v(s_k))$ and $\int_{s_k}^{s_{k+1}} (v(\theta) - v(s_k)) d\theta$ are modeled using

$$v(s_{k+1}) - v(s_k) = \frac{\xi_{k,1}}{N^{\frac{1}{2}}}, \quad \int_{s_k}^{s_{k+1}} (v(\theta) - v(s_k)) d\theta = \frac{1}{2N^{\frac{3}{2}}} \left(\xi_{k,1} + \frac{1}{\sqrt{3}} \xi_{k,2} \right),$$

with $\xi_{k,1}$ and $\xi_{k,2}$ are independent normally $N(0,1)$.

The random variables ΔW_n and ΔZ_n can be reinterpreted as

$$\Delta W_n = W_{t_n+1} - W_{t_n}, \quad \Delta Z_n = \int_{t_n}^{t_{n+1}} (W_s - W_{t_n}) ds,$$

from which, $\int_0^t W_s ds$ can be computed.

Bibliography

[1] Robert A. Adams and John J.F. Fournier. *Sobolev Spaces*, volume 140 of *Pure and Applied Mathematics*. Elsevier, 2003.

[2] Douglas N. Arnold, Franco Brezzi, Bernardo Cockburn, and L. Donatella Marini. Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems. *SIAM Journal on Numerical Analysis*, 39(5):1749–1779, 2002.

[3] L Arnold. *Stochastic differential equations: theory and applications*. 2013. OCLC: 827267008.

[4] Ivo Babuska, Raúl Tempone, and Georgios E. Zouraris. Galerkin Finite Element Approximations of Stochastic Elliptic Partial Differential Equations. *SIAM Journal on Numerical Analysis*, 42(2):800–825, January 2004.

[5] Yuri Bakhtin. The Burgers equation with Poisson random forcing. *The Annals of Probability*, 41(4), July 2013.

[6] C. Bauzet, J. Charrier, and T. Gallouët. Convergence of flux-splitting finite volume schemes for hyperbolic scalar conservation laws with a multiplicative stochastic perturbation. *Mathematics of Computation*, 85(302):2777–2813, February 2016.

[7] Caroline Bauzet. On a time-splitting method for a scalar conservation law with a multiplicative stochastic perturbation and numerical experiments. *Journal of Evolution Equations*, 14(2):333–356, June 2014.

[8] Caroline Bauzet, Guy Vallet, and Petra Wittbold. THE CAUCHY PROBLEM FOR CONSERVATION LAWS WITH A MULTIPLICATIVE STOCHASTIC PERTURBATION. *Journal of Hyperbolic Differential Equations*, 09(04):661–709, December 2012.

[9] Caroline Bauzet, Guy Vallet, and Petra Wittbold. A degenerate parabolic–hyperbolic Cauchy problem with a stochastic force. *Journal of Hyperbolic Differential Equations*, 12(03):501–533, September 2015.

[10] Jeremie Bec and Konstantin Khanin. Burgers Turbulence. *Physics Reports*, 447(1-2):1–66, August 2007. arXiv: 0704.1611.

[11] Dietrich Braess. *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, 3 edition, 2007.

[12] L. Ridgway Scott C. Brenner. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag New York, 2008.

[13] Tianheng Chen and Chi-Wang Shu. Entropy stable high order discontinuous Galerkin methods with suitable quadrature rules for hyperbolic conservation laws. *Journal of Computational Physics*, 345:427–461, September 2017.

[14] V. Cheruvu and J. K. Ryan. A Multi-Wavelet Type Limiter for Discontinuous Galerkin Approximations. 2010.

[15] D. Cioranescu and Patrizia Donato. *An introduction to homogenization*. Number 17 in Oxford lecture series in mathematics and its applications. Oxford University Press, Oxford ; New York, 1999.

