MA 5928
Models and numerical methods for Eulerian and Lagrangian hyperbolic equations

Bruno Després
despres@ann.jussieu.fr
Spring 2019

These notes concern a course given at the TUM during Spring 2019. We consider systems of conservation law and systems of balance laws under the form
\[
\partial_t U + \nabla \cdot f(U) = S
\]
where the space coordinate is \(x \in \mathbb{R}^d\), the unknown is \(U(t, x) \in \mathbb{R}^n\), the flux \(f(U) \in \mathbb{R}^{n \times d}\) is a smooth function from \(\mathbb{R}^n\) into \(\mathbb{R}^{n \times d}\) and \(S(t, x) \in \mathbb{R}^n\) is a given source. More specifically we will consider connections between Eulerian and Lagrangian formulations, the structure of Lagrangian systems, and the derivation of entropy stable Finite Volume schemes.

The textbook for this course is [4]. Recommended references are [7, 15].

1 Models

1.1 Balance law

Consider the one dimensional case \(d = 1\). A quantity of interest is denoted as \(\rho(t, x) \in \mathbb{R}\). It is a function of the time variable \(t \in \mathbb{R}^+\) and of the space variable \(x \in \mathbb{R}\). The integral of \(\rho\) between two given points \(x_0 \in \mathbb{R}\) and \(x_1 \in \mathbb{R}\) is
\[
N(t, x_0, x_1) = \int_{x_0}^{x_1} \rho(t, x)dx, \quad x_0 < x_1.
\]

The variation with respect to the time writes
\[
\frac{d}{dt} N(t, x_0, x_1) = \int_{x_0}^{x_1} \partial_t \rho(t, x)dx.
\] (1)

Introducing a time dependence so that \(x_0(t)\) and \(x_1(t)\) can move, a more general formula for the variation of \(N(x_0(t), x_1(t), t)\) is
\[
\frac{d}{dt} N(t, x_0(t), x_1(t)) = \int_{x_0(t)}^{x_1(t)} \partial_t \rho(t, x)dx + x_1'(t) \rho(t, x_1(t)) - x_0'(t) \rho(t, x_0(t)).
\] (2)
We make the hypothesis that the gain and loss can only be through the endpoints of the interval \([x_0, x_1]\). Moreover we consider fixed endpoints, that is \(x'_0(t) = x'_1(t) = 0\). One writes an additional balance equation during the time interval \(\Delta t > 0\)

\[ N(t + \Delta t, x_0, x_1) = N(t, x_0, x_1) - f(t, x_1)\Delta t + f(t, x_0)\Delta t + o(\Delta t) \]

where \(f(t, x_0)\) (resp. \(f(t, x_1)\)) represents the gains or losses (it depends on the sign) at the endpoints. Pass formerly to the limit \(\Delta t \to 0^+\). One obtains

\[ \frac{d}{dt} N(t, x_0, x_1) + f(t, x_1) - f(t, x_0) = \frac{d}{dt} N(t, x_0, x_1) + \int_{x_0}^{x_1} \partial_x f(t, x)dx = 0. \quad (3) \]

Next combine (1) and (3) to obtain

\[ \int_{x_0}^{x_1} \partial_t \rho(t, x)dx + \int_{x_0}^{x_1} \partial_x f(t, x)dx = 0. \]

Since this integral identity holds for all \(x_0 < x_1\), one obtains the equivalent differential identity

\[ \partial_t \rho(t, x) + \partial_x f(t, x) = 0. \quad (4) \]

This relation is called a conservation law. Even if the formula is symmetric with respect to the time and the spaces variables, one basically considers that the variables do not play the same role. As a consequence the quantity \(\rho\) under the differential time operator will be called the unknown. The other quantity \(f\) is the flux.

The previous method is very general and can be extended immediately in any space dimension. For example one has the conservation law in dimension three \((x, y, z) \in \mathbb{R}^3\)

\[ \partial_t \rho(t, x, y, z) + \partial_x f(t, x, y, z) + \partial_y g(t, x, y, z) + \partial_z h(t, x, y, z) = 0. \]

In the above equation the number of equation is one, while the number of unknown quantities \(\rho, f, g, h\) is four. So it remains to specify the fluxes \(f, g\) and \(h\) in function of the unknown \(\rho\) to obtain a closed equation.
1.2 Traffic flow

