Partial and Stochastic Differential Equations: Theoretical and Numerical Aspects

Ludovic Goudenège, Adam Larat

Partial and Stochastic Differential Equations: Theoretical and Numerical Aspects

Ludovic Goudenège and Adam Larat

January 29, 2020

Abstract

A certain vision of PDEs and SDEs. How mathematics are thouroughly used in some very different ways to allow the transition from physics and reality to models and computational prediction.

Contents

1 Introduction

1.1 From reality to models. Computational prediction

1.1.1 Ordinary Differential Equations (ODEs)

Newton's second law: "*The rate of change of the linear momentum of an object is directly proportional to the external force acting on this object*".

$$\frac{d}{dt}\left(m\vec{\mathbf{v}}\right) = \overrightarrow{\mathrm{F}_{\mathrm{ext}}} \quad \Rightarrow \quad m\frac{d^2\vec{\mathbf{x}}}{dt^2} = \overrightarrow{\mathrm{F}_{\mathrm{ext}}}. \tag{1}$$

Case of pure gravity:

$$\begin{aligned} &\ddot{z} = -g, \qquad z(t=0) = z_0 \ \text{ and } \ \dot{z}(t=0) = v_z^0 \\ \Rightarrow \quad &z(t) = -\tfrac{1}{2}gt^2 + v_z^0 t + z_0. \end{aligned} \tag{2}$$

Definition 1.1 (*Ordinary Differential Equation*)

An Ordinary Differential Equation (ODE) is an equation involving a function (possibly vectorial) of one independant variable, $f : t \mapsto f(t)$, its derivatives $f', f'', \ldots, f^{(n)}, \ldots$ and the variable t itself.

Definition 1.2 (*Order of an ODE*)

An ODE is said to be of order $n \in \mathbb{N}^*$ when it involves only the first n derivatives of f and n is the smallest such number:

$$h(f, f', \ldots, f^{(n)}, t) = 0. \tag{3}$$

Definition 1.3 (*Cauchy Problem*)

*A **Cauchy Problem** is a mathematical problem maid of:*

- *An n-th order ODE, $n \in \mathbb{N}^*$,*
- *A set of initial conditions, which can be considered at $t_0 = 0$ without loss of generality:*

$$f(0) = f_0, \quad f'(0) = f_1, \quad \ldots, \quad f^{(n)}(0) = f_n. \tag{4}$$

Property 1.4

A n-th order ODE can always be reduced to a system of n first order ODEs.

Proof:

$$\left\{\begin{array}{ccc} u_0 & \leftrightarrow & f \\ u_1 & \leftrightarrow & f' \\ & \vdots & \\ u_{n-1} & \leftrightarrow & f^{(n-1)} \end{array}\right. \quad \Longrightarrow \quad \left\{\begin{array}{rcl} (u_0)' - u_1 & = & 0 \\ & \vdots & \\ (u_{n-2})' - u_{n-1} & = & 0 \\ h\left(u_0, u_1, \ldots, u_{n-1}, (u_{n-1})', t\right) & = & 0 \end{array}\right.$$

$$\Longleftrightarrow \qquad \mathscr{H}\left(\mathbf{U}, \mathbf{U}', t\right) = 0 \qquad \blacksquare$$

Definition 1.5 (*Explicit ODE*)

An ODE is said to be ***explicit*** *of order n when it can be written under the form:*

$$f^{(n)} = h((f, f', \ldots, f^{(n-1)}, t) \tag{5}$$

Then, using the same techniqe as in the previous proof, it can obviously be turned into a system of n first order explicit ODEs:

$$\mathbf{U}' = \mathscr{H}\left(\mathbf{U}, t\right), \qquad \left(\begin{array}{ccc} u_0 & \leftrightarrow & f \\ u_1 & \leftrightarrow & f' \\ & \vdots & \\ u_{n-1} & \leftrightarrow & f^{(n-1)} \end{array}\right) \tag{6}$$

Theorem 1.6 (*Cauchy-Lipschitz*)

Consider the initial value problem:

$$\left\{\begin{array}{rcl} \boldsymbol{U}'(t) & = & \mathscr{H}(\boldsymbol{U}(t), t), \quad t \in [-\varepsilon, \varepsilon], \varepsilon > 0, \\ \boldsymbol{U}(0) & = & \boldsymbol{U}_0. \end{array}\right. \tag{7}$$

Suppose $\mathscr{H}$ is ***continuous*** *in t and* ***Lipschitz-continuous*** *in $\mathbf{U}$ in a neighborhood of the initial value $\mathbf{U}_0$. Then, for some value of ε, there exists a unique solution $\mathbf{U}(t)$ of (7) on the interval $[-\varepsilon, \varepsilon]$.*

Corollary 1.7 (*Global existence*)

If $\mathscr{H}$ is moreover globally Lipschitz-continuous on the whole set of possible states $\mathbf{U}$, the solution is global ($t \in \mathbb{R}$).

String at rest between two walls: See Figures 1 and 2 for notations.

Static equilibrium of the infinitesimal section drawn in Figure 2 is given by the equality

$$\rho dx \vec{g} + \vec{\tau}(x) + \vec{\tau}(x + dx) = \vec{0}. \tag{8}$$

By projecting this vectorial relation on the horizontal coodinate, we get

$$|\tau_x(x)| = |\tau_x(x + dx)| = \tau. \tag{9}$$

Then, on the vertical axis, we come with

$$\begin{aligned} -\tau\frac{\mathcal{D}w}{\mathcal{D}x}(x)+\tau\frac{\mathcal{D}w}{\mathcal{D}x}(x+dx)-\rho dxg &= 0 \\ \Rightarrow \qquad \tau dx\frac{d^2w}{dx^2}(x)+\mathcal{O}(dx^2) &= \rho dxg \\ \Rightarrow \qquad \frac{d^2w}{dx^2} &= \kappa\rho g, \end{aligned} \tag{10}$$

where w is the vertical position of the string, ρ is the linear density of the rope, g is gravity and κ is the surface tension coefficient.

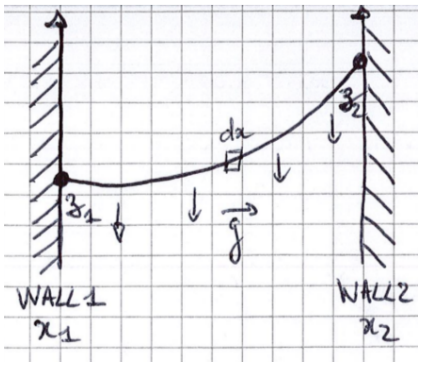


Figure 1: 1D string attached to two walls situated at abscissa x_1 and x_2 and subject to gravity.

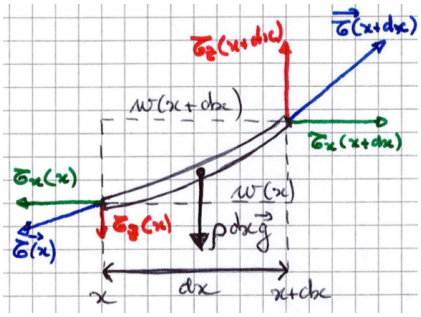


Figure 2: Detail of previous figure 1 on an infinitesimal section. Two forces apply on this section: its weight and the tension of the string on both sides of section.

This equation comes with two constraints which expresse the fact the string it attached on both walls:

$$w(x_1)=z_1, \quad w(x_2)=z_2. \tag{11}$$

The considered problem is not a **Cauchy problem**. However it is well-posed, as the solution is obviously a parabola between the two attachement point, which curvature is driven by the elasticity κ of the rope, its linear density ρ and the intensity g of the gravity.

1.1.2 Partial Differential Equations (PDEs)

A soap membrane at rest On each infinitesimal element of surface $dx \times dy$ apply two forces: it own weight and the tension of the surrounding membrane. By a reasoning analogous to the one for the string at rest (1D), one can prove that the inner membrane tension, also called **surface tension**, is proportional to the total local curvature:

$$\tau = \kappa\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}\right)=\kappa\Delta w.$$

Figure 3: Static string

The fundamental principle of statics then claims:

$$\kappa\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}\right)=\rho g \Leftrightarrow \Delta w = f. \tag{12}$$

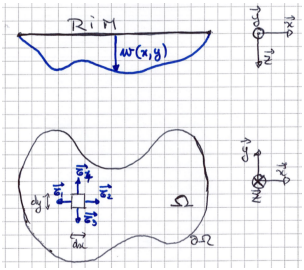


Figure 4: A soap membrane is attached to a rim. It hangs down due to gravity. rim is shown on the bottom figure. Top figure shows a lateral view of the rim. We vizualize the vertical displacement $w(x, y)$ of the membrane due to gravity.

Boundary conditions:

$$w = 0, \text{ on } \Gamma = \partial\Omega.$$

The mathematical problem combining this boundary condition with the Poisson equation (12) is called the **Dirichlet problem**.

Circular bubble: if the rim is a **circle** $\mathscr{C}(0, 1)$, then the solution is obviously

$$w(x, y) = \frac{f}{4}\left(x^2 + y^2 - 1\right)$$

- $\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = f$,
- wherever $x^2 + y^2 = 1, \quad w(x, y) = 0$.

What if the rim is now a potato shape (see Figure 5)? Or worse, it is not regular anymore, like a square or a fractal curve?

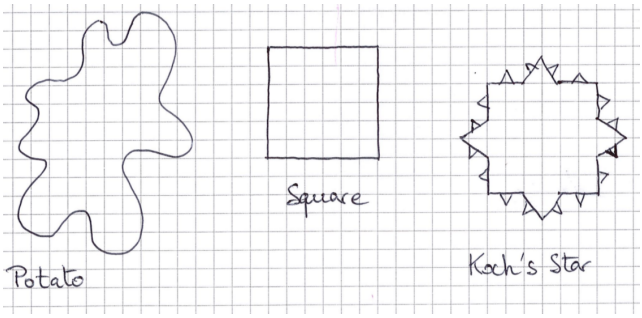


Figure 5: If solving the soap membrane problem seems fairly easy with a circular rim, the question is much harder with any shape, especially when it lacks regularity.

Definition 1.8 (*PDE*)

A Partial Differential Equation (PDE) is a differential equation involving an unknown multivariate (possibly vectorial) function u and its partial derivatives.

In the case of a scalar function, one can write:

$$\overrightarrow{\boldsymbol{\mathcal{F}}}(x_1, \dots, \mathbf{x}_d, u, \frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_d}, \frac{\partial^2 u}{\partial x_i \partial x_j}, \dots) = 0 \tag{13}$$

Definition 1.9 (*Order of a PDE*)

The order of a PDE is defined as the biggest order of the differential operator implied in the PDE. Sometimes this order can be reduced using various technics but there is no general method.

Definition 1.10 (*Quasi-Linear PDE*)

A PDE of order n is said to be ***quasi-linear*** *when it can be written into the form:*

$$\sum_{k=0}^{n} \sum_{i_1,\dots,i_k \in [\![1,d]\!]} A_{i_1 \dots i_k}(x_1, \dots, x_d, u) \frac{\partial^k u}{\partial x_1 \dots \partial x_k} = 0. \tag{14}$$

The coefficients $A_{i_1 \dots i_k}$ may depend on the space variables and u but not on its partial derivatives.

1.1.3 Computational prediction

What is a good numerical method? How do I rely on my computational prediction?

Approximate problem:

- Instead of considering the continuous problem, we discretize it and try to obtain an approximation of the exact solution at a certain number of

degrees of freedom (DoFs). The latter could be certain points in space, coefficients of a decomposition of the solution onto a smart chosen basis, etc.

The algorithm cost is technically reachable:

- Only a finite number of DoFs are available,
- The amount of calculation per DoF is finite and decent.

The continuous problem is well-posed:

- If the problem arising from modelling does not have a solution or have many of them, it is not worth developping a numerical method. One could obtain anything!

