



FACULTADE DE MATEMÁTICAS

Traballo Fin de Grao

# Introducción a las leyes de conservación y a su resolución numérica

Ángel Pita da Veiga Ramonde

Junio, 2024

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



GRAO DE MATEMÁTICAS

Traballo Fin de Grao

# Introducción a las leyes de conservación y a su resolución numérica

Ángel Pita da Veiga Ramonde

Junio, 2024

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA



# Trabajo propuesto

<b>Área del Conocimiento:</b> Matemática Aplicada
<b>Título:</b> Introducción a las leyes de conservación y a su resolución numérica
<b>Breve descripción del contenido</b>
La primera parte de este trabajo está dedicada al estudio de las leyes de conservación, incluyendo la deducción de algunas de las propiedades más importantes de sus soluciones clásicas, débiles y entrópicas. La segunda parte es una introducción al método de volúmenes finitos para la resolución numérica de las leyes de conservación. Se utilizarán las ecuaciones del transporte, Burgers y Saint-Venant para ilustrar los comportamientos de los métodos implementados.



# Índice

<b>Resumen</b>	<b>VIII</b>
<b>Introducción</b>	<b>1</b>
<b>1. Sistemas hiperbólicos de leyes de conservación</b>	<b>3</b>
1.1. Deducción de las leyes de conservación . . . . .	3
1.2. Clasificación de sistemas de leyes de conservación . . . . .	6
1.3. Ejemplos de sistemas de leyes de conservación . . . . .	7
<b>2. Soluciones del problema de Cauchy</b>	<b>11</b>
2.1. Soluciones clásicas . . . . .	11
2.2. Soluciones débiles . . . . .	17
2.3. Soluciones entrópicas . . . . .	24
2.4. El problema de Riemann . . . . .	31
<b>3. Descripción general de los esquemas de volúmenes finitos</b>	<b>37</b>
3.1. Descripción general de los métodos de volúmenes finitos . . . . .	37
3.2. Métodos conservativos . . . . .	39
3.3. Consistencia . . . . .	40
3.4. Métodos entrópicos . . . . .	43
3.5. Monotonía . . . . .	44


---

3.6. Estabilidad y convergencia . . . . .	46
3.7. Estabilidad lineal . . . . .	47
<b>4. Ejemplos de esquemas numéricos</b>	<b>49</b>
4.1. Esquema de Lax-Friedrichs . . . . .	49
4.2. Esquema de Lax-Wendroff . . . . .	53
4.3. Método de Godunov . . . . .	55
4.4. Resolventes aproximadas del problema Riemann . . . . .	59
<b>I. Códigos MATLAB</b>	<b>61</b>
I.1. Códigos relativos a la ecuación del transporte . . . . .	61
I.2. Códigos relativos a la ecuación de Burgers . . . . .	66
I.3. Códigos relativos a las ecuaciones de Saint-Venant . . . . .	73
<b>Bibliografía</b>	<b>83</b>






## Resumen

La modelización matemática con leyes de conservación describe el comportamiento de un sistema teniendo en cuenta la conservación de alguna magnitud, como la masa, la energía o el momento. Estos modelos se utilizan en campos como la física, la ingeniería y la ciencia ambiental para comprender y predecir el comportamiento de sistemas complejos, desde la dinámica de fluidos y el flujo de tráfico hasta las reacciones químicas y las interacciones ecológicas. En este trabajo se estudiarán en primer lugar las propiedades analíticas más importantes de los sistemas de leyes de conservación, viendo los conceptos de solución clásica, solución débil, condiciones de entropía y el problema de Riemann. Puesto que, en general, no es posible obtener la solución exacta de las leyes de conservación, posteriormente, teniendo en cuenta las características y dificultades encontradas, se diseñarán métodos numéricos para aproximar numéricamente las soluciones de estos sistemas. Se hará especial énfasis en el método de Godunov, cuya deducción matemática es consecuencia directa de las propiedades fundamentales de las leyes de conservación. Finalmente, se aportan códigos  MATLAB que implementan los métodos descritos.

## Abstract

Mathematical modeling with conservation laws describes the behavior of a system by considering the conservation of some quantity, such as mass, energy, or momentum. These models are used in fields such as physics, engineering, and environmental science to understand and predict the behavior of complex systems, from fluid dynamics and traffic flow to chemical reactions and ecological interactions. In this work, we will first study the most important analytical properties of systems of conservation laws, including the concepts of classical solutions, weak solutions, entropy conditions, and the Riemann problem. Since it is generally not possible to obtain the exact solution of conservation laws, numerical methods will later be designed to numerically approximate the solutions of these systems, taking into account the characteristics and difficulties encountered. Special emphasis will be placed on the Godunov method, whose mathematical derivation is a direct consequence of the fundamental properties of conservation laws. Finally,  MATLAB codes implementing the described methods will be provided.



# Introducción

Las leyes de conservación son ecuaciones en derivadas parciales hiperbólicas que modelan procesos físicos cuya característica fundamental es que la información relativa a las magnitudes consideradas se transmite a velocidad finita. Esta propiedad hace que sean ampliamente usadas en dinámica de fluidos, por ejemplo, para resolver las *ecuaciones de Euler* de dinámica de los gases o las *ecuaciones de Saint-Venant* que modelan el comportamiento de aguas someras. Aunque los métodos de volúmenes finitos se desarrollaron para resolver estas ecuaciones, se comenzará por aplicarlos a ecuaciones más sencillas, como la ecuación del transporte y la de Burgers, para posteriormente emplear los volúmenes finitos para resolver la ecuación de Saint-Venant en el caso unidimensional.

Este trabajo busca dar una introducción a la resolución numérica de estas ecuaciones y está dividido en dos partes claramente diferenciadas. La primera parte está formada por los Capítulos 1 y 2 y está dedicada a introducir los conceptos básicos de las leyes de conservación, así como hacer un estudio analítico del comportamiento de sus soluciones. La segunda parte del trabajo está formada por los Capítulos 3 y 4 y consiste en la descripción de los métodos numéricos más básicos empleados para resolver algunas de las leyes de conservación unidimensionales más sencillas, así como dar algunas de las propiedades más elementales de estos métodos.

En particular, el Capítulo 1 presenta el concepto de una ley de conservación a partir de su deducción física para luego generalizarlas a sistemas de leyes de conservación. También se explican algunos conceptos básicos necesarios para su clasificación y se presenta una *aparentemente sencilla* ecuación, la ecuación de Burgers, que ejemplificará algunos de los comportamientos propios de las leyes de conservación no lineales.

En el Capítulo 2 se comienza describiendo el primer concepto de solución que se nos puede ocurrir, las soluciones clásicas del problema de Cauchy, soluciones con regularidad suficiente como para poder interpretarlas en el sentido habitual. Se verá como este concepto de *solución clásica* es, en general, insuficiente para tratar con leyes de conservación no lineales, lo cual sugiere generalizarlo para tratar con *soluciones débiles* pasando a una forma integral de la ley de conservación. Sin embargo, con este paso se verá que se pierde la unicidad de la solución.

Por ello, se introducirán condiciones adicionales para poder así hablar de unicidad de *solución entrópica*, que modelará a la solución físicamente plausible del problema subyacente. Después, se particularizan los conceptos desarrollados para hallar las soluciones entrópicas de un problema de Cauchy de gran importancia: el llamado *problema de Riemann*. Se dará de forma explícita la solución a este problema para dos ecuaciones que se utilizarán de modelo en este trabajo: la ecuación del transporte, la ecuación de Burgers, así como para una función arbitraria cóncava (o convexa) y cóncavo-convexa.

En el Capítulo 3 se describen métodos numéricos para la resolución de leyes de conservación en un marco general y se particularizan algunos conceptos de gran importancia en su análisis numérico, tales como la consistencia, estabilidad, convergencia o monotonía. Se buscará, tal y como pretendía Godunov en [9], que estos métodos *repliquen el comportamiento cualitativo* de las soluciones entrópicas de las leyes de conservación.

En el Capítulo 4 se explicitan algunos métodos de especial relevancia tanto desde el punto de vista académico como práctico; como por ejemplo el esquema de Lax-Friedrichs y el método de Godunov. Cabe destacar que la construcción de este último se basa en las soluciones entrópicas, como las calculadas en el Capítulo 2. Se realizarán test numéricos para comprobar el correcto funcionamiento de estos métodos.

Finalmente, en el Anexo I se presentan los códigos que implementan los métodos numéricos explicados en esta memoria.

Para la elaboración de este documento, se ha hecho una revisión bibliográfica de los libros clásicos sobre leyes de conservación. Son especialmente relevantes los dos libros de E. Godlewski y P.-A. Raviart [7] y [8], las notas del curso MA103 del ENSTA [1], así como los libros de R. J. LeVeque [14], P. G. LeFloch [12], y M. E. Vázquez-Cendón [20], cuyas lecciones y consejos, así como los de S. Busto, también han ayudado enormemente a la redacción de este trabajo. En [19], de E. F. Toro, se han consultado numerosos detalles técnicos sobre la implementación de los métodos de volúmenes finitos, así como sobre soluciones de problemas de Riemann. Para el estudio del sistema de Saint-Venant, ha sido de gran utilidad [18], del mismo autor. Por último, para ciertas generalidades sobre métodos numéricos se ha consultado el libro [16] de J. C. Strikwerda y muchos de los detalles de análisis funcional se han extraído de los libros clásicos de H. Brézis [2], K. Yosida [21] y J. B. Conway [5].

# Capítulo 1

## Sistemas hiperbólicos de leyes de conservación

En este capítulo, se deducirá la forma general de un sistema de leyes de conservación hiperbólico. Después, se pasará a concretar el tipo particular de leyes de conservación que se va a tratar en este trabajo, así como incluir algunos ejemplos que se considerarán reiteradamente en capítulos posteriores.

### 1.1. Deducción de las leyes de conservación

Para aportar mayor claridad, se comenzará por el caso escalar. Se supone que existe una función regular  $u: \mathbb{R}^n \times \mathbb{R} \rightarrow \Omega \subseteq \mathbb{R}$  tal que  $u(\mathbf{x}, t)$  es la densidad (es decir, el valor por unidad de volumen) de una cierta magnitud escalar en la posición  $\mathbf{x} \in \mathbb{R}^n$  y en el instante  $t \in \mathbb{R}$ , y donde el abierto  $\Omega \subseteq \mathbb{R}$  es el conjunto de valores que puede tomar la magnitud  $u$ . Se denominará a la función  $u$  como la *variable conservativa*.

Ahora se va a suponer que sobre la magnitud que se está considerando rige un *principio de conservación*, es decir, se está suponiendo la siguiente hipótesis: *la tasa de variación por unidad de tiempo de la magnitud en cualquier recinto coincide con el flujo saliente a través de la frontera de dicho recinto*.

Se va a suponer también que el flujo de la magnitud en el punto  $\mathbf{x} \in \mathbb{R}^n$  y el instante  $t \in \mathbb{R}$  viene dado por  $\mathbf{F}(\mathbf{x}, t)$ , con  $\mathbf{F} \in \mathcal{C}^2(\mathbb{R}^n \times \mathbb{R}, \mathbb{R}^n)$ . Adicionalmente, se va a suponer que existe una función  $\mathbf{f} \in \mathcal{C}^2(\Omega, \mathbb{R}^n)$ , denominada *flujo* (físico), tal que se puede escribir el flujo como  $\mathbf{F}(\mathbf{x}, t) = \mathbf{f}(u(\mathbf{x}, t))$ .

Sea  $\mathcal{D} \subset \mathbb{R}^n$  una región regular arbitraria de  $\mathbb{R}^n$ , es decir, un subconjunto de  $\mathbb{R}^n$  que es la clausura de un abierto conexo cuya frontera es una curva regular a trozos. Entonces, se puede escribir matemáticamente el principio de conservación relativo a  $u$  en  $\mathcal{D}$  como

$$\frac{d}{dt} \int_{\mathcal{D}} u(\mathbf{x}, t) \, d\mathbf{x} = - \oint_{\partial\mathcal{D}} \mathbf{f}(u(\mathbf{s}, t)) \cdot \mathbf{n}(\mathbf{s}) \, ds, \quad \forall t \in \mathbb{R},$$

donde  $\mathbf{n}(\mathbf{s})$  es el vector unitario normal exterior a  $\partial\mathcal{D}$  en el punto  $\mathbf{s} \in \partial\mathcal{D}$ . Utilizando el teorema de Gauss y el teorema de la convergencia dominada, se puede reescribir la ecuación anterior como

$$\int_{\mathcal{D}} [\partial_t u(\mathbf{x}, t) + \operatorname{div} \mathbf{f}(u(\mathbf{x}, t))] \, dx = 0, \quad \forall t \in \mathbb{R},$$

o bien, denotando por  $f_j: \Omega \rightarrow \mathbb{R}$ , con  $j \in \{1, \dots, n\}$ , a cada una de las componentes de  $\mathbf{f}$ ,

$$\int_{\mathcal{D}} \left[ \partial_t u(\mathbf{x}, t) + \sum_{j=1}^n \partial_{x_j} f_j(u(\mathbf{x}, t)) \right] \, dx = 0, \quad \forall t \in \mathbb{R}.$$

Puesto que la región regular  $\mathcal{D} \subset \mathbb{R}^n$  era arbitraria y se está imponiendo el principio de conservación en cualquier recinto, la igualdad anterior se dará si, y solo si, el integrando es nulo en todo punto  $\mathbf{x} \in \mathbb{R}^n$  y todo instante de tiempo  $t \in [0, +\infty)$ . De este razonamiento se obtiene la *forma conservativa* de una ley de conservación escalar, que es la siguiente ecuación en derivadas parciales:

$$\partial_t u(\mathbf{x}, t) + \sum_{j=1}^n \partial_{x_j} f_j(u(\mathbf{x}, t)) \, dx = 0, \quad \forall \mathbf{x} \in \mathbb{R}^n, \quad \forall t \in [0, +\infty). \quad (1.1)$$

El razonamiento obtenido para obtener las ecuaciones anteriores puede extenderse para sistemas de varias variables conservativas. Se va a seguir –de acuerdo con [7] y [20]– la siguiente definición para un sistema de este tipo:

**Definición 1.1** (Sistema de leyes de conservación). Sean  $\mathbf{f}_j: \Omega \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^m$  funciones regulares, para todo  $j \in \{1, \dots, n\}$ . El sistema de leyes de conservación asociado al vector de variables conservativas  $\mathbf{u}: \mathbb{R}^n \times \mathbb{R} \rightarrow \Omega$  viene dado por

$$\partial_t \mathbf{u}(\mathbf{x}, t) + \sum_{j=1}^n \partial_{x_j} \mathbf{f}_j(\mathbf{u}(\mathbf{x}, t)) = \mathbf{0}, \quad \forall \mathbf{x} \in \mathbb{R}^n, \quad \forall t \in \mathbb{R}. \quad (1.2)$$

Al abierto  $\Omega \subseteq \mathbb{R}^m$  se le denomina conjunto de estados y a cada una de las funciones  $\mathbf{f}_j$ , con  $j \in \{1, \dots, n\}$ , *flujos físicos* o, simplemente *flujos*, si no hay ambigüedad con las funciones que después se denominarán *flujos numéricos*.

Si se denota  $\mathbb{F} = (\mathbf{f}_1, \dots, \mathbf{f}_n)^t$ , se puede escribir (1.2) de forma más compacta mediante la expresión

$$\partial_t \mathbf{u}(\mathbf{x}, t) + \operatorname{div} \mathbb{F}(\mathbf{u}(\mathbf{x}, t)) = \mathbf{0}, \quad \forall \mathbf{x} \in \mathbb{R}^n, \quad \forall t \in [0, +\infty),$$

donde se está denotando la divergencia tensorial

$$(\operatorname{div} \mathbb{F})_i = \sum_{j=1}^m \partial_j \mathbb{F}_{ij}, \quad \forall i \in \{1, \dots, n\}.$$

El estudio teórico de este tipo de sistemas es complejo y este trabajo se ceñirá únicamente al estudio de sistemas unidimensionales. Por ello, se particularizará la definición anterior para el caso  $n = 1$ .

**Definición 1.2** (Sistema de leyes de conservación unidimensionales). Sea  $\mathbf{f}: \Omega \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^m$  una función regular. El sistema de leyes de conservación asociado al vector de variables conservativas  $\mathbf{u}: \mathbb{R} \times \mathbb{R} \rightarrow \Omega \subseteq \mathbb{R}^m$  viene dado por

$$\partial_t \mathbf{u}(x, t) + \partial_x \mathbf{f}(\mathbf{u}(x, t)) = \mathbf{0}, \quad \forall x \in \mathbb{R}, \quad \forall t \in \mathbb{R}. \quad (1.3)$$

Al abierto  $\Omega \subseteq \mathbb{R}^m$  se le denomina conjunto de estados y a la función  $\mathbf{f}$  se la denomina flujo físico del sistema. Se dice que la ecuación (1.3) está escrita en *forma conservativa*.

Una forma alternativa de escribir las leyes de conservación es la siguiente: se define la matriz jacobiana del flujo físico  $\mathbf{f}$  como

$$\mathbb{A}(\boldsymbol{\omega}) = (a_{ij}(\boldsymbol{\omega}))_{i,j=1}^m = (\partial_j f_i(\boldsymbol{\omega}))_{i,j=1}^m = \begin{pmatrix} \partial_1 f_1(\boldsymbol{\omega}) & \partial_2 f_1(\boldsymbol{\omega}) & \dots & \partial_m f_1(\boldsymbol{\omega}) \\ \partial_1 f_2(\boldsymbol{\omega}) & \partial_2 f_2(\boldsymbol{\omega}) & \dots & \partial_m f_2(\boldsymbol{\omega}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 f_m(\boldsymbol{\omega}) & \partial_2 f_m(\boldsymbol{\omega}) & \dots & \partial_m f_m(\boldsymbol{\omega}) \end{pmatrix}, \quad \forall \boldsymbol{\omega} \in \Omega, \quad (1.4)$$

donde la entrada  $a_{ij}$  de la matriz  $\mathbb{A}$  viene dada por la derivada parcial de  $f_i$ , la componente  $i$ -ésima del flujo físico, respecto a  $u_j$ , la componente  $j$ -ésima de la variable conservativa, para todo  $i, j \in \{1, \dots, m\}$ .

Usando esta matriz y la regla de la cadena, se puede escribir la ecuación (1.3) del siguiente modo,

$$\partial_t \mathbf{u}(x, t) + \mathbb{A}(\mathbf{u}(x, t)) \partial_x \mathbf{u}(x, t) = \mathbf{0}, \quad \forall x \in \mathbb{R}, \quad \forall t \in \mathbb{R}, \quad (1.5)$$

que recibe el nombre de forma no conservativa. Tal y como se verá más adelante, las formas (1.3) y (1.5) son equivalentes cuando  $\mathbf{u}$  es suficientemente regular. Sin embargo, en general, esto no tiene por qué ser cierto. A continuación, se verán algunos ejemplos de sistemas de leyes de conservación unidimensionales.

**Ejemplo 1.3.** El caso más sencillo que se tratará será el de la *ecuación del transporte*, donde se supone que el flujo de una cierta magnitud conservativa es proporcional a su densidad. Para esta ecuación, el flujo físico viene dado por  $f(\omega) = \mu\omega$ , para todo  $\omega \in \mathbb{R}$ , donde  $\mu \in \mathbb{R}$  es una

constante (que se supone no nula), que se denomina velocidad de propagación. Por lo tanto, la ley de conservación se escribe en este caso como

$$\partial_t u(x, t) + \mu \partial_x u(x, t) = 0, \quad \forall x \in \mathbb{R}, \quad \forall t \in \mathbb{R}.$$

**Ejemplo 1.4.** Las ecuaciones linealizadas de dinámica de gases son, de acuerdo con [19]:

$$\begin{cases} \partial_t \rho + \rho_0 \partial_x v = 0, \\ \partial_t v + \frac{a^2}{\rho_0} \partial_x \rho = 0, \end{cases}$$

donde las incógnitas son la densidad  $\rho$  y la velocidad  $v$ , y donde  $\rho_0 \in \mathbb{R}$  es una constante de referencia de la densidad y  $a \in \mathbb{R}^+$  es la velocidad del sonido. Se puede escribir el sistema anterior como un sistema de leyes de conservación. Para ello, basta definir

$$\mathbf{u}(x, t) = \begin{pmatrix} \rho(x, t) \\ v(x, t) \end{pmatrix}, \quad \mathbf{f}(\mathbf{u}(x, t)) = \begin{pmatrix} \rho_0 v(x, t) \\ \frac{a^2}{\rho_0} \rho(x, t) \end{pmatrix}, \quad \forall x \in \mathbb{R}, \quad \forall t \in \mathbb{R}.$$

## 1.2. Clasificación de sistemas de leyes de conservación

Antes de continuar, se recuerdan los siguientes conceptos básicos de álgebra lineal:

**Definición 1.5** (Autovalor y autovector). Se dice que  $\lambda \in \mathbb{C}$  es un autovalor de una matriz cuadrada  $\mathbb{A} \in \mathcal{M}_{m \times m}(\mathbb{R})$  si existe un  $\mathbf{e} \in \mathbb{R}^m \setminus \{\mathbf{0}\}$  tal que  $\mathbb{A}\mathbf{e} = \lambda\mathbf{e}$  y en este caso se dice que  $\mathbf{e}$  es un autovector de  $\mathbb{A}$  asociado al autovalor  $\lambda$ . Además, la matriz  $\mathbb{A}$  tendrá  $m$  autovalores, no necesariamente distintos. Para cada autovalor  $\lambda_i$  se denotará a su autovector asociado por  $\mathbf{e}_i$ , para todo  $i \in \{1, \dots, m\}$ .

Se van a clasificar las leyes de conservación en base a los autovalores de su matriz jacobiana asociada (1.4).

**Definición 1.6** (Sistema hiperbólico de leyes de conservación). Un sistema de leyes de conservación unidimensionales de la forma (1.3) es hiperbólico en  $(x, t) \in \mathbb{R} \times \mathbb{R}$  si la matriz jacobiana del flujo físico  $\mathbf{f}$  en dicho punto, que se denotaba por  $\mathbb{A}(\mathbf{u}(x, t))$ , tiene  $m$  autovalores reales

$$\lambda_1(\mathbf{u}(x, t)) \leq \lambda_2(\mathbf{u}(x, t)) \leq \dots \leq \lambda_m(\mathbf{u}(x, t))$$

y los autovectores asociados  $\{\mathbf{e}_1(\mathbf{u}(x, t)), \mathbf{e}_2(\mathbf{u}(x, t)), \dots, \mathbf{e}_m(\mathbf{u}(x, t))\}$  son linealmente independientes. El sistema es estrictamente hiperbólico en dicho punto si además los autovalores son todos distintos entre sí. El sistema es hiperbólico (respectivamente, estrictamente hiperbólico) si es hiperbólico en todo  $(x, t) \in \mathbb{R} \times \mathbb{R}$ .

**Definición 1.7** (Sistema elíptico de leyes de conservación). Un sistema de leyes de conservación unidimensionales de la forma (1.3) es elíptico en  $(x, t) \in \mathbb{R} \times \mathbb{R}$  si la matriz jacobiana del flujo numérico  $\mathbf{f}$  en dicho punto, que se denotaba por  $\mathbb{A}(\mathbf{u}(x, t))$ , no tiene autovalores reales. El sistema es elíptico si la matriz  $\mathbb{A}$  no tiene autovalores reales en ningún punto.

Este trabajo se centra únicamente en el estudio de sistemas de leyes de conservación hiperbólicos. Un caso particular de sistemas hiperbólicos de leyes de conservación son los sistemas lineales. El interés de este caso viene de que se puedan resolver estos sistemas de forma exacta y con ello tener problemas de prueba para los métodos numéricos.

**Definición 1.8** (Sistema lineal de leyes de conservación). Un sistema de leyes de conservación unidimensionales de la forma (1.3) es lineal si existe una matriz  $\mathbb{A} \in \mathcal{M}_{m \times m}(\mathbb{R})$  de coeficientes constantes tal que el flujo físico del sistema se pueda escribir como  $\mathbf{f}(\boldsymbol{\omega}) = \mathbb{A}\boldsymbol{\omega}$ , para todo  $\boldsymbol{\omega} \in \Omega$ . De este modo, se puede reescribir el sistema como

$$\partial_t \mathbf{u}(x, t) + \mathbb{A} \partial_x \mathbf{u}(x, t) = \mathbf{0}, \quad \forall x \in \mathbb{R}, \quad \forall t \in \mathbb{R}. \quad (1.6)$$

En este caso, la matriz jacobiana del flujo físico coincide con la propia matriz  $\mathbb{A}$  que define el sistema lineal. Si el sistema sea hiperbólico, la matriz  $\mathbb{A}$  es diagonalizable en  $\mathbb{R}$ , es decir, admite una descomposición  $\mathbb{A} = \mathbb{P}\mathbb{D}\mathbb{P}^{-1}$ , siendo  $\mathbb{P} = (\mathbf{e}_1 | \dots | \mathbf{e}_m)^t \in \mathcal{M}_{m \times m}(\mathbb{R})$  una matriz no singular y  $\mathbb{D} = (d_{ii})_{i=1}^m \in \mathcal{M}_{m \times m}(\mathbb{R})$  la matriz diagonal dada por  $d_{ii} = \lambda_i$ , con  $\lambda_i \in \mathbb{R}$  autovalor con autovector asociado  $\mathbf{e}_i \in \mathbb{R}^m$ , para todo  $i \in \{1, \dots, m\}$ . Esta estructura particular de la matriz  $\mathbb{A}$  será lo que permita detallar explícitamente la solución del sistema lineal.

### 1.3. Ejemplos de sistemas de leyes de conservación

Se van a ver ahora más ejemplos de sistemas de leyes de conservación, viendo además si son hiperbólicos o no.

**Ejemplo 1.9.** La ecuación del transporte que se detallaba en el Ejemplo 1.3 se puede escribir como un sistema lineal de leyes de conservación definiendo la matriz  $\mathbb{A} = (\mu)$ . Trivialmente, es siempre estrictamente hiperbólica.

**Ejemplo 1.10.** En el Ejemplo 1.4 se vio como las ecuaciones linealizadas de dinámica de los gases se pueden escribir como un sistema lineal de leyes de conservación, considerando la matriz

$$\mathbb{A} = \begin{pmatrix} 0 & \rho_0 \\ \frac{a^2}{\rho_0} & 0 \end{pmatrix}.$$

Se va a ver ahora que este sistema lineal es hiperbólico. Para ello, se determinan cuáles son los autovalores de  $\mathbb{A}$  calculando las raíces del polinomio característico. Se tiene que

$$|\mathbb{A} - \lambda I| = \begin{vmatrix} -\lambda & \rho_0 \\ \frac{a^2}{\rho_0} & -\lambda \end{vmatrix} = \lambda^2 - a^2 = 0,$$

por lo tanto la matriz  $\mathbb{A}$  tiene dos autovalores reales y distintos, dados por  $\lambda_1 = -a$  y  $\lambda_2 = a$ . Esto prueba que el sistema es también hiperbólico.

**Ejemplo 1.11.** El siguiente sistema, deducido en [19], se conoce como el modelo de la pequeña perturbación:

$$\begin{cases} \partial_x u - a^2 \partial_y v = 0, \\ \partial_x v - \partial_y u = 0, \end{cases} \quad (1.7)$$

donde  $a^2$  viene dado por

$$a^2 = \frac{1}{M^2 - 1},$$

siendo  $M \in \mathbb{R}^+$  una constante y donde  $u$  y  $v$  modelan pequeñas perturbaciones en las velocidades de las componentes  $x$  e  $y$  (respectivamente) de una partícula. Se puede escribir este sistema de ecuaciones diferenciales como un sistema lineal de leyes de conservación, definiendo

$$\mathbf{u}(x, y) = \begin{pmatrix} u(x, y) \\ v(x, y) \end{pmatrix}, \quad \mathbb{A} = \begin{pmatrix} 0 & -a^2 \\ -1 & 0 \end{pmatrix}.$$

Con esta notación, se puede escribir el modelo de pequeñas perturbaciones como

$$\partial_x \mathbf{u}(x, y) + \mathbb{A} \partial_y \mathbf{u}(x, y) = \mathbf{0}, \quad \forall (x, y) \in \mathbb{R}^2.$$

Se van a estudiar ahora los autovalores de  $\mathbb{A}$ , para ello se calculan las raíces de su polinomio característico:

$$|\mathbb{A} - \lambda I| = \begin{vmatrix} -\lambda & -a^2 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 - a^2 = 0 \iff \lambda^2 = a^2 = \frac{1}{M^2 - 1}.$$

Si  $M > 1$ , entonces  $a^2 > 0$  y la matriz  $\mathbb{A}$  tiene dos autovalores reales y distintos entre si, dados por

$$\lambda_1 = -\sqrt{\frac{1}{M^2 - 1}} \in \mathbb{R}, \quad \lambda_2 = \sqrt{\frac{1}{M^2 - 1}} \in \mathbb{R},$$

por lo que el sistema sería estrictamente hiperbólico. Si  $M < 1$ , los autovalores serían imaginarios puros, por lo que el sistema no sería hiperbólico en ningún punto, si no que sería elíptico.

**Ejemplo 1.12.** Cabe notar que todos los ejemplos anteriores son sistemas de leyes de conservación lineales. Los modelos lineales tienen un interés académico indiscutible ya que se puede calcular la solución exacta explícitamente. Esto proporcionará una herramienta útil para comprobar el buen funcionamiento de los métodos numéricos que se desarrollarán más adelante. Desde

el punto de vista de la modelización, bajo hipótesis de pequeñas perturbaciones, también son de utilidad. Sin embargo, con frecuencia los modelos lineales no son lo suficientemente precisos.

En este ejemplo se va a presentar, siguiendo a [7], una ley de conservación no lineal, escalar y unidimensional, que se utilizará sistemáticamente para probar los métodos numéricos que se planteen, por su simplicidad (dentro del campo de ecuaciones no lineales).

Particularizando en el caso unidimensional y escalar, en el razonamiento inicial que se hizo al comienzo de este capítulo, se supuso que la función  $F \in \mathcal{C}^2(\mathbb{R} \times \mathbb{R}, \mathbb{R})$ , tal que  $F(x, t)$  es el flujo de la magnitud conservativa en la posición  $x \in \mathbb{R}$  e instante  $t \in \mathbb{R}$ , admitía una expresión del tipo  $F(x, t) = f(u(x, t))$ , con  $f \in \mathcal{C}^2(\Omega, \mathbb{R})$ . Sin embargo, a la hora de modelar problemas físicos, es más realista suponer

$$F(x, t) = f(u(x, t)) - \nu \partial_x u(x, t),$$

donde  $\nu \in \mathbb{R}^+$  es un parámetro positivo y, en general, *pequeño*. Suponiendo el flujo de esta forma y razonando de modo análogo a como se hizo anteriormente, la forma general de una ley de conservación escalar y unidimensional sería

$$\partial_t u(x, t) + \partial_x f(u(x, t)) - \nu \partial_{xx} u(x, t) = 0, \quad \forall x \in \mathbb{R}, \quad \forall t \in \mathbb{R}.$$

El término  $-\nu \partial_{xx} u(x, t)$  permite tener en cuenta los efectos de disipación de la energía y de dispersión. Por analogía con la dinámica de gases, se llamará término de *viscosidad*.

Se puede generalizar este razonamiento al caso de  $m$  variables conservativas, escribiendo la forma general de una ley de conservación con viscosidad como

$$\partial_t \mathbf{u}(x, t) + \partial_x \mathbf{f}(\mathbf{u}(x, t)) - \nu \partial_{xx} \mathbf{u}(x, t) = \mathbf{0}, \quad \forall x \in \mathbb{R}, \quad \forall t \in \mathbb{R}.$$

El caso particular en el que el flujo viene dado por  $f(\omega) = \frac{\omega^2}{2}$ ,  $\forall \omega \in \mathbb{R}$ , es el modelo más simple que permite modelar el comportamiento de un fluido viscoso. La ecuación en derivadas parciales correspondiente se escribe en forma no conservativa como

$$\partial_t u(x, t) + u \partial_x u(x, t) - \nu \partial_{xx} u(x, t) = 0, \quad \forall x \in \mathbb{R}, \quad \forall t \in \mathbb{R} \quad (1.8)$$

y es la denominada *ecuación de Burgers* (con viscosidad). El problema de Cauchy asociado a la ecuación (1.8) tiene solución única y regular, que además se puede calcular analíticamente (usando la transformada de Cole-Hopf; para más información, consultar [4] y [11]). La ecuación en derivadas parciales que surge de tomar el límite  $\nu \rightarrow 0$  es, escrita en forma no conservativa,

$$\partial_t u(x, t) + u \partial_x u(x, t) = 0, \quad \forall x \in \mathbb{R}, \quad \forall t \in \mathbb{R}$$

y, en forma conservativa,

$$\partial_t u(x, t) + \partial_x \frac{u(x, t)^2}{2} = 0, \quad \forall x \in \mathbb{R}, \quad \forall t \in \mathbb{R} \quad (1.9)$$

y se denominada *ecuación de Burgers* (sin viscosidad). Es importante señalar que es una ley de conservación en el sentido que se estaba tratando en este capítulo, es decir, en el sentido de la Definición 1.1.

**Ejemplo 1.13.** A continuación se verá un sistema de leyes de conservación no lineal que se utilizará como paradigma de problema vectorial a lo largo de todo este documento. Se trata de un caso concreto de las ecuaciones de Saint-Venant (o de aguas someras) en el caso unidimensional.