Let us consider traffic flow for which the main unknown is the density \( \rho(t,x) \) of vehicles along an infinite road \( x \in \mathbb{R} \). The number of vehicles between \( x_0 \) and \( x_1 \) is by definition \( N(t,x_0,x_1) = \int_{x_0}^{x_1} \rho(t,x) dx \) for \( x_0 < x_1 \). One observes that \( N(t,x_0,x_1) \) is in general a real number, as usual in such models. Let us denote the velocity of the vehicles as \( u(t,x) \). What was referred to as the gain or loss is \( f = \rho u \). One obtains the conservation law

\[
\partial_t \rho + \partial_x \rho u = 0. \tag{5}
\]

It is standard to make a modeling hypothesis: we admit that a reasonable driver adapts the speed of the vehicle in function of the density of vehicles around his own vehicle. More precisely, the denser the traffic the slower the speed. On the contrary one drives fast if the local density is low. The consequence is that the velocity \( u \) is determined as a function of the density \( \rho \). One obtains the equation \( \partial_t \rho + \partial_x f(\rho) = 0 \) with a flux \( f(\rho) = \rho u(\rho) \). The so-called LWR (it stands for Lightill-Whitham-Richards [13]) corresponds to

\[
\rho \mapsto u(\rho) \equiv u_{\text{max}} \left( 1 - \frac{\rho}{\rho_{\text{max}}} \right),
\]

where the constants \( u_{\text{max}} \) and \( \rho_{\text{max}} \) can be determined by basic considerations. Typically the maximal velocity on highways is \( u_{\text{max}} = 130 \text{ km/h} \) in Europe, and the maximal density can be estimated in function of the mean length \( l \) of vehicles: \( \rho = \frac{1}{l} \). The LWR law is represented in figure 2. The conservation law for traffic flow takes the form

\[
\partial_t \rho + \partial_x f(\rho) = 0, \quad f(\rho) = \rho u(\rho) = u_{\text{max}} \left( \rho - \frac{\rho^2}{\rho_{\text{max}}} \right). \tag{6}
\]
The non dimensional version of the equation is obtained for \( u_{\text{max}} = 1 \) and \( \rho_{\text{max}} = 1 \)

\[
\partial_t \rho + \partial_x (\rho - \rho^2) = 0.
\]

Let us define the new unknown

\[
v = \frac{1}{2} (1 - \rho)
\]

It satisfies the Burgers equation

\[
\partial_t v + \frac{1}{2} \partial_x v^2 = 0.
\]

The Burgers equation is non linear. If \( v_1 \) and \( v_2 \) are two solutions of the Burgers equation, then the function \( v_3 = v_1 + v_2 \) is a priori not a solution of the Burgers equation. That is the superposition principle is not true for non linear equations.

The Burgers equation is scale invariant. Let \( v \) be a solution of the Burgers equation and let \( \lambda \in \mathbb{R} \). Then the function \( w = \lambda v \) is solution of \( \partial_s w + \frac{1}{2} \partial_s w^2 = 0 \) after rescaling of the time variable \( s = \lambda t \).

### 1.3 Shallow water

We show hereafter how to derive the shallow water equations also denoted as the Saint Venant equations, starting from minimal assumptions.

Consider the two dimensional cut of a lake or river represented in Figure 3. The two components of the velocity of the fluid are noted \( \mathbf{u} = (u, v) \). The first component is the horizontal velocity and the second is the vertical velocity.

Water is a priori considered as an incompressible fluid, that is the density is constant: \( \rho = \rho_c > 0 \). The incompressibility condition on the velocity fields

\[
\partial_x u + \partial_y v = 0.
\]

Let us note \( h(t, x) \) the height of a water column. The balance law technique yields \( N(t, x_0, x_1) = \rho_c \int_{x_0}^{x_1} h(t, x) dx \). The time variation of \( N(t, x_0, x_1) \) due to the boundaries fluxes is given by

\[
\frac{d}{dt} N(t, x_0, x_1) + \int_{x_0}^{x_1} \left( \partial_x (u h) + \partial_y (v h) \right) dx = 0.
\]

Figure 3: A column of water between \( x_0 \) and \( x_1 \).
\[ f(t, x_1) - f(t, x_0) = 0. \] The left and right fluxes are naturally
\[ f(t, x_0) = \rho_c \int_0^{h(t,x_0)} u(t, x_0, y) dy \quad \text{and} \quad f(t, x_1) = \rho_c \int_0^{h(t,x_1)} u(t, x_1, y) dy. \]