The numerical method converges toward the exact solution:

- One wants to prove that the accuracy of the result increases monotonically with the number of DoFs and that the result becomes exact in the limit on an infinite number on degrees of freedom.
- The method is accurate enough that a good level of prediction is reached within the limited amount of computational power available.

The numerical method is stable:

$$y' + y = 0,\ y(0) = 1 \quad \Rightarrow \quad y(t) = \exp(-t).$$

$$y(t + dt) = y(t) + dty'(t) + \cancel{\mathcal{O}(dt^2)} \quad \Rightarrow \quad y^{n+1} = y^n - dty^n = (1 - dt)y^n.$$

- $dt = 0$: $y(t) = 1 \ldots,$
- $dt \in]0, 1[$: monotone decrease toward zero, first order accurate scheme,
- $dt = 1$ or 2: $y^n \in \{-1, 0, 1\}, \forall n \in \mathbb{N}^*,$
- $dt \in]1, 2[$: oscillatory behavior. Numerician try to avoid this behavior because it can be quite dangerous:

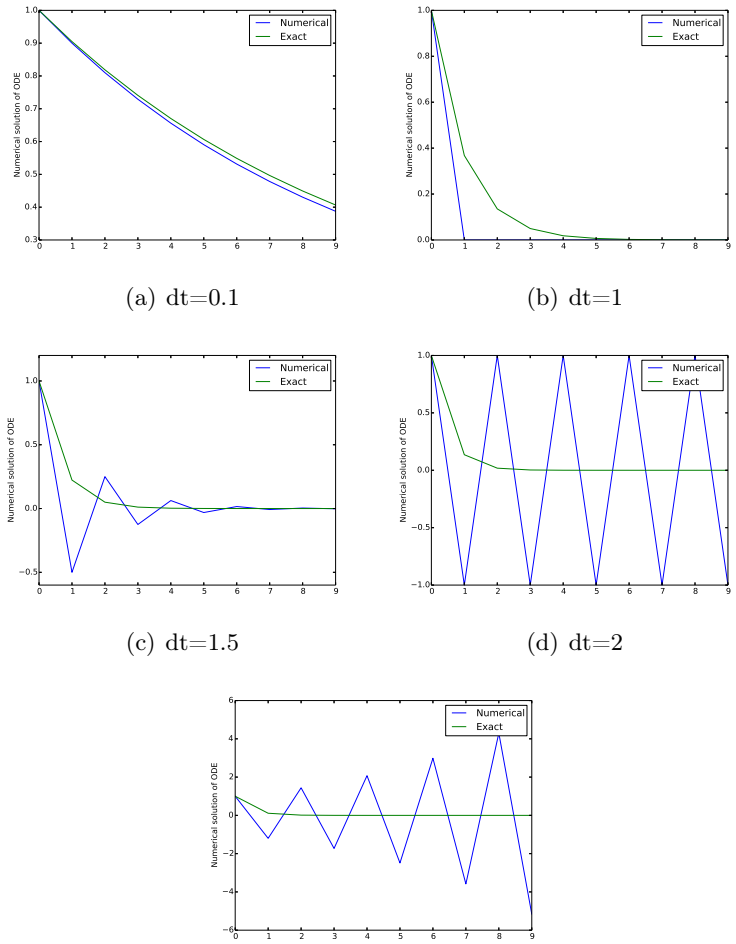
$$\begin{cases} y' + y = 0.1, t > 0, \\ y(0) = 1, \\ y > 0. \end{cases} \tag{15}$$

The last constraint of positivity is automatically verified by the continuous solution of the two first equations ($y(t) = \exp(t) + 0.1$). It

could however be violated by an oscillatory scheme. Imagine if this is an equation on the density or the temperature. What can you say about a result predicting negative energies?

On the other hand, if you look at the solution, we still have a good picture of the trend of the exact solution and engineers do employ those methods in industrial codes. For example, one could filter the first oscillating mode to recover a smoother solution. This is called **post-processing**.

- $dt > 2$: unstable: $|y^n|$ is exponentially increasing.



(a) dt=0.1

(b) dt=1

(c) dt=1.5

(d) dt=2

(e) dt=2.2

Figure 6: Stability for various dt.

1.2 PDE Classification

You may have already seen the classification of second order quasi-linear PDEs

$$A\frac{\partial^2 u}{\partial x^2} + B\frac{\partial^2 u}{\partial x \partial y} + C\frac{\partial^2 u}{\partial y^2} + D\frac{\partial u}{\partial x} + E\frac{\partial u}{\partial y} + Fu = 0, \tag{16}$$

A, B, C not simultaneously all null.

Lets consider the **conic section:**

$$AX^2 + BXY + CY^2 + DX + EY + F = 0. \tag{17}$$

We know that the nature of this curve of the plan is determined by the sign of $\Delta = B^2 - 4AC$:

- $\Delta > 0$: Hyperbola, (16) is said to be hyperbolic,
- $\Delta = 0$: Parabola, (16) is said to be parabolic,
- $\Delta < 0$: Ellipse, (16) is said to be elliptic.

This classification is very restrictive. It is nonetheless mostly the only way the classification of PDEs is taught. Fortunately, there is a much better and more general explanation.

In the general case, we can associate to a PDE a differential operator (say L) and rewrite the equation in the form $Lu = 0$. Computing the eigenvalues of this operator we can classify the PDE in the following sense:

- elliptic : The eigenvalues are all positive or negative,
- parabolic : The eigenvalues are all positive or negative, except one that is zero.
- hyperbolic : There is one positive value and all the rest are negative (or respectively the inverse).
- hybrid : Other cases.

1.3 On the necessity to weaken the notion of derivative

1.3.1 Linear scalar transport equation

$$\begin{cases} \partial_t u + a\partial_x u = 0, & x \in \mathbb{R}, t > 0 \\ u(x,0) = u_0(x), & x \in \mathbb{R}, u_0 \in \mathcal{C}^\infty \end{cases} \tag{18}$$

- $u(x,t) = u_0(x - at)$ is obviously the unique solution of this problem. Uniqueness comes from the linear caracter of the equation.

$$\partial_x u = u_0', \quad \partial_t u = -au_0' \quad \Leftarrow \quad \partial_t u + a\partial_x u = -au_0' + au_0' = 0.$$

- The solution is just a shift in time at velocity a of the initial profile, see Figure 7.

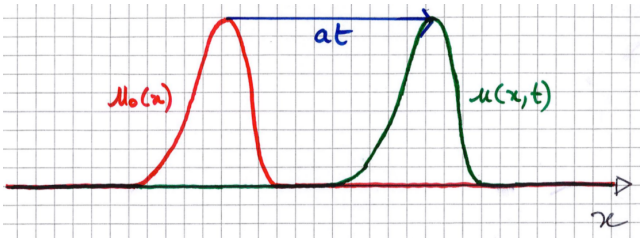


Figure 7: The initial profile $u_0(x)$ is simply advected at velocity a.

1.3.2 Non-Linear Burgers' equation

$$\begin{cases} \partial_t u + u\partial_x u = 0, & x \in \mathbb{R}, t > 0 \\ u(x,0) = u_0(x), & x \in \mathbb{R}, u_0 \in \mathcal{C}^\infty \end{cases} \tag{19}$$

- Now the velocity $a = u$ depends on the height of the solution.
- This is for example the behavior of a wave in the sea. In the middle of the ocean, the height of the perturbation is negligeable compared to the water depth and the propagation of the wave is nearly linear. When approaching the shore, we tend to a more non-linear regime, see figure 8.
- At some finite time T^*, the spatial derivative $\partial_x u$ becomes infinite and the wave breaks on the shore.
- In fact, the mathematical solution does not get multivalued. The discontinuity generated at time T^* start propagating with the flow. This is what is called a **shock**.

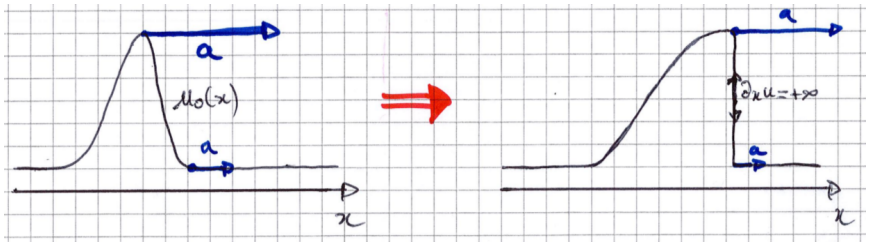


Figure 8: In the case of the Burgers' equation, the propagation of the profile depends on the height. The top of the wave propagates faster than its foot and the wave is suppose to break. In fact, at some point a discontinuity is produced ($\partial_x u = +\infty$) and starts to propagates in the domain. This is what is called a **shock**

We now come to a difficult mathematical problem. Even though the initial condition is maximum regular ($\mathcal{C}^\infty$), the solution of problem (19) becomes discontinuous in finite time and stays so for the rest of the time.

"*How do you define a solution of a differential problem when you know this solution is not differentiable itself?*"

Remark 1.11

In fact, we could ask the same question for the linear advection (18) *with an irregular initial condition:*

$$u_0(x) = \mathcal{H}(x) - \mathcal{H}(x-1) = (x > 0) - (x > 1).$$

Obviously $u_0(x-at)$ *is still a solution (and it is the only one). But under which framework? It is not differentiable everywhere!*

2 Some tools for the theoretical analysis of PDEs

2.1 Algebraical and topological duals

- Let $\mathbb{E}$ be a $\mathbb{K}$-vectorial space, $\mathbb{K}$ being a topological field, possibly a subfield of $\mathcal{C}$. A **linear form** on $\mathbb{E}$ is a morphism of vectorial spaces from $\mathbb{E}$ to $\mathbb{K}$:
$$f : \left\{ \begin{array}{ccc} \mathbb{E} & \longrightarrow & \mathbb{K} \\ x & \longmapsto & f(x) \end{array} \right. , \quad \text{f linear.} \tag{20}$$

- $\mathbb{E}^* = \{\text{f linear form on } \mathbb{E}\}$ is also a $\mathbb{K}$-vectorial space called the **algebraical dual** of $\mathbb{E}$.

- When $\mathbb{E}$ is of finite dimension, it can be proven that
$$\left\{ \begin{array}{ccc} \dim \mathbb{E} & = & \dim \mathbb{E}^* \\ \mathbb{E} & \sim & \mathbb{E}^* \\ \mathbb{E}^{**} & \equiv & \mathbb{E} \end{array} \right.$$

- If $\mathbb{E}$ is a **topological vectorial space** (with a topology compatible with that of the underlying field $\mathbb{K}$. *ie.* : the sum of two vectors is a continuous application from $\mathbb{E} \times \mathbb{E}$ to $\mathbb{E}$ and the scalar multiplication is a continuous application from $\mathbb{K} \times \mathbb{E}$ to $\mathbb{E}$.), one can define the notions of **neighborhood**, of **limit**, and therefore of **continuity**. Its **topological dual** $\mathbb{E}'$ is the set of **continuous linear forms**:
$$\left\{ \begin{array}{ccl} \mathbb{E}' & \subset & \mathbb{E}^* \\ \mathbb{E}' & = & \mathbb{E}^*, \text{ in finite dimension} \\ \mathbb{E}'' & \equiv & \mathbb{E}, \text{ for all Hilbert spaces, thanks to Riesz Theorem.} \end{array} \right.$$

2.2 Some functional spaces

Let $\Omega \subset \mathbb{R}^n$ be an open subset. It could be bounded or not. For each function on Ω, we define its **support** by:
$$\text{Supp}\,(f) = \{x \in \Omega \ / \ f(x) \neq 0\}\,. \tag{21}$$
Its support is compact when it is a compact subset of Ω.

On this open subset we define the following functional spaces:

- $\mathscr{D}(\Omega) = \{\varphi \in \mathcal{C}^\infty(\Omega) \ / \ \text{Supp}\,(\varphi) \text{ is compact.}\}$ is the space of **test functions**.

- $\mathscr{D}'(\Omega) = \{\text{linear continuous forms on } \mathscr{D}(\Omega)\}$ is the space of **distributions**.