[16] Bernardo Cockburn. Discontinuous Galerkin Methods for Convection-Dominated Problems. In Timothy J. Barth and Herman Deconinck, editors, *High-Order Methods for Computational Physics*, Lecture Notes in Computational Science and Engineering, pages 69–224. Springer, Berlin, Heidelberg, 1999.

[17] Bernardo Cockburn and Chi-Wang Shu. Tvb runge-kutta local projection discontinuous galerkin finite element method for conservation laws. ii: General framework. *Mathematics of Computation*, 52:411–435, 1989.

[18] Gerrit van Dijk. *Distribution theory: convolution, fourier transform, and laplace transform*. De Gruyter graduate lectures. De Gruyter, Berlin, 2013. OCLC: ocn834977511.

[19] Roberto Díaz-Adame and Silvia Jerez. Convergence of time-splitting approximations for degenerate convection-diffusion equations with a random source. *Journal of Numerical Mathematics*, 0(0):000010151520200012, September 2020.

[20] Lawrence C. Evans. *Partial differential equations*. Number v. 19 in Graduate studies in mathematics. American Mathematical Society, Providence, R.I, 2nd ed edition, 2010. OCLC: ocn465190110.

[21] Lawrence C. Evans. *An introduction to stochastic differential equations*. American Mathematical Society, Providence, Rhode Island, 2013.

[22] Mircea Grigoriu. *Stochastic systems: uncertainty, quantification, and propagation*. Springer series in reliability engineering. Springer, London, 2012.

[23] Zhao Guo-Zhong, Yu Xi-Jun, and Wu Di. Numerical solution of the Burgers' equation by local discontinuous Galerkin method. *Applied Mathematics and Computation*, 216(12):3671–3679, August 2010.

[24] Richard Haberman. *Applied partial differential equations: with Fourier series and boundary value problems*. PEARSON, Boston, 5th ed edition, 2013.

[25] Jan S. Hesthaven and Tim Warburton. *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*. Number 54 in Texts in applied mathematics. Springer, New York, 2008. OCLC: ocn191889938.

[26] Desmond J. Higham. An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations. *SIAM Review*, 43(3).

[27] Ch Hirsch. *Numerical computation of internal and external flows: fundamentals of computational fluid dynamics*. Elsevier/Butterworth-Heinemann, Oxford ; Burlington, MA, 2nd ed edition, 2007. OCLC: ocn148277909.

[28] H. Holden, editor. *Stochastic partial differential equations: a modeling, white noise functional approach*. Universitext. Springer, London ; New York, 2nd ed edition, 2010. OCLC: ocn320493973.

[29] H. Holden, T. Lindstrøm, B. øksendal, J. Ubøe, and T.S. Zhang. The burgers equation with a noisy force and the stochastic heat equation. *Communications in Partial Differential Equations*, 19(1-2):119–141, January 1994.

[30] Helge Holden and Nils Henrik Risebro. *Conservation laws with a random source*. Number 1993,2 in Preprint series / Institute of Mathematics, University of Oslo Pure mathematics. Inst., Univ, Oslo, 1993. OCLC: 258236290.

[31] Willem Hundsdorfer and Jan Verwer. *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, volume 33 of *Springer Series in Computational Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2003.

[32] A. Jentzen and P. E. Kloeden. The Numerical Approximation of Stochastic Partial Differential Equations. *Milan Journal of Mathematics*, 77(1):205–244, December 2009.

[33] Silvia Jerez and Carlos Parés. Entropy Stable Schemes for Degenerate Convection-Diffusion Equations. *SIAM Journal on Numerical Analysis*, 55(1):240–264, January 2017.

[34] Guang Shan Jiang and Chi-Wang Shu. On a cell entropy inequality for discontinuous Galerkin methods. *Mathematics of Computation*, 62(206):531–531, May 1994.

[35] C. Johnson. *Numerical solution of partial differential equations by the finite element method*. Cambridge University Press, 1998.

[36] Jürgen Jost. *Partial differential equations*. Springer, New York, 2013. OCLC: 1120534993.