Se supone que un fluido incompresible se mueve a velocidad  $v$  por un canal con fondo plano, de ancho 1 m y con fluido con una altura variable  $h$  y tal que  $v$  es constante a lo largo de cada sección del canal. El siguiente sistema de ecuaciones no lineales modela la situación descrita

$$\partial_t \begin{pmatrix} h \\ hv \end{pmatrix} + \partial_x \begin{pmatrix} vh \\ hv^2 + \frac{gh^2}{2} \end{pmatrix} = \mathbf{0}, \quad \forall x \in \mathbb{R}, \quad \forall t \in \mathbb{R}, \quad (1.10)$$

donde  $g$  es la aceleración gravitatoria. Se pueden consultar más detalles del modelo físico, así como de la deducción de las ecuaciones, en [18].

Se puede escribir el sistema (1.10) como una ley de conservación definiendo el vector de magnitudes conservativas y el flujo físico como

$$\mathbf{u}(x, t) = \begin{pmatrix} h(x, t) \\ h(x, t)v(x, t) \end{pmatrix}, \quad \mathbf{f}(\omega_1, \omega_2) = \begin{pmatrix} \omega_2 \\ \frac{\omega_2^2}{\omega_1} + \frac{g\omega_1^2}{2} \end{pmatrix}, \quad \forall x \in \mathbb{R}, \quad \forall t \in \mathbb{R}, \quad \forall (\omega_1, \omega_2) \in \Omega, \quad (1.11)$$

donde el conjunto de estados viene dado por  $\Omega = (0, +\infty) \times \mathbb{R}$ . Este sistema es estrictamente hiperbólico. En efecto, la matriz jacobiana del flujo físico viene dada por

$$\mathbb{A}(\boldsymbol{\omega}) = \begin{pmatrix} 0 & 1 \\ -\frac{\omega_2^2}{\omega_1^2} + g\omega_1 & \frac{2\omega_2}{\omega_1} \end{pmatrix},$$

con autovalores

$$\lambda_1(\boldsymbol{\omega}) = \frac{\omega_2}{\omega_1} - \sqrt{g\omega_1}, \quad \lambda_2(\boldsymbol{\omega}) = \frac{\omega_2}{\omega_1} + \sqrt{g\omega_1}.$$

Nótese que  $\lambda_1(\boldsymbol{\omega}) \neq \lambda_2(\boldsymbol{\omega})$  para todo  $\boldsymbol{\omega} \in \Omega$ . En consecuencia, se concluye que el sistema (1.10) es un sistema de leyes de conservación estrictamente hiperbólico.

## Capítulo 2

# Soluciones del problema de Cauchy

A la hora de resolver sistemas de leyes de conservación, se impone una condición inicial en el instante de tiempo  $t = 0$ . Este es el denominado *problema de Cauchy*, que en el caso unidimensional se escribe

$$\begin{cases} \partial_t \mathbf{u}(x, t) + \partial_x \mathbf{f}(\mathbf{u}(x, t)) = \mathbf{0}, & \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty), \\ \mathbf{u}(x, 0) = \mathbf{u}_0(x), & \forall x \in \mathbb{R}, \end{cases} \quad (2.1)$$

donde  $\Omega \subseteq \mathbb{R}^m$  es un abierto,  $\mathbf{f} \in \mathcal{C}^2(\Omega, \mathbb{R}^m)$  y se supone inicialmente que  $\mathbf{u}_0 \in \mathcal{C}^1(\mathbb{R}, \Omega)$ .

### 2.1. Soluciones clásicas

Puesto que en (2.1) aparecen derivadas de primer orden respecto a la variable  $t$  y a la variable  $x$ , parece razonable pedirle a las soluciones de dicho problema que sean derivables en ambas variables y estas derivadas parciales sean además continuas. Esto motiva la siguiente definición:

**Definición 2.1** (Solución clásica). Se dice que una función  $\mathbf{u}: \mathbb{R} \times [0, T) \rightarrow \Omega \subseteq \mathbb{R}^m$ , con tiempo maximal de definición  $0 < T \leq +\infty$ , es una *solución clásica* del problema de Cauchy (2.1) si es una función  $\mathcal{C}^1(\mathbb{R} \times [0, T), \Omega)$  y satisface (2.1) en todo punto  $(x, t) \in \mathbb{R} \times [0, T)$ . Si no se especifica el tiempo maximal  $T$ , se supone que es una solución clásica en  $\mathbb{R} \times [0, +\infty)$ .

En este capítulo se verá que, en el caso escalar, existe solución clásica del problema de Cauchy (2.1), al menos para  $T$  suficientemente pequeño. Para ello, se empleará el método de las curvas características, que permite calcular la solución exacta de una ley de conservación escalar hasta un tiempo finito en el caso no lineal y en  $[0, +\infty)$  en el caso lineal.

Para dar una idea intuitiva, se comienza tratando la ecuación del transporte, con problema

de Cauchy asociado, escrito en forma no conservativa como

$$\begin{cases} \partial_t u(x, t) + \lambda \partial_x u(x, t) = 0, & \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty), \\ u(x, 0) = u_0(x), & \forall x \in \mathbb{R}, \end{cases} \quad (2.2)$$

donde  $\lambda \in \mathbb{R}$  es una constante (que se llamaba velocidad de propagación) y  $u_0 \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ . Considerando las parciales como operadores lineales, la ecuación en derivadas parciales anterior se escribe

$$(\partial_t + \lambda \partial_x) u(x, t) = 0.$$

Esto sugiere considerar un cambio de variable  $x = \xi + \lambda t$ , y definir una función  $w$  tal que

$$u(x, t) = u(\xi + \lambda t, t) = w(\xi, t),$$

con  $\xi \in \mathbb{R}$ . Entonces, se tiene que

$$\partial_t w(\xi, t) = \lambda \partial_x u(\xi + \lambda t, t) + \partial_t u(\xi + \lambda t, t) = \lambda \partial_x u(\xi + \lambda t, t) - \lambda \partial_x u(\xi + \lambda t, t) = 0.$$

Fijado  $\xi$ ,  $w$  es constante en tiempo, es decir, se tiene que  $u$  es constante a lo largo de la curva  $t \mapsto (\xi + \lambda t, t)$ , lo que implica que se puede calcular la solución en  $(x, t)$  remontando al punto de corte de la curva con el eje  $t = 0$ , pues se conoce la solución en dicho tiempo. Es decir, se tiene que

$$u(x, t) = w(\xi, t) = w(x - \lambda t, 0) = u_0(x - \lambda t).$$

Se dice que la curva  $t \mapsto (x + \lambda t, t)$  es una curva característica de la ecuación del transporte. En la Figura 2.1 se tiene una ilustración de este procedimiento de resolución.

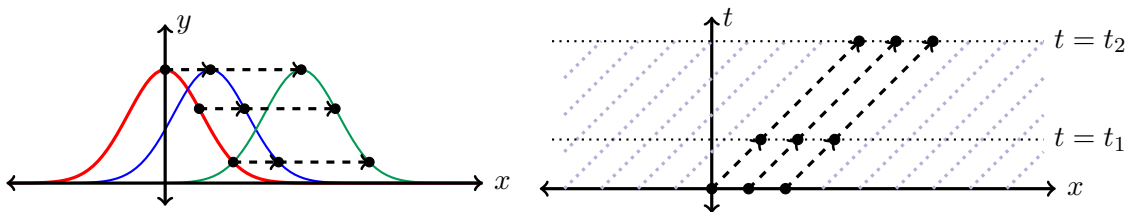


Figura 2.1: Diagrama ilustrativo de la solución de la ecuación del transporte. Figura de la izquierda: en rojo, condición inicial, en azul, solución en tiempo  $t_1$ , en verde, solución en tiempo  $t_2$ ; y las flechas representan el transporte de tres puntos concretos. Figura de la derecha: curvas características del problema. Están resaltadas las características de los puntos transportados en la figura de la izquierda.

Se acaba de ver como calcular la solución exacta de la ecuación del transporte escalar. En este apartado se verá, basándose en [19], como a partir de esta solución, se puede calcular también

de forma explícita la solución clásica de sistemas lineales hiperbólicos de leyes de conservación, en el caso unidimensional. El problema de Cauchy asociado a este sistema viene dado por

$$\begin{cases} \partial_t \mathbf{u}(x, t) + \mathbb{A} \partial_x \mathbf{u}(x, t) = \mathbf{0}, & \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty), \\ \mathbf{u}(x, 0) = \mathbf{u}_0(x), & \forall x \in \mathbb{R}. \end{cases} \quad (2.3)$$

Tal y como se comentó en el capítulo anterior, si se supone que el sistema es hiperbólico, entonces la matriz  $\mathbb{A}$  admite una descomposición del tipo  $\mathbb{A} = \mathbb{P}\mathbb{D}\mathbb{P}^{-1}$ , donde  $\mathbb{P} = (\mathbf{e}_1 | \dots | \mathbf{e}_m)^t \in \mathcal{M}_{m \times m}(\mathbb{R})$  y  $\mathbb{D} = (d_{ii})_{i=1}^m \in \mathcal{M}_{m \times m}(\mathbb{R})$ , con  $d_{ii} = \lambda_i$ , siendo  $\{\lambda_1, \dots, \lambda_m\}$  los autovalores de  $\mathbb{A}$  y  $\mathbf{e}_k$  el autovector asociado al autovalor  $\lambda_k$ , para todo  $k \in \{1, \dots, m\}$ .

Teniendo en cuenta que  $\mathbb{P}$  es no singular, se define el vector de variables características como  $\mathbf{w} = \mathbb{P}^{-1}\mathbf{u}$ . Como la matriz  $\mathbb{A}$  es constante,  $\mathbb{P}$  y  $\mathbb{P}^{-1}$  también lo serán. Por lo tanto, se tiene que

$$\mathbf{u} = \mathbb{P}\mathbf{w} \implies \begin{cases} \partial_t \mathbf{u} = \mathbb{P} \partial_t \mathbf{w}, \\ \partial_x \mathbf{u} = \mathbb{P} \partial_x \mathbf{w}. \end{cases}$$

En consecuencia, se deduce que

$$\partial_t \mathbf{u}(x, t) + \mathbb{A} \partial_x \mathbf{u}(x, t) = \mathbf{0} \implies \mathbb{P} \partial_t \mathbf{w}(x, t) + \mathbb{A}\mathbb{P} \partial_x \mathbf{w}(x, t) = \mathbf{0}.$$

Multiplicando la última igualdad por  $\mathbb{P}^{-1}$  se llega al sistema escrito en *forma característica*:

$$\partial_t \mathbf{w}(x, t) + \mathbb{D} \partial_x \mathbf{w}(x, t) = \mathbf{0}, \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty). \quad (2.4)$$

Se puede escribir (2.4) matricialmente como

$$\begin{pmatrix} \partial_t w_1(x, t) \\ \partial_t w_2(x, t) \\ \vdots \\ \partial_t w_m(x, t) \end{pmatrix} + \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_m \end{pmatrix} \begin{pmatrix} \partial_x w_1(x, t) \\ \partial_x w_2(x, t) \\ \vdots \\ \partial_x w_m(x, t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

con lo que se observa fácilmente que se trata de un sistema desacoplado de  $m$  ecuaciones de la forma

$$\partial_t w_i(x, t) + \lambda_i \partial_x w_i(x, t) = 0, \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty), \quad i \in \{1, \dots, m\},$$

que son ecuaciones del transporte, de las cuales ya se estudió su solución exacta. Poniendo  $\mathbf{w}_0 = \mathbb{P}^{-1}\mathbf{u}_0$ , las soluciones de las  $m$  ecuaciones desacopladas vienen dadas por

$$w_i(x, t) = (\mathbf{w}_0)_i(x - \lambda_i t), \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty), \quad i \in \{1, \dots, m\},$$

por lo que, revirtiendo el cambio de variable, se puede dar la solución de (2.3) como  $\mathbf{u} = \mathbb{P}\mathbf{w}$ .

A continuación, se va generalizar el desarrollo de la ecuación del transporte escalar al caso no lineal. Se considerará ahora el problema de Cauchy escalar, escrito en general como

$$\begin{cases} \partial_t u(x, t) + \partial_x f(u(x, t)) = 0, & \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty), \\ u(x, 0) = u_0(x), & \forall x \in \mathbb{R}, \end{cases} \quad (2.5)$$

suponiendo que  $\Omega \subset \mathbb{R}$  es un abierto,  $u_0 \in \mathcal{C}^1(\mathbb{R}, \Omega)$  y  $f \in \mathcal{C}^2(\Omega, \mathbb{R})$ . Si  $u: \mathbb{R} \times [0, +\infty) \rightarrow \Omega \subseteq \mathbb{R}$  es una solución clásica del problema de Cauchy (2.5), se puede escribir la ecuación en derivadas parciales anterior en forma no conservativa:

$$\partial_t u(x, t) + f'(u(x, t)) \partial_x u(x, t) = 0, \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty). \quad (2.6)$$

Por analogía con lo ocurrido con la ecuación del transporte, se supone que existe una curva  $t \mapsto (X(t), t)$  sobre la cual la solución  $u: \mathbb{R} \times [0, +\infty) \rightarrow \Omega$  es constante. Entonces se tiene que

$$\begin{aligned} 0 &= \frac{d}{dt} u(X(t), t) = \partial_x u(X(t), t) X'(t) + \partial_t u(X(t), t) \\ &= \partial_x u(X(t), t) X'(t) - f'(u(X(t), t)) \partial_x u(X(t), t) = \partial_x u(X(t), t) [X'(t) - f'(u(X(t), t))]. \end{aligned}$$

Por lo tanto, sobre una curva  $t \mapsto (X(t), t)$  del plano  $x/t$  definida por  $X'(t) = f'(u(X(t), t))$ , para todo  $x \in \mathbb{R}$  y  $t \in [0, +\infty)$ , la solución de (2.5) permanece constante. Estas ideas sugieren la siguiente definición:

**Definición 2.2** (Curva característica). Sea una ley de conservación escalar unidimensional escrita en forma no conservativa como (2.6) y sea  $(\bar{x}, \bar{t}) \in \mathbb{R} \times [0, +\infty)$  un punto arbitrario. Se define la curva característica asociada a (2.6) y que pasa por  $(\bar{x}, \bar{t})$  como la curva  $t \mapsto (X(t), t)$  del plano  $x/t$  definida por el problema de valor inicial

$$\begin{cases} X'(t) = f'(u(X(t), t)), \\ X(\bar{t}) = \bar{x}. \end{cases} \quad (2.7)$$

De los razonamientos anteriores, se deduce el siguiente resultado:

**Proposición 2.3.** Si  $t \mapsto (X(t), t)$  es una curva característica, definida en  $I \subseteq [0, +\infty)$  por (2.7), asociada a la ecuación (2.6) que pasa por  $(\bar{x}, \bar{t})$  y  $u \in \mathcal{C}^1(\mathbb{R} \times I, \Omega)$  es una solución de esta ley de conservación, entonces  $u$  es constante a lo largo de la curva característica.

Gracias a esta invariancia de la solución a lo largo de las características, se puede explicitar la solución de (2.7) como

$$X(t) = \bar{x} + f'(u(\bar{x}, \bar{t}))(t - \bar{t}), \quad \forall t \in [0, +\infty). \quad (2.8)$$

En particular, observamos que se trata de rectas. Para que este procedimiento funcione, por cada punto del plano  $(x, t)$  debe pasar una *única* curva característica. En caso contrario, la

solución podría estar multievaluada, lo cual nos llevaría a una contradicción. En la Figura 2.2 se ve qué ocurre cuando no se cumple dicha condición, considerando, por ejemplo, la ecuación de Burgers. A este fenómeno se le denomina *choque*. A partir de un cierto tiempo  $t^*$ , se producirá un choque, es decir, las características se cortan y ya no se puede hablar de solución en el sentido clásico.

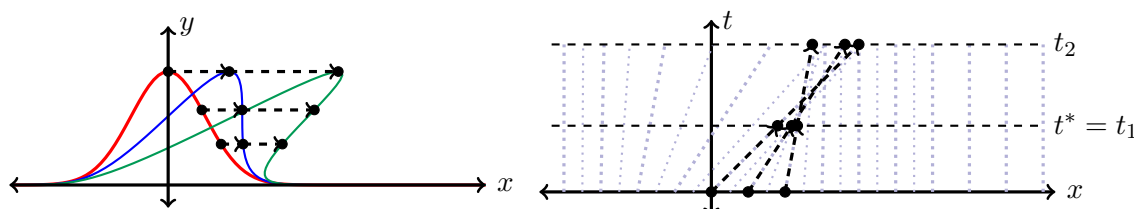


Figura 2.2: Diagrama ilustrativo de la solución de la ecuación de Burgers por el método de las características. Figura de la izquierda: en rojo, condición inicial, en azul, solución en tiempo  $t_1$ , en verde, solución en tiempo  $t_2$ ; y las flechas representan el transporte de tres puntos concretos. Se ve como a partir de  $t_1$ , ya no hay solución, pues el objeto representado ya no es una función. Figura de la derecha: curvas características del problema. Están resaltadas las características de los puntos transportados en la figura de la izquierda.

El siguiente resultado da una condición suficiente para que no haya solución clásica globalmente en todo  $[0, +\infty)$  y determina cómo calcular el tiempo maximal  $t^*$  de existencia de dicha solución clásica.

**Teorema 2.4** ([7], pág. 26). *Si la aplicación  $u_0$  está acotada, el problema de Cauchy (2.5) tiene solución clásica definida globalmente en  $[0, t^*)$ , donde  $t^* \in [0, +\infty]$  viene dado por*

$$t^* = \begin{cases} +\infty & \text{si } f''(u_0(x))u_0'(x) \geq 0, \quad \forall x \in \mathbb{R}, \\ \frac{-1}{\inf_{x \in \mathbb{R}} \{f''(u_0(x))u_0'(x)\}} & \text{en caso contrario.} \end{cases} \quad (2.9)$$

Además,  $t^*$  será el tiempo maximal de definición de dicha solución.

*Demostración.* En primer lugar, se verá que existe solución clásica única para todo  $t \in [0, t^*)$ . Se tiene que, para un  $t \in [0, t^*)$  fijado, la aplicación

$$F: \mathbb{R} \longrightarrow \mathbb{R} \\ \xi \longmapsto F(\xi) = \xi + f'(u_0(\xi))t$$

es una biyección. En efecto, se tiene que

$$F'(\xi) = 1 + f''(u_0(\xi))u_0'(\xi)t \geq 1 - \frac{t}{t^*} > 0, \quad \forall \xi \in \mathbb{R},$$

por lo que  $F$  es creciente. Pero, además, por estar  $u_0$  acotada,

$$\lim_{\xi \rightarrow -\infty} F(\xi) = -\infty, \quad \text{y} \quad \lim_{\xi \rightarrow \infty} F(\xi) = \infty,$$

lo cual prueba que  $F$  es una biyección. En consecuencia, para todo  $(x, t) \in \mathbb{R} \times [0, t^*)$  existe un único  $\xi(x, t)$  tal que

$$x = \xi(x, t) + f'(u_0(\xi(x, t)))t \implies u(x, t) = u_0(\xi(x, t)),$$

lo cual prueba la existencia de solución clásica en  $[0, t^*)$ . Hay que demostrar ahora que el tiempo maximal es igual a  $t^*$ . Se define

$$t_M = \sup \{t \in (0, t^*): \text{ existe solución clásica en } \mathbb{R} \times [0, t)\}.$$

De lo anterior, se deduce que  $t_M \geq t^*$ . Hay que ver ahora que  $t_M \leq t^*$ , lo cual probaría que  $t_M = t^*$ . Sean  $\xi_1, \xi_2 \in \mathbb{R}$  con  $\xi_1 < \xi_2$ , tales que  $f'(u_0(\xi_1)) > f'(u_0(\xi_2))$ . Entonces las rectas características que surgen de los puntos  $(\xi_1, t)$  y  $(\xi_2, t)$  se cortan en el punto  $(\bar{x}, \bar{t})$ , dado por

$$\bar{x} = \xi_1 + f'(u_0(\xi_1))t = \xi_2 + f'(u_0(\xi_2))t, \quad \bar{t} = \frac{\xi_2 - \xi_1}{f'(u_0(\xi_1)) - f'(u_0(\xi_2))}.$$

Si existe solución clásica definida globalmente en el instante  $\bar{t}$ , entonces debe ocurrir que, remontando por las características,

$$u(\bar{x}, \bar{t}) = u_0(\xi_1) = u_0(\xi_2),$$

lo cual no es posible, pues  $f'(u_0(\xi_1)) \neq f'(u_0(\xi_2))$ . Entonces  $\bar{t} \geq t_M$ . Sea ahora  $\xi \in \mathbb{R}$  tal que  $f''(u_0(\xi))u'_0(\xi) < 0$ . Para un  $h > 0$  lo suficientemente pequeño, se tiene que  $f'(u_0(\xi + h)) < f'(u_0(\xi))$ . Tomando  $\xi_1 = \xi$  y  $\xi_2 = \xi + h$ , del razonamiento anterior se deduce que, tomando el paso al límite cuando  $h \rightarrow 0^+$ ,

$$t_M \leq \lim_{h \rightarrow 0^+} \bar{t}(\xi, \xi + h) = \lim_{h \rightarrow 0^+} \frac{h}{f'(u_0(\xi)) - f'(u_0(\xi + h))} = \frac{-1}{f''(u_0(\xi))u'_0(\xi)} \leq t^*,$$

lo cual implica que  $t_M = t^*$ , como se quería ver.  $\square$

**Corolario 2.5.** Si  $u_0 \in C^1(\mathbb{R}, \Omega)$  es una función creciente en todo  $\mathbb{R}$  y  $f \in C^2(\Omega, \mathbb{R})$  es una función estrictamente convexa, entonces el método de las características permite construir una solución clásica del problema de Cauchy (2.5) en todo tiempo  $t \in [0, +\infty)$ .

**Ejemplo 2.6.** Se considera la ecuación de Burgers y se supone la condición inicial viene dada por  $u_0(x) = \exp(-x^2)$ , para todo  $x \in \mathbb{R}$ . Se va a ver hasta qué tiempo existe solución clásica de este problema de Cauchy. En este caso, se tiene que

$$\left. \begin{array}{l} f''(u_0(\xi)) = 1 \\ u'_0(x) = -2xe^{-x^2} \end{array} \right\} \implies g(x) := f''(u_0(x))u'_0(x) = -2xe^{-x^2}, \quad \forall x \in \mathbb{R},$$

y se tiene además que

$$g'(x) = -2e^{-x^2} + 4x^2e^{-x^2} = 0 \iff x \in \left\{ -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right\}.$$

La función  $g$  en dichos puntos toma los valores

$$g\left(\frac{\sqrt{2}}{2}\right) = -\sqrt{2}e^{-\frac{1}{2}}, \quad g\left(-\frac{\sqrt{2}}{2}\right) = \sqrt{2}e^{-\frac{1}{2}},$$

de donde se concluye que

$$t^* = \frac{-1}{\inf_{x \in \mathbb{R}} g(x)} = \frac{-1}{\inf_{x \in \mathbb{R}} \{f''(u_0(x))u_0'(x)\}} = \frac{\sqrt{2}}{2}e^{\frac{1}{2}}.$$

## 2.2. Soluciones débiles

En la sección anterior se vio como, en general, no existe una solución clásica del problema de Cauchy (2.1). En esta sección, se va a extender el concepto de solución clásica para poder trabajar con choques y con soluciones regulares a trozos que podrían presentar discontinuidades.

Antes de continuar, se van a recordar algunos conceptos de análisis funcional imprescindibles para poder continuar. Para ello, se han consultado los libros clásicos de análisis funcional [2], [5] y [21]. Sea el espacio de medida  $(\Omega, \mathcal{M}, \mu)$ , donde  $\Omega$  es un conjunto,  $\mu$  es la medida de Lebesgue y  $\mathcal{M}$  es la  $\sigma$ -álgebra sobre  $\Omega$  de los conjuntos Lebesgue medibles. Para toda función  $f: \Omega \rightarrow \mathbb{R}$  medible se introducen las normas

$$\|f\|_p = \left( \int_{\Omega} |f|^p d\mu \right)^{\frac{1}{p}}, \quad \forall p \in [0, +\infty),$$

$$\|f\|_{\infty} = \inf \{t \in \mathbb{R}: \mu(\{\mathbf{x} \in \Omega: |f(\mathbf{x})| > t\}) = 0\}.$$

Estas normas permiten definir los espacios funcionales

$$L^p(\Omega) = \{f: \Omega \rightarrow \mathbb{R}: f \text{ es } \mathcal{L}\text{-medible y } \|f\|_p < +\infty\}, \quad \forall p \in [0, +\infty],$$

entendiendo el conjunto anterior como un conjunto cociente con la relación de equivalencia

$$f \sim g: \iff \mu(\{\mathbf{x} \in \Omega: f(\mathbf{x}) \neq g(\mathbf{x})\}) = 0.$$

Sin embargo, la pertenencia a estos espacios es, a veces, una condición demasiado exigente, por lo que se introducen los espacios  $L^p$  locales, definidos por

$$L^p_{\text{loc}}(\Omega) = \{f: \Omega \rightarrow \mathbb{R}: f \text{ es } \mathcal{L}\text{-medible y } f_K \in L^p(K), \forall K \Subset \Omega\}, \quad \forall p \in [0, +\infty],$$

donde la notación  $A \Subset B$  significa que  $A$  es compacto y  $A \subset B$ . Además, si  $A \subset \mathbb{R}^n$  es un conjunto abierto, se va a denotar

$$\begin{cases} L^p(\Omega, A) = \{f \in L^p(\Omega)^m : f(\Omega) \subset A\}, \\ L^p_{\text{loc}}(\Omega, A) = \{f \in L^p_{\text{loc}}(\Omega)^m : f(\Omega) \subset A\}, \end{cases} \quad \forall p \in [0, +\infty].$$

Por otra parte, el soporte de una función  $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$  es el conjunto  $\text{sop } f = \overline{\{\mathbf{x} \in \Omega : f(\mathbf{x}) \neq 0\}}$ . Así, si  $\Omega \subset \mathbb{R}^n$  es un conjunto abierto, se definen las funciones de clase  $k$ , con  $0 \leq k \leq \infty$ , y de soporte compacto como

$$\mathcal{C}_0^k(\Omega) = \left\{ f \in \mathcal{C}^k(\Omega) : \text{sop } f \Subset \Omega \right\}.$$

Si  $K \subset \mathbb{R}^n$  es un cerrado, se define

$$\mathcal{C}_0^k(K) = \left\{ f|_K : f \in \mathcal{C}_0^k(\mathbb{R}^n) \right\}.$$

Para algunos resultados de este trabajo, será necesario trabajar con distribuciones, aunque el estudio de éstas no sea un objetivo de este trabajo. Es por ello que las explicaciones siguientes serán muy esquemáticas y, por simplicidad y brevedad, se ha decidido incluir únicamente los contenidos que hacen posible demostrar los resultados expuestos en este documento. Para un estudio más exhaustivo de estos conceptos, se remite a [5] o [3].

Dado un abierto  $\Omega \subset \mathbb{R}^n$ , se define el conjunto  $\mathcal{D}(\Omega) = \mathcal{C}_0^\infty(\Omega)$ . Los elementos  $\varphi$  de  $\mathcal{D}(\Omega)$  se denominan *funciones de ensayo*, o *funciones test*. Puede probarse (ver [3]) que este conjunto es no vacío. Se dice que una sucesión  $(\varphi_k)_{k \in \mathbb{N}}$  de elementos de  $\mathcal{D}(\Omega)$  converge a  $\varphi \in \mathcal{D}(\Omega)$  si existe un compacto  $K \subset \Omega$  tal que  $\text{sop } \varphi_k \subset K$ , para todo  $k \in \mathbb{N}$ , y si para todo multiíndice  $\alpha$ ,  $\partial^\alpha \varphi_k \rightarrow \partial^\alpha \varphi$  uniformemente. Se dice que una forma lineal  $S: \mathcal{D}(\Omega) \rightarrow \mathbb{R}$  es una distribución sobre  $\Omega$  si se cumple que, cuando  $k \rightarrow +\infty$ ,

$$[\varphi_k \rightarrow \varphi] \implies [S(\varphi_k) \rightarrow S(\varphi)].$$

Se denota por  $\mathcal{D}'(\Omega)$  al conjunto de todas las distribuciones sobre  $\Omega$ . La evaluación de cada distribución  $S \in \mathcal{D}'(\Omega)$  en una función de ensayo  $\varphi \in \mathcal{D}(\Omega)$  se denota por  $\langle S, \varphi \rangle := S(\varphi)$ . Se dice que una sucesión de distribuciones  $(S_k)_{k \in \mathbb{N}}$  converge a  $S \in \mathcal{D}'(\Omega)$  si se cumple que  $\langle S_k, \varphi \rangle \rightarrow \langle S, \varphi \rangle$  para toda función de ensayo  $\varphi \in \mathcal{D}(\Omega)$ .

Además, es importante hacer notar que se puede dotar a  $\mathcal{D}(\Omega)$  de una topología tal que dicho espacio sea límite inductivo de espacios de Fréchet, lo cual permitiría identificar a  $\mathcal{D}'(\Omega)$  con el dual topológico de  $\mathcal{D}(\Omega)$ .

**Proposición 2.7.** *Si  $f \in L^1_{\text{loc}}(\Omega)$ , entonces la forma lineal*

$$\varphi \in \mathcal{D}(\Omega) \longrightarrow \int_{\Omega} f \varphi \in \mathbb{R}$$

*es una distribución. Se dice que es la distribución asociada a  $f$  y se denota por  $T_f$ .*

Finalmente, se dirá que una sucesión de funciones  $(f_k)_{k \in \mathbb{N}} \subset L^1_{\text{loc}}(\Omega)$  converge a  $f \in L^1_{\text{loc}}(\Omega)$  en el sentido de las distribuciones si  $T_{f_k} \rightarrow T_f$  en  $\mathcal{D}'(\Omega)$ , es decir, si

$$\lim_{k \rightarrow \infty} \langle T_{f_k}, \varphi \rangle = \langle T_f, \varphi \rangle, \quad \forall \varphi \in \mathcal{D}(\Omega).$$

En esta sección se va a suponer que la condición inicial  $\mathbf{u}_0$  del problema de Cauchy (2.1) ya no es necesariamente continua, si no que pertenece al espacio  $L^\infty_{\text{loc}}(\mathbb{R}, \Omega)$ .

Para extender el concepto de solución clásica a situaciones que permitan discontinuidades, inicialmente se va a dar una ecuación que cumplan todas las soluciones clásicas de (2.1). Se escribirá una versión débil de (2.1) empleando integrales, en la que la función  $u$  no se vea afectada por operadores diferenciales.

**Proposición 2.8.** Sean  $\mathbf{u}_0 \in L^\infty_{\text{loc}}(\mathbb{R}, \Omega)$  y  $\mathbf{f} \in \mathcal{C}^2(\Omega, \mathbb{R}^m)$ . Si  $\mathbf{u} \in \mathcal{C}^1(\mathbb{R} \times [0, +\infty), \Omega)$  es una solución clásica del problema de Cauchy (2.1), entonces se cumple que

$$\int_0^\infty \int_{\mathbb{R}} [\mathbf{u}(x, t) \cdot \partial_t \varphi(x, t) + \mathbf{f}(\mathbf{u}(x, t)) \cdot \partial_x \varphi(x, t)] dx dt + \int_{\mathbb{R}} \mathbf{u}_0(x) \cdot \varphi(x, 0) dx = 0, \quad (2.10)$$

para toda función  $\varphi \in \mathcal{C}_0^1(\mathbb{R} \times [0, +\infty), \mathbb{R}^m)$ .

*Demostración.* Si se cumple (2.1) entonces, para una función  $\varphi \in \mathcal{C}_0^1(\mathbb{R} \times [0, +\infty), \mathbb{R}^m)$  arbitraria, se tiene que

$$\{\partial_t \mathbf{u}(x, t) + \partial_x \mathbf{f}(\mathbf{u}(x, t))\} \cdot \varphi(x, t) = 0, \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty).$$

Integrando sobre  $\mathbb{R} \times [0, +\infty)$ , se llega a que

$$\int_0^\infty \int_{\mathbb{R}} [\partial_t \mathbf{u}(x, t) \cdot \varphi(x, t) + \partial_x \mathbf{f}(\mathbf{u}(x, t)) \cdot \varphi(x, t)] dx dt = 0.$$

Ahora bien, integrando por partes,

$$\int_0^\infty \int_{\mathbb{R}} [\mathbf{u}(x, t) \cdot \partial_t \varphi(x, t) + \mathbf{f}(\mathbf{u}(x, t)) \cdot \partial_x \varphi(x, t)] dx dt = - \int_{\mathbb{R}} \mathbf{u}_0(x) \cdot \varphi(x, 0) dx,$$

llegando así a la identidad del enunciado. □

Este resultado motiva la siguiente definición:

**Definición 2.9** (Solución débil). Sean  $\mathbf{u}_0 \in L^\infty_{\text{loc}}(\mathbb{R}, \Omega)$  y  $\mathbf{f} \in \mathcal{C}^2(\Omega, \mathbb{R}^m)$ . Se dice que una función  $\mathbf{u} \in L^\infty_{\text{loc}}(\mathbb{R} \times [0, +\infty), \Omega)$  es una solución débil del problema de Cauchy (2.1) si cumple la igualdad (2.10) para toda función  $\varphi \in \mathcal{C}_0^1(\mathbb{R} \times [0, +\infty), \mathbb{R}^m)$ .

*Observación 2.10.* Si la igualdad (2.10) se cumple para toda función  $\varphi \in \mathcal{C}_0^\infty(\mathbb{R} \times (0, +\infty), \mathbb{R}^m)$ , entonces se dirá que  $\mathbf{u}$  es una solución de (2.1) en el *sentido de las distribuciones*. Nótese que, en este caso no se tiene en cuenta la condición inicial.