- $\mathscr{S}(\Omega) = \left\{\varphi \in \mathcal{C}^\infty(\Omega) \ / \ \forall(\alpha,\beta) \in \mathbb{R}_+,\ \|x^\alpha \mathrm{D}^\beta \varphi\|_\infty < +\infty\right\}$. The support of φ is no more necessarily compact, but all derivatives of φ decay faster than any polynomial at infinity. This space is called the **Schwartz space**.

- $\mathscr{S}'(\Omega)$, the topological dual of $\mathscr{S}(\Omega)$ is the space of **tempered distributions**. It is essential for the extension of the Fourier transform to less regular functions.

- $\mathrm{L}^p(\Omega) = \left\{f : \Omega \to \mathbb{R} \ / \ \int_\Omega |f|^p \, dX < +\infty\right\},\ p-1 \in \mathbb{R}_+$.
  If Ω is bounded, the series of L^p spaces is decreasing:

$$\mathrm{L}^1 \supset \mathrm{L}^2 \supset \ldots \supset \mathrm{L}^\infty \tag{22}$$

  The topological dual of L^p is given by:

$$(\mathrm{L}^p)' = \mathrm{L}^q, \quad \text{with } \frac{1}{p} + \frac{1}{q} = 1,\ p,q > 1 \tag{23}$$

  In particular

$$\left(\mathrm{L}^2\right)' = \mathrm{L}^2. \tag{24}$$

- $\forall f \in \mathrm{L}^1_{loc}(\Omega)$, one can define the distribution $T_f \in \mathscr{D}'(\Omega)$ by:

$$\forall \varphi \in \mathscr{D}(\Omega), \quad T_f[\varphi] = \langle T_f, \varphi \rangle_{\mathscr{D}',\mathscr{D}} = \int_\Omega f\varphi \, dX \tag{25}$$

  - it is obviously linear,
  - For all sequence $(\varphi_n)_{n\in\mathbb{N}}$ converging to φ in $\mathscr{D}(\Omega)$, there exist K compact such that $\cup_n \mathrm{Supp}(\varphi_n) \subset K$ and

$$|\langle T_f, \varphi_n - \varphi \rangle| \leq \|f\mathbf{1}_K\|_1 \|\varphi_n - \varphi\|_\infty,$$

    which proves the continuous character of the linear form.

- Such a distribution is uniquely defined by f and $\mathrm{L}^1_{loc}(\Omega)$ can then be identified to a subset of $\mathscr{D}'(\Omega)$. In particular, $\mathscr{D}'(\Omega)$ contains $\mathscr{D}(\Omega)$ and all the L^p spaces, since by (22)

$$\mathrm{L}^1_{loc}(\Omega) \supset \mathrm{L}^p(\Omega), \quad \forall p \geq 1.$$

- A classification of these functional spaces is illustrated in Figure 2.2.

Remark 2.1

As one can see, $\mathscr{S}(\Omega)$ includes $\mathscr{D}(\Omega)$ and as a consequence, $\mathscr{D}'(\Omega)$ includes $\mathscr{S}'(\Omega)$. So one can say the following fact: the "smaller" the space

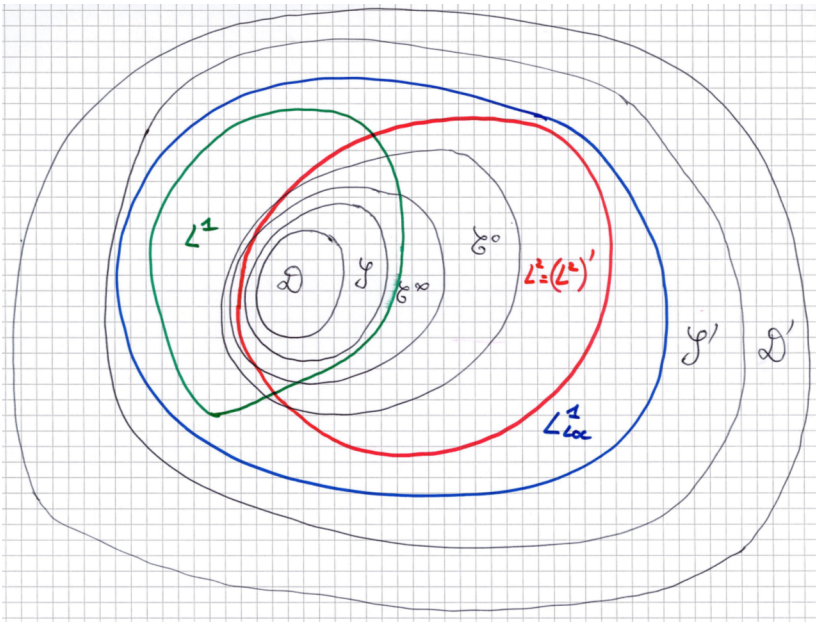


Figure 9: Classification of the functional spaces described above.

> *of test function, the "larger" its topological dual. This transformation occurs across a neutral element:* L^2.

2.3 Weak derivatives

2.3.1 Definition

- Let $T \in \mathscr{D}'(\Omega)$ a distribution. The partial derivative $\partial_{x_i} T$ is the distribution such that

$$\forall \varphi \in \mathscr{D}(\Omega), \quad \langle \partial_{x_i} T, \varphi \rangle_{\mathscr{D}',\mathscr{D}} = -\langle T, \partial_{x_i} \varphi \rangle_{\mathscr{D}',\mathscr{D}}. \tag{26}$$

  - $\partial_{x_i} T$ is a linear form on $\mathscr{D}(\Omega)$,
  - it is continuous: forall sequence $(\varphi_n)_{n\in\mathbb{N}}$ converging to φ in $\mathscr{D}(\Omega)$,

$$|\langle \partial_{x_i} T, \varphi - \varphi_n \rangle_{\mathscr{D}',\mathscr{D}}| = |\langle \mathrm{T}, \partial_{x_i}\varphi - \partial_{x_i}\varphi_n \rangle_{\mathscr{D}',\mathscr{D}}| \longrightarrow 0.$$

- If $f \in \mathcal{C}^1(\Omega)$, this definition matches exactly the definition you know. Indeed, $\partial_{x_i} f \in \mathcal{C}_0(\Omega) \subset \mathrm{L}^1_{loc}(\Omega)$, so that

$$\langle \partial_{x_i} f, \varphi \rangle_{\mathscr{D}',\mathscr{D}} = \int_\Omega \partial_{x_i} f \varphi \, dX = -\int_\Omega f \partial_{x_i}\varphi \, dX + \cancel{\int_{\partial\Omega} f\varphi} = -\langle f, \partial_{x_i}\varphi \rangle_{\mathscr{D}',\mathscr{D}}$$

  The partial derivative of the distribution associated to f is the distribution associated to the partial derivative of f.

2.3.2 Examples

- $f : x \longmapsto |x|$ is $\mathrm{L}^1_{loc}(\Omega)$, so $\forall \varphi \in \mathscr{D}(\Omega)$

$$\begin{aligned}\langle f', \varphi \rangle = -\langle f, \varphi' \rangle &= -\int_{\mathbb{R}} f\varphi' dx = \int_{\mathbb{R}_-} x\varphi' dx - \int_{\mathbb{R}_+} x\varphi' dx, \\ &= -\int_{\mathbb{R}_-} \varphi dx + \cancel{[x\varphi]_{-\infty}^{0}} + \int_{\mathbb{R}_+} \varphi dx - \cancel{[x\varphi]_{0}^{+\infty}}, \\ &= \int_{\mathbb{R}} \left(2\mathbf{1}_{\mathbb{R}_+} - 1\right) \varphi dx = \left\langle 2\mathbf{1}_{\mathbb{R}_+} - 1, \varphi \right\rangle_{\mathscr{D}', \mathscr{D}},\end{aligned}$$

because $(2\mathbf{1}_{\mathbb{R}_+} - 1) \in \mathrm{L}^1_{loc}(\Omega)$ and therefore

$$f' = 2\mathbf{1}_{\mathbb{R}_+} - 1.$$

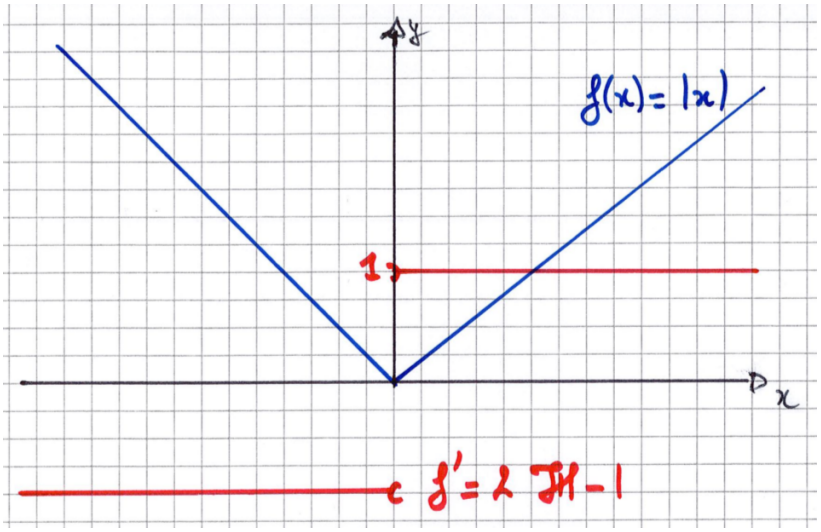


Figure 10: The derivative of the absolute value in the sense of the distribution is a discontinuous $\mathrm{L}^1_{loc}(\Omega)$ function, having the value of the strong derivative wherever the function is differentiable.

- $\mathcal{H} : x \longmapsto \mathbf{1}_{R_+}$, the **Heavyside function** is $\mathrm{L}^1_{loc}(\Omega)$, so $\forall \varphi \in \mathscr{D}(\Omega)$

$$\langle f', \varphi \rangle = -\langle f, \varphi' \rangle = -\int_{\mathbb{R}_+} \varphi' dx = +\varphi(0).$$

f' is then the distribution which to each test function φ associate its value at $x = 0$, $\varphi(0)$. This distribution is called the **Dirac distribution**:

$$f' = \delta_0.$$

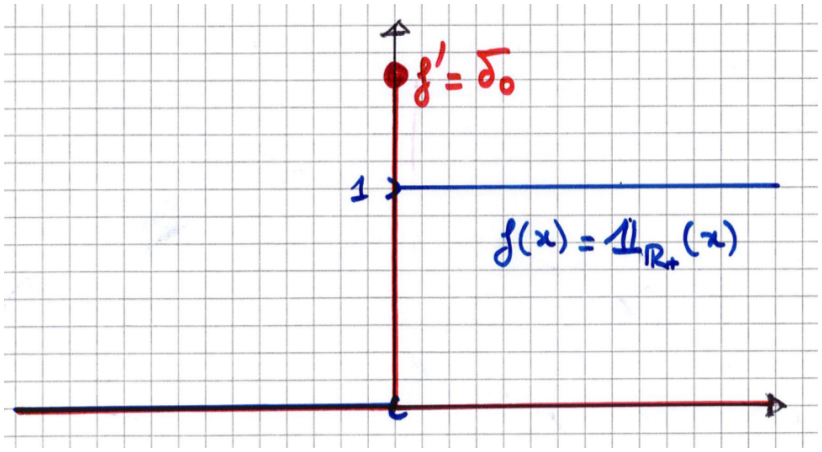


Figure 11: The derivative of the Heavyside function in the sense of the distribution is the Dirac delta function.

- Exercise: prove that δ_0 has no representant in $L^1_{loc}(\mathbb{R})$.