[37] Peter E. Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*. Number 23 in Applications of mathematics. Springer, Berlin ; New York, corr. 3rd print edition, 1999.

[38] Shigeaki Koike. A beginner's guide to the theory of viscosity solutions. 2014.

[39] P. Kotelenez. *Stochastic ordinary and stochastic partial differential equations: transition from microscopic to macroscopic equations*. Number 58 in Stochastic modelling and applied probability. Springer Science+Business Media, New York, 2008. OCLC: ocn172984133.

[40] Lilia Krivodonova. Limiters for high-order discontinuous Galerkin methods. *Journal of Computational Physics*, 226(1):879–896, September 2007.

[41] Vidyadhar G. Kulkarni. *Introduction to modeling and analysis of stochastic systems*. Springer texts in statistics. Springer, New York, 2nd ed edition, 2011. OCLC: ocn606740411.

[42] Siu Kwan Lam, Antoine Pitrou, and Stanley Seibert. Numba: A llvm-based python jit compiler. In *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*, pages 1–6, 2015.

[43] Mikel Landajuela. Burgers Equation. *BCAM Basque Center for Applied Mathematics*, 2011.

[44] Peter D. Lax. *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*. Number 11 in Regional conference series in applied mathematics. Soc. for Industrial and Applied Math, Philadelphia, Pa, 5. print edition, 2000.

[45] Jean-François Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*, volume 274 of *Graduate Texts in Mathematics*. Springer International Publishing, Cham, 2016.

[46] Randall J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhäuser Basel, Basel, 1992.

[47] Yunzhang Li, Chi-Wang Shu, and Shanjian Tang. A Discontinuous Galerkin Method for Stochastic Conservation Laws. *SIAM Journal on Scientific Computing*, 42(1):A54–A86, January 2020.

[48] Yunzhang Li, Chi-Wang Shu, and Shanjian Tang. A local discontinuous Galerkin method for nonlinear parabolic SPDEs. *ESAIM: Mathematical Modelling and Numerical Analysis*, 55:S187–S223, 2021.

[49] Pierre-Louis Lions, Benoît Perthame, and Panagiotis E. Souganidis. Scalar conservation laws with rough (stochastic) fluxes. *Stochastic Partial Differential Equations: Analysis and Computations*, 1(4):664–686, December 2013.

[50] Kun Liu and Béatrice M. Rivière. Discontinuous Galerkin methods for elliptic partial differential equations with random coefficients. *International Journal of Computer Mathematics*, 90(11):2477–2490, November 2013.

[51] J. David Logan. *Applied Mathematics*. John Wiley & Sons, June 2013. Google-Books-ID: uEVYKvk8EnIC.

[52] Simon J. A. Malham and Anke Wiese. An introduction to SDE simulation. *arXiv:1004.0646 [math]*, April 2010. arXiv: 1004.0646.

[53] Andreas Meister and Jens Struckmeier. *Hyperbolic partial differential equations: theory, numerics and applications*. Vieweg, Braunschweig, 1. ed edition, 2002. OCLC: 248081497.

[54] Hiroshi Nakazawa. Stochastic Burgers' equation in the inviscid limit. *Advances in Applied Mathematics*, 3(1):18–42, March 1982.

[55] Stanley Osher and Ronald P Fedkiw. Level Set Methods and Dynamic Implicit Surfaces, 0000. OCLC: 1159713773.

[56] P. Papadopoulos. *Introduction to the Finite Element Method*, volume 1. Panayiotis Papadopoulos, 2009.

[57] Andrea Pascucci. *PDE and martingale methods in option pricing*. Number 2 in Bocconi & Springer Series. Springer [u.a.], Milano, 2011. OCLC: 845771916.

[58] J. Proft and B. Rivière. Discontinuous galerkin methods for convection-diffusion equations with varying and vanishing diffusivity, int. *J. Numer. Anal. Mod*, pages 6–2009.