After simplification by \( \rho_c \)
\[
\frac{d}{dt} \int_{x_0}^{x_1} h(t, x) dx + \int_0^{h(t,x_1)} u(t, x_1, y) dy - \int_0^{h(t,x_0)} u(t, x_0, y) dy = 0.
\]

It is convenient to define the mean horizontal velocity of the column of water as
\[ \bar{u}(t, x) = \frac{\int_0^{h(t,x)} u(t, x, y) dy}{h(t,x)}, \]
so that one can write
\[
\frac{d}{dt} \int_{x_0}^{x_1} h(t, x) dx + h(t, x_1) \bar{u}(t, x_1) - h(t, x_0) \bar{u}(t, x_0) = 0.
\]

Since it holds for all \( x_0 < x_1 \), it yields a first conservation law
\[ \partial_t h + \partial_x (h \bar{u}) = 0. \quad (7) \]

This conservation law is very similar to the traffic flow equation. The difference is that the mean horizontal velocity has no reason to be a function of the height \( h \). It means that we must derive one equation, or more, in order to obtain a closed system.

It is known that this additional second equation exists. It is also a conservation law where the main unknown is \( \bar{u} \). To construct it one can use the formula (2) for the time variation of a mass between moving boundaries under the form
\[ N(t, x_0(t), x_1(t)) = \rho_c \int_{x_0(t)}^{x_1(t)} h(t, x) dx, \]
where the left boundary is defined by \( x_0(0) = X_0 \) and \( x'_0(t) = \bar{u}(t, x_0(t)). \) Similarly the right boundary is defined by \( x_1(0) = X_1 \) and \( x'_1(t) = \bar{u}(t, x_1(t)). \) Therefore
\[
\frac{d}{dt} N(t, x_0(t), x_1(t))
= \rho_c \left( \int_{x_0}^{x_1} \partial_t h(t, x) dx + x'_1(t) h(t, x_1(t)) - x'_0(t) h(t, x_0(t)) \right)
= \rho_c \int_{x_0}^{x_1} \left( \partial_t h(t, x) + \bar{u}(t, x_1(t)) h(t, x_1(t)) - \bar{u}(t, x_0(t)) h(t, x_0(t)) \right) dx
= \rho_c \int_{x_0}^{x_1} \left( \partial_t h(t, x) + \partial_x (h(t, x) \bar{u}(t, x)) \right) dx = 0,
\]
which shows that the mass is constant in the moving interval. It allows the classical analogy where the water column is like an individual particle with constant mass \( N(t, x_0(t), x_1(t)) \) for which the Newton’s law applies. The horizontal inertial momentum of the water column is
\[ I(t, x_0(t), x_1(t)) = \rho_c \int_{x_0(t)}^{x_1(t)} h \bar{u} dx. \]
The total pressure integrated on the vertical line is
\[ p = \frac{g}{2} \rho_c h^2 \]

Figure 4: Sketch of the left force \( F_0 \) and right force \( F_1 \) on the boundaries of the water column.

The sum of forces on boundaries is the right hand side of the Newton's law
\[ \frac{d}{dt} I(t, x_0(t), x_1(t)) = F(t, x_1(t)) + F(t, x_0(t)). \] (8)

The force can be interpreted as the integral on the vertical line of the hydrostatic pressure. Taking attention to the sign, it yields
\[ F(t, x_0) = \int_0^{h(t,x_0)} p(t, x_0, y)dy \] and \[ F(t, x_1) = -\int_0^{h(t,x_1)} p(t, x_1, y)dy. \]

The hydrostatic pressure at altitude \( y \) is proportional to the height of water above \( y \), so
\[ p(t, x, y) = \rho_c \int_y^{h(t,x)} g dy = \rho_c g(h(t, x) - y) \]
and \( g \) is the local gravity constant. So the forces are
\[ F(t, x_0) = \frac{g}{2} \rho_c h^2(t, x_0), \quad F(t, x_1) = -\frac{g}{2} \rho_c h^2(t, x_1). \]

Plug in (8), and obtain
\[ \frac{d}{dt} \int_{x_0(t)}^{x_1(t)} h u dx + \frac{1}{\rho_c} \int_{x_0(t)}^{x_1(t)} \partial_x F dx = \frac{d}{dt} \int_{x_0(t)}^{x_1(t)} h u dx + \frac{g}{2} \int_{x_0(t)}^{x_1(t)} \partial_x h^2 dx = 0. \]