Como consecuencia de la Proposición (2.8), si  $\mathbf{u}$  es una solución clásica del problema de Cauchy, entonces también será una solución débil. Ahora bien, aunque el recíproco no es cierto, existe un resultado que nos da una condición suficiente para que esto ocurra.

**Proposición 2.11.** *Si  $\mathbf{u} \in \mathcal{C}^1(\mathbb{R} \times [0, +\infty), \Omega)$  es una solución débil del problema de Cauchy (2.1), entonces también es una solución clásica.*

*Demostración.* Si  $\mathbf{u} \in \mathcal{C}^1(\mathbb{R} \times [0, +\infty), \Omega)$  es una solución débil de (2.1), entonces cumple la igualdad (2.10) para toda función  $\varphi \in \mathcal{C}_0^1(\mathbb{R} \times [0, +\infty), \mathbb{R}^m)$ . En consecuencia, también cumplirá dicha igualdad para toda función  $\varphi \in \mathcal{C}_0^\infty(\mathbb{R} \times (0, +\infty), \mathbb{R}^m)$ . Por lo tanto,

$$\int_0^\infty \int_{\mathbb{R}} [\mathbf{u} \cdot \partial_t \varphi + \mathbf{f}(\mathbf{u}) \cdot \partial_x \varphi] dx dt = 0.$$

Esto significa que la igualdad

$$\partial_t \mathbf{u}(x, t) + \partial_x \mathbf{f}(\mathbf{u}(x, t)) = \mathbf{0}$$

se verifica en el *sentido de las distribuciones*. Pero como  $\mathbf{u}$  es regular, entonces la igualdad anterior también se cumplirá en el sentido clásico. Hay que ver ahora que  $\mathbf{u}$  también cumple la condición inicial. Sea  $\varphi \in \mathcal{C}_0^1(\mathbb{R} \times [0, +\infty), \mathbb{R}^m)$ . Entonces, integrando la igualdad (2.10) por partes, se tiene que

$$\begin{aligned} & - \int_0^\infty \int_{\mathbb{R}} [\partial_t \mathbf{u}(x, t) + \partial_x \mathbf{f}(\mathbf{u}(x, t))] \cdot \varphi(x, t) dx dt + \int_{\mathbb{R}} [\mathbf{u}_0(x) - \mathbf{u}(x, 0)] \cdot \varphi(x, 0) dx \\ & = \int_{\mathbb{R}} [\mathbf{u}_0(x) - \mathbf{u}(x, 0)] \cdot \varphi(x, 0) dx = 0. \end{aligned}$$

Como la función  $\varphi$  era arbitraria, se deduce que  $\mathbf{u}(x, 0) = \mathbf{u}_0(x)$  para todo  $x \in \mathbb{R}$ , como se quería probar.  $\square$

En particular, se tiene interés en estudiar las propiedades de las soluciones débiles que sean de clase  $\mathcal{C}^1(\mathbb{R} \times [0, +\infty), \Omega)$  a trozos. Antes de continuar, se van a aclarar algunos términos de notación.

**Definición 2.12** (Función  $\mathcal{C}^1$  a trozos). Sea  $\mathcal{O}$  un abierto acotado de  $\mathbb{R} \times [0, +\infty)$ . Se dice que una función  $\mathbf{u}: \mathbb{R} \times [0, +\infty) \rightarrow \Omega \subseteq \mathbb{R}^m$  es de clase  $\mathcal{C}^1$  a trozos sobre  $\mathcal{O}$  si existe un conjunto finito de curvas  $\{\Sigma_1, \dots, \Sigma_p\}$  de la forma

$$\Sigma_i \equiv t \mapsto (\sigma_i(t), t), \quad t \in (a_i, b_i), \quad \forall i \in \{1, \dots, p\},$$

donde  $\sigma_i \in \mathcal{C}^1((a_i, b_i), \mathbb{R})$  y  $0 \leq a_i < b_i \leq +\infty$ , de modo que  $\mathbf{u}|_{\mathcal{O}}$  es la restricción de una función  $\mathcal{C}^1(\mathbb{R} \times [0, +\infty), \Omega)$  en cada componente conexa de  $\mathcal{O} \setminus \{\Sigma_1, \dots, \Sigma_p\}$ . Se dice que  $\mathbf{u}$  es de clase  $\mathcal{C}^1$  a trozos si es de clase  $\mathcal{C}^1$  a trozos sobre cualquier abierto acotado  $\mathcal{O}$  de  $\mathbb{R} \times [0, +\infty)$ .

**Definición 2.13.** Sea  $\mathbf{u}: \mathbb{R} \times [0, +\infty) \rightarrow \Omega \subseteq \mathbb{R}^m$  una función de clase  $\mathcal{C}^1$  a trozos y sea  $\Sigma \equiv t \mapsto (\sigma(t), t), t \in (a, b)$  una de sus curvas de discontinuidad en el plano  $(x, t)$ . Se define el vector normal orientado a la curva  $\Sigma$  como

$$\mathbf{n}_\Sigma(t) = \begin{pmatrix} 1 \\ -\sigma'(t) \end{pmatrix} \in \mathbb{R}^2, \quad \forall t \in (a, b).$$

Cabe notar que este vector *apunta* hacia el sentido creciente del eje  $x$ , como se puede ver en la Figura 2.3. Se definen los valores de  $\mathbf{u}$  a la izquierda y a la derecha de la curva de discontinuidad  $\Sigma$  como

$$\begin{cases} \mathbf{u}_\Sigma^-(t) = \lim_{\varepsilon \rightarrow 0^+} \mathbf{u}((\sigma(t), t) - \varepsilon \mathbf{n}_\Sigma(t)), \\ \mathbf{u}_\Sigma^+(t) = \lim_{\varepsilon \rightarrow 0^+} \mathbf{u}((\sigma(t), t) + \varepsilon \mathbf{n}_\Sigma(t)), \end{cases} \quad \forall t \in (a, b).$$

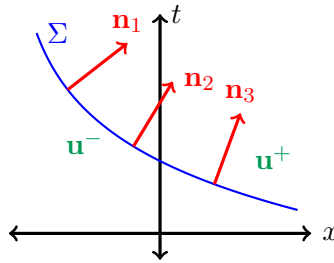


Figura 2.3: Diagrama ilustrativo del concepto de vector normal a una curva de discontinuidad, así como de los valores de la función a la izquierda y derecha de dicha curva.

Para que una función de clase  $\mathcal{C}^1$  a trozos pueda ser solución débil, no puede tener un *salto* arbitrario a lo largo de sus curvas de discontinuidad; si no que debe cumplir cierta condición que establece el siguiente teorema.

**Teorema 2.14** ([7], pág 29). *Sea  $\mathbf{u}: \mathbb{R} \times [0, +\infty) \rightarrow \Omega \subseteq \mathbb{R}^m$  una función de clase  $\mathcal{C}^1$  a trozos. Entonces  $\mathbf{u}$  es una solución débil de (2.1) si, y sólo si, satisface la condición inicial y cumple los siguientes puntos:*

(I)  $\mathbf{u}$  es una solución clásica del problema (2.1) allí donde sea regular.

(II) A lo largo de toda curva de discontinuidad  $\Sigma \equiv t \mapsto (\sigma(t), t)$ , con  $t \in (a, b)$ , se cumple que

$$\sigma'(t)(\mathbf{u}_\Sigma^+(t) - \mathbf{u}_\Sigma^-(t)) = \mathbf{f}(\mathbf{u}_\Sigma^+(t)) - \mathbf{f}(\mathbf{u}_\Sigma^-(t)), \quad \forall t \in (a, b). \quad (2.11)$$

*Demostración.* Se van a probar ambas implicaciones:

( $\implies$ ) Sea  $\mathbf{u}: \mathbb{R} \times [0, +\infty) \rightarrow \Omega \subseteq \mathbb{R}^m$  una función de clase  $\mathcal{C}^1$  a trozos y sea  $\Sigma \equiv t \mapsto (\sigma(t), t)$ , con  $t \in (a, b)$ , una curva de discontinuidad de  $\mathbf{u}$ . El punto (I) se deduce inmediatamente de

la Proposición 2.11. Para probar el otro punto, se fija un  $t_0 \in (a, b)$  y se toma un  $\varepsilon \in \mathbb{R}^+$  lo suficientemente pequeño, de modo que  $K = B((\sigma(t_0), t_0), \varepsilon)$  sea una bola abierta tal que  $\Sigma$  divida a esta bola en dos componentes conexas,  $K$  esté contenida en el semiplano  $t > 0$  y  $\Sigma \cap K$  sea la única curva de discontinuidad que pase por  $K$ . En consecuencia,  $\mathbf{u}|_{K \setminus \Sigma} \in \mathcal{C}^1(K \setminus \Sigma, \Omega)$ . Se denominará  $K^+$  a la componente a la que apunta el vector  $\mathbf{n}_\Sigma(t_0)$  y  $K^-$  a la otra componente.

Sea  $\varphi \in \mathcal{C}_0^1(K, \mathbb{R}^m)$ . Como  $\mathbf{u}$  es una solución débil, entonces cumple la igualdad (2.10) y, en consecuencia,

$$\int_K [\mathbf{u} \cdot \partial_t \varphi + \mathbf{f}(\mathbf{u}) \cdot \partial_x \varphi] = \int_{K^-} [\mathbf{u} \cdot \partial_t \varphi + \mathbf{f}(\mathbf{u}) \cdot \partial_x \varphi] + \int_{K^+} [\mathbf{u} \cdot \partial_t \varphi + \mathbf{f}(\mathbf{u}) \cdot \partial_x \varphi] = 0. \quad (2.12)$$

Aplicando el teorema de Green en  $K^-$  y  $K^+$  y recordando que se estaba tomando el vector normal  $\mathbf{n}_\Sigma = (1, -\sigma') \equiv (1, -s)$ , se deduce que

$$\begin{aligned} \int_{K^-} [\mathbf{u} \cdot \partial_t \varphi + \mathbf{f}(\mathbf{u}) \partial_x \varphi] &= - \int_{K^-} [\partial_t \mathbf{u} + \partial_x \mathbf{f}(\mathbf{u})] \cdot \varphi + \oint_{K \cap \Sigma} [-s \mathbf{u}_\Sigma^- + \mathbf{f}(\mathbf{u}_\Sigma^-)] \cdot \varphi \\ \int_{K^+} [\mathbf{u} \cdot \partial_t \varphi + \mathbf{f}(\mathbf{u}) \partial_x \varphi] &= - \int_{K^+} [\partial_t \mathbf{u} + \partial_x \mathbf{f}(\mathbf{u})] \cdot \varphi - \oint_{K \cap \Sigma} [-s \mathbf{u}_\Sigma^+ + \mathbf{f}(\mathbf{u}_\Sigma^+)] \cdot \varphi \end{aligned}$$

Por ser  $\mathbf{u}$  solución clásica allí donde es de clase  $\mathcal{C}^1$ , entonces las integrales de los miembros derechos sobre  $K^-$  y  $K^+$  son nulas. Así, juntando las igualdades anteriores con (2.12), se tiene que

$$\int_K [\mathbf{u} \cdot \partial_t \varphi + \mathbf{f}(\mathbf{u}) \cdot \partial_x \varphi] = \oint_{K \cap \Sigma} [-s (\mathbf{u}_\Sigma^+ - \mathbf{u}_\Sigma^-) + (\mathbf{f}(\mathbf{u}_\Sigma^+) - \mathbf{f}(\mathbf{u}_\Sigma^-))] \cdot \varphi = 0.$$

Puesto que el  $\varphi \in \mathcal{C}_0^1(K, \mathbb{R}^m)$  era arbitrario, de la igualdad anterior se deduce que la condición (2.11) se debe cumplir a lo largo de todo  $K \cap \Sigma$ . Como el  $t_0 \in (a, b)$  y la curva  $\Sigma$  también eran cualesquiera, entonces queda probado el punto (II).

( $\Leftarrow$ ) De las igualdades vistas en la otra implicación, se deduce que si  $\mathbf{u}$  cumple los puntos (I) y (II), entonces cumple la igualdad (2.10) para toda función  $\varphi \in \mathcal{C}_0^1(\mathbb{R} \times [0, +\infty), \mathbb{R}^m)$ .

□

*Observación 2.15.* La condición (2.11) se denomina *condición de Rankine-Hugoniot* y se escribe de forma abreviada como

$$s[[\mathbf{u}]] = [[\mathbf{f}(\mathbf{u})]], \quad (2.13)$$

donde  $s(t) = \sigma'(t)$  para todo  $t \in [a, b]$ . Cabe notar que, en el caso escalar, se puede usar esta igualdad para despejar la velocidad del choque:

$$s(t) = \sigma'(t) = \frac{[[f(u)]]}{[[u]]}, \quad \forall t \in [a, b]. \quad (2.14)$$

**Corolario 2.16.** Si  $u \in C^0(\mathbb{R} \times [0, +\infty), \Omega)$  es una función de clase  $C^1$  a trozos y es una solución clásica del problema de Cauchy (2.1) donde es regular, entonces es una solución débil de dicho problema.

A continuación, se va a ver como reescribir la condición de Rakine-Hugoniot para algunos de los ejemplos que se estudiaron en el capítulo anterior.

**Ejemplo 2.17.** En la ecuación del transporte, el flujo viene dado por  $f(\omega) = \mu\omega$ , para todo  $\omega \in \mathbb{R}$ , con  $\mu \in \mathbb{R}$ . Si  $\Sigma \equiv t \mapsto (\sigma(t), t)$ , con  $t \in (a, b)$ , es una curva de discontinuidad, suponiendo además que  $u$  es discontinua a lo largo de  $\Sigma$ , es decir, que  $u_{\Sigma}^+(t) \neq u_{\Sigma}^-(t)$ , para todo  $t \in (a, b)$ , entonces, fijado un  $t \in (a, b)$ , se tiene que

$$\sigma'(t)(u_{\Sigma}^+(t) - u_{\Sigma}^-(t)) = f(u_{\Sigma}^+(t)) - f(u_{\Sigma}^-(t)) \iff \sigma'(t) = \mu,$$

con lo cual la condición (2.13) se puede reescribir como  $s = \mu$ . Esto significa que las posibles curvas de discontinuidad de la ecuación del transporte son precisamente las curvas características.

**Ejemplo 2.18.** El flujo de la ecuación de Burgers sin viscosidad viene dado por  $f(\omega) = \frac{\omega^2}{2}$ , para todo  $\omega \in \mathbb{R}$ . Sea  $\Sigma$  una curva de discontinuidad en las mismas condiciones que en el ejemplo anterior. Entonces, fijado un  $t \in (a, b)$ , se tiene que

$$\begin{aligned} \sigma'(t)(u_{\Sigma}^+(t) - u_{\Sigma}^-(t)) &= f(u_{\Sigma}^+(t)) - f(u_{\Sigma}^-(t)) \iff \sigma'(t)(u_{\Sigma}^+(t) - u_{\Sigma}^-(t)) = \frac{1}{2} \{ (u_{\Sigma}^+(t))^2 - (u_{\Sigma}^-(t))^2 \} \\ \iff \sigma'(t) &= \frac{(u_{\Sigma}^+(t))^2 - (u_{\Sigma}^-(t))^2}{2 \{ u_{\Sigma}^+(t) - u_{\Sigma}^-(t) \}} \iff \sigma'(t) = \frac{u_{\Sigma}^+(t) + u_{\Sigma}^-(t)}{2}, \end{aligned}$$

de modo que la condición (2.13) se puede reescribir como  $s = \frac{1}{2}(u^+ + u^-)$ .

Gracias al concepto de solución débil, se pueden extender las soluciones clásicas para el caso discontinuo pero, a cambio, se pierde la unicidad de solución, tal y como se muestra en el siguiente ejemplo.

**Ejemplo 2.19.** Considérese el problema de Cauchy asociado a la ecuación de Burgers sin viscosidad y a la siguiente condición inicial:

$$u_0(x) = \begin{cases} 0 & \text{si } x \leq 0, \\ 1 & \text{si } x > 0, \end{cases} \quad \forall x \in \mathbb{R}.$$

Tomando un punto  $(\xi, 0)$  del plano  $(x, t)$ , entonces la curva característica que pasa por el punto  $(\xi, 0)$  es

$$\begin{cases} x = \xi & \text{si } \xi \leq 0, \\ x = \xi + t & \text{si } \xi > 0. \end{cases}$$

Por tanto, usando el método de las características, se puede construir una solución clásica en todo el plano  $(x, t)$  menos en  $0 \leq x < t$ . En dicha región, la función  $(x, t) \mapsto \frac{x}{t}$  es solución

de la ecuación en derivadas parciales de (2.1). Juntando estas funciones, se obtiene una función continua

$$u(x, t) = \begin{cases} 0 & \text{si } x \leq 0, \\ \frac{x}{t} & \text{si } 0 < x \leq t, \\ 1 & \text{si } x > t, \end{cases} \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty).$$

Puesto que esta función es una solución clásica allí donde es regular y además es continua, entonces, por el Corolario 2.16, es una solución débil del problema de Cauchy. Sin embargo, esta solución débil no es única, si no que, de hecho, hay infinitas. Sea un  $\alpha \in (0, 1)$  y la función

$$u_\alpha(x, t) = \begin{cases} 0 & \text{si } x \leq \frac{\alpha}{2}t, \\ \alpha & \text{si } \frac{\alpha}{2}t < x \leq \frac{t}{2}, \\ 1 - \alpha & \text{si } \frac{t}{2} < x \leq (1 - \frac{\alpha}{2})t, \\ 1 & \text{si } x > (1 - \frac{\alpha}{2})t, \end{cases} \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty).$$

Todas estas funciones son discontinuas, verifican la ecuación en derivadas parciales allí donde son regulares, cumplen la condición inicial y verifican la condición de Rankine-Hugoniot en sus discontinuidades. En consecuencia, son soluciones débiles del problema de Cauchy (2.1) para la ecuación de Burgers.

### 2.3. Soluciones entrópicas

Para intentar solucionar los problemas que puede ocasionar la pérdida de unicidad, se va a presentar un problema auxiliar que permita seleccionar de entre todas las soluciones aquella que es físicamente aceptable. En el Ejemplo 1.12, se vio que añadir un término de viscosidad a una ley de conservación hace que se modelen los problemas físicos subyacentes de manera más realista. Es por tanto, que en vez de buscar las soluciones del problema de Cauchy (2.1), se pueden buscar las del problema

$$\begin{cases} \partial_t \mathbf{u}(x, t) + \partial_x \mathbf{f}(\mathbf{u}(x, t)) - \nu \partial_{xx} \mathbf{u}(x, t) = \mathbf{0}, & \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty), \\ \mathbf{u}(x, 0) = \mathbf{u}_0(x), & \forall x \in \mathbb{R}, \end{cases} \quad (2.15)$$

con  $\nu \in \mathbb{R}^+$ , generalmente pequeño y con  $\Omega \subseteq \mathbb{R}^m$  un abierto,  $\mathbf{u}_0 \in L^\infty(\mathbb{R}, \Omega)$  y  $\mathbf{f} \in \mathcal{C}^2(\Omega, \mathbb{R}^m)$ , con la esperanza de que este problema tenga solución única. Se utilizará (2.15) para dar un criterio para seleccionar de entre todas las soluciones débiles de (2.1) aquella que se aproxime mejor a la solución físicamente aceptable, que se llama *solución límite con viscosidad*: de entre todas las soluciones del problema original, se seleccionará aquella que es límite de las soluciones con viscosidad cuando  $\nu \rightarrow 0$ .

De aquí en adelante, se supondrá que el problema (2.15) tiene solución única  $\mathbf{u}_\nu \in L^\infty(\mathbb{R} \times [0, +\infty), \Omega)$  y que es además muy regular en  $\mathbb{R} \times (0, +\infty)$ . Para ver la prueba de esto, en el caso escalar, se remite a [7]. Este resultado en el caso general sigue siendo un problema abierto.

Antes de continuar, se recuerdan los siguientes conceptos sobre *convexidad*, extraídos de [17]:

**Definición 2.20** (Conjunto convexo). Se dice que un conjunto  $\Omega \subseteq \mathbb{R}^m$  es convexo si, para todo  $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ ,

$$\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2 \in \Omega, \quad \forall \lambda \in [0, 1].$$

**Definición 2.21** (Función convexa). Sea  $\Omega \subseteq \mathbb{R}^m$  un conjunto convexo no vacío. Se dice que una función  $f: \Omega \rightarrow \mathbb{R}$  es convexa si, para todo  $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$  y todo  $\lambda \in (0, 1)$ ,

$$f(\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2).$$

Si la desigualdad anterior es estricta para todo  $\mathbf{x}_1 \neq \mathbf{x}_2$ , se dice que  $f$  es estrictamente convexa. Finalmente, si  $-f$  es convexa (respectivamente, estrictamente convexa), se dice que  $f$  es cóncava (respectivamente, estrictamente cóncava).

Se van a utilizar más adelante las siguientes caracterizaciones de la convexidad a partir de la derivada primera o de la derivada segunda:

**Teorema 2.22** ([17], pág 42). Sea  $\Omega \subseteq \mathbb{R}^m$  un conjunto abierto, convexo y no vacío y sea  $f: \Omega \rightarrow \mathbb{R}$  una función diferenciable. Entonces  $f$  es convexa si, y sólo si,

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad \forall \mathbf{x}, \mathbf{y} \in \Omega.$$

Además,  $f$  es estrictamente convexa si, y sólo si,

$$f(\mathbf{y}) > f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{y} - \mathbf{x}), \quad \forall \mathbf{x}, \mathbf{y} \in \Omega, \mathbf{x} \neq \mathbf{y}.$$

**Teorema 2.23** ([17], pág 43). Sea  $\Omega \subseteq \mathbb{R}^m$  un conjunto abierto, convexo y no vacío y sea  $f \in \mathcal{C}^2(\Omega, \mathbb{R})$ . Se cumple que

- La función  $f$  es convexa si, y sólo si, su matriz hessiana es semidefinida positiva en todo punto de  $\Omega$ .
- La función  $f$  es estrictamente convexa si su matriz hessiana es definida positiva en todo punto de  $\Omega$ .

A continuación, se define el concepto de entropía matemática, que ayudará a seleccionar la solución límite con viscosidad de manera práctica. Para ello, de aquí en adelante se va a suponer que  $\Omega \subseteq \mathbb{R}^m$  es convexo.

**Definición 2.24** (Entropía). Sean  $\mathcal{U} \in \mathcal{C}^2(\Omega, \mathbb{R})$  y  $\mathcal{F} \in \mathcal{C}^1(\Omega, \mathbb{R})$ , con  $\Omega \subseteq \mathbb{R}^m$  convexo. Se dice que  $(\mathcal{U}, \mathcal{F})$  es una entropía para el problema (2.1) si  $\mathcal{U}$  es una función estrictamente convexa y se cumple que  $\mathcal{F}'(\boldsymbol{\omega}) = \mathcal{U}'(\boldsymbol{\omega})\mathbb{A}(\boldsymbol{\omega})$  para todo  $\boldsymbol{\omega} \in \Omega$ , donde se está denotando  $\mathcal{U}' = \nabla \mathcal{U}^t$ ,  $\mathcal{F}' = \nabla \mathcal{F}^t$  y  $\mathbb{A}$  a la matriz jacobiana de  $\mathbf{f}$ .

**Teorema 2.25** ([7], pág. 42). Sea  $(\mathbf{u}_\nu)_{\nu \in (0, +\infty)}$  una colección de soluciones regulares de (2.15) tales que

$$(i) \text{ Existe } C \geq 0, \text{ independiente de } \nu, \text{ tal que } \|\mathbf{u}_\nu\|_\infty \leq C, \quad \forall \nu \in (0, +\infty),$$

$$(ii) \text{ Existe una función } \mathbf{u} \in L^\infty(\mathbb{R} \times [0, +\infty), \Omega) \text{ tal que } \mathbf{u}_\nu \longrightarrow \mathbf{u} \text{ casi por doquier, cuando } \nu \longrightarrow 0^+.$$

Entonces  $\mathbf{u}$  es una solución del problema de Cauchy (2.1) en el sentido de las distribuciones y además, para toda entropía  $(\mathcal{U}, \mathcal{F})$  de dicho problema, se cumple la siguiente desigualdad, denominada condición de entropía,

$$\partial_t \mathcal{U}(\mathbf{u}) + \partial_x \mathcal{F}(\mathbf{u}) \leq 0 \quad (2.16)$$

en el sentido de las distribuciones, es decir, que para toda función  $\varphi \in C_0^\infty(\mathbb{R} \times (0, +\infty), \mathbb{R})$ , con  $\varphi \geq 0$  se cumple que

$$\int_0^\infty \int_{\mathbb{R}} (\mathcal{U}(\mathbf{u}) \partial_t \varphi + \mathcal{F}(\mathbf{u}) \partial_x \varphi) \geq 0.$$

*Esquema de la demostración.* Para todo  $\nu \in (0, +\infty)$ , por ser  $\mathbf{u}_\nu$  solución de (2.15), se tiene que

$$\mathcal{U}'(\mathbf{u}_\nu(x, t)) \partial_t \mathbf{u}_\nu(x, t) + \mathcal{U}'(\mathbf{u}_\nu(x, t)) \partial_x \mathbf{f}(\mathbf{u}_\nu(x, t)) - \nu \mathcal{U}'(\mathbf{u}_\nu(x, t)) \partial_{xx} \mathbf{u}_\nu(x, t) = 0,$$

para todo  $x \in \mathbb{R}$  y  $t \in [0, +\infty)$ . Utilizando la regla de la cadena y las propiedades de la entropía, se llega a que

$$\partial_t \mathcal{U}(\mathbf{u}_\nu(x, t)) + \partial_x \mathcal{F}(\mathbf{u}_\nu(x, t)) - \nu \mathcal{U}'(\mathbf{u}_\nu(x, t)) \partial_{xx} \mathbf{u}_\nu(x, t) = 0, \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty).$$

Por otra parte, se tiene que, para todo  $x \in \mathbb{R}$  y  $t \in [0, +\infty)$ ,

$$\nu \mathcal{U}'(\mathbf{u}_\nu(x, t)) \partial_{xx} \mathbf{u}_\nu(x, t) = \nu \partial_{xx} \mathcal{U}(\mathbf{u}_\nu) - \nu (\partial_x \mathbf{u}_\nu(x, t))^t \mathcal{U}''(\mathbf{u}_\nu(x, t)) \partial_x \mathbf{u}_\nu(x, t).$$

Por la convexidad estricta de  $\mathcal{U}$ ,

$$\mathcal{U}'(\mathbf{u}_\nu(x, t)) \partial_{xx} \mathbf{u}_\nu(x, t) \leq \partial_{xx} \mathcal{U}(\mathbf{u}_\nu(x, t)), \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty),$$

por lo que se llega a que

$$\partial_t \mathcal{U}(\mathbf{u}_\nu(x, t)) + \partial_x \mathcal{F}(\mathbf{u}_\nu(x, t)) \leq \nu \partial_{xx} \mathcal{U}(\mathbf{u}_\nu(x, t)), \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty).$$

De las hipótesis y del teorema de la convergencia dominada, se deduce que  $\mathbf{u}_\nu \rightarrow \mathbf{u}$  y  $\mathbf{f}(\mathbf{u}_\nu) \rightarrow \mathbf{f}(\mathbf{u})$  en el sentido de las distribuciones. Además, por la continuidad del operador derivada en el espacio de las distribuciones,

$$\partial_t \mathbf{u}_\nu \longrightarrow \partial_t \mathbf{u}, \quad \partial_x \mathbf{f}(\mathbf{u}_\nu) \longrightarrow \partial_x \mathbf{f}(\mathbf{u}),$$

de donde se concluye que (2.1) se cumple en el sentido de las distribuciones. Utilizando los mismos argumentos, se llega a que

$$\partial_t \mathcal{U}(\mathbf{u}_\nu) \longrightarrow \partial_t \mathcal{U}(\mathbf{u}), \quad \partial_x \mathcal{F}(\mathbf{u}_\nu) \longrightarrow \partial_x \mathcal{F}(\mathbf{u}), \quad \nu \partial_{xx} \mathcal{U}(\mathbf{u}_\nu) \longrightarrow 0,$$

lo cual permite concluir.  $\square$

El siguiente resultado relaciona los conceptos de solución débil y condición de entropía:

**Teorema 2.26** ([7], pág, 44). *Sea  $\mathbf{u}: \mathbb{R} \times [0, +\infty) \longrightarrow \Omega \subseteq \mathbb{R}^m$  una función de clase  $\mathcal{C}^1$  a trozos que sea una solución débil del problema de Cauchy (2.18). Entonces  $\mathbf{u}$  cumple la condición de entropía 2.16 si, y sólo si, para toda entropía  $(\mathcal{U}, \mathcal{F})$ , se tiene que, a lo largo de toda curva de discontinuidad  $\Sigma \equiv t \longmapsto (\sigma(t), t)$ , con  $t \in (a, b)$ ,*

$$\sigma'(t)(\mathcal{U}(\mathbf{u}_\Sigma^+(t)) - \mathcal{U}(\mathbf{u}_\Sigma^-(t))) \geq \mathcal{F}(\mathbf{u}_\Sigma^+(t)) - \mathcal{F}(\mathbf{u}_\Sigma^-(t)), \quad \forall t \in (a, b). \quad (2.17)$$

*Demostración.* La prueba es análoga a la del Teorema 2.14.  $\square$

El siguiente resultado asegura que toda solución clásica cumple la condición de entropía.

**Proposición 2.27.** *Si  $\mathbf{u} \in \mathcal{C}^1(\mathbb{R} \times [0, +\infty), \Omega)$  es una solución clásica del problema de Cauchy (2.1), entonces cumple la condición de entropía (2.16) para toda entropía  $(\mathcal{U}, \mathcal{F})$ ; donde, además, se alcanza la igualdad.*

*Demostración.* Por ser  $\mathbf{u}$  solución clásica, entonces cumple la ley de conservación expresada en forma no conservativa. Multiplicando esta expresión por la derivada de la función  $\mathcal{U}'$  se tiene que

$$\mathcal{U}'(\mathbf{u}(x, t)) \partial_t \mathbf{u}(x, t) + \mathcal{U}'(\mathbf{u}(x, t)) f'(\mathbf{u}(x, t)) \partial_x \mathbf{u}(x, t) = 0, \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty),$$

de donde se deduce que

$$\partial_t \mathcal{U}(\mathbf{u}(x, t)) + \partial_x \mathcal{F}(\mathbf{u}(x, t)) = 0, \quad \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty).$$

Por tanto, se cumple (2.16) en el sentido clásico y, en consecuencia, en el de las distribuciones.  $\square$

Como se deduce del Teorema 2.26, al contrario de lo que ocurre con las soluciones clásicas, no se puede asegurar que las soluciones débiles del problema de Cauchy (2.1) cumplan la condición de entropía. Esto motiva la siguiente definición.

**Definición 2.28** (Solución entrópicas). Se dice que una solución débil  $\mathbf{u} \in L^\infty(\mathbb{R} \times [0, +\infty), \Omega)$  del problema de Cauchy (2.1) es una solución entrópica si cumple la condición de entropía (2.16) para toda entropía  $(\mathcal{U}, \mathcal{F})$  para (2.1).

De aquí en adelante se considera, por simplicidad, el caso escalar y unidimensional. Por ello, se tratará el siguiente problema de Cauchy:

$$\begin{cases} \partial_t u(x, t) + \partial_x f(u(x, t)) = 0, & \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty), \\ u(x, 0) = u_0(x), & \forall x \in \mathbb{R}, \end{cases} \quad (2.18)$$

donde  $\Omega \subseteq \mathbb{R}$  es un conjunto convexo,  $u_0 \in L^\infty(\mathbb{R}, \Omega)$  y  $f \in \mathcal{C}^2(\Omega, \mathbb{R})$ .

El siguiente resultado permitirá relacionar las soluciones débiles con las entrópicas, en el caso en el que sean funciones de clase  $\mathcal{C}^1$  a trozos.

**Teorema 2.29** ([7], pág. 46). *Sea  $u: \mathbb{R} \times [0, +\infty) \rightarrow \Omega \subseteq \mathbb{R}$  una función de clase  $\mathcal{C}^1$  a trozos, solución débil del problema de Cauchy (2.18) con función de flujo  $f: \Omega \rightarrow \mathbb{R}$  estrictamente convexa. Si la función  $u$  cumple la condición de entropía (2.16) para una entropía  $(\mathcal{U}_0, \mathcal{F}_0)$ , entonces la función  $u$  cumple la condición de entropía (2.16) para toda entropía  $(\mathcal{U}, \mathcal{F})$ .*

*Demostración.* Sea  $(\mathcal{U}, \mathcal{F})$  una entropía para (2.18) y sea  $\Sigma \equiv t \mapsto (\sigma(t), t)$ , con  $t \in (a, b)$ , una curva de discontinuidad de  $u$ . Por simplicidad, se supone que  $\Sigma$  es la única curva de discontinuidad de  $u$ . Si  $u$  es continua a lo largo de  $\Sigma$ , entonces, en virtud del Teorema 2.26, es solución entrópica.