2.4 Weak formulation of a PDE

Let's come back to the soap membrane problem. Let Ω be a compact subset of $\mathbb{R}^2$ and $\mathring{\Omega}$ it open interior. The static shape of the membrane under simple

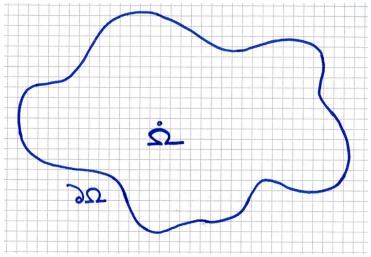


Figure 12: Domaine of study Ω for the Poisson equation $\Delta w = \bar{f}$. $\mathring{\Omega}$ is its open iterior, $\partial\Omega$ is the outside boundary.

gravity is the solution of the following problem:

$$\begin{cases} \Delta w &= \bar{f}, & X \in \mathring{\Omega} \\ w &= 0, & X \in \partial\Omega \end{cases} \tag{27}$$

$\bar{f}$ is indicated with a bar to emphasize it is a constant function in our problem. It could be different.

Let's look at the weaker definition of this problem. Let $\varphi \in \mathscr{D}(\Omega)$, a test

function. A solution of (27) in the sense of distributions would verify

$$\begin{array}{rcl} \left\langle \Delta w - \bar{f}, \varphi \right\rangle_{\mathscr{D}',\mathscr{D}} & = & 0 \\ \Leftrightarrow \quad -\left\langle \partial_x w, \partial_x \varphi \right\rangle_{\mathscr{D}',\mathscr{D}} - \left\langle \partial_y w, \partial_y \varphi \right\rangle_{\mathscr{D}',\mathscr{D}} - \left\langle \bar{f}, \varphi \right\rangle_{\mathscr{D}',\mathscr{D}} & = & 0 \end{array} \tag{28}$$

If the test function φ is taken in $\mathscr{D}(\Omega)$, the solution w is sought in $\mathscr{D}'(\Omega)$ which is a very large space and one cannot really conclude. The next part of the reasoning is to widden the space of test functions until the space of solution is restricted enough to be able to conclude.

Remark 2.2

Beware of keeping the constraints at the boundary $\partial\Omega$, which in this case is

- $\varphi = 0$, *because* $Supp(\varphi) \subsetneq \mathring{\Omega}$,
- $w = 0$, *because of the boundary condition.*

For this problem (27), the good answer is L^2. Then, for the duality bracket of (28) to be well defined, we need:

$$\left\{ \begin{array}{l} \partial_x \varphi \in \mathrm{L}^2, \quad \partial_y \varphi \in \mathrm{L}^2, \quad \varphi \in \mathrm{L}^2 \\ \partial_x w \in \left(\mathrm{L}^2\right)' = \mathrm{L}^2, \quad \partial_y w \in \left(\mathrm{L}^2\right)' = \mathrm{L}^2. \end{array} \right.$$

$\bar{f}$ being constant and Ω being bounded, it is not a problem.

2.5 Sobolev spaces

- We define

$$\forall \alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} \in \mathbb{N}^d, \qquad |\alpha| = \sum_{i=1}^{d} \alpha_i. \tag{29}$$

$$\forall \alpha \in \mathbb{N}^d, \forall u \in \mathscr{D}'(\Omega), \qquad \mathcal{D}^\alpha u = \frac{\partial^{|\alpha|} u}{\partial^{\alpha_1} x_1 \ldots \partial^{\alpha_d} x_d} \tag{30}$$

$$\mathrm{H}^k(\Omega) = \left\{ u \in \mathrm{L}^2 \ / \ \forall \alpha \in \mathbb{N}^d, |\alpha| \leq k, \mathcal{D}^\alpha u \in \mathrm{L}^2 \right\} \tag{31}$$

$$\mathrm{H}_0^k(\Omega) = \left\{ u \in \mathrm{H}^k(\Omega) \ / \ u_{|\partial\Omega} = 0 \right\} \tag{32}$$

- The weak formulation of problem (27) reads:

Find $w \in H_0^1(\Omega)$, such that $\forall \varphi \in H_0^1(\Omega)$:

$$\int_\Omega \overrightarrow{\nabla w} \cdot \overrightarrow{\nabla \varphi} \, dX + \int_\Omega \bar{f} \varphi \, dX = 0 \tag{33}$$

2.6 Lax-Milgram theorem and well-posedness

Definition 2.3

A mathematical problem is well-posed in the sense of Hadamard, if:

- *it admits a unique solution,*
- *this solution depends continuously on the parameters of the problem (the source term and the boundary conditions in our case)*

Theorem 2.4 (*Lax-Milgram*)

Let H be a Hilbert space, $a : H \times H \to \mathbb{K}$ a bilinear form which is:

- ***continuous**: $\exists C > 0, \forall (u, v) \in H^2, \quad |a(u, v)| \leq C\|u\|_H \|v\|_H$,*
- ***coercive**: $\exists \alpha > 0, \forall u \in H, \quad a(u, u) \geq \alpha \|u\|_H$,*

and $\Phi \in H'$. Then the problem

$$\text{"Find } u \in H, \text{ such that } \forall v \in H$$

$$a(u, v) = \langle \Phi, v \rangle_{H',H}\text{"}$$

admits a unique solution.

Moreover, we have the following estimation

$$\|u\|_H \leq \frac{\|\Phi\|_{H'}}{\alpha} \tag{34}$$

Proof: Similar to Riesz Theorem ■

2.7 Finite element approximation

- Let $\mathcal{M}_h$ be a triangular meshing of the domain. To each node i of the mesh situated in $\mathring{\Omega}$, we associate the unique piecewise linear continuous function with value 1 at X_i and 0 at all the other nodes.

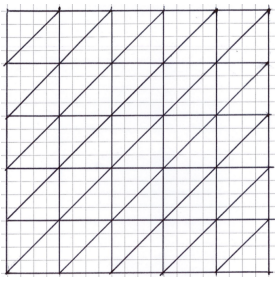


Figure 13: Structured mesh on the unit square. Meshes do not have to be structured. But they usually have to be conformal.

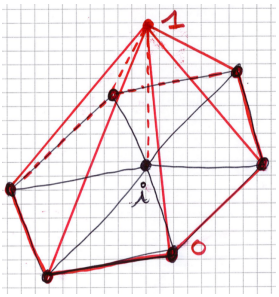


Figure 14: φ_i is defined as the unique continuous function, which restriction to every triangle of the mesh is linear, and which takes value 1 at i and 0 at every other node of the mesh.

- The family $\left\{\varphi_i, \quad X_i \in \mathring{\Omega}\right\}$ spans a subspace of $\mathrm{H}_0^1(\Omega)$ of piecewise linear continuous functions taking all possible values at the nodes X_i. We denote it by H_h. Because the φ_i's are linearly independant, it is a basis of H_h.

- In this context, we approximate (33) by the following reduced problem

$$\text{"}\textit{Find } w_h \in H_h \textit{ such that } \forall \varphi_h \in H_h$$
$$\int_\Omega \overrightarrow{\nabla w_h} \cdot \overrightarrow{\nabla \varphi_h}\, dX + \int_\Omega \bar{f}\varphi_h\, dX = 0\text{"}$$

  By extension of the properties of H in H_h, this reduced problem verifies the conditions of the theorem of Lax-Milgram and the discrete problem is well-posed.

- Considering that $\{\varphi_i, i \in [\![1, \mathbb{N}]\!]\}$ is a basis of H_h and that every function of H_h writes

$$w = \sum_{j=1}^{N} w_j \varphi_j, \quad \boldsymbol{W} = (w_1, \ldots, w_N)^t \in \mathbb{R}^N, \tag{35}$$

  the discrete problem is equivalent to

$$\int_\Omega \left(\sum_{j=1}^{N} w_j \varphi_j\right) \varphi_i\, dX + \int_\Omega \bar{f}\varphi_i dX = 0, \quad \forall i \in [\![1, N]\!],$$
$$\Leftrightarrow \quad \sum_{j=1}^{N} w_j \left(\int_\Omega \varphi_i \varphi_j dX\right) = -\int_\Omega \bar{f}\varphi_i dX, \quad \forall i \in [\![1, N]\!],$$
$$\Leftrightarrow \quad A\boldsymbol{W} = B,$$

  with

$$A_{ij} = \int_\Omega \varphi_i \varphi_j dX, \quad \boldsymbol{W} = \begin{pmatrix} w_1 \\ \ldots \\ w_N \end{pmatrix}, \quad \text{and} \quad B_i = -\int_\Omega \bar{f}\varphi_i dX.$$

Since the problem is well-posed in H_h, we already know the matrix A is invertible and the numerical problem comes to a $N \times N$ linear system!

3 Numerical discretization of hyperbolic conservation laws

3.1 What is a hyperbolic conservation law?

3.1.1 Description

Let D be an open subset of $\mathbb{R}^m$, and $\mathbf{U}$ a vector of m variables

$$\mathbf{U} = (u_1, \ldots, u_m)^t \in \mathrm{D}. \tag{36}$$

$\mathbf{U}$ is called the *state variable* or the *vector of conserved quantities*. It is a function of space and time, with values in D:

$$\mathbf{U} : \mathbb{R}^d \times \mathbb{R}^+ \longrightarrow \mathrm{D}.$$

We call *system of m conservation laws*, the differential system

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}_i(\mathbf{U})}{\partial x_i} = 0, \quad \mathbf{X} = (x_1, \ldots, x_d)^t \in \mathbb{R}^d, \quad t \geq 0 \tag{37}$$

where $\mathbf{F}_i,\ i = 1, \ldots, d$, are called the *flux-functions*. They are smooth functions from D into $\mathbb{R}^m$. We also introduce the *flux-vector* $\overrightarrow{\boldsymbol{\mathcal{F}}} = (\mathbf{F}_1, \ldots, \mathbf{F}_d)$, which enables us to rewrite equation (37) into an equivalent more compact and general form

$$\frac{\partial \mathbf{U}}{\partial t} + \overrightarrow{\nabla}.\overrightarrow{\boldsymbol{\mathcal{F}}}(\mathbf{U}) = 0, \quad \mathbf{X} \in \mathbb{R}^d, \quad t \geq 0. \tag{37}$$

Furthermore, if we suppose that the *flux-functions* are differentiable and the solution $\mathbf{U}$ regular enough, the system can be put into a so-called *quasi-linear form*

$$\frac{\partial \mathbf{U}}{\partial t} + \vec{\boldsymbol{\lambda}}.\overrightarrow{\nabla \mathbf{U}} = 0, \quad \mathbf{X} \in \mathbb{R}^d, \quad t \geq 0 \tag{38}$$

with $\vec{\boldsymbol{\lambda}} = \left(\dfrac{\partial \mathbf{F}_1}{\partial \mathbf{U}}, \ldots, \dfrac{\partial \mathbf{F}_d}{\partial \mathbf{U}} \right)$, the *flux Jacobians*.

System (37) expresses the conservation of the quantities $u_1, \ldots, u_m$: if Ω is an arbitrary sub-domain of $\mathbb{R}^d$ and $\vec{\mathbf{n}}$ is the outward unit normal to $\partial\Omega$, the boundary of Ω, it follows from the integration of (37) over Ω and the Gauss-Ostrogradsky formula that:

$$\frac{d}{dt}\left(\int_\Omega \mathbf{U} d\mathbf{X} \right) + \int_{\partial\Omega} \overrightarrow{\boldsymbol{\mathcal{F}}}(\mathbf{U}) . \vec{\mathbf{n}}\, ds = 0. \tag{39}$$

This means that the time variations of $\int_\Omega \mathbf{U} d\mathbf{X}$ (the total amount of $\mathbf{U}$ in Ω) is equal to the average flux $\overrightarrow{\boldsymbol{\mathcal{F}}}(\mathbf{U}) \cdot \vec{\mathbf{n}}$ entering Ω. Because the flux entering Ω is also the flux going out of $\mathbb{R}^d \backslash \Omega$, the quantities $u_1, \ldots, u_m$ are conserved inside the whole space.