The formula (2) and the definitions of the velocities \( x'_0(t) \) and \( x'_1(t) \) show that
\[ \frac{d}{dt} \int_{x_0(t)}^{x_1(t)} h u dx = \int_{x_0(t)}^{x_1(t)} \partial_t (h u) dx + (h u^2)(t, x_1(t)) - (h u^2)(t, x_0(t)) \]
\[ \int_{x_0(t)}^{x_1(t)} \left[ \partial_t (h \Pi) + \partial_x (h \Pi^2) \right] dx. \]

So
\[ \int_{x_0(t)}^{x_1(t)} \partial_t (h \Pi) dx + \int_{x_0(t)}^{x_1(t)} \partial_x \left( h \Pi^2 + \frac{g}{2} \xi^2 \right) dx = 0. \]

Since \( x_0(t) \) and \( x_1(t) \) are arbitrary, it yields a second conservation law
\[ \partial_t (h \Pi) + \partial_x \left( h \Pi^2 + \frac{g}{2} \xi^2 \right) = 0. \]  

(9)

Finally we obtain the shallow water system defined by equations (7) and (9)
\[
\begin{cases}
\partial_t h + \partial_x (h \Pi) = 0; \\
\partial_t (h \Pi) + \partial_x \left( h \Pi^2 + \frac{g}{2} \xi^2 \right) = 0, \quad g > 0.
\end{cases}
\]  

(10)

This is a closed system of two equations with two unknowns.

### 1.4 Compressible gas dynamics

One can construct the system of compressible non viscous gas dynamics with the same method as for the traffic flow and shallow water equations. However an additional hypothesis of thermodynamical nature is needed: it relates the macroscopic coefficient of the pressure law to its microscopic features, using thermodynamical considerations. The reader interested by the physical foundations of thermodynamics will find much more in classical physical textbooks such as [1, 2] or in many mathematical textbooks at the intersection of mathematics and continuum mechanics, see [3, 16].

The thermodynamic assumption is that the pressure of a gas can be written as a function of two independent parameters which are the density \( \rho \) and the temperature \( T \), that is
\[ p = p(\rho, T). \]

We admit that the temperature can be computed with respect to the density and another thermodynamical quantity referred to as the internal energy \( \varepsilon \). Denote the velocity of the gas as \( u \). The total energy per unit mass is the sum of the internal (or potential energy) and of the kinetic energy \( e = \varepsilon + \frac{1}{2} |u|^2 \). A standard pressure law\(^1\) is the one of a perfect polytropic gas, with \( \gamma = 1.4 \) for air,
\[ p = (\gamma - 1) \rho \varepsilon, \quad \varepsilon = C_v T, \quad C_v > 0, \quad \gamma > 1. \]  

(13)

\(^1\)Many other pressure law exist. For example the stiffened gas pressure law writes
\[ p = (\gamma - 1) \rho \varepsilon - \gamma \Pi. \]  

(11)

Water is not a gas of course, however it can be modeled quite accurately taking \( \gamma = 5, 5 \) and \( \Pi = 4921, 15 \) bars. Another example is the van der Walls pressure law
\[ p = \frac{a \varepsilon}{\tau - b} - \frac{c}{\tau^2}, \quad a, b, c > 0, \quad \tau = \frac{1}{\rho}, \quad \gamma = \frac{1}{\rho}. \]  

(12)

where the variable \( \tau \) denotes the specific volume. The van der Walls pressure law is used for phase transitions.
Whatever is the pressure law, the equations can be constructed with the balance law method. For simplicity we consider the one dimensional configuration depicted in figure 5. An elementary (infinitesimal) quantity of gas is contained in the interval \([x_0(t), x_1(t)]\). The points can move, that is \(x'(t, X) = u(t, x(t, X))\), \(x(0, X) = X\).

\[
x_0(t + \Delta t) \quad x_1(t + \Delta t)
\]

\[
t + \Delta t
\]

\[
x_0(t) \quad x_1(t)
\]

\[
t
\]

Figure 5: Elementary quantity of gas in a moving interval

At any time, the mass in the interval is \(N(x_0(t), x_1(t), t) = \int_{x_0(t)}^{x_1(t)} \rho(t, x)dx\).