Por tanto, se supondrá que  $u^-(t) \neq u^+(t)$  para todo  $t \in (a, b)$ . Utilizando la condición de Rankine-Hugoniot, se puede reescribir (2.17) a lo largo de  $\Sigma$  como

$$\frac{f(u_\Sigma^+(t)) - f(u_\Sigma^-(t))}{u_\Sigma^+(t) - u_\Sigma^-(t)} (\mathcal{U}(u_\Sigma^+(t)) - \mathcal{U}(u_\Sigma^-(t))) - (\mathcal{F}(u_\Sigma^+(t)) - \mathcal{F}(u_\Sigma^-(t))) \geq 0, \quad \forall t \in (a, b). \quad (2.19)$$

Para cada  $t \in (a, b)$ , se define la función

$$\zeta: \Omega \rightarrow \mathbb{R}$$

$$\omega \mapsto \zeta(\omega) = \begin{cases} \frac{f(\omega) - f(u_\Sigma^-(t))}{\omega - u_\Sigma^-(t)} (\mathcal{U}(\omega) - \mathcal{U}(u_\Sigma^-(t))) - (\mathcal{F}(\omega) - \mathcal{F}(u_\Sigma^-(t))) \geq 0 & \text{si } \omega \neq u_\Sigma^-(t), \\ 0 & \text{si } \omega = u_\Sigma^-(t). \end{cases}$$

Cabe notar que comprobar (2.19) equivale a comprobar que  $\zeta(u_\Sigma^+(t)) \geq 0$  para todo  $t \in (a, b)$ . En primer lugar, hay que ver que  $\zeta$  es una función estrictamente decreciente de  $\omega$ . Se tiene que, para todo  $t \in (a, b)$  y todo  $\omega \neq u_\Sigma^-(t)$ ,

$$\zeta'(\omega) = \frac{f'(\omega)(\omega - u_\Sigma^-(t)) - f(\omega) + f(u_\Sigma^-(t))}{(\omega - u_\Sigma^-(t))^2} (\mathcal{U}(\omega) - \mathcal{U}(u_\Sigma^-(t))) + \frac{f(\omega) - f(u_\Sigma^-(t))}{\omega - u_\Sigma^-(t)} \mathcal{U}'(\omega) - \mathcal{F}'(\omega).$$

De la definición de entropía y agrupando términos, se sigue que se puede reescribir, para todo  $t \in (a, b)$  y todo  $\omega \neq u_\Sigma^-(t)$ ,

$$\zeta'(\omega) = \frac{[(u_\Sigma^-(t) - \omega)f'(\omega) - f(u_\Sigma^-(t)) + f(\omega)] [\mathcal{U}(u_\Sigma^-(t)) - \mathcal{U}(\omega) - \mathcal{U}'(\omega)(u_\Sigma^-(t) - \omega)]}{(\omega - u_\Sigma^-(t))^2}.$$

Como se había supuesto que  $f$  y  $\mathcal{U}$  eran funciones estrictamente convexas, se deduce que

$$\left. \begin{aligned} f(u_{\Sigma}^-(t)) - f(\omega) &> (u_{\Sigma}^-(t) - \omega)f'(\omega) \\ \mathcal{U}(u_{\Sigma}^-(t)) - \mathcal{U}(\omega) &> (u_{\Sigma}^-(t) - \omega)\mathcal{U}'(\omega) \end{aligned} \right\} \implies \zeta'(\omega) < 0, \quad \forall \omega \in \Omega \setminus \{u_{\Sigma}^-(t)\},$$

lo cual prueba que  $\zeta$  es estrictamente decreciente en  $\Omega \setminus \{u_{\Sigma}^-(t)\}$ . En consecuencia, se deduce que, para toda entropía  $(\mathcal{U}, \mathcal{F})$ ,

$$\zeta_{(\mathcal{U}, \mathcal{F})}(u_{\Sigma}^+(t)) > 0 \iff u_{\Sigma}^+(t) < u_{\Sigma}^-(t), \quad \forall t \in (a, b).$$

Por hipótesis, existe una entropía  $(\mathcal{U}_0, \mathcal{F}_0)$  para la cual  $u$  cumple la condición de entropía (2.16), de donde se deduce que

$$\zeta_{(\mathcal{U}_0, \mathcal{F}_0)}(u_{\Sigma}^+(t)) > 0 \iff u_{\Sigma}^+(t) < u_{\Sigma}^-(t) \iff \zeta_{(\mathcal{U}, \mathcal{F})}(u_{\Sigma}^+(t)) > 0,$$

lo que prueba que se tiene que  $\zeta_{(\mathcal{U}, \mathcal{F})}(u_{\Sigma}^+(t)) \geq 0$ , por lo que  $u$  cumple (2.16) para la entropía  $(\mathcal{U}, \mathcal{F})$ , como se quería probar.  $\square$

**Corolario 2.30.** *Si  $f$  es estrictamente convexa, entonces  $u: \mathbb{R} \times [0, +\infty) \rightarrow \Omega \subseteq \mathbb{R}$  es una solución entrópica del problema de Cauchy (2.18) si, y solo si, a lo largo de toda curva de discontinuidad  $\Sigma \equiv t \mapsto (\sigma(t), t)$ , con  $t \in (a, b)$ , se tiene que*

$$u_{\Sigma}^+(t) < u_{\Sigma}^-(t), \quad \forall t \in (a, b). \quad (2.20)$$

**Corolario 2.31.** *Si  $u: \mathbb{R} \times [0, +\infty) \rightarrow \Omega \subseteq \mathbb{R}$  es una función de clase  $\mathcal{C}^1$  a trozos, es continua y es una solución clásica del problema de Cauchy (2.18) allí donde es regular, entonces es una solución entrópica de dicho problema.*

Otro criterio más general para comprobar si una solución débil de clase  $\mathcal{C}^1$  a trozos es entrópica o no y que no demostraremos, es el siguiente:

**Teorema 2.32** ([15], pág. 166). *Sea  $u: \mathbb{R} \times [0, +\infty) \rightarrow \Omega \subseteq \mathbb{R}$  una función de clase  $\mathcal{C}^1$  a trozos que sea una solución débil del problema de Cauchy (2.18) con función de flujo  $f \in \mathcal{C}^2(\Omega, \mathbb{R})$  arbitraria. Entonces,  $u$  es solución entrópica de dicho problema si, y sólo si, para toda curva de discontinuidad  $\Sigma \equiv t \mapsto (\sigma(t), t)$ , con  $t \in (a, b)$ , se tiene que*

$$\frac{f(u^*(t)) - f(u_{\Sigma}^+(t))}{u^*(t) - u_{\Sigma}^+(t)} \leq \frac{f(u_{\Sigma}^+(t)) - f(u_{\Sigma}^-(t))}{u_{\Sigma}^+(t) - u_{\Sigma}^-(t)} \leq \frac{f(u^*(t)) - f(u_{\Sigma}^-(t))}{u^*(t) - u_{\Sigma}^-(t)},$$

para todo  $u^*(t)$  entre  $u_{\Sigma}^-(t)$  y  $u_{\Sigma}^+(t)$ , para todo  $t \in (a, b)$ .

Este resultado tiene una interpretación geométrica interesante. Suponiendo que  $u$  tiene una única curva de discontinuidad  $\Sigma$  y que  $u_{\Sigma}^+ < u_{\Sigma}^-$  (respectivamente  $u_{\Sigma}^+ > u_{\Sigma}^-$ ). Entonces  $u$  es entrópica, de acuerdo con el Teorema 2.32, si el grafo de  $f$  entre  $u_{\Sigma}^+$  y  $u_{\Sigma}^-$  está por debajo (respectivamente, por encima) del segmento que une  $(u_{\Sigma}^+, f(u_{\Sigma}^+))$  con  $(u_{\Sigma}^-, f(u_{\Sigma}^-))$ .

Se puede ver un diagrama ilustrativo de esta condición en la Figura 2.4.

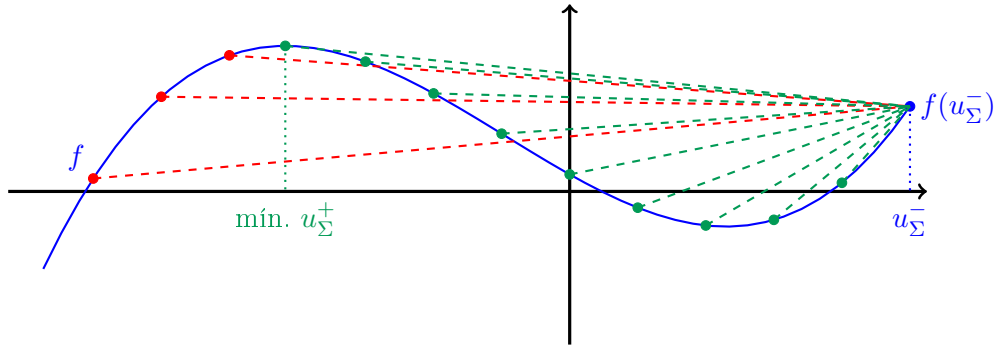


Figura 2.4: Interpretación geométrica del Teorema 2.32. Los puntos verdes se corresponden con valores de  $f(u_{\Sigma}^+)$  tales que  $u$  es solución entrópica. En rojo, los valores de  $f(u_{\Sigma}^+)$  para los cuales esto no se cumple.

Como consecuencia directa del Teorema 2.32, se tiene la siguiente condición necesaria para que una solución sea entrópica:

**Teorema 2.33** ([6], pág. 19). *Sea  $u: \mathbb{R} \times [0, +\infty) \rightarrow \Omega \subseteq \mathbb{R}$  una función de clase  $C^1$  a trozos que sea una solución débil del problema de Cauchy (2.18) con función de flujo  $f \in C^2(\Omega, \mathbb{R})$  arbitraria. Si  $u$  es solución entrópica de dicho problema, entonces, para toda curva de discontinuidad  $\Sigma \equiv t \mapsto (\sigma(t), t)$ , con  $t \in (a, b)$ , se tiene que*

$$f'(u_{\Sigma}^+(t)) \leq \sigma'(t) \leq f'(u_{\Sigma}^-(t)), \quad \forall t \in (a, b).$$

*Si  $f$  es estrictamente cóncava o estrictamente convexa, las desigualdades anteriores se convierten en estrictas y, además, se cumple que*

- *Si  $f$  es estrictamente convexa, comprobar la desigualdad anterior equivale a comprobar  $u_{\Sigma}^+(t) < u_{\Sigma}^-(t)$ , para todo  $t \in (a, b)$ .*
- *Si  $f$  es estrictamente cóncava, comprobar la desigualdad anterior equivale a comprobar  $u_{\Sigma}^-(t) < u_{\Sigma}^+(t)$ , para todo  $t \in (a, b)$ .*

Como última pieza, se incluye el siguiente resultado de existencia y unicidad de solución entrópica del problema de Cauchy escalar (2.18).

**Teorema 2.34** ([7], pág. 84). *El problema de Cauchy (2.18) admite una solución entrópica única  $u \in L^\infty(\mathbb{R} \times [0, +\infty), \Omega)$  tal que*

$$\|u(\cdot, t)\|_\infty \leq \|u_0\|_\infty, \quad \text{ctp } t \in [0, +\infty).$$

*Además, si  $u$  y  $v$  son las soluciones entrópicas del problema de Cauchy relativo a las condiciones iniciales  $u_0 \in L^\infty(\mathbb{R}, \Omega)$  y  $v_0 \in L^\infty(\mathbb{R}, \Omega)$ , respectivamente, entonces se cumple que*

$$[u_0(x) \leq v_0(x), \quad \text{ctp } x \in \mathbb{R}] \implies [u(x, t) \leq v(x, t), \quad \text{ctp } x \in \mathbb{R}, \quad \forall t \in [0, +\infty)].$$

La prueba, por su extensión y complejidad, no se incluye en este trabajo. Puede consultarse su desarrollo en [7].

## 2.4. El problema de Riemann

Esta sección está dedicada a la búsqueda de soluciones entrópicas para un problema unidimensional con un tipo de condiciones iniciales muy concretas: es el llamado *problema de Riemann*. La utilidad de dichas soluciones quedará clara en el Capítulo 4, donde se verá que son una pieza clave del método de Godunov.

Se considera una ley de conservación unidimensional con condición inicial dada por

$$\mathbf{u}_0(x) = \begin{cases} \mathbf{u}_L & \text{si } x \leq 0, \\ \mathbf{u}_R & \text{si } x > 0, \end{cases} \quad (2.21)$$

donde  $\mathbf{u}_L, \mathbf{u}_R \in \mathbb{R}^m$  son constantes conocidas como estados izquierdo y derecho, respectivamente, del problema de Riemann (1.3)-(2.21). Este problema tiene gran interés a la hora de estudiar el comportamiento tanto analítico como numérico de las leyes de conservación. En primer lugar, porque modelan algunas situaciones físicas de interés. Por ejemplo, para las ecuaciones de aguas someras, modelan el problema de rotura de una presa o, para las ecuaciones de Euler, el problema de choque en un tubo.

En segundo lugar, porque para esta estructura de condiciones iniciales, aparecen numerosos fenómenos como choques o rarefacciones, de los que se hablará un poco más adelante en este capítulo. Finalmente, otra razón para estudiar este tipo de problemas es que se pueden calcular las soluciones exactas para estas condiciones iniciales para muchas leyes de conservación. En esta sección se detallará el procedimiento que permite obtener las soluciones entrópicas en ciertos casos escalares y sencillos.

Todas estas características hacen que los problemas de Riemann sean muy útiles a la hora de comprobar si los métodos numéricos que se diseñan para las leyes de conservación funcionan correctamente.

En todos los problemas que se consideran en este trabajo, la solución de los problemas de Riemann correspondientes será una función *autosimilar*, es decir, son soluciones de la forma

$$\mathbf{u}(x, t) = \mathbf{v} \left( \frac{x}{t} \right), \quad \forall x \in \mathbb{R}, \quad \forall t \in (0, +\infty),$$

donde  $\mathbf{v}: \mathbb{R} \rightarrow \Omega \subseteq \mathbb{R}^m$  es una función regular a trozos. Por esta razón, se denotará a la solución del problema de Cauchy (1.3)-(2.21) por  $\mathbf{w}_R(\frac{x}{t}; \mathbf{u}_L, \mathbf{u}_R)$ . Pueden verse más detalles que conducen a que se imponga esta hipótesis en [8].

En lo que sigue, por simplicidad, se considera únicamente el caso escalar, es decir, el problema de Cauchy (2.5)–(2.21). Puesto que se cumplen las hipótesis del Teorema 2.34, tiene sentido plantearse buscar las soluciones entrópicas de dicho problema. El siguiente resultado teórico tendrá gran utilidad a la hora de encontrar esa solución, además de garantizar en el caso escalar la hipótesis de autosimilaridad que se está suponiendo en este trabajo.

**Proposición 2.35.** *Existe una función  $v: \mathbb{R} \rightarrow \Omega \subseteq \mathbb{R}$  tal que*

$$u(x, t) = v\left(\frac{x}{t}\right), \quad \forall x \in \mathbb{R}, \quad \forall t \in (0, +\infty),$$

donde  $u$  es la solución del problema de Riemann (2.5)–(2.21). Por esta razón, se dice que  $u$  es autosimilar. Además, en los puntos  $\xi \in \mathbb{R}$  donde  $v$  es regular, se tiene que, o bien  $v'(\xi) = 0$  o bien  $f'(v(\xi)) = \xi$ .

*Demostración.* En primer lugar, dado un  $(x, t) \in \mathbb{R} \times (0, +\infty)$  y un  $k > 0$ , entonces está claro que  $u(x, t) = u(kx, kt)$ , lo cual implica la existencia de una función  $v$  que cumple la propiedad del enunciado. Por otra parte, se toma un  $(x, t) \in \mathbb{R} \times (0, +\infty)$  donde  $u$  sea regular, al menos continuamente diferenciable. Entonces, se tiene que

$$\partial_t u(x, t) + \partial_x f(u(x, t)) = -\frac{x}{t^2} v'\left(\frac{x}{t}\right) + \frac{1}{t} f'\left(v\left(\frac{x}{t}\right)\right) v'\left(\frac{x}{t}\right) = 0.$$

En consecuencia,

$$\frac{1}{t} f'\left(v\left(\frac{x}{t}\right)\right) v'\left(\frac{x}{t}\right) = \frac{x}{t^2} v'\left(\frac{x}{t}\right) \implies \frac{1}{t} v'\left(\frac{x}{t}\right) \left[ f'\left(v\left(\frac{x}{t}\right)\right) - \frac{x}{t} \right] = 0,$$

lo cual implica que se da al menos una de las dos ecuaciones del enunciado.  $\square$

En primer lugar, si  $u_L = u_R$ , es evidente que la solución es dicha constante, para todo  $t \in [0, +\infty)$ .

Considerando el resto de casos, se comienza por emplear el método de las características. Se considera un punto  $(\bar{x}, \bar{t})$  arbitrario del plano  $x/t$ . Denotando por  $\xi$  al punto de corte de su recta característica con el eje  $t = 0$ , entonces, recordando (2.8), se tiene que la curva característica que pasa por  $(\bar{x}, \bar{t})$  viene dada por es

$$X(t) = \xi + f'(u_0(\xi))t = \begin{cases} \xi + f'(u_L)t & \text{si } \xi \leq 0, \\ \xi + f'(u_R)t & \text{si } \xi > 0. \end{cases} \quad (2.22)$$

Emplear este método tiene sentido cuando las rectas características no se corten. Si  $f$  es estrictamente convexa, esto ocurre si  $u_L < u_R$  y si  $f$  es estrictamente cóncava, si  $u_R < u_L$ . Por tanto, en estos casos, se tiene que

$$u(\bar{x}, \bar{t}) = u(X(\bar{t}), \bar{t}) = u(X(0), 0) = u(\xi, 0) = u_0(\xi),$$

lo que significa que, despejando  $\xi$  como función de  $x$  y  $t$  de las rectas características, se puede dar la solución en algunas regiones como:

$$u(x, t) = u_L \quad \text{si } x \leq f'(u_L)t, \quad u(x, t) = u_R \quad \text{si } x > f'(u_R)t. \quad (2.23)$$

Sin embargo, no es posible así de obtener la solución allí donde  $x \in (f'(u_L)t, f'(u_R)t]$ . Se puede solucionar esto en el caso en el que  $f'$  es invertible en  $[u_L, u_R]$ . Esto lleva a concluir dando una solución del problema (2.5)–(2.21) como:

$$u(x, t) = \begin{cases} u_L & \text{si } x \leq f'(u_L)t, \\ (f')^{-1} \left( \frac{x}{t} \right) & \text{si } f'(u_L)t < x < f'(u_R)t, \\ u_R & \text{si } x \geq f'(u_R)t. \end{cases} \quad (2.24)$$

*Observación 2.36.* La función  $f'$  será invertible si  $f''$  no se anula en ningún punto, es decir, si  $f$  es estrictamente convexa o estrictamente cóncava.

Este tipo de solución continua se denomina *onda de rarefacción*. Por la forma de construir esta solución, es claro que, utilizando el Corolario 2.16, se trata de una solución débil de (2.5)–(2.21), sin embargo, se debe comprobar que sea la solución entrópica. Para ello, se utilizará el Corolario 2.31 y bastará con ver que la solución (2.24) es continua para todo  $t \in (0, +\infty)$ . Pero esto es obvio, puesto que  $u$  es trivialmente continua en

$$\mathbb{R} \times (0, +\infty) \setminus \{(x, t) \in \mathbb{R} \times (0, +\infty) : x = f'(u_L)t \vee x = f'(u_R)t\}$$

y además

$$\begin{aligned} \lim_{x \rightarrow (f'(u_L)t)^+} u(x, t) &= \lim_{x \rightarrow f'(u_L)t} (f')^{-1} \left( \frac{x}{t} \right) = (f')^{-1} (f'(u_L)) = u_L = u(x, f'(u_L)t), \\ \lim_{x \rightarrow (f'(u_R)t)^-} u(x, t) &= \lim_{x \rightarrow f'(u_R)t} (f')^{-1} \left( \frac{x}{t} \right) = (f')^{-1} (f'(u_R)) = u_R = u(x, f'(u_R)t). \end{aligned}$$

Se concluye así que (2.24) es la única solución entrópica de (2.5)–(2.21) cuando  $u_L < u_R$ , si  $f$  es convexa, y cuando  $u_R < u_L$  si  $f$  es cóncava. En el caso de la ecuación de Burgers, (2.24) se escribe como

$$u(x, t) = \begin{cases} u_L & \text{si } x \leq u_L t, \\ \frac{x}{t} & \text{si } u_L t < x < u_R t, \\ u_R & \text{si } x \geq u_R t. \end{cases} \quad (2.25)$$

Se puede ver una representación gráfica de estas soluciones en la Figura 2.5.

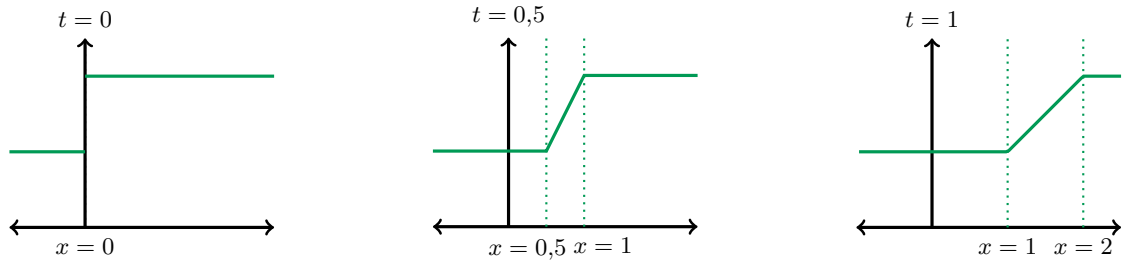


Figura 2.5: Solución exacta del problema de Riemann con estados iniciales  $u_L = 1$  y  $u_R = 2$  para la ecuación de Burgers.

Cuando no es posible utilizar el método de las características, se tendrá que abordar el problema de manera alternativa. En  $t = 0$  existe una única curva de discontinuidad, en  $x = 0$ . De la condición de Rankine-Hugoniot se puede despejar la velocidad de la única curva de discontinuidad mediante la ecuación

$$s(t) = \frac{[[f(u)]]}{[[u]]} = \frac{f(u_R) - f(u_L)}{u_R - u_L}, \quad \forall t \in (0, +\infty),$$

de donde se deduce que dicha curva viene dada por  $\Sigma \equiv t \mapsto (st, t)$ , para todo  $t \in (0, +\infty)$ . Por lo tanto, puesto que la función

$$u(x, t) = \begin{cases} u_L & \text{si } x \leq st, \\ u_R & \text{si } x > st, \end{cases} \quad (2.26)$$

es solución de (2.5)–(2.21) allí donde es regular y, además cumple la condición de Rankine-Hugoniot, entonces es solución débil de la ley de conservación. Este tipo de soluciones se conoce como *onda de choque*. Además, es inmediato ver que esta solución es entrópica a partir del Teorema (2.30). En consecuencia, para este caso, (2.26) es la única solución entrópica de (2.5)–(2.21) en el caso de estar tratando con una onda de choque. Si se está considerando la ecuación de Burgers, (2.26) toma la expresión

$$u(x, t) = \begin{cases} u_L & \text{si } x \leq \frac{1}{2}(u_L + u_R)t, \\ u_R & \text{si } x > \frac{1}{2}(u_L + u_R)t. \end{cases} \quad (2.27)$$

De nuevo, se puede ver una ilustración de estas soluciones en la Figura 2.6.

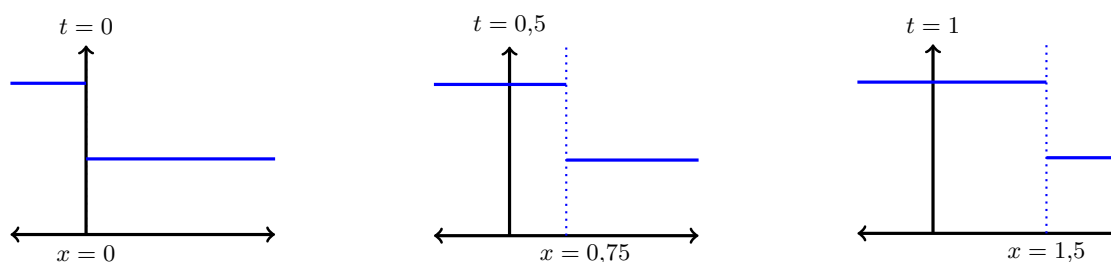


Figura 2.6: Solución exacta del problema de Riemann con estados iniciales  $u_L = 2$  y  $u_R = 1$  para la ecuación de Burgers.

De las explicaciones anteriores se deduce el siguiente teorema:

**Teorema 2.37** ([12], pág. 32). *Si la función de flujo  $f$  es estrictamente convexa (respectivamente, estrictamente cóncava), entonces el problema de Riemann escalar (2.5)–(2.21) admite una única solución entrópica y autosimilar, dada por*

- Una onda de rarefacción de la forma (2.24) si  $u_L < u_R$  (respect.  $u_R < u_L$ ).
- Una onda de choque de la forma (2.26) si  $u_R < u_L$  (respect.  $u_L < u_R$ ).

Un estudio pormenorizado del caso general es muy complicado y, en muchas ocasiones, imposible. Por tanto, siguiendo [12], se incluye únicamente otro estudio de las soluciones del problema de Riemann en el caso en el que el flujo  $f$  es cóncavo–convexo, es decir, que cumple las siguientes propiedades:

$$\omega f''(\omega) > 0, \quad \forall \omega \neq 0, \quad f'''(0) \neq 0 \quad \text{y} \quad \lim_{|\omega| \rightarrow \infty} f'(\omega) = +\infty. \quad (\text{HCC})$$

Dado un  $u_L \geq 0$  fijo, se estudiará, aplicando el Teorema 2.32, cómo son las soluciones del problema de Riemann (2.5)–(2.21) para todo  $u_R \neq u_L$ . En primer lugar, se debe hacer notar que existe un punto de gran relevancia: el punto  $u^*$  más alejado de  $u_L$  tal que el segmento que une  $(u^*, f(u^*))$  con  $(u_L, f(u_L))$  permanece por encima del gráfico de  $f$  entre  $u^*$  y  $u_L$ . Puede verse que dicho punto es la única solución de la ecuación

$$\frac{f(u_L) - f(u^*)}{u_L - u^*} = f'(u^*). \quad (2.28)$$

Necesariamente,  $u^* < u_L$  y también, de la definición de concavidad y convexidad,  $u^* < 0$ . Ahora bien, atendiendo a la interpretación geométrica del Teorema 2.32, se tienen tres opciones:

- Si  $u_R > u_L$ , entonces el segmento que une  $(u_R, f(u_R))$  con  $(u_L, f(u_L))$  permanece por encima de la gráfica de  $f$  entre dicho puntos, y entonces, por el mismo razonamiento que se utilizó en Teorema 2.37, se deduce que la solución entrópica es una onda de rarefacción de la forma (2.24). Nótese que en  $[u_L, u_R]$ , la función  $f'$  es invertible.

- Si  $u_L > u_R$  y  $u^* < u_R$ , entonces el segmento que une  $(u_R, f(u_R))$  con  $(u_L, f(u_L))$  permanece por encima de la gráfica de  $f$  entre dicho puntos, y entonces, por el mismo razonamiento antes, se deduce que la solución entrópica es una onda de rarefacción de la forma (2.26).
- Por último, si  $u_R < u^*$ , entonces el segmento que une  $(u_R, f(u_R))$  con  $(u_L, f(u_L))$  corta a la gráfica de  $f$ . Sin embargo, se puede observar que se pueden concatenar dos segmentos: uno entre  $(u_R, f(u_R))$  y  $(u^*, f(u^*))$  que permanece por debajo de la gráfica y otro entre  $(u^*, f(u^*))$  y  $(u_L, f(u_L))$ , por encima de dicha gráfica. De este modo, razonando de modo análogo al Teorema 2.37, se puede deducir que la solución entrópica consiste en dos ondas: un choque entre  $u_L$  y  $u^*$  y una rarefacción entre  $u^*$  y  $u_R$ .

Finalmente, se puede comprobar fácilmente que la solución así construida es entrópica: es solución débil por cumplir la ecuación en derivadas parciales en todo punto donde es regular y es además solución entrópica porque se verifican globalmente las hipótesis del Teorema 2.32.

**Teorema 2.38** ([12], pág. 33). *Si la función de flujo  $f$  es cóncava-convexa (i.e., cumple las hipótesis (HCC)) y  $u_L > 0$ , si se denota por el estado  $u^*$  a la única solución de la ecuación (2.28), entonces el problema de Riemann escalar (2.5)–(2.21) admite una única solución entrópica y autosimilar, dada por*

- Una onda de rarefacción entre  $u_L$  y  $u_R$  si  $u_R \geq u_L$ .
- Una onda de choque si  $u_R \in [u^*, u_L)$ .
- Una onda de choque entre  $u_L$  y  $u^*$  y una onda de rarefacción entre  $u^*$  y  $u_R$  si  $u_R < u^*$ .

En la Figura 2.7 se puede ver una interpretación geométrica de estas casuísticas.

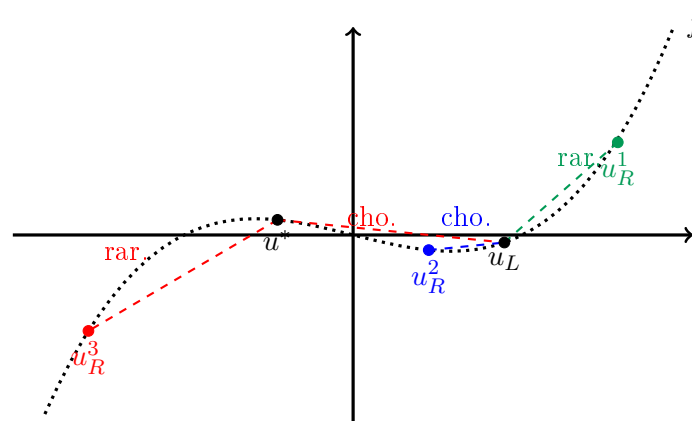


Figura 2.7: Interpretación geométrica de la deducción de la solución entrópica del problema de Riemann para una función cóncava-convexa.

## Capítulo 3

# Descripción general de los esquemas de volúmenes finitos

En los capítulos anteriores, se ha visto que, en general, es muy complicado hallar la solución entrópica de una ley de conservación. Por esta razón surge la necesidad de diseñar *métodos numéricos* para aproximarlas. Para ello, es fundamental *discretizar* el problema. Puesto que este trabajo está centrado en leyes de conservación unidimensionales, en lo que sigue se considerará el siguiente problema:

$$\begin{cases} \partial_t \mathbf{u}(x, t) + \partial_x \mathbf{f}(\mathbf{u}(x, t)) = 0, & \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty), \\ \mathbf{u}(x, 0) = \mathbf{u}_0(x), & \forall x \in \mathbb{R}, \end{cases} \quad (3.1)$$

donde  $\Omega \subseteq \mathbb{R}^m$  es un abierto,  $\mathbf{f} \in \mathcal{C}^2(\Omega, \mathbb{R}^m)$  y  $\mathbf{u}_0 \in L^\infty(\mathbb{R}, \Omega)$ .

### 3.1. Descripción general de los métodos de volúmenes finitos

En primer lugar, se va a considerar una malla  $\mathcal{T} = (\mathcal{C}_i)_{i \in \mathbb{Z}}$  de  $\mathbb{R}$ , que debe cumplir las siguientes propiedades:

- (I)  $\mathcal{C}_i = (a_i, b_i)$  para todo  $i \in \mathbb{Z}$ , con  $a_i, b_i \in \mathbb{R}$  y  $a_i < b_i$ .
- (II)  $\mathcal{C}_i \cap \mathcal{C}_j \neq \emptyset$  para todo  $i, j \in \mathbb{Z}$  tales que  $i \neq j$ .
- (III)  $\bigcup_{i \in \mathbb{Z}} \overline{\mathcal{C}_i} = \mathbb{R}$ .
- (IV)  $\mathcal{C}_i$  y  $\mathcal{C}_{i+1}$  son consecutivos, esto es, para cada  $i \in \mathbb{Z}$ ,  $b_i = a_{i+1}$ .

Se llama *volumen finito* a cada una de las celdas  $\mathcal{C}$  de  $\mathcal{T}$  y se dice que la norma de  $\mathcal{T}$  viene dada por el supremo de las *medidas* (longitudes  $\Delta x_i$  en este caso unidimensional) de sus volúmenes finitos, es decir, por

$$\|\mathcal{T}\| := \Delta x = \sup_{i \in \mathbb{Z}} \Delta x_i = \sup_{i \in \mathbb{Z}} \{b_i - a_i\}.$$

Se denota por  $x_i$  al punto medio de cada volumen finito  $\mathcal{C}_i$ , es decir,  $x_i = \frac{1}{2}(a_i + b_i)$ , para todo  $i \in \mathbb{Z}$ . Por otra parte, dado un paso de tiempo  $\Delta t > 0$ , se considerarán los instantes  $t_n = (n-1)\Delta t$  para todo  $n \in \mathbb{N}$ .

El método de los volúmenes finitos propone aproximar la solución de (3.1) en el nodo  $x_i$  y en el tiempo  $t_n$  como una aproximación del promedio de la solución en el volumen finito  $\mathcal{C}_i$ , es decir, propone tomar aproximar  $\mathbf{u}(x_i, t_n)$  por

$$\mathbf{u}_i^n \approx \frac{1}{\Delta x_i} \int_{\mathcal{C}_i} \mathbf{u}(x, t_n) dx, \quad \forall i \in \mathbb{Z}, \quad \forall n \in \mathbb{N}. \quad (3.2)$$

Como se están tomando aproximaciones en subintervalos de  $\mathbb{R}$ , tiene sentido aproximar también

$$[\mathbf{u}(x, t_n) \approx \mathbf{u}^n(x_i) = \mathbf{u}_i^n, \quad \text{si } x \in \mathcal{C}_i], \quad \forall x \in \mathbb{R}, \quad \forall n \in \mathbb{N}.$$

Así, para cada paso de tiempo, se denotará por  $\mathbf{u}^n$  a una aproximación de  $\mathbf{u}(\cdot, t_n)$  usando un método de volúmenes finitos. Puesto que el promedio de la solución en cada volumen finito no puede ser siempre calculado de forma exacta, dando aproximaciones de estos promedios surgen los distintos esquemas de volúmenes finitos para determinar los  $\mathbf{u}_i^n$ . Esta idea intuitiva se puede ver en la Figura 3.1.