Definition 3.1 (*Hyperbolicity*)

An differential operator

$$\mathcal{D} = \partial_t \cdot + \sum_{i=1}^{d} A_i(\mathbf{X}, t)\partial_i \cdot \tag{40}$$

is called

- *hyperbolic, if the matrices*

$$\mathscr{A}(\boldsymbol{\xi}) = \sum_i A_i \xi_i \tag{41}$$

  are diagonalizable with real eigenvalues for all $\boldsymbol{\xi}$ in $S^{d-1} = \left\{\boldsymbol{x} \in \mathbb{R}^d;\ \|\boldsymbol{x}\|_{L^2} = 1\right\}$,

- *constantly hyperbolic, if moreover the multiplicities of the eigenvalues remain constant as $\boldsymbol{\xi}$ covers the sphere S^{d-1},*

- *strictly hyperbolic, in the special case where all eigenvalues are real and simple for every $\boldsymbol{\xi}$.*

Remark 3.2 (*Stability L^2*)

In the case when matrices A_i may depend on time but not on the space variable $\mathbf{X}$, we can show that these conditions of hyperbolicityt imply stability in L^2.

To do so, we search the solutions of this problem for an initial condition taken in the set of tempered distributions, $\mathscr{S}'(\mathbb{R}^d)$. On this space, the Fourier transform is defined as the adjoint of the Fourier transform on the Schwartz class, $\mathscr{S}(\mathbb{R}^d)$. Then, if $\boldsymbol{\xi} \in \mathbb{R}^d$ is the Fourier variable in space and $\hat{\mathbf{U}}$ the Fourier transform of $\mathbf{U}$, the equation becomes:

$$\frac{\partial \hat{\boldsymbol{U}}}{\partial t} = -i\mathscr{A}(\boldsymbol{\xi})\hat{\boldsymbol{U}}, \quad \forall \boldsymbol{\xi} \in \mathbb{R}^d, t \in [0; T]\,, \tag{42}$$

where we have used the notation $\mathscr{A}(\boldsymbol{\xi}) = \sum_i A_i \xi_i$.

Because the Fourier transform is an isometry of L^2, operator (40) is well-posed in $L^2(\mathbb{R}^2)$ if and only if

$$\sup_{\boldsymbol{\xi} \in \mathbb{R}^d} \|\exp(-i\mathscr{A}(\boldsymbol{\xi}))\| < +\infty. \tag{43}$$

This is in particular true when the system is hyperbolic, ie. when the spatial part of the differential operator is diagonalizable with real eigenvalues.

Examples of hyperbolic conservation laws

- 1D Euler equations

$$\left\{\begin{array}{rcrcl} \partial_t\rho & + & \partial_x\rho u & = & 0, \\ \partial_t\rho u & + & \partial_x(\rho u^2+p) & = & 0, \\ \partial_t\rho E & + & \partial_x(\rho u H) & = & 0, \end{array}\right. \tag{44}$$

  where ρ is the density, u the velocity, p the pressure given by a equation of state ($p=(\gamma-1)\rho e$ for polytropic perfect gases), $E=e+\frac{1}{2}u^2$ is the total energy and $H=E+p/\rho$ is the total enthalpy.

- Multi-D Euler equations

$$\left\{\begin{array}{rcrcl} \partial_t\rho & + & \overrightarrow{\nabla}.(\rho\vec{\mathbf{u}}) & = & 0, \\ \partial_t\rho\vec{\mathbf{u}} & + & \overrightarrow{\nabla}.(\rho\vec{\mathbf{u}}\otimes\vec{\mathbf{u}}+p\mathbb{I}) & = & 0, \\ \partial_t\rho E & + & \overrightarrow{\nabla}.(\rho\vec{\mathbf{u}}H) & = & 0. \end{array}\right. \tag{45}$$

In the following, we restrict our study to 1D scalar conservation laws:

$$\left\{\begin{array}{l} \partial_t u+\partial_x f(u)=0, \qquad x\in\mathbb{R}, t>0, \\ u(t=0,x)=u_0(x). \end{array}\right. \tag{46}$$

$$u:\mathbb{R}\times\mathbb{R}^+\to\mathbb{R}, \quad f\in\mathcal{C}^1(\mathbb{R}), \quad u_0\in \mathrm{L}^\infty(\mathbb{R}).$$

3.1.2 Characteristic curves

Definition 3.3 (*Characteristic Curves*)

Let us consider the family of parametric curves

$$\left\{\begin{array}{rcl} \dfrac{\mathcal{D}X}{\mathcal{D}t} & = & f'\left(u(X(t),t)\right), \\ \mathbf{X}(0) & = & x_0\in\mathbb{R} \end{array}\right. \tag{47}$$

The solutions of (46) *are constant along these curves and these curves are therefore straight lines. We call these parametric curves the "characteristics" of equation* (46).

Proof: Along the characteristics, we can define

$$v: t\longmapsto u(X(t),t).$$

Then, by the differentiation chain rule we have:

$$v'(t)=\partial_t u+\frac{\mathcal{D}X}{\mathcal{D}t}\partial_x u=\partial_t u+f'(u)\partial_x u=0.$$

Thus, u is constant along the characteristics and so is $f'=\dfrac{\mathcal{D}X}{\mathcal{D}t}$. The characteristics are therefore straight lines. ■

Examples:

- Linear advection, $f'(u) = a \in \mathbb{R}$:

$$\partial_t u + a\partial_x u = 0.$$

We have seen in Section 1.3.1 that $u(x,t) = u_0(x - at)$ is a solution of the system. The method of characteristics proves it is the only one, see Figure 15.

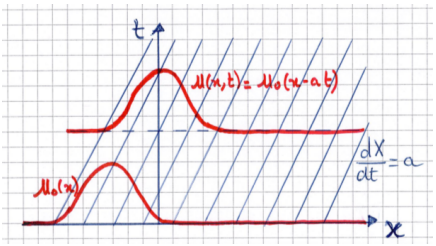


Figure 15: The existence of characteristic curves $X = at$ imposes $u_0(x - at)$ as a unique solution.

- Burgers' equation, $f'(u) = u$:

$$\begin{aligned} &\partial_t u + u\partial_x u = 0 \\ &u_0(x) = \begin{cases} 1, & x < 0 \\ 1 - x, & 0 \leq x \leq 1 \\ 0, & x > 1 \end{cases} \end{aligned} \tag{48}$$

The solution steepens from $t = 0$ to $t = 1$. At $t = 1$ all the characteristics which originate in the interval $[0, 1]$ cross at the same point in the plan (t, x) and the solution is multi-valued at $x = 1$, see Figure 16.

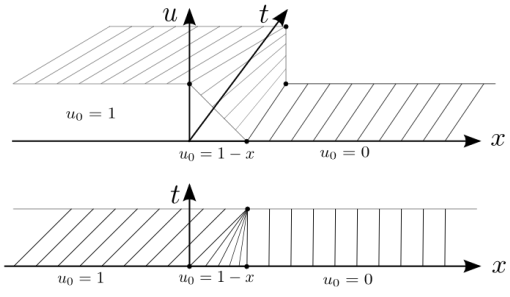


Figure 16: Solution of the 1D scalar Burgers equation with initial conditions (48). All the characteristics meet at point $(1, 1)$ and the solution cannot stay continuous there. A shock has been created.

3.1.3 Weak formulation

As in the introductory section, we see that the solution cannot stay continuous, whatever the regularity of the initial condition (in our example, u_0 is not maximum regular, but it regularization into a $\mathcal{C}^\infty$ function would not

change the conclusion). To move on, we then need to consider the weak formulation of equation (46).

We define the set of $\mathcal{C}^1$ test functions with compact support in $\mathbb{R}_x \times \mathbb{R}_t^+$

$$\mathcal{C}_c^1\left(\mathbb{R}_x \times \mathbb{R}_t^+\right) = \left\{ f \in \mathcal{C}^1\left(\mathbb{R}_x \times \mathbb{R}_t^+\right), \quad \text{Supp}\,(f) \text{ is compact.} \right\} \tag{49}$$

Note that since $\mathbb{R}_t^+$ is closed in $t = 0$, the test function do not necessarily have to vanish along the line $t = 0$. But its support is compact in x and bounded in t.

With these test functions, we now formally multiply (46) by any $\varphi \in \mathcal{C}_c^1\left(\mathbb{R}_x \times \mathbb{R}_t^+\right)$ and integrate by part to obtain the following definition of its weak solutions:

Definition 3.4 (*Weak Solutions*)

Let $u_0 \in L^\infty(R)$. Then $u \in L^\infty(\mathbb{R}_x \times \mathbb{R}_t^+)$ is a weak solution of problem (46) if for any $\varphi \in \mathcal{C}_c^1\left(\mathbb{R}_x \times \mathbb{R}_t^+\right)$, we have

$$\int_0^\infty \int_{\mathbb{R}_x} \left(u.\frac{\partial \varphi}{\partial t} + f(u).\frac{\partial \varphi}{\partial x} \right) dx\, dt + \int_{\mathbb{R}_x} u_0(x).\varphi(x)\, dx = 0. \tag{50}$$

Remark 3.5

As for the Poisson problem, by integration by part it is obvious that a "classical" $\mathcal{C}^1$ solutions of problem (46) are naturally solutions in the weak sense of Definition 3.4.

3.1.4 Rankine-Hugoniot condition

Now that we have a weaker formulation of the differential equation (46) allowing solutions in $\mathrm{L}^\infty(R)$, the following theorem gives a compatibility condition for the discontinuities of u:

Theorem 3.6 (*Rankine-Hugoniot*)

We consider equation (46) where $f \in \mathcal{C}^1$ and $u_0 \in L^\infty$ and piecewise differentiable. Then, $u \in L^\infty$ is a piecewise $\mathcal{C}^1$ solution in the sense of distributions on $\mathbb{R}^x \times \mathbb{R}_t^+$ if and only if:

(i) u is a classical solution of (46) at all the points where it is $\mathcal{C}^1$;

*(ii) at all the discontinuities, u satisfies the jump condition, called **Rankine-Hugoniot condition**:*

$$(f(u_l) - f(u_r)) = \sigma\,(u_l - u_r)\,, \tag{51}$$

where u_l and u_r are respectively the limit left and right states of the discontinuity and σ is the velocity at which this discontinuity moves.

Proof: TODO!!! ■

Example: Burgers' flux, $f(u) = u^2/2$. In the case of the Burgers' flux, jump condition (51) alway reads:

$$\sigma = \frac{(u_l)^2/2 - (u_r)^2/2}{u_l - u_r} = \frac{u_l - u_r}{2}. \tag{52}$$

Then, we can finish example (48). At $(x,t) = (1,1)$, the characteristics coming on both sides give limit states $u_l = 1$ and $u_r = 0$. Therefore, the velocity of the shock is $\sigma = \frac{1}{2}$ and the solution is yet completely determined, see Figure 17.

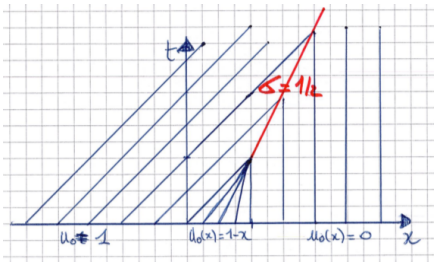


Figure 17: Complete solution of problem (48). Shock generated at point $(x,t) = (1,1)$ propagates at speed $\sigma = \frac{1}{2}$.

3.1.5 Non-Uniqueness of the weak solution

Let us consider the following so-called scalar Riemann problem for the Burgers' equation:

$$\begin{cases} \dfrac{\partial u}{\partial t} + u\dfrac{\partial u}{\partial x} = 0, \ (x,t) \in \mathbb{R}_x \times \mathbb{R}_t^+, \\ u_0(x) = \begin{cases} u_l, & x < 0, \\ u_r, & x > 0. \end{cases} \end{cases} \tag{53}$$

- We suppose that $u_l \geq u_r$. The initial characteristics cross immediately and the solution is given by two uniform states u_l and u_r separated by a shock. The Rankine-Hugoniot condition imposes the velocity of propagation of this discontinuity: $s = (u_l + u_r)/2$.

$$u(x,t) = \begin{cases} u_l, & x < st, \\ u_r, & x > st. \end{cases} \tag{54}$$

- When $u_l < u_r$, because the characteristic curves are never intersecting, one can build a continuous solution:

$$u(x,t) = \begin{cases} u_l, & x < u_l t, \\ x/t, & u_l t < x < u_r t, \\ u_r, & x > u_r t. \end{cases} \tag{55}$$

But for any a between u_r and u_l, we have a family of admissible solutions:

$$u(x,t) = \begin{cases} u_l, & x < s_1 t, \\ -a, & s_1 t < x < 0, \\ a, & 0 < x < s_2 t, \\ u_r, & x > s_2 t. \end{cases} \tag{56}$$

with discontinuities propagating at speeds $s_1 = 0.5(u_l - a)$ and $s_2 = 0.5(u_r + a)$.