[59] Jianxian Qiu. and Chi-Wang Shu. Runge–Kutta Discontinuous Galerkin Method Using WENO Limiters. *SIAM Journal on Scientific Computing*, 26(3):907–929, January 2005.

[60] Deep Ray and Jan S. Hesthaven. An artificial neural network as a troubled-cell indicator. *Journal of Computational Physics*, 367:166–191, August 2018.

[61] W H Reed and T R Hill. Triangular mesh methods for the neutron transport equation. 10 1973.

[62] Béatrice Rivière. *Discontinuous Galerkin methods for solving elliptic and parabolic equations: theory and implementation*. Frontiers in applied mathematics. SIAM, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2008. OCLC: ocn226292048.

[63] V. Thomée S. Larsson. *Partial Differential Equations with Numerical Methods*. Springer Berlin Heidelberg, 2003.

[64] D. Serre. *Systems of conservation laws*. Cambridge University Press, Cambridge ; New York, 1999.

[65] Chi-Wang Shu. TVB uniformly high-order schemes for conservation laws. *Mathematics of Computation*, 49(179):105–105, September 1987.

[66] Tianheng Chen & Chi-Wang Shu. Review Article:Review of Entropy Stable Discontinuous Galerkin Methods for Systems of Conservation Laws on Unstructured Simplex Meshes. *CSIAM Transactions on Applied Mathematics*, 1(1):1–52, June 2020.

[67] David Stirzaker. *Stochastic processes and models*. Oxford University Press, Oxford ; New York, 2005.

[68] Robert S. Strichartz. *A guide to distribution theory and Fourier transforms*. Studies in advanced mathematics. CRC Press, Boca Raton, 1994.

[69] Zheng Sun. Convolution Neural Network Shock Detector for Numerical Solution of Conservation Laws. *Communications in Computational Physics*, 28(5):2075–2108, June 2020.

[70] Eitan Tadmor. The Numerical Viscosity of Entropy Stable Schemes for Systems of Conservation Laws. I. *Mathematics of Computation*, 49(179):91–103, 1987.

[71] Pablo A. Tassi and Carlos A. Vionnet. Discontinuous Galerkin Method For The One Dimensional Simulation Of Shallow Water Flows. *Mecánica Computacional*, 0(0):1474–1488, 2003.

[72] J. A. Trangenstein. *Numerical solution of hyperbolic partial differential equations*. Cambridge University Press, Cambridge ; New York, 2009. OCLC: ocn145388396.

[73] John B. Walsh. An introduction to stochastic partial differential equations. In P. L. Hennequin, editor, *École d'Été de Probabilités de Saint Flour XIV - 1984*, volume 1180, pages 265–439. Springer Berlin Heidelberg, 1986.

[74] Xinyue Yu and Chi-Wang Shu. Multi-layer perceptron estimator for the total variation bounded constant in limiters for discontinuous galerkin methods. *Division of Applied Mathematics Brown University*, 2021.

[75] Mengping Zhang and Chi-Wang Shu. AN ANALYSIS OF THREE DIFFERENT FORMULATIONS OF THE DISCONTINUOUS GALERKIN METHOD FOR DIFFUSION EQUATIONS. *Mathematical Models and Methods in Applied Sciences*, 13(03):395–413, March 2003.

[76] Zhongqiang Zhang. *Numerical methods for stochastic partial differential equations with white noise*. Springer Science+Business Media, New York, NY, 2017.

[77] Hongqiang Zhu, Yue Cheng, and Jianxian Qiu. A Comparison of the Performance of Limiters for Runge-Kutta Discontinuous Galerkin Methods. *Advances in Applied Mathematics and Mechanics*, 5(03):365–390, June 2013.

[78] B. K. Øksendal. *Stochastic differential equations: an introduction with applications*. Universitext. Springer, Berlin ; New York, 6th ed., corr. 4th print edition, 2007. OCLC: ocn166267310.