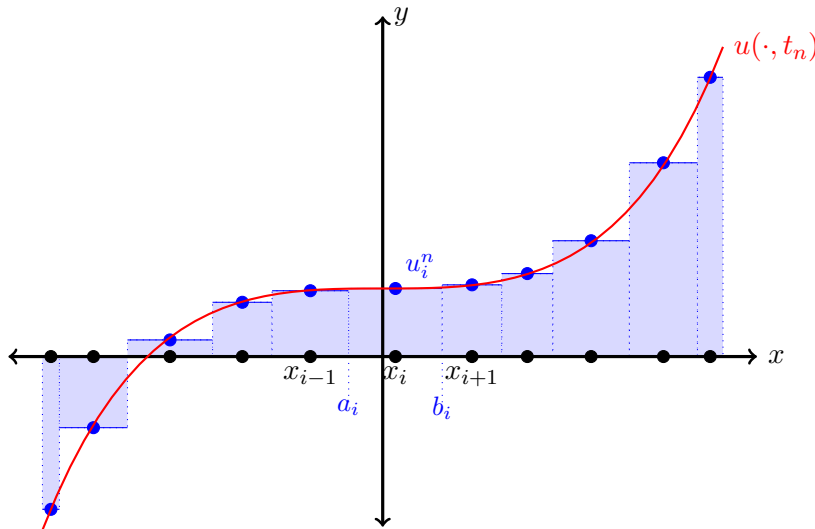


Figura 3.1: Ilustración del concepto de volumen finito y de las aproximaciones  $u_i^n$  en el caso escalar. En rojo, solución para un tiempo  $t_n$ . En azul, promedio  $u_i^n$  de la solución en cada volumen finito. La solución numérica será una aproximación de estos promedios.

La ventaja de trabajar con estos promedios en vez de con los valores puntuales de la solución es que así se está teniendo un enfoque más próximo a la formulación débil (formulación integral) de la ley de conservación, que ya se ha visto que tiene *mejores propiedades*, en el sentido de que se puede hablar existencia de soluciones débiles y unicidad de solución entrópica.

Este trabajo se centrará especialmente en los esquemas numéricos de un paso de tiempo que para la actualización del promedio de cada volumen finito involucren únicamente al volumen finito anterior y siguiente a él, es decir, esquemas de la forma

$$\mathbf{u}_i^{n+1} = \mathcal{H}(\mathbf{u}_{i-1}^n, \mathbf{u}_i^n, \mathbf{u}_{i+1}^n), \quad \forall i \in \mathbb{Z}, \quad \forall n \in \mathbb{N}, \quad (3.3)$$

donde  $\mathcal{H}$  es una función regular y donde los  $u_i^1$  se obtienen a partir de la condición inicial. Esta estructura de los esquemas tiene sentido por la propia naturaleza del problema, ya que la velocidad de propagación de la información es finita.

Además, aunque los contenidos que se explicarán en este trabajo pueden, en general, desarrollarse para mallas  $\mathcal{T}$  arbitrarias, por simplicidad y brevedad se trabajará únicamente en mallas *uniformes*, es decir, mallas tales que  $\|\Delta x_i\| = \Delta x$ , para todo  $i \in \mathbb{Z}$ .

## 3.2. Métodos conservativos

En general, los métodos numéricos que se buscan deben replicar el comportamiento de las leyes de conservación. En particular, deben ser *conservativos*, es decir, que si  $\{I_1, I_1 + 1, \dots, I_2\}$  son los subíndices correspondientes a volúmenes finitos contiguos, entonces  $\sum_{i=I_1}^{I_2} \Delta x \mathbf{u}_i^n$  debe variar únicamente por efecto del flujo en la frontera, esto es, en  $a_{I_1}$  y  $b_{I_2}$ . Se va a ver cómo formular matemáticamente este concepto utilizando leyes de conservación para deducir cómo son los esquemas *de interés* de entre todos los que se pueden definir siguiendo (3.3).

Integrando la ley de conservación (3.1) en un volumen finito  $\mathcal{C}_i$  se obtiene que

$$\partial_t \int_{\mathcal{C}_i} \mathbf{u}(x, t) dx = \mathbf{f}(\mathbf{u}(a_i, t)) - \mathbf{f}(\mathbf{u}(b_i, t)), \quad \forall t \in [0, +\infty).$$

Integrando la expresión anterior en el intervalo de tiempo  $[t_n, t_{n+1}]$ , para un cierto  $n \in \mathbb{N}$ ,

$$\int_{\mathcal{C}_i} \mathbf{u}(x, t_{n+1}) dx - \int_{\mathcal{C}_i} \mathbf{u}(x, t_n) dx = \int_{t_n}^{t_{n+1}} \mathbf{f}(\mathbf{u}(a_i, t)) dt - \int_{t_n}^{t_{n+1}} \mathbf{f}(\mathbf{u}(b_i, t)) dt.$$

Reagrupando términos y dividiendo por  $\Delta x$ , se tiene que

$$\frac{1}{\Delta x} \int_{\mathcal{C}_i} \mathbf{u}(x, t_{n+1}) dx = \frac{1}{\Delta x} \int_{\mathcal{C}_i} \mathbf{u}(x, t_n) dx - \frac{1}{\Delta x} \left\{ \int_{t_n}^{t_{n+1}} \mathbf{f}(\mathbf{u}(b_i, t)) dt - \int_{t_n}^{t_{n+1}} \mathbf{f}(\mathbf{u}(a_i, t)) dt \right\}.$$

Formalmente, esto sugiere la siguiente definición:

**Definición 3.1** (Esquema conservativo). Un esquema numérico (3.3) se escribe en forma conservativa si admite una escritura del tipo

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^n - \frac{\Delta t}{\Delta x} [\mathbf{g}(\mathbf{u}_i^n, \mathbf{u}_{i+1}^n) - \mathbf{g}(\mathbf{u}_{i-1}^n, \mathbf{u}_i^n)], \quad \forall i \in \mathbb{Z}, \quad \forall n \in \mathbb{N}, \quad (3.4)$$

siendo  $\mathbf{g}$  una función regular, llamada *flujo numérico* del método.

*Observación 3.2.* En ocasiones, se utilizará la notación abreviada

$$\mathbf{g}_{i-\frac{1}{2}}^n = \mathbf{g}(u_{i-1}^n, u_i^n), \quad \mathbf{g}_{i+\frac{1}{2}}^n = \mathbf{g}(u_i^n, u_{i+1}^n), \quad \forall i \in \mathbb{Z}, \quad \forall n \in \mathbb{N},$$

y son una aproximación del flujo físico de la ley de conservación en  $\mathbf{u}(a_i, t_n)$  y  $\mathbf{u}(b_i, t_n)$ . Variando la expresión de este flujo numérico se obtienen los distintos esquemas de volúmenes finitos.

Es importante hacer notar que los esquemas de este tipo son *conservativos* en el sentido anterior: la magnitud conservativa sólo varía como efecto del flujo en la frontera. De la expresión (3.4) es fácil ver que, dados  $I_1, I_2 \in \mathbb{Z}$  con  $I_1 < I_2$ , se cumple que

$$\Delta x \mathbf{u}_i^{n+1} = \Delta x \mathbf{u}_i^n - \Delta t [\mathbf{g}(\mathbf{u}_i^n, \mathbf{u}_{i+1}^n) - \mathbf{g}(\mathbf{u}_{i-1}^n, \mathbf{u}_i^n)], \quad I_1 \leq i \leq I_2,$$

de donde finalmente se llega a que

$$\begin{aligned} \Delta x \sum_{i=I_1}^{I_2} \mathbf{u}_i^{n+1} &= \Delta x \sum_{i=I_1}^{I_2} \mathbf{u}_i^n - \Delta t \sum_{i=I_1}^{I_2} [\mathbf{g}(\mathbf{u}_i^n, \mathbf{u}_{i+1}^n) - \mathbf{g}(\mathbf{u}_{i-1}^n, \mathbf{u}_i^n)] \\ &= \Delta x \sum_{i=I_1}^{I_2} \mathbf{u}_i^n - \Delta t (\mathbf{g}_{I_2+\frac{1}{2}}^n - \mathbf{g}_{I_1-\frac{1}{2}}^n). \end{aligned}$$

### 3.3. Consistencia

Antes de poder estudiar si el esquema numérico (3.3) aproxima correctamente una solución débil del problema (3.1), se debe estudiar su consistencia, concepto relacionado con los errores de truncamiento. Dado un esquema conservativo, para cada  $i \in \mathbb{Z}$  y  $n \in \mathbb{N}$ , se define

$$\boldsymbol{\tau}_i^n = \frac{1}{\Delta t} (\bar{\mathbf{u}}_i^{n+1} - \bar{\mathbf{u}}_i^n) + \frac{1}{\Delta x} (\bar{\mathbf{g}}_{i+1}^n - \bar{\mathbf{g}}_{i-1}^n), \quad (3.5)$$

donde  $\bar{\mathbf{u}}: \mathbb{R} \times [0, +\infty) \rightarrow \Omega \subseteq \mathbb{R}^m$  es una solución regular de (3.1) y se está denotando

$$\bar{\mathbf{u}}_i^n = \bar{\mathbf{u}}(x_i, t_n), \quad \text{y} \quad \bar{\mathbf{g}}_{i+1}^n = \mathbf{g}(\bar{\mathbf{u}}_i^n, \bar{\mathbf{u}}_{i+1}^n).$$

**Definición 3.3** (Esquema consistente). Un esquema numérico (3.3) escrito en forma conservativa es consistente si

$$\lim_{\Delta x \rightarrow 0, \Delta t \rightarrow 0} \sup_{i \in \mathbb{Z}, n \in \mathbb{N}} \|\boldsymbol{\tau}_i^n\|_2 = 0$$

para toda solución *regular* de (3.1). Se dice que (3.3) es de orden  $k_1$  en espacio y  $k_2$  en tiempo (o simplemente de orden  $k$  si  $k_1 = k_2$ ) si

$$\sup_{i \in \mathbb{Z}, n \in \mathbb{N}} \|\tau_i^n\|_2 = \mathcal{O}(\Delta x^{k_1}) + \mathcal{O}(\Delta t^{k_2}).$$

**Proposición 3.4** ([1], pág. 137). *Un esquema numérico (3.3) conservativo es consistente si, y solo si, existe un  $\omega_0 \in \mathbb{R}^m$  tal que*

$$\mathbf{g}(\omega, \omega) = \mathbf{f}(\omega) + \omega_0, \quad \forall \omega \in \Omega.$$

*El método tiene además orden de consistencia al menos uno.*

*Demostración.* Si  $\bar{\mathbf{u}}: \mathbb{R} \times [0, +\infty) \rightarrow \Omega \subseteq \mathbb{R}^m$  es una solución regular de (3.1), entonces de la fórmula de Taylor se deduce que, para todo  $i \in \mathbb{Z}$  y  $n \in \mathbb{N}$ ,

$$\begin{cases} \bar{\mathbf{u}}_i^{n+1} &= \bar{\mathbf{u}}_i^n + \Delta t \partial_t \bar{\mathbf{u}}(x_i, t_n) + \mathcal{O}(\Delta t^2), \\ \bar{\mathbf{u}}_{i-1}^n &= \bar{\mathbf{u}}_i^n - \Delta x \partial_x \bar{\mathbf{u}}(x_i, t_n) + \mathcal{O}(\Delta x^2), \\ \bar{\mathbf{u}}_{i+1}^n &= \bar{\mathbf{u}}_i^n + \Delta x \partial_x \bar{\mathbf{u}}(x_i, t_n) + \mathcal{O}(\Delta x^2). \end{cases} \quad (3.6)$$

De estos desarrollos y de aplicar la fórmula de Taylor de nuevo se sigue que

$$\begin{cases} \mathbf{g}(\bar{\mathbf{u}}_i^n, \bar{\mathbf{u}}_{i+1}^n) &= \mathbf{g}(\bar{\mathbf{u}}_i^n, \bar{\mathbf{u}}_i^n) + \Delta x \partial_{\omega_2} \mathbf{g}(\bar{\mathbf{u}}_i^n, \bar{\mathbf{u}}_i^n) \partial_x \bar{\mathbf{u}}(x_i, t_n) + \mathcal{O}(\Delta x^2), \\ \mathbf{g}(\bar{\mathbf{u}}_{i-1}^n, \bar{\mathbf{u}}_i^n) &= \mathbf{g}(\bar{\mathbf{u}}_i^n, \bar{\mathbf{u}}_i^n) - \Delta x \partial_{\omega_1} \mathbf{g}(\bar{\mathbf{u}}_i^n, \bar{\mathbf{u}}_i^n) \partial_x \bar{\mathbf{u}}(x_i, t_n) + \mathcal{O}(\Delta x^2). \end{cases} \quad (3.7)$$

Sustituyendo (3.6) y (3.7) en (3.5), se obtiene

$$\tau_i^n = \partial_t \bar{\mathbf{u}}(x_i, t_n) + [\partial_{\omega_1} \mathbf{g}(\bar{\mathbf{u}}_i^n, \bar{\mathbf{u}}_i^n) + \partial_{\omega_2} \mathbf{g}(\bar{\mathbf{u}}_i^n, \bar{\mathbf{u}}_i^n)] \partial_x \bar{\mathbf{u}}(x_i, t_n) + \mathcal{O}(\Delta x) + \mathcal{O}(\Delta t).$$

En consecuencia, se tiene que

$$\tau_i^n = \mathcal{O}(\Delta x) + \mathcal{O}(\Delta t) \iff \partial_t \bar{\mathbf{u}}(x_i, t_n) + [\partial_{\omega_1} \mathbf{g}(\bar{\mathbf{u}}_i^n, \bar{\mathbf{u}}_i^n) + \partial_{\omega_2} \mathbf{g}(\bar{\mathbf{u}}_i^n, \bar{\mathbf{u}}_i^n)] \partial_x \bar{\mathbf{u}}(x_i, t_n) = \mathbf{0}$$

y esta última igualdad sólo se da si se cumple la formulación no conservativa (1.5) de la ley de conservación, es decir, si

$$\mathbb{A}(\bar{\mathbf{u}}_i^n) = \partial_{\omega_1} \mathbf{g}(\bar{\mathbf{u}}_i^n, \bar{\mathbf{u}}_i^n) + \partial_{\omega_2} \mathbf{g}(\bar{\mathbf{u}}_i^n, \bar{\mathbf{u}}_i^n)$$

para todo  $i \in \mathbb{Z}$  y  $n \in \mathbb{N}$ , lo cual implica la igualdad del enunciado.  $\square$

El concepto de esquema numérico consistente permite introducir el siguiente resultado, de importancia fundamental en este trabajo:

**Teorema 3.5** (de Lax-Wendroff. [13], pág. 240). *Sea  $(\mathcal{T}_k)_{k \in \mathbb{N}}$  una sucesión de mallas uniformes de  $\mathbb{R}$  y una sucesión de pasos de tiempo  $\Delta t_k$  tales que  $\Delta x_k \rightarrow 0$  y  $\Delta t_k \rightarrow 0$  cuando  $k \rightarrow +\infty$ . Para cada  $k \in \mathbb{N}$ , se denotará por  $\mathbf{u}_k$  a la función constante a trozos que es solución aproximada del problema (3.1) con un esquema numérico (3.3) conservativo y consistente, usando la partición  $\mathcal{T}_k$  y el paso de tiempo  $\Delta t_k$ . Si  $(\mathbf{u}_k)_{k \in \mathbb{N}}$  es tal que existe un  $C > 0$  tal que  $\|\mathbf{u}_k\|_\infty < C$ , para todo  $k \in \mathbb{N}$ ; y además  $(\mathbf{u}_k)_{k \in \mathbb{N}}$  converge en  $L^1_{loc}(\mathbb{R} \times [0, +\infty), \Omega)$  a una función  $\mathbf{u}: \mathbb{R} \rightarrow \Omega \subseteq \mathbb{R}^m$ , entonces  $\mathbf{u}$  es una solución débil de la ley de conservación.*

*Demostración.* Sea  $\varphi \in \mathcal{C}_0^1(\mathbb{R} \times [0, +\infty), \mathbb{R}^m)$  una función de ensayo arbitraria. Para cada  $i \in \mathbb{Z}$ ,  $n \in \mathbb{N}$  y  $k \in \mathbb{N}$ , se denotará:

- Por  $(x_{k,i}, t_{k,n})$  a cada uno de los nodos inducidos por la malla  $\mathcal{T}_k$  y el paso temporal  $\Delta t_k$ .
- Por  $\mathbf{u}_{k,i}^n$  a cada aproximación numérica en el nodo  $(x_{k,i}, t_{k,n})$ .
- Por  $\Phi_{k,i}^n$  a la función  $\varphi$  evaluada en el nodo  $(x_{k,i}, t_{k,n})$ .

Multiplicando la expresión (3.4) por  $(\Phi_{k,i}^n)^t$ , se tiene que

$$(\Phi_{k,i}^n)^t \mathbf{u}_i^{n+1} = (\Phi_{k,i}^n)^t \mathbf{u}_i^n - \frac{\Delta t_k}{\Delta x_k} (\Phi_{k,i}^n)^t \left( \mathbf{g}_{k,i+\frac{1}{2}}^n - \mathbf{g}_{k,i-\frac{1}{2}}^n \right),$$

para todo  $i \in \mathbb{Z}$ ,  $n \in \mathbb{N}$  y  $k \in \mathbb{N}$ . Sumando en todos los nodos y todos los pasos de tiempo, se tiene que

$$\sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} (\Phi_{k,i}^n)^t \left( \mathbf{u}_{k,i}^{n+1} - \mathbf{u}_{k,i}^n \right) = -\Delta t_k \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} \frac{(\Phi_{k,i}^n)^t}{\Delta x_k} \left( \mathbf{g}_{k,i+\frac{1}{2}}^n - \mathbf{g}_{k,i-\frac{1}{2}}^n \right),$$

para todo  $k \in \mathbb{N}$ . Utilizando *sumas por partes*, es decir, reordenando los sumatorios anteriores de manera conveniente, y el hecho de que  $\varphi$  tiene soporte compacto, se llega a que

$$-\sum_{i \in \mathbb{Z}} (\Phi_{k,i}^1)^t \mathbf{u}_{k,i}^1 - \sum_{n \geq 2} \sum_{i \in \mathbb{Z}} \left( \Phi_{k,i}^n - \Phi_{k,i}^{n-1} \right)^t \mathbf{u}_{k,i}^n = \Delta t_k \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} \left( \frac{\Phi_{k,i+1}^n - \Phi_{k,i}^n}{\Delta x_k} \right)^t \mathbf{g}_{k,i-\frac{1}{2}}^n.$$

Multiplicando por el paso espacial  $\Delta x$  de cada malla y reordenando términos,

$$\begin{aligned} -\Delta x_k \sum_{i \in \mathbb{Z}} (\Phi_{k,i}^1)^t \mathbf{u}_{k,i}^1 &= \Delta t_k \Delta x_k \sum_{n \geq 2} \sum_{i \in \mathbb{Z}} \left( \frac{\Phi_{k,i}^n - \Phi_{k,i}^{n-1}}{\Delta t_k} \right)^t \mathbf{u}_{k,i}^n \\ &\quad + \Delta t_k \Delta x_k \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} \left( \frac{\Phi_{k,i+1}^n - \Phi_{k,i}^n}{\Delta x_k} \right)^t \mathbf{g}_{k,i-\frac{1}{2}}^n. \end{aligned}$$

Ahora bien, por el evidente parecido de estos términos con sumas de Riemann, se llega a que, si se puede aproximar  $\mathbf{u}_0$  por las funciones definidas a trozos  $\mathbf{u}_{k,i}^1$ , entonces

$$\lim_{k \rightarrow \infty} \Delta x_k \sum_{i \in \mathbb{Z}} (\Phi_{k,i}^1)^t \mathbf{u}_{k,i}^1 = \int_{\mathbb{R}} \varphi(x, 0) \mathbf{u}_0(x) dx,$$

y también

$$\lim_{k \rightarrow \infty} \Delta t_k \Delta x_k \sum_{n \geq 2} \sum_{i \in \mathbb{Z}} \left( \frac{\Phi_{k,i}^n - \Phi_{k,i}^{n-1}}{\Delta t_k} \right)^t \mathbf{u}_{k,i}^n = \int_0^{+\infty} \int_{\mathbb{R}} \partial_t \varphi(x, t) \mathbf{u}(x, t) dx dt.$$

Finalmente, por la convergencia en  $L^1_{\text{loc}}(\mathbb{R} \times [0, +\infty), \Omega)$  de  $\mathbf{u}_k$  a  $\mathbf{u}$ , se llega a que

$$\lim_{k \rightarrow \infty} \Delta t_k \Delta x_k \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} \left( \frac{\Phi_{k,i+1}^n - \Phi_i^n}{\Delta x_k} \right)^t \mathbf{g}_{k,i-\frac{1}{2}}^n = \int_0^\infty \int_{\mathbb{R}} \partial_x \varphi(x, t) \mathbf{f}(\mathbf{u}(x, t)) dx dt.$$

Juntando todo, se tiene que

$$\int_0^\infty \int_{\mathbb{R}} [\mathbf{u}(x, t) \cdot \partial_t \varphi(x, t) + \mathbf{f}(\mathbf{u}(x, t)) \partial_x \varphi(x, t)] dx dt + \int_{\mathbb{R}} \mathbf{u}_0(x) \varphi(x, 0) dx = 0.$$

Puesto que la función test  $\varphi \in \mathcal{C}_0^1(\mathbb{R}, \mathbb{R}^m)$  era arbitraria, esto prueba que  $\mathbf{u}$  es una solución débil del problema (3.1), tal y como se quería comprobar.  $\square$

Se debe resaltar que este resultado no afirma que un esquema consistente y conservativo converja. Únicamente afirma que, en caso de que converja, lo hace a una solución débil de la ley de conservación.

### 3.4. Métodos entrópicos

Aunque el Teorema (3.5) asegura, bajo ciertas hipótesis, que la aproximación del método (3.4) converge a una solución débil, nada garantiza *a priori* que esta solución sea también físicamente aceptable, es decir, entrópica. Se deben buscar más condiciones que nos aseguren que la sucesión de soluciones aproximadas converja y, en caso de hacerlo, lo hagan a una solución entrópica.

**Definición 3.6** (Esquema entrópico). Se dice que el esquema numérico (3.3) es entrópico si, para toda entropía  $(\mathcal{U}, \mathcal{F})$  del problema (3.1), existe una función  $\Psi: \Omega \times \Omega \rightarrow \mathbb{R}$ , llamada flujo numérico entrópico, tal que  $\Psi(\boldsymbol{\omega}, \boldsymbol{\omega}) = \mathcal{F}(\boldsymbol{\omega})$ , para todo  $\boldsymbol{\omega} \in \Omega$  y además

$$\mathcal{U}(\mathbf{u}_i^{n+1}) \leq \mathcal{U}(\mathbf{u}_i^n) - \frac{\Delta t}{\Delta x} [\Psi(\mathbf{u}_i^n, \mathbf{u}_{i+1}^n) - \Psi(\mathbf{u}_{i-1}^n, \mathbf{u}_i^n)], \quad \forall i \in \mathbb{Z}, \quad n \in \mathbb{N}. \quad (3.8)$$

Gracias a esta definición, se tiene el siguiente resultado:

**Teorema 3.7** ([7], pág. 144). *Si el esquema numérico (3.3), bajo las hipótesis del Teorema (3.5), es entrópico, entonces la función  $\mathbf{u}$  límite de la sucesión  $(\mathbf{u}_k)_{k \in \mathbb{N}}$  es solución entrópica de la ley de conservación.*

*Esquema de la demostración.* Utilizando un razonamiento análogo al de la demostración del Teorema (3.5), es fácil ver que para cualquier entropía  $(\mathcal{U}, \mathcal{F})$  del problema (3.1) tal que existe una función  $\Psi$  tal que se cumpla la condición (3.8), entonces el límite  $\mathbf{u}$  de la sucesión  $(\mathbf{u}_k)_{k \in \mathbb{N}}$  cumple la condición de entropía (2.16). Puesto que la entropía  $(\mathcal{U}, \mathcal{F})$  era arbitraria, se deduce que  $\mathbf{u}$  es la solución entrópica del problema (3.1).  $\square$

Sin embargo, esta condición es muy difícil de comprobar en la práctica, por lo que se buscará algún criterio que permita comprobar si un esquema es entrópico, al menos en el caso escalar.

### 3.5. Monotonía

En este apartado estará centrado únicamente en el caso escalar, donde una importante propiedad de los esquemas numéricos es la monotonía. Se debe entender esta propiedad en el sentido de que se están buscando soluciones que *repliquen* el comportamiento de las soluciones entrópicas exactas, como se desprende del Teorema 2.34.

**Definición 3.8** (Esquema monótono). Se dice que el esquema numérico (3.3) es monótono si, para todo  $n \in \mathbb{N}$ , se tiene que

$$[u_i^n \geq v_i^n, \quad \forall i \in \mathbb{Z}] \implies [u_i^{n+1} \geq v_i^{n+1}, \quad \forall i \in \mathbb{Z}],$$

donde  $u$  y  $v$  son soluciones débiles del problema (3.1).

Es sencillo analizar si un esquema es monótono estudiando la función  $\mathcal{H}$  asociada a dicho esquema, gracias al siguiente criterio, de prueba inmediata:

**Proposición 3.9.** *El esquema numérico (3.3) es monótono si, y sólo si,  $\mathcal{H}$  es creciente en cada una de sus componentes.*

La importancia de los esquemas monótonos queda claro por el siguiente resultado:

**Teorema 3.10** ([7], pág. 144). *Si un esquema numérico (3.3) es conservativo, consistente y monótono, entonces es entrópico.*

Otra importante propiedad de los esquemas monótonos es que su orden de convergencia está muy limitado. Mientras que la Proposición 3.4 aseguraba que los esquemas conservativos y consistentes tienen orden al menos uno, el siguiente resultado garantiza que tienen orden exactamente uno.

**Teorema 3.11** ([1], pág. 148). *Si un esquema numérico (3.3) es conservativo, consistente y monótono, entonces es de orden exactamente uno (en tiempo y en espacio); excepto en el caso de la ecuación del transporte con velocidad de propagación  $\lambda = \frac{\Delta x}{\Delta t}$  o  $\lambda = -\frac{\Delta x}{\Delta t}$ , o en el caso donde  $f$  es una constante, donde el esquema es exacto.*

*Demostración.* En primer lugar, se debe hacer notar que, si el esquema (3.3) es conservativo, entonces la función  $\mathcal{H}$  que lo define se escribe

$$\mathcal{H}(\omega_1, \omega_2, \omega_3) = \omega_2 - \frac{\Delta t}{\Delta x} [g(\omega_2, \omega_3) - g(\omega_1, \omega_2)], \quad \forall \omega_1, \omega_2, \omega_3 \in \Omega.$$

Por otra parte, de operar con desarrollos de Taylor de modo análogo a lo que se hizo en la Proposición 3.4, se obtiene que, si  $\bar{u}$  es una solución regular de la ley de conservación,

$$\tau_i^n = \frac{\Delta t}{2} \partial_t^2 \bar{u}(x_i, t_n) + \frac{\Delta x}{2} \partial_x [(\partial_2 g(\bar{u}, \bar{u}) - \partial_1 g(\bar{u}, \bar{u})) \partial_x \bar{u}](x_i, t_n) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta t^2).$$

Por ser  $\bar{u}$  regular, se tiene que, derivando la ley de conservación,

$$\partial_t^2 \bar{u} = \partial_t (-\partial_x f(\bar{u})) = -\partial_x (f'(\bar{u}) \partial_t \bar{u}) = \partial_x (f'(\bar{u}) \partial_x f(\bar{u})) = \partial_x ((f'(\bar{u}))^2 \partial_x \bar{u}).$$

En consecuencia, el error de truncamiento se puede reescribir como

$$\tau_i^n = \frac{\Delta t}{2} \partial_x (\beta(\bar{u}) \partial_x \bar{u})(x_i, t_n) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta t^2),$$

para lo cual se define

$$\beta(\bar{u}) = (f'(\bar{u}))^2 - \frac{\Delta x}{\Delta t} (\partial_1 g(\bar{u}, \bar{u}) - \partial_2 g(\bar{u}, \bar{u})).$$

Además, por hipótesis, el esquema es monótono. Así, en virtud de la Proposición 3.9, se tiene que, para todo  $\omega \in \Omega$ ,

$$\begin{cases} \partial_{\omega_1} \mathcal{H}(\omega, \omega, \omega) = \frac{\Delta t}{\Delta x} \partial_1 g(\omega, \omega) \geq 0, \\ \partial_{\omega_2} \mathcal{H}(\omega, \omega, \omega) = 1 - \frac{\Delta t}{\Delta x} [\partial_1 g(\omega, \omega) - \partial_2 g(\omega, \omega)] \geq 0, \\ \partial_{\omega_3} \mathcal{H}(\omega, \omega, \omega) = -\frac{\Delta t}{\Delta x} \partial_2 g(\omega, \omega) \geq 0, \end{cases}$$

de donde se deducen las desigualdades

$$\partial_1 g(\omega, \omega) \geq 0, \quad \partial_2 g(\omega, \omega) \leq 0, \quad \partial_1 g(\omega, \omega) - \partial_2 g(\omega, \omega) \leq \frac{\Delta x}{\Delta t}, \quad \forall \omega \in \Omega. \quad (3.9)$$

Por otra parte, de la Proposición 3.4 se tiene que

$$\partial_1 g(\omega, \omega) + \partial_2 g(\omega, \omega) = f'(\omega), \quad \forall \omega \in \Omega. \quad (3.10)$$

Así, de (3.9) y (3.10) se deducen las desigualdades

$$(f'(\omega))^2 \leq (\partial_1 g(\omega, \omega) - \partial_2 g(\omega, \omega))^2 \leq \frac{\Delta x}{\Delta t} (\partial_1 g(\omega, \omega) - \partial_2 g(\omega, \omega)) \implies \beta(\omega) \leq 0.$$

De este modo, para que el esquema sea de orden al menos dos, se debe cumplir que  $\beta(\omega) = 0$  para todo  $\omega \in \Omega$ . Pero de los razonamientos anteriores, se obtiene que esto únicamente ocurre en uno de los siguientes casos:

$$\left[ \partial_1 g(\omega, \omega) = \frac{\Delta x}{\Delta t} \quad \text{y} \quad \partial_2 g(\omega, \omega) = 0 \right] \quad \text{o} \quad \left[ \partial_2 g(\omega, \omega) = -\frac{\Delta x}{\Delta t} \quad \text{y} \quad \partial_1 g(\omega, \omega) = 0 \right],$$

lo cual, atendiendo a la igualdad (3.10), corresponde con

$$f'(\omega) = \frac{\Delta x}{\Delta t} \quad \text{o bien} \quad f'(\omega) = -\frac{\Delta x}{\Delta t}, \quad \forall \omega \in \Omega.$$

En ambos casos se trata de la ecuación del transporte, con velocidad de propagación  $\lambda = \frac{\Delta x}{\Delta t}$  y  $\lambda = -\frac{\Delta x}{\Delta t}$ , respectivamente. En estos dos casos, el esquema resultante produce aproximaciones

$$u_i^{n+1} = u_{i-1}^n \quad \text{y} \quad u_i^{n+1} = u_{i+1}^n,$$

lo cual significa que el esquema es exacto si las aproximaciones  $u^1$  de la condición inicial  $u_0$  son exactas. En el resto de casos, el esquema es de orden exactamente uno.  $\square$

En resumen, tener un esquema numérico conservativo, consistente y monótono asegura que dicho esquema sea entrópico. Sin embargo, no podrá tener orden mayor que uno.

### 3.6. Estabilidad y convergencia

En las secciones anteriores se establecieron condiciones bajo las cuales los esquemas numéricos son consistentes y también condiciones necesarias para que, en caso de converger, lo hagan a soluciones entrópicas de la ley de conservación. Sin embargo, esto no garantiza en ningún caso la convergencia, si no que se debe hacer la comprobación adicional de que los métodos sean *estables*. Antes de dar resultados, se va a aclarar la notación que se empleará:

**Definición 3.12** (Norma  $L^2$  discreta). Sea  $\mathbf{h}: \mathbb{R} \rightarrow \mathbb{R}^m$  una función constante en cada uno de los volúmenes finitos de  $\mathcal{T}$ . Poniendo  $\mathbf{h}_i$  como el valor de  $\mathbf{h}$  sobre el volumen finito  $\mathcal{C}_i$ , se define la norma  $L^2$  discreta de  $\mathbf{h}$  sobre  $\mathcal{T}$  como

$$\|\mathbf{h}\|_2 = \left( \Delta x \sum_{i \in \mathbb{Z}} \|h_i\|_2^2 \right)^{\frac{1}{2}} = \left( \Delta x \sum_{i \in \mathbb{Z}} \sum_{k=1}^m |(h_i)_m|^2 \right)^{\frac{1}{2}}.$$

**Definición 3.13** (Esquema  $L^2$ -estable). Se dice que el esquema numérico (3.3) es  $L^2$ -estable si existe una constante  $\kappa > 0$  independiente de  $\Delta t$  tal que

$$\|\mathbf{u}^n\|_2 \leq \kappa \|\mathbf{u}^1\|_2, \quad \forall n \in \mathbb{N}.$$

**Definición 3.14** (Esquema convergente). Se dice que el esquema numérico (3.3) es convergente si, para cualquier condición inicial  $\mathbf{u}_0 \in L^\infty(\mathbb{R}, \Omega)$ , si la inicialización  $\mathbf{u}^1$  del esquema converge a  $\mathbf{u}_0$  cuando  $\|\mathcal{T}\| \rightarrow 0$ , entonces  $\mathbf{u}^n$  converge a una solución débil del problema (3.1) cuando  $\|\mathcal{T}\| \rightarrow 0$  y  $\Delta t \rightarrow 0$ .