3.1.6 Entropy solution

The mathematical problem of existence and uniqueness of the solution of problem (46) has come to a dead end. We have seen that some well chosen cases do not admit classical solutions. Then, we have extended the space of existence of the solutions to a larger class of functions and eventually obtained a possible infinity of solutions. But realistic problems admit only one reproducible solution. Now, we are looking for a sorting criterion that would select the phisically relevant solution among the set of possible *weak solutions*. This criterion is based on the concept of *entropy*.

Hyperbolic conservation laws propagate the information perfectly, at perfectly defined velocities given by the eigenvalues of the differential operator. In really, there is always a dissipation phenomenon which scatters the phase velocities around the group velocity: no real problem is perfectly reversible. Let us consider the following one-dimensional scalar regularized problem, $\varepsilon > 0$ being a small viscous parameter:

$$\frac{\partial u_\varepsilon}{\partial t} + \text{ div}\,(f(u_\varepsilon)) = \varepsilon \Delta u_\varepsilon, \tag{57}$$

with initial condition $u_\varepsilon(x,0) = u_0^\varepsilon \to u_0$ when $\varepsilon \to 0$. We still suppose that u_ε takes its value in D, a sub-domain of $\mathbb{R}$ ($m = 1$). If f is regular enough (Lipschitz), it has been shown that for any positive ε, for any initial condition $u_{0\varepsilon} \in \mathrm{L}^2$, equation (57) admits a unique solution. This result is partly demonstrated in [3].

If we now consider a sequence of ε going to zero, and a sequence of solutions of (57) such that :

a) $\exists C \in \mathbb{R},\ \|u_\varepsilon\|_\infty \leq C$, independently of ε;

b) $u_\varepsilon \underset{\varepsilon \to 0}{\longrightarrow} u$ almost everywhere in $\mathbb{R}^2 \times [0; +\infty[$,

then u is a weak solution of (37) and it verifies, in the sense of distributions, any inequality of the form:

$$\frac{\partial}{\partial t} S(u) + \text{ div}\,(\mathscr{G}(u)) \leq 0, \tag{58}$$

where

(i) $S : \mathrm{D} \to \mathbb{R}$ is a smooth convex function;

(ii) $\mathscr{G}$ is a scalar smooth functions such that

$$S'(u)f'(u) = \mathscr{G}'(u). \tag{59}$$

$(S, \mathscr{G})$ is called *a pair of Entropy-Flux*, S an *entropy function* and $\mathscr{G}$ an *entropy flux*. This result may also be extended to systems, see [3] page 27.

If we now take relation (46) and multiply it formally by $S'(u)$, quick calculation shows that $S(u)$ satisfies an additional conservation relation

$$\frac{\partial}{\partial t} S(u) + \partial_x \mathscr{G}(u) = 0, \quad x \in \mathbb{R}, \quad t > 0. \tag{60}$$

So a **mathematical entropy** is:

- a convex function of the state variables,
- conserved in the subdomains where the solution is smooth, with a associated entropy flux $\mathscr{G}$,
- a signed function across discontinuities, which forbids certain transformation such as those in (56).

Shocks are then **irreversible transformations** characterized by the pair of Entropy-Flux.

The next important result is available in the scalar case for *entropy solutions*. It is the main result of chapter 2 of [3] were one can find a complete and rigorous demonstration.

Theorem 3.7 (*Kruzhkov*)

A weak solution u of a scalar conservation law with a bounded initial condition $u_0 \in L^\infty(\Omega)$, verifying relation (58) for any pair of Entropy-Flux $(S, \mathscr{G})$ is unique and called ***the entropy solution****. Moreover this solution is bounded*

$$\forall T > 0, \quad u \in L^\infty(\Omega \times [0; T]).$$

We were looking for the solution of a sort of idealistic problem (without viscosity), and we found that the only relevant solution is the one coming from the physics. By "*the one coming from the physics*", we mean the solution being the limit of a sequence of solutions of an associated more realistic perturbed problem for a decreasing viscosity coefficient ε. But we do not have to construct such a sequence of realistic solutions in order to find our sought solution. We can simply sort the solutions of the idealistic problem with an entropy criterion. Entropy is then a set of additional conservation relations the solution of problem (46) has to verify.

Remark 3.8 (*Physical interpretation*)

Another point of view is the following: we have started with a system verifying just the first principle of thermodynamics (conservation of the variables), and could find either no solutions (in the class of regular ones) or an infinity (in a weaker class of functions). But by looking at the physics intrinsic to the problem, we found the system of conservation laws is well-posed when it comes with an additional entropy condition. This is the second principle of thermodynamics. This proves that the mathematical problem is strongly bound to the physical one.

3.1.7 Maximum Principle

We can go further in the analysis of the solution and show that the entropy solution of a *conservation law* respects a *maximum principle*. This prevents the sudden appearence of a new global extrema in the solution. This property is very important from a numerical point of view, because one would need it to be transposed to the solution of the numerical scheme used and hence ensure the L^∞ stability of the scheme and prevent the approximated solution to explode in finite time. The next theorem comes from [3], where it is explained and demonstrated in details. It is true only in the scalar case but for any dimension of the spatial domain. It claims the *entropy solution* is bounded in L^∞ norm and monotonically depends on the initial condition.

Theorem 3.9

Let u_0 belong to $L^\infty(\mathbb{R}^2)$. Then the unique entropy solution u of problem

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} + \frac{\partial g(u)}{\partial y} = 0, \qquad & \boldsymbol{x} = (x, y) \in \mathbb{R}^2, \quad t \geq 0 \\ u(\mathbf{x}, 0) = u_0(\boldsymbol{x}), \qquad & \boldsymbol{x} \in \mathbb{R}^2 \end{aligned}$$

with smooth scalar fluxes f and g, belongs to $L^\infty(\mathbb{R}^2 \times [0, T])$. This solution satisfies for almost all $t \geq 0$,

i)

$$\|u(., t)\|_{L^\infty(\mathbb{R}^2)} \leq \|u_0\|_{L^\infty(\mathbb{R}^2)}$$

ii) If v is also the entropy solution of (61) associated with initial condition v_0, we have

$$u_0 \geq v_0 \text{ a.e.} \qquad \Longrightarrow \qquad u(., t) \geq v(., t) \text{ a.e.}$$

3.1.8 Boundary Conditions

Boundary conditions for hyperbolic conservation laws is a relatively tough and yet unsolved topic. It is quite easy to see that in the case of one dimensional scalar conservation laws, we need to impose boundary conditions only on boarders of the computational domain where the characteristics enter the domain. See Figure 18. In a more general context it is however not easy,

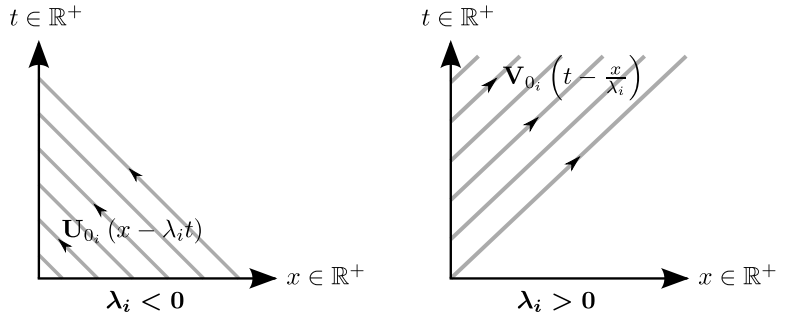


Figure 18: Illustration of boundary condition problem by looking at the linear advection problem on $\mathbb{R}^+$. If the advection velocity (here λ_i) is negative, the solution at $x = 0$ depends on the initial condition and therefore cannot be impose. On the other hand, when the advection velocity is positive, the solution above the straight line $X = \lambda_i t$ depends on the boundary function imposed at $x = 0$.

especially with physical boundary conditions such as "slip wall", "farfield", "free stream", etc. The context of this class is to short to enter these consideration and in the following section we consider test cases which allow us to simply ignore these boundary constraints.

3.2 Numerical Methods for hyperbolic conservation Laws

In the following, we consider two problems:

- the linear advection of a smooth profile:

$$\begin{cases} \partial_t u + \partial_x u = 0, \quad x \in \mathbb{R}/\mathbb{N} \\ u_0(x) = \begin{cases} \cos^4(2\pi x), & \frac{1}{4} \leq x \leq \frac{3}{4} \\ 0, & \text{otherwise,} \end{cases} \end{cases} \tag{61}$$

  which exact solution is $u_0(x - t)$.

- the burger equation with shock:

$$\begin{cases} \partial_t u + \partial_x \frac{u^2}{2} = 0, \quad x \in \mathbb{R} \\ u(x, t = 0) = \begin{cases} 1, & x < \frac{1}{4}, \\ -2x + \frac{3}{2}, & \frac{1}{4} \leq x \leq \frac{3}{4} \\ 0, & x > \frac{3}{4}. \end{cases} \end{cases} \tag{62}$$

It creates a shock at time $t = 0.5$, propagating at speed 0.5.

We solve these two problems on the fixed interval $\mathbf{I} = [0, 1]$ that we subdivide in N sub-intervals

$$\mathbf{I}_i = [x_i, x_{i+1}], \quad i \in [\![0, N-1]\!], \quad \text{where} \quad x_i = \frac{i}{N} = i\Delta x. \tag{63}$$

We thus obtain $N + 1$ nodes and N associated cells at which we wish to approximate the solutions of problems (61) and (62).

3.2.1 Finite Differences

Finite Difference discretization of a PDE is essentially based on the Taylor developments:

$$u_{i+1}^n = u(x_{i+1}, t^n) = u_i^n + \Delta x\, u'(x_i, t^n) + \frac{\Delta x^2}{2} u''(x_i, t^n) + \mathcal{O}(\Delta x^3) \tag{64a}$$

$$u_{i-1}^n = u(x_{i-1}, t^n) = u_i^n - \Delta x\, u'(x_i, t^n) + \frac{\Delta x^2}{2} u''(x_i, t^n) + \mathcal{O}(\Delta x^3) \tag{64b}$$

Same developements can be made in the time direction. From which we obtain that

$$\frac{u_{i+1} - u_i}{\Delta x}, \quad \frac{u_i - u_{i-1}}{\Delta x}, \quad \frac{u_i^{n+1} - u_i^n}{\Delta t},$$

are approximation of order 1 of the spatial and time derivatives and

$$\frac{u_{i+1} - u_{i-1}}{2\Delta x} \quad \text{and} \quad \frac{u_i^{n+1} - u_i^{n-1}}{2\Delta t}$$

are second order approximations. Higher order developments can be lead but these are the only consistent approximations of the first order derivatives on the compact 3-points stencil $\{x_{i-1}, x_i, x_{i+1}\}$. It is enough to illustrate the main features of a numerical scheme for hyperbolic conservation laws.