The total impulse is \(I(x_0(t), x_1(t), t) = \int_{x_0(t)}^{x_1(t)} \rho(t, x)u(t, x)dx\). The mechanical forces on the edges of the moving interval are \(f = p\) on \(x_0(t)\) and \(f = -p\) on \(x_1(t)\). Using again the method described before, we obtain two equations

\[
\begin{align*}
\partial_t \rho + \partial_x (\rho u) &= 0, \\
\partial_t (\rho u) + \partial_x (\rho u^2 + p) &= 0.
\end{align*}
\]

Nevertheless this system is not closed. In order to construct an additional equation, we consider the total energy in the moving interval

\[
E(x_0(t), x_1(t), t) = \int_{x_0(t)}^{x_1(t)} \rho(t, x)e(t, x)dx
\]

where the total energy per unit mass is the sum of the internal energy and the kinetic energy \(e = \varepsilon + \frac{1}{2}u^2\). Energy considerations can be used to determine the work exerted by the forces. Consider a small time interval \(\Delta t\). By definition the work \(w\) is the product of the force time the length of the interval on which it is exerted. That is \(w = \pm pu\Delta t\). One obtains

\[
E(x_0(t + \Delta t), x_1(t + \Delta t), t + \Delta t) = E(x_0(t), x_1(t), t) - \Delta t p(t, x_1(t))u(t, x_1(t)) + \Delta t p(t, x_0(t))u(t, x_0(t)) + o(\Delta t).
\]

Passing formally to the limit for \(\Delta t \rightarrow 0\), one gets

\[
\frac{d}{dt}E(x_0(t), x_1(t), t) + p(t, x_1(t))u(t, x_1(t)) - p(t, x_0(t))u(t, x_0(t)) = 0.
\]
Combine with formula (2) and get \( f_{\epsilon(t)}^{x(t)} \partial_t (pe) dx + f_{x(t)}^{\epsilon(t)} \partial_x (pue + pu) = 0 \). Since it is true for any pair \((x_0(t), x_1(t))\) it yields a new conservation law

\[
\partial_t (pe) + \partial_x (pue + pu) = 0.
\]

One finally obtains the system of compressible non viscous gas dynamics, also denoted as the system of Euler equations

\[
\begin{align*}
\partial_t \rho + \partial_x (pu) &= 0, \\
\partial_t (pu) + \partial_x (p(u^2 + p) + pv) &= 0, \\
\partial_t (pv) + \partial_x (p(uv) + p) &= 0, \\
\partial_t (pe) + \partial_x (pue + pu) &= 0.
\end{align*}
\]

(14)

This system is closed since the pressure \( p \) can be calculated as a function of \( \rho \) and \( \varepsilon = e - \frac{1}{2} u^2 \).

In two dimensions, one obtains by tensorization

\[
\begin{align*}
\partial_t \rho + \partial_x (pu) + \partial_y (pv) + \partial_z (pw) &= 0, \\
\partial_t (pu) + \partial_x (p(u^2 + p) + pv) + \partial_y (p) &= 0, \\
\partial_t (pv) + \partial_x (p(uv) + p) + \partial_y (p) &= 0, \\
\partial_t (pe) + \partial_x (pue + pu) + \partial_y (pue + pv) + \partial_z (pue + pv) &= 0.
\end{align*}
\]

(15)

The difference is mostly in the velocity field which is a vector \( u = (u, v) \) with two components. The pressure \( p \) is function of the density \( \rho \) and internal energy \( \varepsilon = e - \frac{1}{2} (u^2 + v^2) \). In dimension three, one obtains readily

\[
\begin{align*}
\partial_t \rho + \partial_x (pu) + \partial_y (pv) + \partial_z (pw) &= 0, \\
\partial_t (pu) + \partial_x (p(u^2 + p) + pv) + \partial_y (p(uw) + pv) &= 0, \\
\partial_t (pv) + \partial_x (p(uv) + p) + \partial_y (p(uw) + pv) &= 0, \\
\partial_t (pe) + \partial_x (pue + pu) + \partial_y (pue + pv) + \partial_z (pue + pv) &= 0,
\end{align*}
\]

(16)

where the vectorial velocity is \( u = (u, v, w) \) and the pressure \( p \) is function of the density \( \rho \) and internal energy \( \varepsilon = e - \frac{1}{2} (u^2 + v^2 + w^2) \).