Si se considera un sistema lineal de leyes de conservación, el siguiente resultado da una condición necesaria y suficiente para que un método sea convergente:

**Teorema 3.15** (de equivalencia de Lax-Richtmyer. [16], pág. 26). *Para un sistema lineal de leyes de conservación, si el esquema numérico (3.3) es consistente, entonces es convergente si, y sólo si, es  $L^2$ -estable.*

Sin embargo, en el caso general, no existen resultados que aseguren la convergencia de los métodos numéricos. Únicamente se tienen condiciones suficientes para el caso escalar (ver [14], pág. 164), para lo cual se debe definir un nuevo concepto de estabilidad, relacionada con la variación total de cada una de las aproximaciones en tiempo. En líneas generales, se puede probar que, bajo ciertas hipótesis, a medida que se reducen la malla espacial y el paso temporal, las sucesiones de soluciones aproximadas son de variación acotada. Por un argumento de compacidad, se puede extraer una subsucesión convergente a una cierta función, que será solución débil de la ley de conservación como consecuencia del Teorema de Lax-Wendroff.

### 3.7. Estabilidad lineal

En el caso lineal, se puede estudiar la  $L^2$ -estabilidad directamente, empleando análisis de Fourier, tal y como se describirá más adelante. Sin embargo, en el caso no lineal, esto es especialmente complicado, por lo que se recurre a la linealización.

Dado el esquema numérico conservativo (3.4), se define el *esquema linealizado* respecto al nodo  $(x_k, t_s)$  como

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - \frac{\Delta t}{\Delta x} [(\partial_1 \mathbf{g}(\mathbf{u}_k^s, \mathbf{u}_k^s) - \partial_2 \mathbf{g}(\mathbf{u}_k^s, \mathbf{u}_k^s)) \mathbf{w}_i^n + \partial_2 \mathbf{g}(\mathbf{u}_k^s, \mathbf{u}_k^s) \mathbf{w}_{j+1}^n - \partial_1 \mathbf{g}(\mathbf{u}_k^s, \mathbf{u}_k^s) \mathbf{w}_{i-1}^n], \quad (3.11)$$

para todo  $i \in \mathbb{Z}$  y  $n \in \mathbb{N}$ . Puesto que aplicar el análisis de Fourier al esquema no lineal no es posible, se estudiará la  $L^2$ -estabilidad del esquema linealizado en todos los nodos, con la esperanza de que las condiciones de estabilidad que aparezcan sean también válidas para el caso no lineal.

Esto sugiere la siguiente definición.

**Definición 3.16** (Esquema linealmente estable). Se dice que el esquema numérico (3.4) es linealmente estable si (3.11) es  $L^2$ -estable para todo  $k \in \mathbb{Z}$ ,  $s \in \mathbb{N}$ .

Se va a describir cómo estudiar la estabilidad lineal mediante análisis de Fourier. En primer lugar, se puede reescribir la expresión (3.11) como

$$\mathbf{w}_i^{n+1} = \sum_{r=-1}^1 \mathbb{M}_r \mathbf{w}_{i+r}^n, \quad \forall i \in \mathbb{Z}, \quad \forall n \in \mathbb{N},$$

donde  $\mathbb{M}_r \in \mathcal{M}_{m \times m}(\mathbb{R})$  son matrices constantes que dependen de  $k$  y  $s$ . De la igualdad anterior se deduce que

$$\mathbf{w}^{n+1}(x) = \sum_{r=-1}^1 \mathbb{M}_r \mathbf{w}^n(x + r\Delta x), \quad \forall x \in \mathbb{R}, \quad \forall n \in \mathbb{N}. \quad (3.12)$$

Por otra parte, si  $\mathbf{w}^n \in L^2(\mathbb{R}, \mathbb{R}^m)$ , denotando por  $j$  a la unidad imaginaria, se define la *transformada de Fourier* de cada  $\mathbf{w}^{n+1}$  como

$$\hat{\mathbf{w}}^n(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathbf{w}^n(x) e^{-j\xi x} dx, \quad \forall \xi \in \mathbb{R}.$$

A partir de (3.12), se tiene que, para todo  $\xi \in \mathbb{R}$ ,

$$\begin{aligned} \hat{\mathbf{w}}^{n+1}(\xi) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathbf{w}^{n+1}(x) e^{-j\xi x} dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \sum_{r=-1}^1 \mathbb{M}_r \mathbf{w}^n(x + r\Delta x) e^{-j\xi x} dx \\ &= \frac{1}{\sqrt{2\pi}} \sum_{r=-1}^1 \mathbb{M}_r \int_{\mathbb{R}} \mathbf{w}^n(x + r\Delta x) e^{-j\xi x} dx \\ &= \frac{1}{\sqrt{2\pi}} \sum_{r=-1}^1 \mathbb{M}_r \int_{\mathbb{R}} \mathbf{w}^n(x + r\Delta x) e^{j\xi r\Delta x} e^{-j\xi(x+r\Delta x)} dx \\ &= \sum_{r=-1}^1 \mathbb{M}_r e^{j\xi r\Delta x} \left( \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathbf{w}^n(x + r\Delta x) e^{-j\xi(x+r\Delta x)} dx \right) = \underbrace{\left( \sum_{r=-1}^1 \mathbb{M}_r e^{j\xi r\Delta x} \right)}_{\mathbb{G}(\xi)} \hat{\mathbf{w}}^n(\xi). \end{aligned}$$

En consecuencia, se tiene que, aplicando el razonamiento anterior  $n - 1$  veces,

$$\hat{\mathbf{w}}^n(\xi) = (\mathbb{G}(\xi))^{n-1} \hat{\mathbf{w}}^1(\xi), \quad \forall \xi \in \mathbb{R}.$$

Recordando la identidad de Parseval, se deduce que

$$\|\mathbf{w}^n\|_2 = \|\hat{\mathbf{w}}^n\|_2 \leq \|\mathbb{G}(\xi)\|^{n-1} \|\hat{\mathbf{w}}^1\|_2 = \|\mathbb{G}(\xi)\|^{n-1} \|\mathbf{w}^1\|_2, \quad \forall \xi \in \mathbb{R}, \quad \forall n \in \mathbb{N}.$$

De todo este razonamiento, se llega a la conclusión de que, el esquema (3.4) es linealmente estable si, y sólo si,  $\|\mathbb{G}(\xi)\| \leq 1$  para todo  $\xi \in \mathbb{R}$ .

## Capítulo 4

# Ejemplos de esquemas numéricos

En este capítulo se ejemplificarán los conceptos explicados en el capítulo anterior con múltiples esquemas numéricos de tipo volúmenes finitos que se consideran fundamentales para la resolución de leyes de conservación unidimensionales.

### 4.1. Esquema de Lax-Friedrichs

El primero de los métodos para resolver el problema (3.1) que se presentará es el llamado *esquema de Lax-Friedrichs*. Este método propone dar la siguiente actualización de un paso de tiempo

$$\mathbf{u}_i^{n+1} = \frac{1}{2} (\mathbf{u}_{i-1}^n + \mathbf{u}_{i+1}^n) - \frac{\Delta t}{2\Delta x} (\mathbf{f}(\mathbf{u}_{i+1}^n) - \mathbf{f}(\mathbf{u}_{i-1}^n)), \quad \forall i \in \mathbb{Z}, \quad \forall n \in \mathbb{N}, \quad (4.1)$$

que puede escribirse en forma conservativa definiendo el flujo numérico

$$\mathbf{g}(\omega_1, \omega_2) = \frac{\mathbf{f}(\omega_1) + \mathbf{f}(\omega_2)}{2} - \frac{\Delta x}{2\Delta t} (\omega_2 - \omega_1), \quad \forall \omega_1, \omega_2 \in \Omega.$$

La Proposición 3.4 asegura que el método es de orden al menos uno. A continuación, se va a estudiar la estabilidad empleando el análisis de Fourier descrito en la Sección 3.7. En este caso, el esquema linealizado (3.11) es escribe, para  $k \in \mathbb{Z}$  y  $s \in \mathbb{N}$  fijados,

$$\mathbf{w}_i^{n+1} = \left[ \frac{\Delta t}{2\Delta x} \mathbb{A}(\mathbf{w}_k^s) + \frac{1}{2} \mathbb{I} \right] \mathbf{w}_{i-1}^n + \left[ -\frac{\Delta t}{2\Delta x} \mathbb{A}(\mathbf{w}_k^s) + \frac{1}{2} \mathbb{I} \right] \mathbf{w}_{i+1}^n, \quad \forall i \in \mathbb{Z}, \quad n \in \mathbb{N}. \quad (4.2)$$

Definiendo

$$\mathbb{M}_{-1} = \frac{\Delta t}{2\Delta x} \mathbb{A}(\mathbf{w}_k^s) + \frac{1}{2} \mathbb{I}, \quad \mathbb{M}_0 = 0, \quad \mathbb{M}_1 = -\frac{\Delta t}{2\Delta x} \mathbb{A}(\mathbf{w}_k^s) + \frac{1}{2} \mathbb{I},$$

se puede escribir el esquema linealizado como  $\mathbf{w}_i^{n+1} = \sum_{r=-1}^1 \mathbb{M}_r \mathbf{w}_i^n$ . Así, el factor de amplificación es

$$\begin{aligned} \mathbb{G}(\xi) &= \sum_{r=-1}^1 \mathbb{M}_r e^{j\xi r \Delta x} = \left[ \frac{\Delta t}{2\Delta x} \mathbb{A}(\mathbf{w}_k^s) + \frac{1}{2} \mathbb{I} \right] e^{-j\xi \Delta x} + \left[ -\frac{\Delta t}{2\Delta x} \mathbb{A}(\mathbf{w}_k^s) + \frac{1}{2} \mathbb{I} \right] e^{j\xi \Delta x} \\ &= \frac{\Delta t}{2\Delta x} \mathbb{A}(\mathbf{w}_k^s) \left( e^{-j\xi \Delta x} - e^{j\xi \Delta x} \right) + \frac{1}{2} \left( e^{-j\xi \Delta x} + e^{j\xi \Delta x} \right) \mathbb{I} = \cos(\xi \Delta x) \mathbb{I} - \frac{\Delta t}{\Delta x} \mathbb{A}(\mathbf{w}_k^s) \sin(\xi \Delta x) j. \end{aligned}$$

Se va a estudiar bajo qué condiciones se cumple la condición de estabilidad lineal  $\|\mathbb{G}(\xi)\| \leq 1$ , para todo  $\xi \in \mathbb{R}$ . Se tiene que, para un  $\boldsymbol{\omega} \in \Omega$  con  $\|\boldsymbol{\omega}\|_2 = 1$ ,

$$\begin{aligned} \|\mathbb{G}(\xi)\boldsymbol{\omega}\|_2 &= \sqrt{\sum_{p=1}^m \left[ \omega_p^2 \cos^2(\xi \Delta x) + \left( \frac{\Delta t}{\Delta x} \right)^2 (\mathbb{A}(\mathbf{w}_k^s)\boldsymbol{\omega})_p^2 \sin^2(\xi \Delta x) \right]} \\ &= \sqrt{1 - \sin^2(\xi \Delta x) + \left( \frac{\Delta t}{\Delta x} \right)^2 \sum_{p=1}^m (\mathbb{A}(\mathbf{w}_k^s)\boldsymbol{\omega})_p^2 \sin^2(\xi \Delta x)} \\ &= \sqrt{1 + \left[ \left( \frac{\Delta t}{\Delta x} \right)^2 \|\mathbb{A}(\mathbf{w}_k^s)\boldsymbol{\omega}\|_2^2 - 1 \right] \sin^2(\xi \Delta x)}, \end{aligned}$$

de donde se deduce que

$$\|\mathbb{G}(\xi)\| = \sup_{\boldsymbol{\omega} \in \Omega, \|\boldsymbol{\omega}\|_2=1} \|\mathbb{G}(\xi)\boldsymbol{\omega}\|_2 = \sqrt{1 + \left[ \left( \frac{\Delta t}{\Delta x} \right)^2 \|\mathbb{A}(\mathbf{w}_k^s)\|_2^2 - 1 \right] \sin^2(\xi \Delta x)}.$$

Finalmente, se tiene que el esquema linealizado en  $k \in \mathbb{Z}$  y  $s \in \mathbb{N}$  es estable si

$$\|\mathbb{G}(\xi)\| \leq 1 \iff \frac{\Delta t}{\Delta x} \|\mathbb{A}(\mathbf{u}_k^s)\| \leq 1,$$

por lo que la condición necesaria y suficiente para que el esquema de Lax-Friedrichs sea linealmente estable es que

$$\sup_{k \in \mathbb{Z}} \|\mathbb{A}(\mathbf{u}_k^s)\| \frac{\Delta t}{\Delta x} \leq 1, \quad \forall s \in \mathbb{N}. \quad (4.3)$$

A la condición (4.4) se la denomina *condición de Courant–Friedrichs–Lewy* o *condición CFL*. En la práctica, lo que se hace es tomar un  $\mu \in (0, 1]$  tal que el paso de tiempo  $\Delta t$  se ajusta de modo que  $\sup_{k \in \mathbb{Z}} \|\mathbb{A}(\mathbf{u}_k^s)\| \frac{\Delta t}{\Delta x} = \mu \leq 1$ , para todo  $s \in \mathbb{N}$ . A este coeficiente  $\mu$  se le denomina *número de Courant*.

*Observación 4.1.* En el caso escalar, esta condición de estabilidad se escribe

$$\sup_{k \in \mathbb{Z}} |f'(\mathbf{u}_k^s)| \frac{\Delta t}{\Delta x} \leq 1, \quad \forall s \in \mathbb{N}. \quad (4.4)$$

Aunque, tal y como se acaba de ver, estudiar la estabilidad lineal del esquema de Lax-Friedrichs no es muy complicado, comprobar que es entrópico sí es muy complicado en el caso

general, por lo que únicamente se hará dicha comprobación para leyes de conservación escalares. En dicho caso, utilizando la Proposición 3.9, el hecho de que

$$\partial_1 \mathcal{H}(\omega_1, \omega_2, \omega_3) = \frac{1}{2} + \frac{\Delta t}{2\Delta x} f'(\omega_1), \quad \partial_2 \mathcal{H}(\omega_1, \omega_2, \omega_3) = 0, \quad \partial_3 \mathcal{H}(\omega_1, \omega_2, \omega_3) = \frac{1}{2} - \frac{\Delta t}{2\Delta x} f'(\omega_3),$$

para todo  $\omega_1, \omega_2, \omega_3 \in \Omega$ . La condición (4.4) es suficiente para que  $\partial_k \mathcal{H}(\omega_1, \omega_2, \omega_3) \geq 0$ ,  $k \in \{1, 2, 3\}$ , por lo tanto, bajo la condición CFL (4.4), el esquema es monótono y, por tanto, entrópico (por la Proposición 3.9) y de orden exactamente uno (por el Teorema 3.11).

A la hora de implementar los métodos numéricos de este capítulo se han considerado tests relativos a las ecuaciones del transporte, de Burgers y de Saint-Venant explicadas a lo largo de este trabajo. Se busca resolver en intervalos espaciales acotados de la forma  $[a, b]$ , para tiempos en  $[0, t_f]$  y con condiciones iniciales o bien de tipo Riemann (de dos estados), que es una expresión como (2.21), o bien su extensión a un caso de tres estados, que es una condición inicial de la forma

$$\mathbf{u}_0(x) = \begin{cases} \mathbf{u}_L & \text{si } x \leq 0, \\ \mathbf{u}_C & \text{si } 0 < x \leq 1, \\ \mathbf{u}_R & \text{si } 1 < x, \end{cases} \quad (4.5)$$

donde  $\mathbf{u}_L, \mathbf{u}_C, \mathbf{u}_R \in \mathbb{R}^m$  son constantes denominadas *estados iniciales* del problema de Riemann.

*Observación 4.2.* Aunque la teoría se esté desarrollando para el caso no acotado, en la práctica se debe acotar el problema diferencial tanto en espacio como en tiempo. La acotación en tiempo no es problemática, mientras que la espacial sí es más delicada. Esto se debe a que los esquemas numéricos que se están considerando son de la forma (3.3), por lo que, para poder calcular  $\mathbf{u}_i^{n+1}$ , se necesita el valor aproximado de sus volúmenes *vecinos* y para el primer y último volumen no se tiene a uno de estos vecinos. Una opción para resolver este problema es definir *volúmenes fantasma*  $\mathcal{C}_0$  y  $\mathcal{C}_{I+1}$  donde, en cada paso de tiempo se toman  $\mathbf{u}_0^n = \mathbf{u}_1^n$  y  $\mathbf{u}_{I+1}^n = \mathbf{u}_I^n$ , lo cual hace posible realizar las actualizaciones (3.3). Esta opción se corresponde de forma razonable con la realidad física que se pretende modelar. Se pueden ver más detalles en [19].

A continuación, se presentan algunas ejecuciones del esquema de Lax-Friedrichs para distintas ecuaciones. En la Figura 4.1 se resuelve para la ecuación del transporte, como ejemplo de ecuación lineal y escalar, en la Figura 4.2, se resuelve para la ecuación de Burgers, no lineal y escalar, y en la 4.3 se resuelve para las ecuaciones de Saint-Venant, ejemplo de sistema no lineal. Aunque los resultados no parezcan especialmente malos, en la práctica, el esquema de Lax-Friedrichs no se emplea por su *alta disipación* (ver [7] y [20] para más detalles). Se puede apreciar este inconveniente especialmente en sistemas no lineales. Esto se debe a que cuánto más se aleje el número de Courant  $\mu$  de 1, la disipación aumentará y, en el caso no lineal, se toman cotas a la baja de  $\mu$  para garantizar la estabilidad.

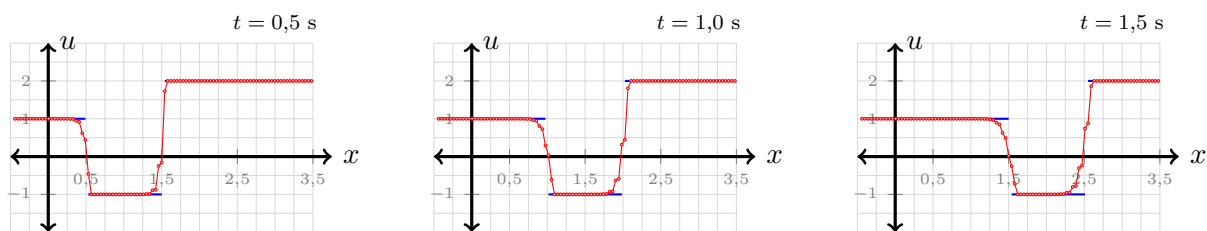


Figura 4.1: Resolución numérica de la ecuación del transporte con  $\lambda = 1$  para una condición inicial de tipo Riemann con tres estados:  $u_L = 1$ ,  $u_C = -1$  y  $u_R = 2$ . Se está resolviendo en el dominio computacional  $[a, b] = [-0,5, 3,5]$ , con  $I = 100$  y  $\mu = 0,95$ . En azul, solución exacta y, en rojo, aproximación mediante el esquema de Lax-Friedrichs.

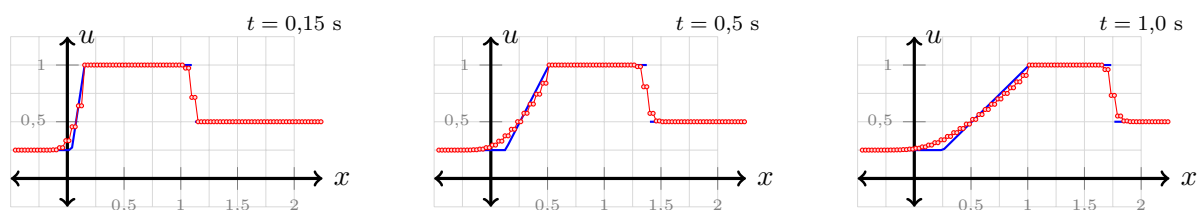


Figura 4.2: Resolución numérica de la ecuación de Burgers para una condición inicial de tipo Riemann con tres estados:  $u_L = 0,25$ ,  $u_C = 1$  y  $u_R = 0,5$ . Se está resolviendo en el dominio computacional  $[a, b] = [-0,5, 2,25]$ , con  $I = 100$  y  $\mu = 1$ . En azul, solución exacta y, en rojo, aproximación mediante el esquema de Lax-Friedrichs.

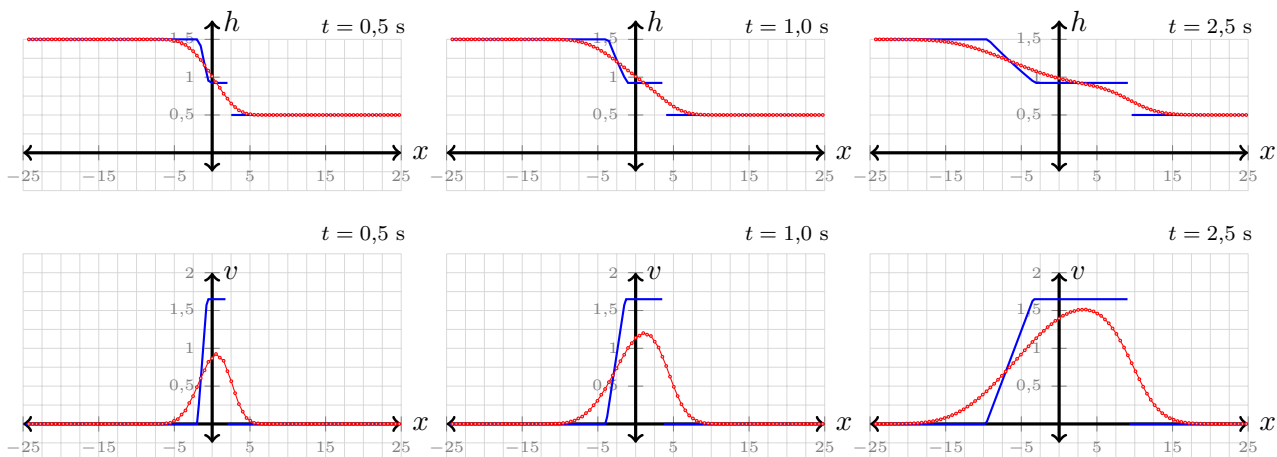


Figura 4.3: Resolución numérica de las ecuaciones de Saint Venant para el problema de rotura de una presa, con  $h_L = 1,5$  m,  $h_R = 0,5$  m,  $v_L = v_R = 0$  m/s. Se está resolviendo en el dominio computacional  $[a, b] = [-0,5, 2,25]$ , con  $I = 100$  y  $\Delta t = 0,03$  s. En azul, solución exacta y, en rojo, aproximación mediante el esquema de Lax-Friedrichs.

## 4.2. Esquema de Lax-Wendroff

El esquema de Lax-Wendroff, a diferencia del resto de esquemas tratados en este trabajo, es de orden dos. Las actualizaciones de este esquema (de un paso de tiempo) vienen dadas por

$$\mathbf{u}_i^{n+1} = \mathbf{u}_i^n - \frac{\Delta t}{2\Delta x} [\mathbf{f}(\mathbf{u}_{i+1}^n) - \mathbf{f}(\mathbf{u}_{i-1}^n)] + \frac{(\Delta t)^2}{2(\Delta x)^2} \left\{ \mathbb{A} \left( \frac{1}{2}(\mathbf{u}_i^n + \mathbf{u}_{i+1}^n) \right) [\mathbf{f}(\mathbf{u}_{i+1}^n) - \mathbf{f}(\mathbf{u}_i^n)] \right\} \\ - \frac{(\Delta t)^2}{2(\Delta x)^2} \left\{ \mathbb{A} \left( \frac{1}{2}(\mathbf{u}_{i-1}^n + \mathbf{u}_i^n) \right) [\mathbf{f}(\mathbf{u}_i^n) - \mathbf{f}(\mathbf{u}_{i-1}^n)] \right\},$$

para todo  $i \in \mathbb{Z}$  y  $n \in \mathbb{N}$ . Se trata de un esquema conservativo, con flujo numérico dado por

$$\mathbf{g}(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2) = \frac{1}{2} [\mathbf{f}(\boldsymbol{\omega}_1) + \mathbf{f}(\boldsymbol{\omega}_2)] - \frac{\Delta t}{2\Delta x} \mathbb{A} \left( \frac{1}{2}(\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2) \right) [\mathbf{f}(\boldsymbol{\omega}_2) - \mathbf{f}(\boldsymbol{\omega}_1)], \quad \forall \boldsymbol{\omega}_2, \boldsymbol{\omega}_1 \in \Omega.$$

La linealización de este esquema resulta en (4.2), por lo tanto (4.3) es también la condición de estabilidad lineal del esquema de Lax-Wendroff. En [7] y [8] puede verse la deducción de este esquema, así como los detalles de la prueba de que, en caso de converger, es de orden dos en tiempo y en espacio.

Además, cabe notar que este método no es entrópico. Para comprobarlo, basta con considerar la ecuación de Burgers y una aproximación  $u^1$  de una condición inicial  $u^0$  tal que

$$u_i^1 = \begin{cases} u_L & \text{si } x_i \leq 0, \\ u_R & \text{si } x_i > 0, \end{cases}$$

donde  $u_L = -u_R$ . Puede comprobarse fácilmente por inducción que, para todo  $i \in \mathbb{Z}$  y  $n \in \mathbb{N}$ ,

$$g(u_i^n, u_{i+1}^n) = g(u_{i-1}^n, u_i^n) = f(u_i^n) \implies u_i^{n+1} = u_i^n.$$

En consecuencia, la solución obtenida es estacionaria. Si  $u_L \geq u_R$ , por los desarrollos de la Sección 2.4 se puede comprobar ésta es la solución entrópica del problema. Sin embargo, si  $u_L < u_R$ , la solución entrópica es una onda de rarefacción, que no es la solución estacionaria calculada por el esquema. Por lo tanto, el esquema de Lax-Wendroff no es entrópico.

Se muestran ejecuciones del esquema de Lax-Wendroff en la Figura 4.4 para el caso lineal y escalar, en la Figura 4.5 para el caso no lineal y escalar y en la Figura 4.6 para un sistema de ecuaciones no lineales. Se puede comprobar como, comparativamente con el esquema de Lax-Friedrichs, este método tiene un orden mayor, ya que allí donde la solución es regular, el esquema de Lax-Wendroff se aproxima mejor a la solución exacta. Además, ya no se tiene el problema de una gran disipación, como ocurría en el caso del esquema de Lax-Friedrichs.

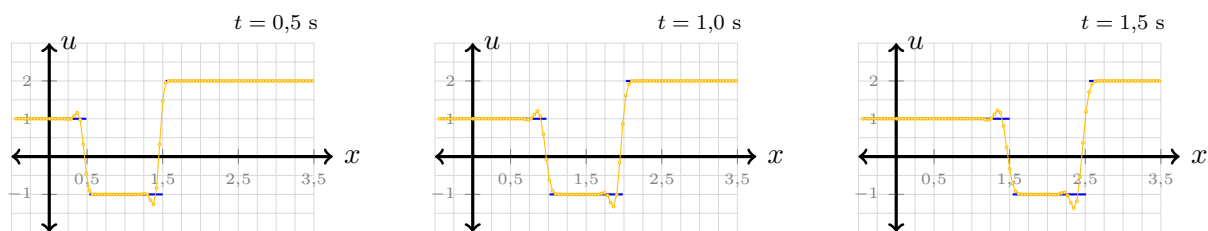


Figura 4.4: Resolución numérica de la ecuación del transporte con  $\lambda = 1$  para una condición inicial de tipo Riemann con tres estados:  $u_L = 1$ ,  $u_C = -1$  y  $u_R = 2$ . Se está resolviendo en el dominio computacional  $[a, b] = [-0,5, 3,5]$ , con  $I = 100$  y  $\mu = 0,9$ . En azul, solución exacta y, en amarillo, aproximación mediante el esquema de Lax-Wendroff.

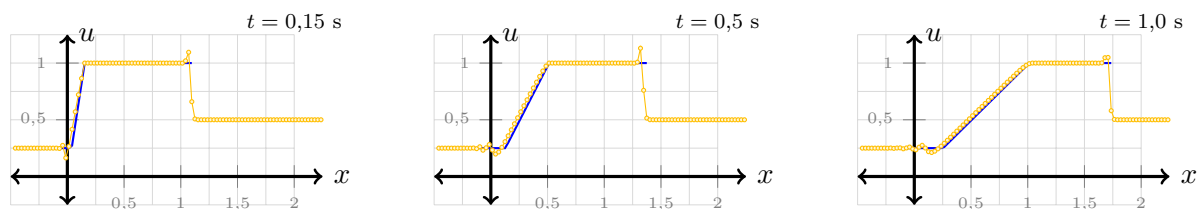


Figura 4.5: Resolución numérica de la ecuación de Burgers para una condición inicial de tipo Riemann con tres estados:  $u_L = 0,25$ ,  $u_C = 1$  y  $u_R = 0,5$ . Se está resolviendo en el dominio computacional  $[a, b] = [-0,5, 2,25]$ , con  $I = 100$  y  $\mu = 0,9$ . En azul, solución exacta y, en amarillo, aproximación mediante el esquema de Lax-Wendroff.

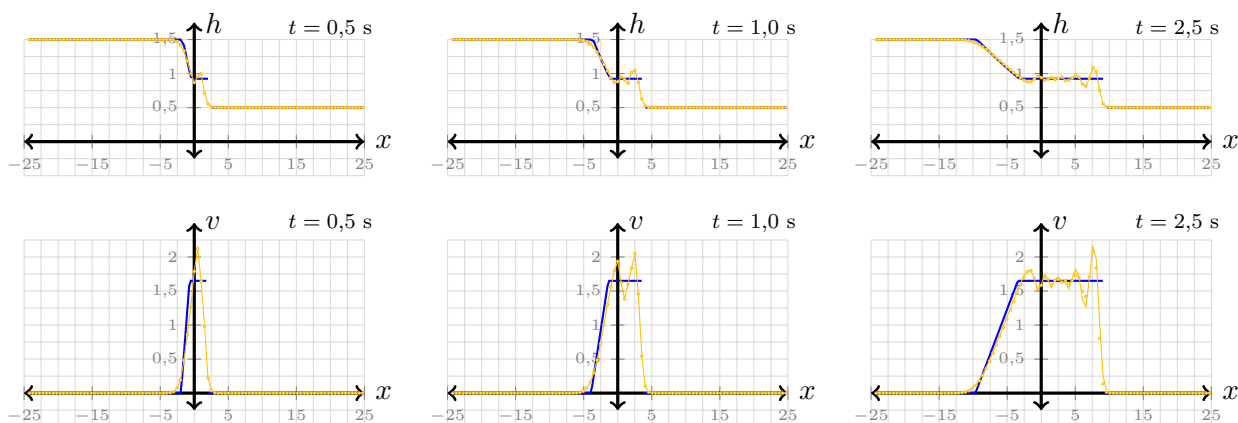


Figura 4.6: Resolución numérica de las ecuaciones de Saint Venant para el problema de rotura de una presa, con  $h_L = 1,5$  m,  $h_R = 0,5$  m,  $v_L = v_R = 0$  m/s. Se está resolviendo en el dominio computacional  $[a, b] = [-0,5, 2,25]$ , con  $I = 100$  y  $\Delta t = 0,03$  s. En azul, solución exacta y, en amarillo, aproximación mediante el esquema de Lax-Wendroff.

Sin embargo, como contrapartida a esta mejora del orden, cerca de estas discontinuidades

aparecen importantes oscilaciones que aumentan considerablemente el error respecto de la solución exacta. Una posible solución a este problema es el uso de *limitadores*, para obtener así un método de orden dos sin estas oscilaciones espurias. Puede leerse más sobre este tema en [13].

### 4.3. Método de Godunov

La interpretación de los esquemas presentados anteriormente tienen mucha similitud con los de métodos de diferencias finitas. En cambio, el esquema de Godunov se deduce directamente de las propiedades de las leyes de conservación.