Explicit Schemes:

$$u_i^{n+1} = u_i^n - \frac{a\Delta t}{\Delta x}\left(u_i^n - u_{i-1}^n\right) \qquad \text{Upwind} \tag{65a}$$

$$u_i^{n+1} = u_i^n - \frac{a\Delta t}{2\Delta x}\left(u_{i+1}^n - u_{i-1}^n\right) \qquad \text{Centered} \tag{65b}$$

$$u_i^{n+1} = u_i^n - \frac{a\Delta t}{\Delta x}\left(u_{i+1}^n - u_i^n\right) \qquad \text{Downwind} \tag{65c}$$

From now on, the non-dimensional ratio $\frac{a\Delta t}{\Delta x}$ will be noted α. These explicit schemes rewrite

$$u_i^{n+1} = (1 - \alpha)u_i^n + \alpha u_{i-1}^n \tag{66a}$$

$$u_i^{n+1} = \frac{\alpha}{2}u_{i-1}^n + u_i^n - \frac{\alpha}{2}u_{i+1}^n \tag{66b}$$

$$u_i^{n+1} = (1 + \alpha)u_{i+1}^n - \alpha u_{i+1}^n \tag{66c}$$

From (66), we see immediatly that under an upper constraint on the time step:

$$\alpha = \frac{a\Delta t}{\Delta x} \leq 1, \tag{67}$$

the update of u_i^{n+1} is a convex combination of the values of u_i in the stencil at time t^n. Then, the solution respect a discrete version of the maximum principle and the discrete solution is stable in L^∞. Such a scheme is called **monotone**.

On the contrary, schemes (66) and (66) show negative coefficients. We can then always think about situation which will violate the discrete maximum principle

$$\min\left(u_{i-1}^n, u_i^n, u_{i+1}^n\right) \leq u_i^{n+1} \leq \max\left(u_{i-1}^n, u_i^n, u_{i+1}^n\right). \tag{68}$$

These schemes are not monotone. In fact it is worse than that and we will show now that these schemes are not stable in the L^2 norm.

Von Neumann's Analysis: Let's look at what happens to a certain Fourier mode of the discrete solution:

$$u(x,t) = \hat{u}(t)\exp(\mathbf{i}kx) \quad \longrightarrow \quad u_i^n = \hat{u}^n \exp(\mathbf{i}kx_i).$$

In the case of the upwind scheme we have

$$\begin{aligned} & \hat{u}^{n+1} = & (1-\alpha)\hat{u}^n + \alpha e^{-\mathbf{i}k\Delta x}\hat{u}^n \\ \Rightarrow \quad & \frac{\hat{u}^{n+1}}{\hat{u}^n} = & 1 - \alpha + \alpha e^{-\mathbf{i}k\Delta x} \end{aligned}$$

This last ratio is the dispersion relation for the considered mode k or $\xi = k\Delta x$. In particular, we can get the amplification factor of this mode:

$$|\mathrm{G}(\xi)|^2 = \frac{|\hat{u}^{n+1}|^2}{|\hat{u}^n|^2}. \tag{69}$$

In the case of the upwind scheme, we have:

$$|\mathrm{G}(\xi)|^2 = (1-\alpha+\alpha\cos\xi)^2 + \alpha^2\sin^2\xi = (1-\alpha)^2 + 2\alpha(1-\alpha)\cos\xi + \alpha^2 \tag{70}$$

Since $-1 \leq \cos\xi \leq 1$, we see that

$$\begin{cases} 1 - 2\alpha \leq |\mathrm{G}(\xi)|^2 \leq 1, & \text{if } \alpha \in [0,1] \\ 1 \leq |\mathrm{G}(\xi)|^2 \leq |1-2\alpha|, & \text{otherwise} \end{cases} \tag{71}$$

This confirm that the explicit upwind scheme is stable under condition (67). If $\alpha > 1$, some modes start to grow exponentially and the solution blows up in finite time.

We do the same work for the explicit centered scheme and get:

$$\begin{aligned} \mathrm{G}(\xi) &= \alpha/2 e^{-\mathbf{i}\xi} + 1 - \alpha/2 e^{\mathbf{i}\xi} = 1 - \mathbf{i}\sin\xi, \\ \Rightarrow |\mathrm{G}(\xi)|^2 &= 1 + \sin^2\xi. \end{aligned}$$

This scheme is then unconditionally unstable! Same conclusion can be drawn for the explicit downstream scheme, since it is equivalent to the upwind scheme with opposite velocity ($\alpha \to -\alpha$).

Implicit schemes

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + a\frac{u_i^{n+1} - u_{i-1}^{n+1}}{\Delta x} = 0 \qquad \text{Upwind} \tag{72a}$$

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + a\frac{u_{i+1}^{n+1} - u_{i-1}^{n+1}}{2\Delta x} = 0 \qquad \text{Centered} \tag{72b}$$

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + a\frac{u_{i+1}^{n+1} - u_i^{n+1}}{\Delta x} = 0 \qquad \text{Downwind} \tag{72c}$$

These schemes are called **implicit** because the solution at next time step is the solution of an implicit equation. In this example, since the continuous equation is linear, this equation is simply a linear system:

$$(1+\alpha)u_i^{n+1} - \alpha u_{i-1}^{n+1} = u_i^n \tag{73a}$$

$$\alpha/2 u_{i+1}^{n+1} + u_i^{n+1} - \alpha/2 u_{i-1}^{n+1} = u_i^n \tag{73b}$$

$$\alpha u_{i+1}^{n+1} + (1-\alpha)u_i^{n+1} = u_i^n \tag{73c}$$

Von Neumann analysis shows that the implicit upwind scheme is unconditionally stable, as is the implicit centered one. As can be expected, the centered version is more accurate since the spatial derivative is approximated at a better precision. The implicit downwind scheme is stable under condition $\alpha \geq 1$. It is then of no use in practice because it is very dissipative.

3.2.2 Stabilized Finite Elements

We consider here a Finite Element approximation in space of the linear advection equation. To do so, the solution is decomposed on the Finite Element basis functions:

$$u_h(t,x) = \sum_j u_j(t)\varphi_j(x). \tag{74}$$

Then we proceed to the variational formulation:

$$\int_\Omega (\partial_t u + a\partial_x u = 0) \times \varphi_i \Leftrightarrow \mathcal{M} d_t U + a\mathcal{K}U = 0,$$

where $U(t) = (u_0(t), \dots, u_N(t))^t$ is the vector of unknowns,

$$\mathcal{M}_{ij} = \int_\Omega \varphi_i \varphi_j dx = \frac{\Delta x}{6} \begin{pmatrix} 4 & 1 & & & (0) \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ (0) & & & 1 & 4 \end{pmatrix} \tag{75}$$

is called the **mass matrix** and

$$\mathcal{K}_{ij} = \int_\Omega \varphi_i \partial_x \varphi_j dx = \frac{1}{2} \begin{pmatrix} 0 & 1 & & & (0) \\ -1 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ (0) & & & -1 & 0 \end{pmatrix} \tag{76}$$

is called the **stiffness matrix**.

We have yet obtained a implicit ODE on the vectorial unknown U that we can try to integrate with a adapted method for ODE integration. One may in particular try the **forward Euler** scheme

$$d_t U = \frac{U^{n+1} - U^n}{\Delta t} + \mathcal{O}\left(\Delta x^2\right),$$

and get the following numerical method

$$\mathcal{M} U^{n+1} = \mathcal{M} U^n - a \Delta t \mathcal{K} U^n. \tag{77}$$

Von Neumann analysis gives

$$\frac{\Delta x}{6}(4 + 2\cos\xi)\hat{\mathbf{U}}^{n+1} = \frac{\Delta x}{6}(4 + 2\cos\xi)\hat{\mathbf{U}}^n - a\Delta t \mathbf{i} \sin\xi \hat{\mathbf{U}}^n$$

$$\Rightarrow |\mathrm{G}(\xi)|^2 = 1 - 3\mathbf{i}\alpha \frac{\sin\xi}{2 + \cos\xi} \tag{78}$$

and the method is even more unstable than the explicit centered scheme.

Remark 3.10

*If one replaces the mass matrix $\mathcal{M}$ by an equivalent diagonal matrix $\Delta x \mathbb{I}$, what is called **mass lumping**, we recover the explicit centered scheme.*

Stabilization

3.2.3 Finite Volume Methods

Instead of considering the degrees of freedom at the nodes of the mesh, Finite Volume methods attach them to the cells of the mesh. In general, these DoFs allow a local polynomial reconstruction of the solution. At first order, one may wish in particular to have a numerical scheme dealing with the time evolution of the average values of the conserved quantities in each cell:

$$\bar{\mathbf{U}}_i = \frac{1}{|\mathbf{I}_i|}\int_{\mathbf{I}_i} \mathbf{U} dx. \tag{79}$$

This can be achieved by a variational formulation with the caracteristic functions of each cells as basis functions: χ_i. Indeed, these caracteristic functions span the space of constant per cells solutions.

$$\int_\Omega (\partial_t u + \partial_x f(u) = 0) \times \chi_i \quad \Rightarrow \quad |\mathbf{I}_i|\frac{d}{dt}\bar{\mathbf{U}}_i + \int_{\mathbf{I}_i} \partial_x f(u) = 0$$

This last integral is crucial. Indeed, one could say it has the value of the difference of the fluxes at the extremities of the interval. But the solution is not defined at the interfaces between two cells because it is spanned by the caracteristic functions. Then the values of the flux at the interfaces is replace by a certain function of the two neighboring averaged values: these are the **numerical fluxes**. The first order finite volume scheme then reads:

$$d_t \bar{u}_i + \frac{1}{\Delta x_i}\left(\mathbf{F^*}(\bar{u}_{i+1}, \bar{u}_i) - \mathbf{F^*}(\bar{u}_i, \bar{u}_{i-1})\right) = 0 \tag{80}$$

As in the case of Finite Elements, we have come to an ODE on $\bar{u}_i$ that we can integrate by our favorite ODE scheme. Forward Euler in particular gives the update:

$$\bar{u}_i^{n+1} = \bar{u}_i^n - \frac{\Delta t}{\Delta x}\left(\mathbf{F^*}(\bar{u}_{i+1}, \bar{u}_i) - \mathbf{F^*}(\bar{u}_i, \bar{u}_{i-1})\right). \tag{81}$$

Lax-Friedrichs Flux

$$\mathbf{F^*}(\bar{u}_{i+1}, \bar{u}_i) = \frac{1}{2}\left(f(\bar{u}_{i+1}) + f(\bar{u}_i)\right) - \frac{\alpha_i}{2}\left(\bar{u}_{i+1} - \bar{u}_i\right). \tag{82}$$

Coefficient α_i is interface dependant. It has to be larger than $f'(\bar{u}_{i+1})$ and $f'(\bar{u}_i)$. Its choice will affect the constraint of stability on the time step. The larger the α_i, the more dissipative the method.

Lax-Wendroff Flux

$$\mathbf{F^*}(\bar{u}_{i+1}, \bar{u}_i) = f\left(\bar{u}_{i+\frac{1}{2}}^{n+\frac{1}{2}}\right), \quad \text{where} \quad \bar{u}_{i+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1}{2}\left(\bar{u}_{i+1} + \bar{u}_i\right) - \frac{\alpha}{2}\left(f(\bar{u}_{i+1}) + f(\bar{u}_i)\right). \tag{83}$$

This scheme is second order in space and time, stable and consistent with the entropy inequalities. However it is not monotonous and can generate spurious oscillations. Since the scheme is stable, these oscillations do not degenerate but they can affect very much the quality of the solution.

Mac-Cormack Method

$$\mathbf{F^*}(\bar{u}_{i+1}, \bar{u}_i) = \frac{1}{2}\left(f(\bar{u}_{i+1}) + f(\bar{u}_i^*)\right), \quad \text{where} \quad \bar{u}_i^* = \bar{u}_i - \alpha\left(f(\bar{u}_{i+1}) - f(\bar{u}_i)\right). \tag{84}$$

Godunov Method

3.2.4 Discontinuous Galerkin Methods

3.2.5 Residual Distribution Schemes

3.2.6 Time Integration

3.3 Positivity and convex constraints preservation

4 Numerical treatment of some SDEs

We want to introduce the differential equations with stochastic terms and describe their links with the ODEs and PDEs. We will speak about stochastic differential equations driven by stochastic process W. In a very formal sense, we are interested in an equation under the form

$$y' = h(y) + \dot{W} \tag{85}$$

where $\dot{W}$ is some sort of derivative of the stochastic process W. But we will see that many stochastic processes are not even differentiable, thus the equation is not well-posed. Like in the previous sections, we will skip this difficulty by introducing some weak formulation of the equation (85) essentially by using a new notion of integration named the Itô's integral.