De la expresión (3.2) se seguía que  $\mathbf{u}_i^n$  era una aproximación del valor promedio de  $\mathbf{u}(\cdot, t_n)$  en el volumen finito  $\mathcal{C}_i$ . Ahora bien, si a partir de la función  $\mathbf{u}(\cdot, t_n)$  se pudiese calcular de forma exacta  $\mathbf{u}(\cdot, t_{n+1})$ , entonces integrando esta función en  $\mathcal{C}_i$ , se obtendría la aproximación  $\mathbf{u}_i^{n+1}$ . Se puede escribir este razonamiento como

$$\mathbf{u}_i^{n+1} \approx \frac{1}{\Delta x} \int_{\mathcal{C}_i} \mathbf{u}(x, t_{n+1}) dx = \frac{1}{\Delta x} \int_{\mathcal{C}_i} \mathbf{w}^n(x, \Delta t) dx,$$

donde  $\mathbf{w}^n$  es la solución de la ley de conservación que tiene a la función  $\mathbf{u}(\cdot, t_n)$ , como condición inicial. Sin embargo, en los métodos de volúmenes finitos no se tienen las funciones  $\mathbf{u}(\cdot, t_n)$ , si no que se tienen funciones constantes  $\mathbf{u}^n(\cdot)$  a trozos sobre cada volumen. Con este contexto, Godunov propuso el esquema numérico

$$\mathbf{u}_i^{n+1} = \frac{1}{\Delta x} \int_{\mathcal{C}_i} \mathbf{w}^n(x, \Delta t) dx, \quad \forall i \in \mathbb{Z}, \quad n \in \mathbb{N},$$

donde  $\mathbf{w}^n$  es la solución entrópica exacta del problema

$$\begin{cases} \partial_t \mathbf{w}^n(x, t) + \partial_x \mathbf{f}(\mathbf{w}^n(x, t)) = 0, & \forall x \in \mathbb{R}, \quad \forall t \in [0, +\infty), \\ \mathbf{w}^n(x, 0) = \mathbf{u}^n(x), & \forall x \in \mathbb{R}. \end{cases} \quad (4.6)$$

Aunque puede parecer que resolver este problema es tan complicado como resolver el problema original, se probará más adelante que, para un  $\Delta t > 0$  lo suficientemente pequeño, la función  $\mathbf{w}^n$  viene dada por

$$\mathbf{w}^n(x, t) = \begin{cases} \mathbf{w}_R \left( \frac{x - a_i}{t}; \mathbf{u}_{i-1}^n, \mathbf{u}_i^n \right) & \text{si } x \in [a_i, x_i), \\ \mathbf{w}_R \left( \frac{x - b_i}{t}; \mathbf{u}_i^n, \mathbf{u}_{i+1}^n \right) & \text{si } x \in [x_i, b_i), \end{cases} \quad (4.7)$$

donde por  $\mathbf{w}_R \left( \frac{x}{t}; \omega_1, \omega_2 \right)$  se denotaba a la solución (autosimilar) del problema de Riemann con estados iniciales  $\omega_1, \omega_2 \in \Omega$ . En efecto, el problema (4.6) se puede interpretar como una concatenación de problemas de tipo Riemann de la forma

$$\begin{cases} \partial_t \mathbf{w}^n(x, t) + \partial_x \mathbf{f}(\mathbf{w}^n(x, t)) = 0, & \text{si } x \in [x_k, x_{k+1}], \quad \forall t \in [0, \Delta t), \\ \mathbf{w}^n(x, 0) = \begin{cases} \mathbf{u}_k^n & \text{si } x < b_k, \\ \mathbf{u}_{k+1}^n & \text{si } x > b_k. \end{cases} \end{cases} \quad (P_k)$$

Para que tenga sentido concatenar estos problemas, las discontinuidades de cada uno de los problemas de Riemann no deben interactuar entre ellas. Es suficiente que, para cada  $k \in \mathbb{Z}$ , el problema  $(P_k)$  cumpla que

$$\mathbf{w}^n(x_k, t) = \mathbf{u}_k^n, \quad \mathbf{w}^n(x_{k+1}, t) = \mathbf{u}_{k+1}^n, \quad \forall t \in (0, \Delta t).$$

Observando que  $x_k - b_k = -\frac{\Delta x}{2}$  y  $x_k - a_k = \frac{\Delta x}{2}$  y utilizando que las soluciones de  $(P_k)$  son autosimilares, se deduce que comprobar las igualdades anteriores equivale a comprobar que

$$\mathbf{w}_R\left(-\frac{\Delta x}{2t}; \mathbf{u}_k^n, \mathbf{u}_{k+1}^n\right) = \mathbf{u}_k^n, \quad \mathbf{w}_R\left(\frac{\Delta x}{2t}; \mathbf{u}_k^n, \mathbf{u}_{k+1}^n\right) = \mathbf{u}_{k+1}^n, \quad \forall t \in (0, \Delta t). \quad (4.8)$$

A partir de los desarrollos de la Sección 2.4, se deduce que, en el caso escalar, para que se den las dos igualdades anteriores, debe ocurrir que

$$\frac{\Delta x}{2\Delta t} \geq \max\{|f'(\omega)|: \omega \text{ entre } u_k^n \text{ y } u_{k+1}^n\}.$$

Por lo tanto, en el caso escalar, para que tenga sentido emplear el método de Godunov, basta con tomar  $\Delta t > 0$  tal que

$$\max_{\omega \in \Omega}(|f'(\omega)|) \frac{2\Delta t}{\Delta x} \leq 1.$$

En el caso vectorial, esta condición se puede generalizar como

$$s_{\max} \frac{2\Delta t}{\Delta x} \leq 1, \quad (4.9)$$

donde  $s_{\max}$  es la mayor velocidad de propagación (en valor absoluto) de los sistemas  $(P_k)$ . Se pueden consultar más detalles de este desarrollo en [8] y [13].

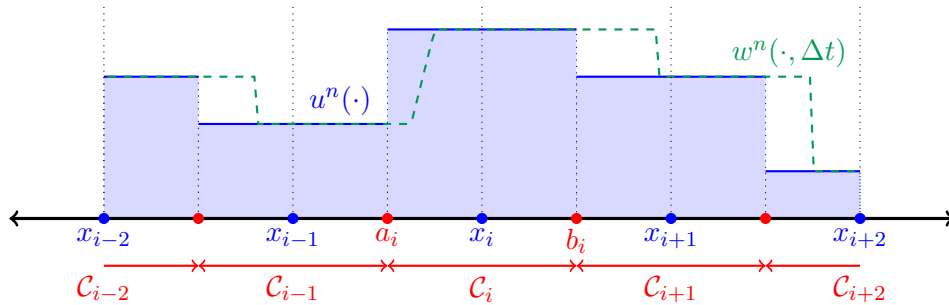


Figura 4.7: Ilustración de la idea de la solución del problema (4.6) en el caso escalar. La línea azul continua es la condición inicial  $u^n$ , la línea verde es la solución  $w^n$  para un paso de tiempo  $\Delta t$ . Las líneas discontinuas verticales representan los cambios de definición de cada problema de Riemann de (4.7).

En la Figura 4.7 se tiene un ejemplo ilustrativo de cómo son las soluciones de (4.7). Nótese que para emplear este método, es fundamental conocer de forma exacta la solución entrópica de

los problemas de Riemann asociados a la ley de conservación que se esté resolviendo. Es por esta razón por lo que se hicieron los desarrollos de la Sección 2.4.

Ahora bien, aunque este esquema pueda parecer más complejo que los anteriores, está claro que es también conservativo. En efecto, por ser  $\mathbf{w}^n$  solución de la ley de conservación, se tiene que

$$\int_{\mathcal{C}_i} \int_{t_n}^{t_{n+1}} [(\partial_t \mathbf{w}^n(x, t) + \partial_x \mathbf{f}(\mathbf{w}^n(x, t))) dt] dx = 0.$$

Utilizando la regla de Barrow, de la igualdad anterior se deduce que

$$\int_{a_i}^{b_i} [\mathbf{w}^n(x, t_{n+1}) - \mathbf{w}^n(x, t_n)] dx + \int_{t_n}^{t_{n+1}} [\mathbf{f}(\mathbf{w}^n(b_i, t)) - \mathbf{f}(\mathbf{w}^n(a_i, t))] dt = 0,$$

de donde, finalmente, a partir de (4.7) se deduce que

$$\begin{aligned} \Delta x (\mathbf{u}_i^{n+1} - \mathbf{u}_i^n) + \Delta t [\mathbf{f}(\mathbf{w}_R(0; \mathbf{u}_i^n, \mathbf{u}_{i+1}^n)) - \mathbf{f}(\mathbf{w}_R(0; \mathbf{u}_{i-1}^n, \mathbf{u}_i^n))] &= 0 \\ \implies \mathbf{u}_i^{n+1} = \mathbf{u}_i^n - \frac{\Delta t}{\Delta x} [\mathbf{f}(\mathbf{w}_R(0; \mathbf{u}_i^n, \mathbf{u}_{i+1}^n)) - \mathbf{f}(\mathbf{w}_R(0; \mathbf{u}_{i-1}^n, \mathbf{u}_i^n))], \end{aligned}$$

para todo  $i \in \mathbb{Z}$  y  $n \in \mathbb{N}$ , lo cual significa que el esquema de Godunov es conservativo, con flujo numérico dado por

$$\mathbf{g}(\omega_1, \omega_2) = \mathbf{f}(\mathbf{w}_R(0; \omega_1, \omega_2)), \quad \forall \omega_1, \omega_2 \in \Omega.$$

Por ser conservativo, y de la Proposición 3.4 se sigue que el esquema de Godunov es consistente y de orden al menos uno. Por otra parte, en el caso escalar, el esquema es también monótono, ya que la resolución exacta de problemas de Cauchy es monótona y el operador  $f \mapsto \frac{1}{\Delta x} \int_{\mathcal{C}_i} f$  es monótono, para toda función integrable. En consecuencia, se concluye que el esquema de Godunov es monótono, entrópico y de orden exactamente uno.

Finalmente, se debe hacer notar que hay pocos resultados de estabilidad del método de Godunov, especialmente en el caso vectorial. En la práctica, simplemente se toma un paso de tiempo que cumpla la condición (4.9), donde el  $s_{\max}$  se aproxima, no necesariamente de forma muy precisa.

Se han realizado los mismos ejemplos de prueba a los que se sometieron los esquemas de Lax-Friedrichs y Lax-Wendroff, obteniendo las Figuras 4.8, 4.9 y 4.10. Se puede apreciar como, a pesar de ser un método de orden uno, no es tan disipativo como el esquema de Lax-Friedrichs y, además, no presenta las oscilaciones del esquema de Lax-Wendroff. Incluso en un sistema de ecuaciones no lineales como el de la Figura 4.10, el resultado es satisfactorio.

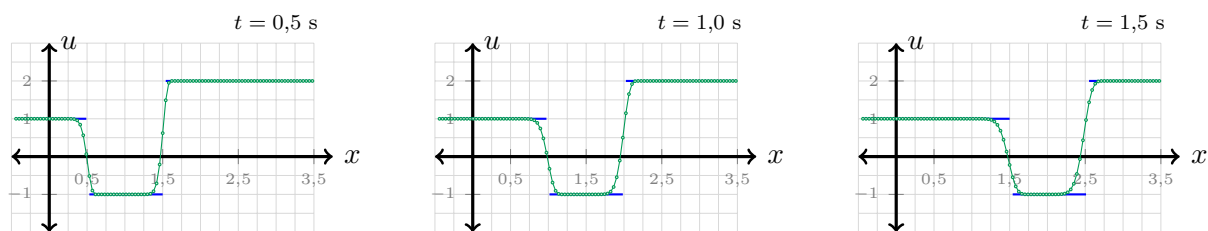


Figura 4.8: Resolución numérica de la ecuación del transporte con  $\lambda = 1$  para una condición inicial de tipo Riemann con tres estados:  $u_L = 1$ ,  $u_C = -1$  y  $u_R = 2$ . Se está resolviendo en el dominio computacional  $[a, b] = [-0,5, 3,5]$ , con  $I = 100$  y  $\Delta t = 0,035$  s. En azul, solución exacta y, en verde, aproximación mediante el esquema de Godunov.

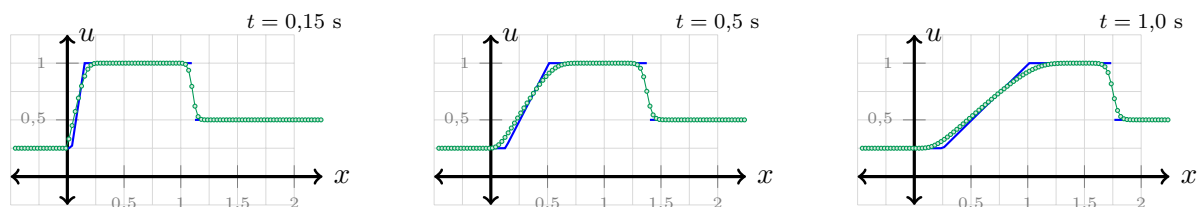


Figura 4.9: Resolución numérica de la ecuación de Burgers para una condición inicial de tipo Riemann con tres estados:  $u_L = 0,25$ ,  $u_C = 1$  y  $u_R = 0,5$ . Se está resolviendo en el dominio computacional  $[a, b] = [-0,5, 2,25]$ , con  $I = 100$  y  $\Delta t = 0,015$  s. En azul, solución exacta y, en verde, aproximación mediante el método de Godunov.

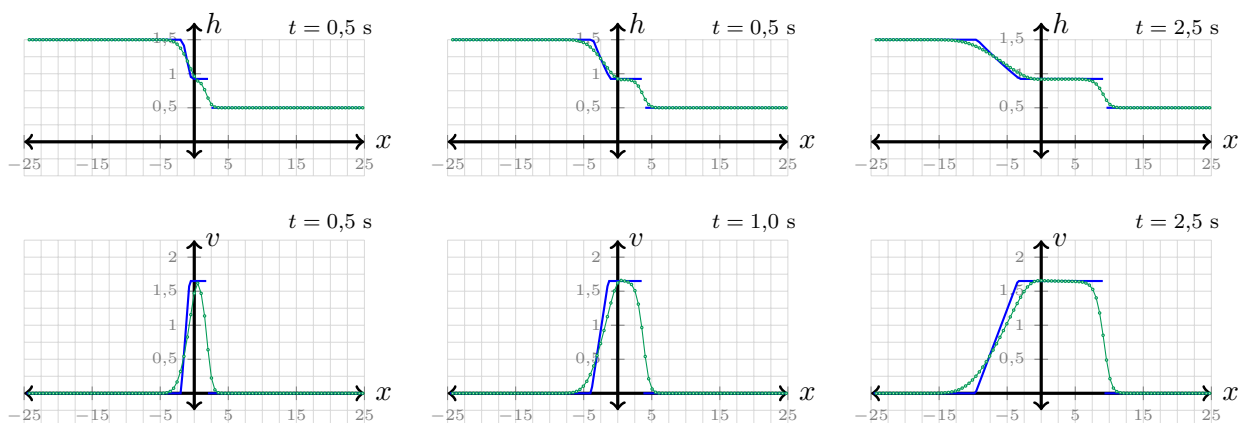


Figura 4.10: Resolución numérica de las ecuaciones de Saint Venant para el problema de rotura de una presa, con  $h_L = 1,5$  m,  $h_R = 0,5$  m,  $v_L = v_R = 0$  m/s. Se está resolviendo en el dominio computacional  $[a, b] = [-0,5, 2,25]$ , con  $I = 100$  y  $\Delta t = 0,03$  s. En azul, solución exacta y, en verde, aproximación mediante el método de Godunov.

*Observación 4.3.* Para resolver las ecuaciones de Saint-Venant, se ha implementado el método

de Godunov a partir de las soluciones exactas del problema de Riemann extraídas de [18]. Estas soluciones exactas son también la solución representada en las Figuras 4.3, 4.6 y 4.11.

#### 4.4. Resolventes aproximadas del problema Riemann

Un inconveniente del método de Godunov es que precisa calcular la solución exacta de un problema de Riemann. Aunque para ciertas ecuaciones, como la ecuación del transporte o la ecuación de Burgers, sea posible tener una expresión sencilla de dicha solución, esto, en general, no es posible. Además, en ocasiones, como para las ecuaciones de Saint-Venant, se tienen las soluciones exactas de los problemas de Riemann pero son muy costosas computacionalmente. Es por ello que una idea para simplificar este método es sustituir en la expresión del flujo numérico la solución exacta  $\mathbf{w}_R(0; \boldsymbol{\omega}_1, \boldsymbol{\omega}_2)$  por una solución aproximada del problema de Riemann con estados  $\boldsymbol{\omega}_1$  y  $\boldsymbol{\omega}_2$ .

Un ejemplo significativo de este tipo de esquemas es el llamado  $Q$ -esquema de Van Leer, donde el flujo numérico viene dado por

$$\mathbf{g}(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2) = \frac{1}{2} [\mathbf{f}(\boldsymbol{\omega}_1) + \mathbf{f}(\boldsymbol{\omega}_2)] - \frac{1}{2} \left| \mathbb{A} \left( \frac{1}{2} (\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2) \right) \right| (\boldsymbol{\omega}_2 - \boldsymbol{\omega}_1), \quad \forall \boldsymbol{\omega}_1, \boldsymbol{\omega}_2 \in \Omega, \quad (4.10)$$

donde por el valor absoluto de una matriz  $\mathbb{M}$  diagonalizable se entiende

$$|\mathbb{M}| = \mathbb{P} \text{diag}(|\lambda_k|)_{k=1}^m \mathbb{P}^{-1},$$

siendo  $\lambda_k$ , con  $k \in \{1, \dots, m\}$ , los autovalores de la matriz  $\mathbb{M}$  y  $\mathbb{P}$  la matriz de paso de la diagonalización. La matriz  $\mathbb{A}$  es diagonalizable en todo  $\Omega$  por estar suponiendo que el sistema de leyes de conservación (3.1) es estrictamente hiperbólico. No se entrará en los detalles de cómo se deduce este esquema, ni en sus propiedades de estabilidad. Estos aspectos pueden consultarse en [10].

Aunque este esquema es, aparentemente, mucho más sencillo que el método de Godunov, produce muy buenos resultados, y muy similares a dicho método, tal y como se puede ver en la Figura 4.11. Sin embargo, este esquema no es entrópico, lo que evidencian numerosos casos patológicos, del que se puede ver un ejemplo en la Figura 4.12. Esto se debe a que se produce una pérdida de la viscosidad numérica en los puntos donde la matriz jacobiana  $\mathbb{A}$  se anula. Como consecuencia de esto, aunque el esquema numérico converja, no lo hace a una solución entrópica, por lo que se debería descartar este resultado.

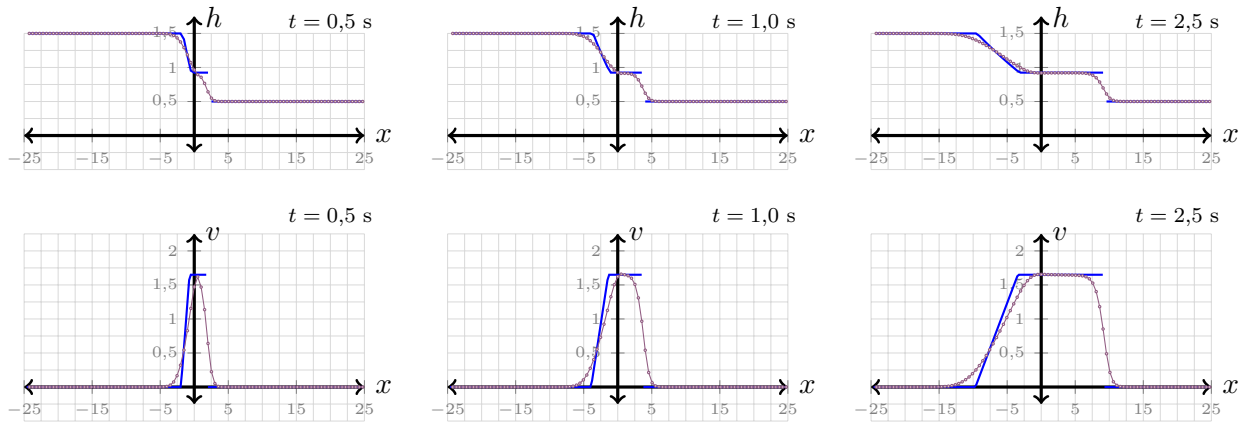


Figura 4.11: Resolución numérica de las ecuaciones de Saint Venant para el problema de rotura de una presa, con  $h_L = 1,5$  m,  $h_R = 0,5$  m,  $v_L = v_R = 0$  m/s. Se está resolviendo en el dominio computacional  $[a, b] = [-0,5, 2,25]$ , con  $I = 100$  y  $\Delta t = 0,03$  s. En azul, solución exacta y, en morado, aproximación mediante el Q-esquema de Van Leer.

Para solucionar este problema, se propone la regularización de Rusanov. Es un método de resolventes aproximadas del problema de Riemann, conservativo, con flujo numérico dado por

$$\mathbf{g}(\omega_1, \omega_2) = \frac{1}{2} [\mathbf{f}(\omega_1) + \mathbf{f}(\omega_2)] - \frac{1}{2} \alpha_{\text{RUS}}(\omega_1, \omega_2) (\omega_2 - \omega_1), \quad \forall \omega_1, \omega_2 \in \Omega, \quad (4.11)$$

donde la constante  $\alpha_{\text{RUS}}(\omega_1, \omega_2)$  viene dada por

$$\alpha_{\text{RUS}}(\omega_1, \omega_2) = \max \{ \max \{ |\lambda_k(\omega_1)| : k \in \{1, \dots, m\} \}, \max \{ |\lambda_k(\omega_2)| : k \in \{1, \dots, m\} \} \},$$

siendo  $\lambda_k$ , con  $k \in \{1, \dots, m\}$ , los autovalores de la matriz  $\mathbf{A}$ . Pueden consultarse más detalles en [20]. En la Figura 4.12 se muestra cómo el esquema con la regularización de Rusanov sí converge a la solución entrópica buscada.

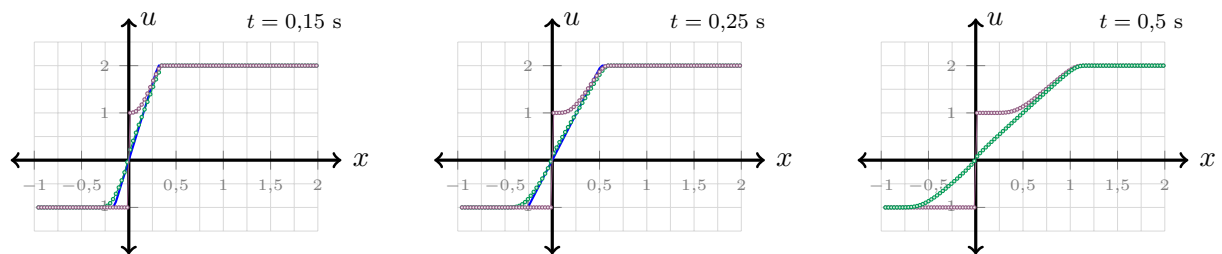


Figura 4.12: Resolución numérica de la ecuación de Burgers para una condición inicial de tipo Riemann con estados  $u_L = -1$  y  $u_R = 2$ . Se está resolviendo en el dominio computacional  $[a, b] = [-1, 2]$ , con  $I = 100$  y  $\mu = 0,85$ . En azul, solución exacta y, en morado, aproximación mediante el Q-esquema de Van Leer, y en verde, con regularización de Rusanov.

## Anexo I

# Códigos MATLAB desarrollados

### I.1. Códigos relativos a la ecuación del transporte

```
1 % transporte_ppal.m
2
3 % Condicion inicial de tipo Riemann
4 uL = 1; uR = 2;
5 u0 = @(x) uL.*(x<=0) + uR.*(x>0);
6
7 % Discretizacion espacial
8 a = -0.5; b = 3.5;
9 I = 100;
10 x = linspace(a, b, I)';
11
12 % Discretizacion de u0 en cada volumen
13 [U0, xm, dx] = disct_u0(u0,x);
14
15 % Determinamos el numero de volúmenes finitos
16 I = length(U0);
17
18 % Discretizacion temporal
19 T = 0.5; CFL = 0.95;
20 dt = CFL*min(dx); N = ceil(T/dt);
21
22 % Esquemas numericos para volúmenes finitos
23 U_LXF = lax_friedrichs(U0, dx, dt, N);
24 U_LWE = lax_wendroff(U0, dx, dt, N);
25 U_GOD = godunov(U0, dx, dt, N);
26 U_VAN = van_leer(U0, dx, dt, N);
27 U_RUS = rusanov(U0, dx, dt, N);
28
```

```

29 ylimC = [min([uL uC uR])-1, max([uL uC uR])+1];
30
31 % Graficamos los resultados finales
32 figure(2)
33 for n = 1:N
34     Ue = u0(xm - n*dt);
35     plot(xm, Ue, 'x');
36     ylim(ylimC);
37     hold on
38     title("Metodo de volúmenes finitos, t = " + n*dt)
39     plot(xm,U_LXF(1:I,n),'o');
40     plot(xm,U_LWE(1:I,n),'o');
41     plot(xm,U_GOD(1:I,n),'*');
42     plot(xm,U_VAN(1:I,n),'o');
43     plot(xm,U_RUS(1:I,n),'o');
44     legend("Solucion exacta", "Metodo de Lax-Friedrichs", "Metodo de Lax-
        Wendroff", ...
        "Metodo de Godunov", "Q-esquema de Van Leer", "Reg. de Rusanov");
45     pause(0.1)
46     hold off
47 end
48

```

```

1 function [U0, xm, dx] = disct_u0(u0,x)
2
3     % Discretizacion de la condicion inicial
4     % Entrada -----
5     % u0: condicion inicial
6     % x: discretizacion espacial
7     % Salida -----
8     % U0: promedio de u0 en cada celda
9     % xm: puntos medios de los volúmenes finitos
10    % dx: longitud de cada celda
11
12    % Damos el no. de volúmenes finitos
13    I = length(x) - 1;
14
15    % Longitud de cada volumen finito
16    dx = x(2:I+1) - x(1:I);
17
18    % Definimos los puntos medios
19    xm = 0.5*(x(2:I+1) + x(1:I));
20
21    % Definimos la estimacion en cada intervalo
22    U0 = zeros(I,1);
23    for i = 1:I
24        U0(i) = integral(u0, x(i), x(i+1))/dx(i);
25    end

```

```
26
27 end
```

```
1 function U = lax_friedrichs(U0, dx, dt, N)
2
3     % Esquema de Lax-Friedrichs para volúmenes finitos
4     % Entrada -----
5     % U0: condición inicial discretizada
6     % dx: longitud de cada celda
7     % dt: longitud del paso de tiempo
8     % N: no. de pasos de tiempo
9     % Salida -----
10    % U: matriz con los resultados numéricos,
11    % donde  $U(i,n) = U_i^n$  (aproximación de  $u(x_i, t_n)$ )
12
13    I = length(U0);
14
15    % Inicializamos
16    U = zeros(I,N);
17    U(:,1) = U0;
18
19    % Iteraciones del esquema
20    for n = 1:N-1
21        U(1,n+1) = 0.5*(U(1,n) + U(2,n)) - 0.5*dt*(U(2,n) - U(1,n))./dx(1);
22        U(2:I-1,n+1) = 0.5*(U(1:I-2,n) + U(3:I,n)) - 0.5*dt*(U(3:I,n) - U(1:I-2,
23            n))./dx(2:I-1);
24        U(I,n+1) = 0.5*(U(I-1,n) + U(I,n)) - 0.5*dt*(U(I,n) - U(I-1,n))./dx(I);
25    end
26 end
```

```
1 function U = lax_wendroff(U0, dx, dt, N)
2
3     % Esquema de Lax-Wendroff para volúmenes finitos
4     % Entrada -----
5     % U0: condición inicial discretizada
6     % dx: longitud de cada celda
7     % dt: longitud del paso de tiempo
8     % N: no. de pasos de tiempo
9     % Salida -----
10    % U: matriz con los resultados numéricos,
11    % donde  $U(i,n) = U_i^n$  (aproximación de  $u(x_i, t_n)$ )
12
13    I = length(U0);
14
15    % Inicializamos
16    U = zeros(I,N);
```

```

17     U(:,1) = U0(:);
18
19     % Iteraciones del esquema
20     for n = 1:N-1
21         Un = U(:,n);
22         U(1,n+1) = Un(1) - dt*(g(Un(1),Un(2), dx(1), dt) - g(Un(1),Un(1), dx(1),
23             dt))/dx(1);
24         for i = 2:I-1
25             U(i,n+1) = Un(i) - dt*(g(Un(i), Un(i+1), dx(i), dt) - g(Un(i-1), Un(
26                 i), dx(i), dt))/dx(i);
27         end
28         U(I,n+1) = Un(I) - dt*(g(Un(I),Un(I), dx(I), dt) - g(Un(I-1),Un(I), dx(I)
29             ), dt))/dx(I);
30     end
31 end
32
33 function Rg = g(u1,u2,dx,dt)
34     f = @(t) t;
35     Rg = 0.5*(f(u1)+f(u2)) - 0.5*(dt/dx)*( 1*(f(u2) - f(u1)));
36 end

```

```

1 function U = godunov(U0, h, dt, N)
2
3     % Esquema de Godunov para volúmenes finitos
4     % Entrada -----
5     % U0: condicion inicial discretizada
6     % dx: longitud de cada celda
7     % dt: longitud del paso de tiempo
8     % N: no. de pasos de tiempo
9     % Salida -----
10    % U: matriz con los resultados numericos,
11    % donde U(i,n) = U_i^n (aproximacion de u(x_i,t_n))
12
13    I = length(U0);
14
15    g = @(u,v) riemann_solver(u, v);
16
17    % Inicializamos
18    U = zeros(I,N);
19    U(:,1) = U0;
20
21    % Iteraciones del esquema
22    for n = 1:N-1
23        U(1,n+1) = U(1,n) - dt*(g(U(1,n),U(2,n)) - g(U(1,n),U(1,n)))/h(1);
24        for i = 2:I-1
25            U(i,n+1) = U(i,n) - dt*(g(U(i,n),U(i+1,n)) - g(U(i-1,n),U(i,n)))/h(

```

```

        i);
26     end
27     U(I,n+1) = U(I,n) - dt*(g(U(I,n),U(I,n)) - g(U(I-1,n),U(I,n)))/h(I);
28
29     end
30
31 end

```

```

1 function Re = riemann_solver(uL, uR)
2
3     s = 1;
4     Re = (0<=s).*uL + (0>s).*uR;
5
6 end

```

```

1 function U = van_leer(U0, dx, dt, N)
2
3     % Esquema de Van Leer para volúmenes finitos
4     % Entrada -----
5     % U0: condicion inicial discretizada
6     % dx: longitud de cada celda
7     % dt: longitud del paso de tiempo
8     % N: no. de pasos de tiempo
9     % Salida -----
10    % U: matriz con los resultados numericos,
11    % donde U(i,n) = U_i^n (aproximacion de u(x_i,t_n))
12
13    I = length(U0);
14
15    % Inicializamos
16    U = zeros(I,N);
17    U(:,1) = U0(:);
18
19    % Iteraciones del esquema
20    for n = 1:N-1
21
22        U(1,n+1) = U(1,n) - dt*(g(U(1,n),U(2,n)) - g(U(1,n),U(1,n)))/dx(1);
23        for i = 2:I-1
24            U(i,n+1) = U(i,n) - dt*(g(U(i,n),U(i+1,n)) - g(U(i-1,n), U(i,n)))/dx
                (i);
25        end
26        U(I,n+1) = U(I,n) - dt*(g(U(I,n), U(I,n)) - g(U(I-1,n),U(I,n)))/dx(I);
27    end
28
29 end
30
31 function Rg = g(u1,u2)

```

```

32     Rg = 0.5*(u1+u2) - 0.5*abs(1).*(u2-u1);
33 end

```

```

1  function U = rusanov(U0, dx, dt, N)
2
3     % Esquema de Van Leer (con reg. de Rusanov) para volúmenes finitos
4     % Entrada -----
5     % U0: condiciin inicial discretizada
6     % dx: longitud de cada celda
7     % dt: longitud del paso de tiempo
8     % N: no. de pasos de tiempo
9     % Salida -----
10    % U: matriz con los resultados numericos,
11    % donde U(i,n) = U_i^n (aproximacion de u(x_i,t_n))
12
13    I = length(U0);
14
15    % Inicializamos
16    U = zeros(I,N);
17    U(:,1) = U0(:);
18
19    % Iteraciones del esquema
20    for n = 1:N-1
21        U(1,n+1) = U(1,n) - dt*(g(U(1,n),U(2,n)) - g(U(1,n),U(1,n)))/dx(1);
22        for i = 2:I-1
23            U(i,n+1) = U(i,n) - dt*(g(U(i,n),U(i+1,n)) - g(U(i-1,n), U(i,n)))/dx
                (i);
24        end
25        U(I,n+1) = U(I,n) - dt*(g(U(I,n), U(I,n)) - g(U(I-1,n),U(I,n)))/dx(I);
26    end
27
28 end
29
30 function Rg = g(u1,u2)
31     Rg = 0.5*(u1+u2) - 0.5*max(abs(1),abs(1)).*(u2-u1);
32 end

```

## I.2. Códigos relativos a la ecuación de Burgers

```

1  % burgers_ppal.m
2
3  % Condicion inicial de tipo Riemann (tres estados)
4  uL = -1;
5  uC = 1;

```

```

6  uR = 0.5;
7
8  u0 = @(x) uL.*(x<=0) + uC.*(x>0 & x<=1) + uR.*(x>1);
9
10 % Discretizacion espacial
11 a = -2;
12 b = 2;
13 I = 100;
14 x = linspace(a, b, I)';
15
16 % Discretizacion de u0 en cada volumen
17 [U0, xm, dx] = disct_u0(u0,x);
18
19 % Determinamos el numero de volúmenes finitos
20 I = length(U0);
21
22 % Discretizacion temporal
23 T = 1;
24 CFL = 0.9;
25 dt = CFL*min(dx)/max(abs(U0));
26 N = ceil(T/dt);
27
28 % Esquemas numericos para volúmenes finitos
29 U_LXF = lax_friedrichs(U0, dx, dt, N);
30 U_LWE = lax_wendroff(U0, dx, dt, N);
31 U_GOD = godunov(U0, dx, dt, N);
32 U_VAN = van_leer(U0, dx, dt, N);
33 U_RUS = rusanov(U0, dx, dt, N);
34
35 ylimC = [min([uL uC uR])-1, max([uL uC uR])+1];
36
37 % Graficamos los resultados finales
38 figure(2)
39 for n = 1:N
40     Ue = u_exact3(xm, n*dt, uL, uC, uR);
41     plot(xm, Ue, 'x');
42     ylim(ylimC);
43     hold on
44     title("Metodo de volúmenes finitos, t = " + n*dt)
45     plot(xm,U_LXF(1:I,n),'o');
46     plot(xm,U_LWE(1:I,n),'o');
47     plot(xm,U_GOD(1:I,n),'o');
48     plot(xm,U_VAN(1:I,n),'o');
49     plot(xm,U_RUS(1:I,n),'o');
50     legend("Solucion exacta", "Metodo de Lax-Friedrichs", "Metodo de Lax-
        Wendroff", ...