4.1 Random process and Brownian motion

Definition 4.1

A random process is a temporal function taking possibly different values at fixed time t as a random variable.

Its complete understanding needs the theory of probability and stochastic processes. But we can consider simple processes and skip all the general theory. The most classical random variable with density is the Gaussian random variable. It is completely determined by two given real numbers named the mean and the variance.

Definition 4.2

We say that G is a Gaussian random variable of mean μ and standard deviation σ if the density of G is given by the Gaussian function

$$\mathbb{P}(G \in \mathrm{d}x) = \frac{1}{\sigma\sqrt{2\pi}} e^{(\mu-x)^2/2\sigma^2} \mathrm{d}x$$

A Brownian motion is a very complex (probabilistic) process defined on $[0, 1]$. But it is a universal process which is completely defined by only three properties (see [10])

Definition 4.3

A stochastic process

$$B : \begin{array}{ccc} [0,1] & \to & \mathbb{R} \\ t & \mapsto & B(t) = B_t \end{array}$$

is a Brownian motion if and only if it verifies the following properties:

- *$B(0) = 0$,*

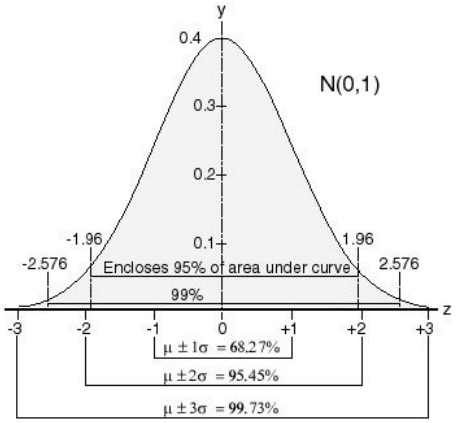


Figure 19: Gaussian curve

- *$B(t) - B(s)$ is a gaussian random variable of mean 0 and variance $\sigma^2 = t - s$,*
- *the increments are independent.*

Given the existence of this random process we can solve our first stochastic differential equation (SDE) driven by the stochastic process $W := B$ a Brownian motion. Roughly speaking we are looking for a random process y such that

$$y' = c + \dot{W}$$

where $'$ and $\dot{}$ have certainly the "same" sense since acting on similar objects. The solution is obviously $y(t) = y_0 + ct + W_t$ where y_0 is the initial condition of the SDE ($y(0) = y_0$). This is a Gaussian process with independent increment but this is not a Brownian motion. Actually we speak about a drifted Brownian motion.

Now can we solve an other more complex SDE ?

$$y' = -by + \dot{W} \tag{86}$$

The solution without $\dot{W}$ is $y(t) = y_0 e^{-bt}$. By a classical method named the variation of the constant, we expect that the solution has the form:

$$y(t) = y_0 e^{-bt} + e^{-bt} \int_0^t e^{bs} \dot{W}_s \mathrm{d}s. \tag{87}$$

But what is the sense of the derivative of the integral part ? Is this term sufficiently smooth according to the regularity of $\dot{W}$? Since $\dot{W}$ is certainly not a continuous function, the integral part is certainly neither well defined

nor differentiable. But, if $'$ and $\dot{}$ are the corresponding operators of integral operators $\int$, perhaps we can have a formulation in the form

$$\int_0^t y'(s)\mathrm{d}s = -b\int_0^t y(s)\mathrm{d}s + \int_0^t \dot{W}_s \Leftrightarrow y(t) = -b\int_0^t y(s)\mathrm{d}s + W_t. \quad (88)$$

In this case, since W has more regularity than $\dot{W}$ and y has certainly sufficient regularity, we can look for a solution of this "more regular" equation.

4.2 Stochastic integral

An Itô's integral is some sort of a Riemann's integral with respect to a random process with possibly infinite quadratic variation. Skiping the details, suppose that we can construct this integral as a "limiting process"[1] using the following formula

$$\int (X, \mathrm{d}W) = \lim_{h\to 0} \sum_{i=0}^{N-1} X(t_i)(W_{t_{i+1}} - W_{t_i}), \quad (89)$$

where $0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = 1$ (with N depending on h) is a partition of the interval $[0, 1]$ with a maximum size step h converging to 0. The limit is actually taken over all the possible partitions of maximum step size h. This limit has possibly no sense in a classical sense and we need probabilistic notion of convergence in order to fully describe the limiting process. Given this new integral, we will say that y is an Itô process if it has a good decomposition.

Definition 4.4

y is an Itô process if there exists two process b and σ (sufficiently smooth and integrable) such that

$$y(t) = y(0) + \int_0^t b(s)\mathrm{d}s + \int_0^t (\sigma, \mathrm{d}W).$$

In this case, we can compute the "derivative" of y writing

$$dy = b(t)\mathrm{d}t + \sigma(t)\mathrm{d}W.$$

Now, using this definition, we can give a more correct writing of a stochastic differential equation.

[1] I skip voluntarily all the details about the notion of convergence or the martingale structure which are completely out of the scope of this introduction about partial differential equations. The interested reader can find many courses about SDE in the classical literature see [11].

4.3 Stochastic differential equations and PDEs

We will say that the solution to the formal equation $y' = b + \sigma\dot{W}$ is the Itô process characterized by b and σ. We will rewrite this equation under the form.

$$\begin{cases} \mathrm{d}y &= b\mathrm{d}t + \sigma\mathrm{d}W, \\ y(0) &= y_0. \end{cases}$$

Thanks to this new definition of a solution, we can search a solution to a stochastic differential equation of a more complex form

$$\begin{cases} \mathrm{d}y &= b(t, y(t))\mathrm{d}t + \sigma(t, y(t))\mathrm{d}W, \\ y(0) &= y_0. \end{cases} \tag{90}$$

Theorem 4.5

Suppose that $|b(t,0)|$ and $|\sigma(t,0)|$ are square integrable (i.e. in $\mathbb{L}^2(\mathbb{R}^+)$) and that there exists $K > 0$ such that:

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq K|x - y|.$$

Then for all initial condition $y_0 \in \mathbb{L}^2$ and for all $T > 0$ there exists a unique solution to the SDE (90) on $[0,T]$.

We come back to the equation (86). It verifies the hypotheses of the theorem thus we know that we have a unique solution. With our new notations we expect that the solution is the following Itô process:

$$y(t) = y(0) - by_0 \int_0^t e^{-b(t-s)}\mathrm{d}s + \int_0^t (e^{-b(t-s)}, \mathrm{d}W)$$

or simply

$$y(t) = y_0 e^{-bt} + e^{-bt} \int_0^t (e^{bs}, \mathrm{d}W),$$

so we have

$$\mathrm{d}y = -by_0 e^{-bt}\mathrm{d}t - be^{-bt}\mathrm{d}t \int_0^t (e^{bs}, \mathrm{d}W) + e^{-bt}e^{bt}\mathrm{d}W_t = -by\mathrm{d}t + \mathrm{d}W_t.$$

Finally, in order to conclude this part about theoretical description of stochastic differential equation, we will describe a crucial link between stochastic differential equations and the classical theory of partial differential equations. This link is fully described by the Feynman-Kac formula.

Theorem 4.6 (*Feyman-Kac formula*)

Consider the reverse Cauchy problem with a classical PDE

$$\begin{cases} 0 &= \partial_t u + \mu(t,x)\partial_x u + \frac{1}{2}\sigma^2(t,x)\partial_{xx} u, \\ u(T,x) &= u_T(x), \end{cases}$$

and $y^{t,x}$ the Itô process solution of

$$\begin{cases} \mathrm{d}y &= \mu(t,y)\mathrm{d}t + \sigma(t,y)\mathrm{d}W, \\ y(t) &= x, \end{cases}$$

with an "initial" condition at time t and with W a Brownian motion then

$$u(t,x) = \mathbb{E}[u_T(y^{t,x}(T))].$$

where $\mathbb{E}$ denotes the expectation of the process (i.e. the mean of all random realizations).

4.4 Numerical treatment of stochastic differential equations

Thanks to the Feyman-Kac formula, the simulation of a solution to a stochastic differential equation can lead to the solution of a partial differential equation. Moreover the simulation of a stochastic differential equation is very closed to the simulation of an ordinary differential equation since it does not imply partial derivative but only differential operators. In this section we will describe some stochastic schemes.

The simpler scheme is the Euler-Maruyama method.

Definition 4.7

The Euler-Maruyama scheme for the stochastic differential equation

$$\begin{cases} \mathrm{d}y &= \mu(t,y)\mathrm{d}t + \sigma(t,y)\mathrm{d}W, \\ y(0) &= x, \end{cases}$$

is the numerical recursive procedure defined for a step size $dt > 0$

$$\begin{aligned} T_0 &= 0, \\ Y_0 &= x, \\ T_{n+1} &= T_n + dt, \\ Y_{n+1} &= Y_n + \mu(T_n, Y_n)dt + \sigma(T_n, Y_n)(W(T_{n+1}) - W(T_n)). \end{aligned}$$

Since the increments of a Brownian motion are independent and are Gaussian random variables, we know that the quantities $(W(T_{n+1}) - W(T_n))$ are in fact independent realizations of a Gaussian random variable of mean 0 and variance $T_{n+1} - T_n = dt$. The simplified Euler-Maruyama scheme becomes

$$\begin{aligned} Y_0 &= x, \\ Y_{n+1} &= Y_n + \mu(ndt, Y_n)dt + \sigma(ndt, Y_n)\sqrt{dt}G_n, \end{aligned}$$

where G_n are independent Gaussian random variables $\mathcal{N}(0,1)$.

5 Conclusion

These lectures are only a very short introduction to the theory of partial and stochastic differential equations. Here are the notes of the two authors for the lecture given in september 2014 at the "Institut des Hautes Études Scientifiques" in Bures-sur-Yvette.

The authors are aware that many objects need a complete theory to be fully understandable but they have tried to illustrate the principal aspects of these domains of research for the students. They have voluntarily skipped many details. For this reason many definitions, properties or theorem are not complete. We apologize to the attentive readers for this.

Anyway, we hope that this document is a good introduction for the beginners and that it will encourage the students to pursue in this area of research.

6 References

Theory of PDEs

[1] Lawrence C. Evans. *Partial Differential Equations: Second Edition*. American Mathematical Society, 2010.

Finite Element Methods

[2] A. Ern and J-L. Guermond. *Theory and Practice of Finite Elements*. Vol. 159. Applied Mathematical Sciences. Springer-Verlag, 2004.

Hyperbolic Conservation Laws

[3] E. Godlewski and P.A. Raviart. *Hyperbolic systems of conservation laws*. Ellipses, Paris, 1991.

[4] E. Godlewski and P.-A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws*. Vol. 118. New York: Springer-Verlag, 1996.

[5] D. Serre. *Systems of conservation laws I - Hyperbolicity, Entropies, Shock waves*. Cambridge University Press, 1999.

[6] E.F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*. New York: Springer, 1999.

[7] R. J. LeVeque. *Finite volume methods for hyperbolic problems*. Cambridge Texts in Applied Mathematics. Cambridge: Cambridge University Press, 2002, pp. xx+558.

[8] C. Hirsch. *Numerical computation of Internal and External flows: the fundamentals of Computational Fluid Dynamics*. Elsevier, June 2007.

[9] Jan S. Hestaven and Tim Warburton. *Nodal Discontinuous Galerkin Methods*. Springer-Verlag, 2008.

Stochastic Differential Equations

[10] P. Mörters and Y. Peres. *Brownian Motion*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2010.

[11] R. Durrett. *Stochatic calculus*. Probability and Stochastics Series. CRC press.