```

```

51     "Metodo de Godunov", "Q-esquema de Van Leer", "Q-esquema de Van Leer [
        Rusanov]");
52     pause(0.1)
53     hold off
54 end

```

```

1  function Ue = u_exact2(x, t, uL, uR)
2
3     xt = x./t;
4
5     s = 0.5*(uL + uR);
6     if uL >= uR
7         Ue = uL.*(xt<=s) + uR.*(xt>s);
8     elseif uL < uR
9         Ue = uL.*(xt<=uL) + uR.*(xt>uR) + xt.*(uL < xt & xt <= uR);
10    end
11
12 end

```

```

1  function Ue = u_exact3(x, t, uL, uC, uR)
2
3     Ue = 0.*x;
4
5     if uL <= uC && uC <= uR
6         Ue = uL.*(x<=uL.*t) + (x./t).*(uL.*t<x & x<=uC.*t) + uC.*(uC.*t<x & x<=
            uC.*t + 1) + ((x-1)./t).*(uC.*t + 1 < x & x <= uR.*t + 1) + uR.*(x>
            uR.*t+1);
7     end
8
9     if uL >= uC && uC >= uR
10        te = 2/(uL-uR);
11        xi = 0.5*(uL + uR).*t + (uC - uR)./(uL - uR);
12        xi1 = 0.5*(uL + uC).*t;
13        xi2 = 0.5*(uC + uR).*t + 1;
14        Ue = uL.*(t<te & x<=xi1) + uC.*(t<te & xi1<x & x<=xi2) + uR.*(t<te & x>
            xi2) + ...
            uL.*(t>=te & x<=xi) + uR.*(t>=te & x>xi);
15    end
16
17
18    if uL < uC && uC >= uR && uL <= uR
19        te = 2/(uC-uR);
20        xi = 0.5*(uC + uR).*t + 1;
21        xit = sqrt(2*(uC - uR).*t) + uR.*t;
22        Ue = uL.*(t<te & x<=uL.*t) + (x./t).*(t<te & uL.*t<x & x<=uC.*t) + uC.*(
            t<te & x>uC.*t & x<=xi) + uR.*(t<te & x>xi) + ...
            uL.*(t>=te & x<=uL.*t) + (x./t).*(t>=te & uL.*t<x & x<=xit) + uR.*(t
            >=te & x>xit);
23

```

```

24     end
25
26     if uL <= uC && uC >= uR && uL >= uR
27         te = 2/(uC-uR);
28         tee = 2*(uC-uR)/((uL-uR)^2);
29         xi = 0.5*(uC + uR).*t + 1;
30         xit = sqrt(2*(uC - uR).*t) + uR.*t;
31         xitt = 0.5*(uL+uR).*t + (uC-uR)./(uL-uR);
32         Ue = uL.*(t<te & x<=uL.*t) + (x./t).*(t<te & uL.*t<x & x<=uC.*t) + uC.*(
33             t<te & x>uC.*t & x<=xi) + uR.*(t<te & x>xi) + ...
34             uL.*(t>=te & t < tee & x<=uL.*t) + (x./t).*(t>=te & t < tee & uL.*t<
35                 x & x<=xit) + uR.*(t>=te & t < tee & x>xit) + ...
36             uL.*(t>=tee & x<=xitt) + uR.*(t>=tee & x>xitt);
37     end
end

```

```

1  function U = lax_friedrichs(U0, dx, dt, N)
2
3      % Esquema de Lax-Friedrichs para volúmenes finitos
4      % Entrada -----
5      % U0: condición inicial discretizada
6      % dx: longitud de cada celda
7      % dt: longitud del paso de tiempo
8      % N: no. de pasos de tiempo
9      % Salida -----
10     % U: matriz con los resultados numéricos,
11     % donde U(i,n) = U_i^n (aproximación de u(x_i,t_n))
12
13     I = length(U0);
14
15     f = @(u) u.^2/2;
16
17     % Inicializamos
18     U = zeros(I,N);
19     U(:,1) = U0;
20
21     % Iteraciones del esquema
22     for n = 1:N-1
23         U(1,n+1) = 0.5*(U(1,n) + U(2,n)) - 0.5*dt*(f(U(2,n)) - f(U(1,n)))./dx(1)
24             ;
25         U(2:I-1,n+1) = 0.5*(U(1:I-2,n) + U(3:I,n)) - 0.5*dt*(f(U(3:I,n)) - f(U
26             (1:I-2,n)))./dx(2:I-1);
27         U(I,n+1) = 0.5*(U(I-1,n) + U(I,n)) - 0.5*dt*(f(U(I,n)) - f(U(I-1,n)))./
28             dx(I);
29     end
end

```

28 `end`

```

1  function U = lax_wendroff(U0, dx, dt, N)
2
3      % Esquema de Lax-Wendroff para volúmenes finitos
4      % Entrada -----
5      % U0: condición inicial discretizada
6      % dx: longitud de cada celda
7      % dt: longitud del paso de tiempo
8      % N: no. de pasos de tiempo
9      % Salida -----
10     % U: matriz con los resultados numéricos,
11     % donde  $U(i,n) = U_i^n$  (aproximación de  $u(x_i, t_n)$ )
12
13     I = length(U0);
14
15     % Inicializamos
16     U = zeros(I,N);
17     U(:,1) = U0(:);
18
19     % Iteraciones del esquema
20     for n = 1:N-1
21         Un = U(:,n);
22         U(1,n+1) = Un(1) - dt*(g(Un(1),Un(2), dx(1), dt) - g(Un(1),Un(1), dx(1),
23             dt))/dx(1);
24         for i = 2:I-1
25             U(i,n+1) = Un(i) - dt*(g(Un(i), Un(i+1), dx(i), dt) - g(Un(i-1), Un(
26                 i), dx(i), dt))/dx(i);
27         end
28         U(I,n+1) = Un(I) - dt*(g(Un(I),Un(I), dx(I), dt) - g(Un(I-1),Un(I), dx(I)
29             ), dt))/dx(I);
30     end
31 end
32
33 function Rg = g(u1,u2,dx,dt)
34     f = @(t) t.^2/2;
35     Rg = 0.5*(f(u1)+f(u2)) - 0.25*(dt/dx)*(f(u2) - f(u1))*(u1+u2);
36 end

```

```

1  function U = godunov(U0, h, dt, N)
2
3      % Esquema de Godunov para volúmenes finitos
4      % Entrada -----
5      % U0: condición inicial discretizada
6      % dx: longitud de cada celda
7      % dt: longitud del paso de tiempo

```

```

8      % N: no. de pasos de tiempo
9      % Salida -----
10     % U: matriz con los resultados numericos,
11     % donde U(i,n) = U_i^n (aproximacion de u(x_i,t_n))
12
13     I = length(U0);
14
15     g = @(u,v) 0.5.*u_exact2(0, 1, u, v).^2;
16
17     % Inicializamos
18     U = zeros(I,N);
19     U(:,1) = U0;
20
21     % Iteraciones del esquema
22     for n = 1:N-1
23         U(1,n+1) = U(1,n) - dt*(g(U(1,n),U(2,n)) - g(U(1,n),U(1,n)))./h(1);
24         for i = 2:I-1
25             U(i,n+1) = U(i,n) - dt*(g(U(i,n),U(i+1,n)) - g(U(i-1,n),U(i,n)))./h(
                i);
26         end
27         U(I,n+1) = U(I,n) - dt*(g(U(I,n),U(I,n)) - g(U(I-1,n),U(I,n)))./h(I);
28     end
29
30 end

```

```

1  function U = van_leer(U0, dx, dt, N)
2
3      % Esquema de Van Leer para volúmenes finitos
4      % Entrada -----
5      % U0: condicion inicial discretizada
6      % dx: longitud de cada celda
7      % dt: longitud del paso de tiempo
8      % N: no. de pasos de tiempo
9      % Salida -----
10     % U: matriz con los resultados numericos,
11     % donde U(i,n) = U_i^n (aproximacion de u(x_i,t_n))
12
13     I = length(U0);
14
15     % Inicializamos
16     U = zeros(I,N);
17     U(:,1) = U0(:);
18
19     % Iteraciones del esquema
20     for n = 1:N-1
21         U(1,n+1) = U(1,n) - dt*(g(U(1,n),U(2,n)) - g(U(1,n),U(1,n)))/dx(1);
22         for i = 2:I-1

```

```

23         U(i,n+1) = U(i,n) - dt*(g(U(i,n),U(i+1,n)) - g(U(i-1,n), U(i,n)))/dx
           (i);
24     end
25     U(I,n+1) = U(I,n) - dt*(g(U(I,n), U(I,n)) - g(U(I-1,n),U(I,n)))/dx(I);
26 end
27
28 end
29
30 function Rg = g(u1,u2)
31     Rg = 0.25*(u1.^2+u2.^2) - 0.25*abs(u1+u2).*(u2-u1);
32 end

```

```

1 function U = rusanov(U0, dx, dt, N)
2
3     % Esquema de Van Leer (con reg. de Rusanov) para volúmenes finitos
4     % Entrada -----
5     % U0: condicion inicial discretizada
6     % dx: longitud de cada celda
7     % dt: longitud del paso de tiempo
8     % N: no. de pasos de tiempo
9     % Salida -----
10    % U: matriz con los resultados num4ricos,
11    % donde U(i,n) = U_i^n (aproximacion de u(x_i,t_n))
12
13    I = length(U0);
14
15    % Inicializamos
16    U = zeros(I,N);
17    U(:,1) = U0(:);
18
19    % Iteraciones del esquema
20    for n = 1:N-1
21        U(1,n+1) = U(1,n) - dt*(g(U(1,n),U(2,n)) - g(U(1,n),U(1,n)))/dx(1);
22        for i = 2:I-1
23            U(i,n+1) = U(i,n) - dt*(g(U(i,n),U(i+1,n)) - g(U(i-1,n), U(i,n)))/dx
                (i);
24        end
25        U(I,n+1) = U(I,n) - dt*(g(U(I,n), U(I,n)) - g(U(I-1,n),U(I,n)))/dx(I);
26    end
27
28 end
29
30 function Rg = g(u1,u2)
31     Rg = 0.25*(u1.^2+u2.^2) - 0.5*max(abs(u1),abs(u2)).*(u2-u1);
32 end

```

## I.3. Códigos relativos a las ecuaciones de Saint-Venant

```

1  % saint_venant_ppal.m
2
3  % Condicion inicial de tipo Riemann
4  hL = 1.5;
5  hR = 0.5;
6  vL = 0;
7  vR = 0;
8
9  h0 = @(x) hL.*(x<=0) + hR.*(x>0);
10 q0 = @(x) h0(x).*(vL.*(x<=0) + vR.*(x>0));
11
12 % Discretizacion espacial
13 a = -25;
14 b = 25;
15 I = 100;
16 x = linspace(a, b, I)';
17
18 % Discretizacion de u0 en cada volumen
19 [U0, xm, dx] = disct_u0(h0, q0, x);
20
21 % Determinamos el numero de volúmenes finitos
22 I = length(U0);
23
24 % Discretizacion temporal
25 T = 2.5;
26 CFL = 0.95;
27 dt = 0.05;
28 N = ceil(T/dt);
29
30 % Esquemas numericos para volúmenes finitos
31 U_LXF = lax_friedrichs(U0, dx, dt, N);
32 U_WEN = lax_wendroff(U0, dx, dt, N);
33 U_GOD = godunov(U0, dx, dt, N);
34 U_VAN = van_leer(U0, dx, dt, N);
35 U_RUS = rusanov(U0, dx, dt, N);
36
37 % Pasamos la segunda variable de conservativa a primitiva
38 U_LXF(:, :, 2) = U_LXF(:, :, 2) ./ max(U_LXF(:, :, 1), 1e-4 + 0.*U_LXF(:, :, 2));
39 U_WEN(:, :, 2) = U_WEN(:, :, 2) ./ max(U_WEN(:, :, 1), 1e-4 + 0.*U_WEN(:, :, 2));
40 U_GOD(:, :, 2) = U_GOD(:, :, 2) ./ max(U_GOD(:, :, 1), 1e-4 + 0.*U_GOD(:, :, 2));
41 U_VAN(:, :, 2) = U_VAN(:, :, 2) ./ max(U_VAN(:, :, 1), 1e-4 + 0.*U_VAN(:, :, 2));
42 U_RUS(:, :, 2) = U_RUS(:, :, 2) ./ max(U_RUS(:, :, 1), 1e-4 + 0.*U_RUS(:, :, 2));
43
44 % Determinamos los limites de graficacion
45 ylimH = [min(min(U_LXF(:, :, 1))) - 0.25, max(max(U_LXF(:, :, 1)))];

```

```

46 ylimV = [min(min(U_VAN(:, :, 2))), max(max(U_VAN(:, :, 2)))]];
47
48 % Graficamos los resultados finales
49 for n = 1:N
50     figure(1)
51     UE = riemann_solver(xm/(n*dt), hL, hR, vL, vR);
52     plot(xm, UE(1:I, 1), '-');
53     hold on
54     plot(xm, U_LXF(1:I, n, 1), 'o-');
55     plot(xm, U_WEN(1:I, n, 1), 'o-');
56     plot(xm, U_GOD(1:I, n, 1), 'o-');
57     plot(xm, U_VAN(1:I, n, 1), 'o-');
58     plot(xm, U_RUS(1:I, n, 1), 'o-');
59     title("Ecuaciones de Saint-Venant (altura h), t = " + n*dt)
60     ylim(ylimH);
61     legend("Solucion exacta", "Metodo de Lax-Friedrichs", ...
62           "Metodo de Lax-Wendroff", "Metodo de Godunov", ...
63           "Q-esquema de Van Leer", "Q-esquema con Rusanov")
64     hold off
65     figure(2)
66     plot(xm, UE(1:I, 2) ./ UE(1:I, 1), '-');
67     hold on
68     plot(xm, U_LXF(1:I, n, 2), 'o-');
69     plot(xm, U_WEN(1:I, n, 2), 'o-');
70     plot(xm, U_GOD(1:I, n, 2), 'o-');
71     plot(xm, U_VAN(1:I, n, 2), 'o-');
72     plot(xm, U_RUS(1:I, n, 2), 'o-');
73     title("Ecuaciones de Saint-Venant (velocidad v), t = " + n*dt)
74     ylim(ylimV);
75     legend("Solucion exacta", "Metodo de Lax-Friedrichs", ...
76           "Metodo de Lax-Wendroff", "Metodo de Godunov", ...
77           "Q-esquema de Van Lee", "Q-esquema con Rusanov")
78     hold off
79     pause(0.01)
80 end

```

```

1 function [U0, xm, dx] = disct_u0(h0, q0, x)
2
3     % Discretizacion de la condicion inicial
4
5     I = length(x) - 1;
6
7     dx = x(2:I+1) - x(1:I);
8
9     xm = zeros(I, 1);
10    xm(1:I) = 0.5*(x(1:I)+x(2:I+1));
11

```

```

12     % Definimos la estimacion en cada intervalo
13     U0 = zeros(I,2);
14     for i = 1:I
15         U0(i,1) = integral(h0, x(i), x(i+1))/dx(i);
16         U0(i,2) = integral(q0, x(i), x(i+1))/dx(i);
17     end
18
19 end

```

```

1 function h1 = newton(f, h0, hL, hR)
2
3     tol = 1e-4;
4     g = 9.81;
5
6     % Definimos la derivada de f
7     gK = @(H,h) sqrt(0.5*g*((H+h)/H*h));
8     fprima = @(H,h) (H<=h)*(g/sqrt(g*h)) + (H>h)*(gK(H,h) - g*(H-h)/(4*H^2*gK(H,
9         h)));
10
11     % Algoritmo iterativo
12     for n = 1:1000
13         h1 = h0 - f(h0)/(fprima(h0,hL) + fprima(h0,hR));
14         param = 0.5*abs(h1-h0)/(h1+h0);
15         h0 = h1;
16         if param < tol
17             break;
18         end
19     end
20 end

```

```

1 function Re = riemann_solver(xt, hL, hR, uL, uR)
2
3     % Definimos la constante gravitatoria
4     g = 9.81;
5
6     aL = sqrt(g*hL); aR = sqrt(g*hR);
7
8     % Calculamos los valores estrella
9     Phi = @(H,h) (H<=h)*(2*(sqrt(g*H) - sqrt(g*h))) + (H>h)*(H-h)*sqrt(0.5*g*(h+
10         H)/(H*h));
11
12     Q = @(t) Phi(t,hL) + Phi(t,hR) + uR - uL;
13     h0 = (0.5*(aL+aR) - 0.25*(uR-uL))^2/g;
14     h = newton(Q, h0, hL, hR);
15     u = 0.5*(uL+uR) + 0.5*(Phi(h,hR) - Phi(h,hL));
16     a = sqrt(g*h);

```

```

16
17 % Clasificamos las ondas
18 if h > hL
19     q1= sqrt(0.5*h*(h+hL)/(hL^2));
20     SL = uL - aL*q1;
21     if h > hR
22         q2 = sqrt(0.5*h*(h+hR)/(hR^2));
23         SR = uR + aR*q2;
24         % Damos los valores numericos
25         Re(:,1) = 0.*xt + (xt<=SL).*hL + (xt>SL & xt <=SR).*h + hR.*(xt>SR);
26         Re(:,2) = 0.*xt + (xt<=SL).*uL + (xt>SL & xt <=SR).*u + uR.*(xt>SR);
27     elseif h <= hR
28         STR = u + a;
29         SHR = uR + aR;
30         % Damos los valores numericos
31         Re(:,1) = 0.*xt + (xt<=SL).*hL + (xt>SL & xt <=STR).*h + ...
32             (xt>STR & xt <=SHR).*(((-uR+2*aR+xt)/3).^2)/g + hR.*(xt>SHR);
33         Re(:,2) = 0.*xt + (xt<=SL).*uL + (xt>SL & xt <=STR).*u + ...
34             (xt>STR & xt <=SHR).*(uR-2*aR+2*xt)/3 + uR.*(xt>SHR);
35     end
36 elseif h <= hL
37     SHL = uL - aL;
38     STL = u - a;
39     if h > hR
40         q2 = sqrt(0.5*h*(h+hR)/(hR^2));
41         SR = uR + aR*q2;
42         % Damos los valores numericos
43         Re(:,1) = 0.*xt + (xt<=SHL).*hL + ...
44             (xt>SHL & xt <=STL).*((uL+2*aL-xt)/3).^2)/g + ...
45             (xt>STL & xt <=SR).*h + hR.*(xt>SR);
46         Re(:,2) = 0.*xt + (xt<=SHL).*uL + ...
47             (xt>SHL & xt <=STL).*(uL+2*aL+2*xt)/3 + ...
48             (xt>STL & xt <=SR).*u + uR.*(xt>SR);
49     elseif h <= hR
50         STR = u + a;
51         SHR = uR + aR;
52         Re(:,1) = 0.*xt + (xt<=SHL).*hL + ...
53             (xt>SHL & xt <=STL).*((uL+2*aL-xt)/3).^2)/g + ...
54             (xt>STL & xt <=STR).*h + ...
55             (xt>STR & xt <=SHR).*(((-uR+2*aR+xt)/3).^2)/g + ...
56             hR.*(xt>SHR);
57         Re(:,2) = 0.*xt + (xt<=SHL).*uL + ...
58             (xt>SHL & xt <=STL).*(uL+2*aL+2*xt)/3 + ...
59             (xt>STL & xt <=STR).*u + ...
60             (xt>STR & xt <=SHR).*(uR-2*aR+2*xt)/3 + ...
61             uR.*(xt>SHR);
62     end

```

```

63     end
64
65     % Cambiamos a las magnitudes conservativas
66     Re(:,2) = Re(:,2).*Re(:,1);
67
68 end

```

```

1  function U = lax_friedrichs(U0, dx, dt, N)
2
3     % Esquema de Lax-Friedrichs para volúmenes finitos
4     % Entrada -----
5     % U0: condicion inicial discretizada
6     % dx: longitud de cada celda
7     % dt: longitud del paso de tiempo
8     % N: no. de pasos de tiempo
9     % Salida -----
10    % U: matriz con los resultados numericos,
11    % donde U(i,n) = U_i^n (aproximacion de u(x_i,t_n))
12
13    I = length(U0);
14
15    f = @(w) [w(:,2); (w(:,2).^2)./(w(:,1)) + 0.5*9.81*(w(:,1).^2)];
16
17    % Inicializamos
18    U = zeros(I,N,2);
19    U(:,1,1:2) = U0(:,1:2);
20
21    % Iteraciones del esquema
22    for n = 1:N-1
23        U1 = U(:,n,1); U2 = U(:,n,2);
24        U(1,n+1,:) = 0.5*([U1(2), U2(2)]' + [U1(1), U2(1)]') - 0.5*dt*(f([U1(2),
25            U2(2)]) - f([U1(1), U2(1)]))/dx(1);
26        for i = 2:I-1
27            U(i,n+1,:) = 0.5*([U1(i+1), U2(i+1)]' + [U1(i-1), U2(i-1)]') - 0.5*
28                dt*(f([U1(i+1), U2(i+1)]) - f([U1(i-1), U2(i-1)]))/dx(i);
29        end
30        U(I,n+1,:) = 0.5*([U1(I), U2(I)]' + [U1(I-1), U2(I-1)]') - 0.5*dt*(f([U1
31            (I), U2(I)] - f([U1(I-1), U2(I-1)]))/dx(I);
32    end
33 end

```

```

1  function U = lax_wendroff(U0, dx, dt, N)
2
3     % Esquema de Lax-Wendroff para volúmenes finitos
4     % Entrada -----
5     % U0: condicion inicial discretizada

```

```

6      % dx: longitud de cada celda
7      % dt: longitud del paso de tiempo
8      % N: no. de pasos de tiempo
9      % Salida -----
10     % U: matriz con los resultados numericos,
11     % donde U(i,n) = U_i^n (aproximacion de u(x_i,t_n))
12
13     I = length(U0);
14
15     % Inicializamos
16     U = zeros(I,N,2);
17     U(:,1,1:2) = U0(:,1:2);
18
19     % Iteraciones del esquema
20     for n = 1:N-1
21         U1 = U(:,n,1); U2 = U(:,n,2);
22         U(1,n+1,:) = [U1(1); U2(1)] - dt*(g([U1(1); U2(1)], [U1(2); U2(2)], dx(1)
23             , dt) - g([U1(1); U2(1)], [U1(1); U2(1)], dx(1), dt))/dx(1);
24         for i = 2:I-1
25             U(i,n+1,:) = [U1(i); U2(i)] - dt*(g([U1(i); U2(i)], [U1(i+1); U2(i+1)
26                 ], dx(i), dt) - g([U1(i-1); U2(i-1)], [U1(i); U2(i)], dx(i), dt))
27                 /dx(i);
28         end
29         U(I,n+1,:) = [U1(I); U2(I)] - dt*(g([U1(I); U2(I)], [U1(I); U2(I)], dx(I)
30             , dt) - g([U1(I-1); U2(I-1)], [U1(I); U2(I)], dx(I), dt))/dx(I);
31     end
32
33     end
34
35     function Rg = g(u1,u2,h,dt)
36         g = 9.81;
37         f = @(t) [t(2); (t(2).^2)./(t(1)) + 0.5*g*(t(1).^2)];
38         w1 = 0.5*(u1(1) + u2(1));
39         w2 = 0.5*(u1(2) + u2(2));
40         Rg = 0.5*(f(u1)+f(u2)) - 0.5*(dt/h)*[0 1; -w2.^2./w1.^2 + g*w1 2*w2./w1]*(f(
41             u2)-f(u1));
42     end

```

```

1     function U = godunov(U0, dx, dt, N)
2
3     % Esquema de Godunov para volúmenes finitos
4     % Entrada -----
5     % U0: condicion inicial discretizada
6     % dx: longitud de cada celda
7     % dt: longitud del paso de tiempo
8     % N: no. de pasos de tiempo
9     % Salida -----

```

```

10     % U: matriz con los resultados numericos,
11     % donde U(i,n) = U_i^n (aproximacion de u(x_i,t_n))
12
13     I = length(U0);
14
15     % Inicializamos
16     U = zeros(I,N,2);
17     U(:,1,1:2) = U0(:,1:2);
18
19     % Iteraciones del esquema
20     for n = 1:N-1
21         U1 = U(:,n,1); U2 = U(:,n,2);
22         U(1,n+1,:) = [U1(1); U2(1)] - dt*(g([U1(1); U2(1)], [U1(2); U2(2)]) - g([
23             U1(1); U2(1)], [U1(1); U2(1)]))/dx(1);
24         for i = 2:I-1
25             U(i,n+1,:) = [U1(i); U2(i)] - dt*(g([U1(i); U2(i)], [U1(i+1); U2(i+1)
26                 ]) - g([U1(i-1); U2(i-1)], [U1(i); U2(i)]))/dx(i);
27         end
28         U(I,n+1,:) = [U1(I); U2(I)] - dt*(g([U1(I); U2(I)], [U1(I); U2(I)]) - g([
29             U1(I-1); U2(I-1)], [U1(I); U2(I)]))/dx(I);
30     end
31
32     function Rg = g(u1,u2)
33         g = 9.81;
34         f = @(t) [t(2); (t(2).^2)./(t(1)) + 0.5*g*(t(1).^2)];
35         hL = u1(1); hR = u2(1);
36         uL = u1(2)/hL; uR = u2(2)/hR;
37         RE = riemann_solver(0, hL, hR, uL, uR)';
38         Rg = f([RE(1), RE(2)]);
39     end

```

```

1     function U = van_leer(U0, dx, dt, N)
2
3     % Esquema de Van Leer para volúmenes finitos
4     % Entrada -----
5     % U0: condicion inicial discretizada
6     % dx: longitud de cada celda
7     % dt: longitud del paso de tiempo
8     % N: no. de pasos de tiempo
9     % Salida -----
10    % U: matriz con los resultados numericos,
11    % donde U(i,n) = U_i^n (aproximacion de u(x_i,t_n))
12
13    I = length(U0);
14

```

```

15     % Inicializamos
16     U = zeros(I,N,2);
17     U(:,1,1:2) = U0(:,1:2);
18
19     % Iteraciones del esquema
20     for n = 1:N-1
21         U1 = U(:,n,1); U2 = U(:,n,2);
22         U(1,n+1,:) = [U1(1); U2(1)] - dt*(g([U1(1); U2(1)],[U1(2); U2(2)]) - g([
23             U1(1); U2(1)],[U1(1); U2(1)]))/dx(1);
24         for i = 2:I-1
25             U(i,n+1,:) = [U1(i); U2(i)] - dt*(g([U1(i); U2(i)],[U1(i+1); U2(i+1)
26                 ]) - g([U1(i-1); U2(i-1)],[U1(i); U2(i)]))/dx(i);
27         end
28         U(I,n+1,:) = [U1(I); U2(I)] - dt*(g([U1(I); U2(I)],[U1(I); U2(I)]) - g([
29             U1(I-1); U2(I-1)],[U1(I); U2(I)]))/dx(I);
30     end
31
32 function Rg = g(u1,u2)
33     g = 9.81;
34     f = @(t) [t(2); (t(2).^2)./(t(1)) + 0.5*g*(t(1).^2)];
35     u = 0.5*(u1+u2);
36     lambda1 = u(2)./u(1) - sqrt(g*u(1));
37     lambda2 = u(2)./u(1) + sqrt(g*u(1));
38     Lambda = [abs(lambda1), 0; 0, abs(lambda2)];
39     P = [1,1; lambda1, lambda2];
40     Pin = [-lambda2, 1; lambda1, -1]./(lambda1-lambda2);
41     Q = P*abs(Lambda)*Pin;
42     Rg = 0.5*(f(u1) + f(u2)) - 0.5*Q*(u2-u1);
43 end

```

```

1 function U = rusanov(U0, dx, dt, N)
2
3     % Esquema de Rusanov para volúmenes finitos
4     % Entrada -----
5     % U0: condicion inicial discretizada
6     % dx: longitud de cada celda
7     % dt: longitud del paso de tiempo
8     % N: no. de pasos de tiempo
9     % Salida -----
10    % U: matriz con los resultados numericos,
11    % donde U(i,n) = U_i^n (aproximacion de u(x_i,t_n))
12
13    I = length(U0);
14
15    % Inicializamos

```

```

16     U = zeros(I,N,2);
17     U(:,1,1:2) = U0(:,1:2);
18
19     % Iteraciones del esquema
20     for n = 1:N-1
21         U1 = U(:,n,1); U2 = U(:,n,2);
22         U(1,n+1,:) = [U1(1); U2(1)] - dt*(g([U1(1); U2(1)], [U1(2); U2(2)]) - g([
                U1(1); U2(1)], [U1(1); U2(1)]))/dx(1);
23         for i = 2:I-1
24             U(i,n+1,:) = [U1(i); U2(i)] - dt*(g([U1(i); U2(i)], [U1(i+1); U2(i+1)
                ]) - g([U1(i-1); U2(i-1)], [U1(i); U2(i)]))/dx(i);
25         end
26         U(I,n+1,:) = [U1(I); U2(I)] - dt*(g([U1(I); U2(I)], [U1(I); U2(I)]) - g([
                U1(I-1); U2(I-1)], [U1(I); U2(I)]))/dx(I);
27     end
28
29 end
30
31 function Rg = g(u1,u2)
32     g = 9.81;
33     f = @(t) [t(2); (t(2).^2)./(t(1)) + 0.5*g*(t(1).^2)];
34     lambda1_r = max(abs(u1(2)./u1(1) - sqrt(g*u1(1))), abs(u2(2)./u2(1) - sqrt(g
        *u2(1))));
35     lambda2_r = max(abs(u1(2)./u1(1) + sqrt(g*u1(1))), abs(u2(2)./u2(1) + sqrt(g
        *u2(1))));
36     Lambda = [lambda1_r, 0; 0, lambda2_r];
37     u = 0.5*(u1+u2);
38     lambda1 = u(2)./u(1) - sqrt(g*u(1));
39     lambda2 = u(2)./u(1) + sqrt(g*u(1));
40     P = [1,1; lambda1, lambda2];
41     Pin = [-lambda2, 1; lambda1, -1]./(lambda1-lambda2);
42     Q = P*abs(Lambda)*Pin;
43     Rg = 0.5*(f(u1) + f(u2)) - 0.5*Q*(u2-u1);
44 end

```



# Bibliografía

- [1] BONNET BEN-DHIA, A.-S., FLISS, S., JOLY, P., Y MOIREAU, P. Introduction aux équations aux dérivées partielles et à leur approximation numérique. *ENSTA Paris* (2024).
- [2] BREZIS, H. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer, 2010.
- [3] CASAS RENTERIA, E. *Introducción a las ecuaciones en derivadas parciales*. Universidad de Cantabria (Servicio de publicaciones), 1992.
- [4] COLE, J. D. On a Quasi-Linear Parabolic Equation Occurring in Aerodynamics. *Quarterly of Applied Mathematics* 9, 3 (1951), 225–236.
- [5] CONWAY, J. B. *A Course in Functional Analysis*, vol. 96 of *Grad. Texts Math*. Springer, Cham, 1985.
- [6] COULOMBEL, J.-F. *Equations hyperboliques non-linéaires*. Université de Lille, 2009.
- [7] GODLEWSKI, E., Y RAVIART, P.-A. *Hyperbolic Systems of Conservation Laws*. No. 3/4 in *Mathématiques & applications*. Ellipses, Paris, 1991.
- [8] GODLEWSKI, E., Y RAVIART, P.-A. *Numerical Approximation of Hyperbolic Systems of Conservation Laws*. Springer, 1996.
- [9] GODUNOV, S. K. Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics. *Математический журнал* 47(89), 3 (1959), 271–306.
- [10] HARTEN, A., LAX, P. D., Y VAN LEER, B. On Upstream Differencing and Godunov-Type Schemes for Hyperbolic Conservation Laws. *Upwind and High-Resolution Schemes* (1997), 53–79.
- [11] HOPF, E. The Partial Differential Equation  $u_t + uu_x = \mu u_{xx}$ . *Communications on Pure and Applied Mathematics* 3 (1950), 201–230.

- 
- [12] LEFLOCH, P. G. *Hyperbolic Systems of Conservation Laws*. Birkhäuser Basel, 2002.
- [13] LEVEQUE, R. J. *Finite-Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2002.
- [14] LEVEQUE, R. J. *Numerical Methods for Conservation Laws*. Lectures in Mathematics, ETH Zürich. Springer Basel AG, 2008.
- [15] OLEINIK, O. Discontinuous Solutions of Non-Linear Differential Equations. *American Mathematical Society Translations*, 2 (1957), 95–172.
- [16] STRIKWERDA, J. C. *Finite Difference Schemes and Partial Differential Equations, Second Edition*. SIAM, 1989.
- [17] SUN, W., Y YUAN, Y.-X. *Optimization Theory and Methods*. Springer Optimization and Its Applications. Springer New York, NY, 2006.
- [18] TORO, E. F. *Shock-Capturing Methods for Free-Surface Shallow Flows*. John Wiley and Sons, Chichester & etc, 2001.
- [19] TORO, E. F. *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*. Springer, 2009.
- [20] VAZQUEZ-CENDÓN, M. E. *Solving Hyperbolic Equations with Finite Volume Methods*. Springer, 2015.
- [21] YOSIDA, K. *Functional Analysis*. Birkhäuser Basel, 1965.