### 1.5 Canonical form of a system of conservation laws

Denote the space coordinate \( x = (x_1, \ldots, x_d) \) in \( \mathbb{R}^d \). All the previous systems of conservation laws can be written under the canonical form

\[
\partial_t U(x, t) + \nabla \cdot f(U(x, t)) = 0
\]

where \( U : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^n \) is the unknown. The flux \( f : \mathbb{R}^n \to \mathbb{R}^{n \times d} \) is matrix valued, that is \( f(U) = (f_{ij}(U))_{1 \leq i \leq n, 1 \leq j \leq d} \). One can also particularize the columns of \( f \): \( f(U) = (f_1(U), \ldots, f_d(U)) \) where \( f_j(U) \in \mathbb{R}^n \) for \( 1 \leq j \leq d \). The divergence of \( f \) is

\[
\nabla \cdot f(U) = \sum_{j=1}^d \partial_x f_j(U) = \begin{pmatrix}
\sum_{j=1}^d \partial_{x_j} f_1(U) \\
\vdots \\
\sum_{j=1}^d \partial_{x_j} f_n(U)
\end{pmatrix} \in \mathbb{R}^n.
\]
For example the system (15) corresponds to \( d = 2 \) and \( n = 4 \)
\[
U = \begin{pmatrix}
\rho \\
\rho u \\
\rho v \\
\rho e
\end{pmatrix} \in \mathbb{R}^4 \quad \text{and} \quad f(U) = \begin{pmatrix}
\rho u \\
\rho u^2 + p \\
\rho uv \\
\rho e + pu
\end{pmatrix} \in \mathbb{R}^{4 \times 2}.
\]

1.6 Lagrange coordinates

The traffic flow model, the shallow water equations and the Euler system are written in what is called Eulerian coordinates. Eulerian coordinates correspond to the coordinates of a fixed observer.

On the contrary Lagrange coordinates are, in some sense, attached to the local flow velocity. Local means that the change of coordinates is different from point to another because the velocity takes different values in different parts of the fluid. Therefore the Lagrangian coordinates can be identified to the Eulerian coordinates at another time which is usually and arbitrarily the initial time, that is \( t_{\text{ini}} = 0 \). In dimension \( d = 1 \) it takes the form
\[
x(t = 0, X) = X \quad \text{and} \quad \partial_t x(t, X) = u(t, X), t).
\]

Our goal is to use this transformation to rewrite the original Eulerian equations in Lagrangian coordinates.

The algebra attached to the Euler-to-Lagrange transformation is not completely evident as stressed in [6, 7, 15]. We will use the so-called Piola identities to detail this transformation. The following presentation is close to the one from [3]. The Piola identities are also called the geometric conservation laws.

1.7 General change of coordinates in balance laws

Consider a transformation
\[
x \mapsto \varphi(x) = X = (X_1, \ldots, X_d) \in \mathbb{R}^d,
\]
from \( \mathbb{R}^d \) to \( \mathbb{R}^d \). This transformation is smooth with continuous derivatives, typically is \( C^1 \). The Jacobian matrix of the transformation is
\[
\nabla \varphi = \left( \frac{\partial X_i}{\partial x_j} \right)_{1 \leq i, j \leq d} = \begin{pmatrix}
\frac{\partial X_1}{\partial x_1} & \frac{\partial X_1}{\partial x_2} & \cdots & \frac{\partial X_1}{\partial x_d} \\
\frac{\partial X_2}{\partial x_1} & \frac{\partial X_2}{\partial x_2} & \cdots & \frac{\partial X_2}{\partial x_d} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial X_d}{\partial x_1} & \frac{\partial X_d}{\partial x_2} & \cdots & \frac{\partial X_d}{\partial x_d}
\end{pmatrix}.
\]

Assuming this matrix is non singular, \( \det(\nabla \varphi) \neq 0 \), the transformation is locally invertible. The inverse transformation is
\[
X \mapsto \psi(X) = x = (x_1, \ldots, x_d) \in \mathbb{R}^d,
\]
so that 
\[ \psi(\varphi(x)) = x \text{ and } \varphi(\psi(X)) = X. \]

Differentiation shows the Jacobian matrices are inverse one to the other
\[ \nabla \psi(\varphi(x)) \nabla \varphi(x) = I_d \text{ and } \nabla \varphi(\psi(X)) \nabla \psi(X) = I_d. \]

This relation is often written on a more compact form: \( \nabla \psi = (\nabla \varphi)^{-1} \). To go further we need the comatrix.

**Definition 1 (Comatrix).** The comatrix \( \text{com}(M) \in \mathbb{R}^{d \times d} \) of a matrix \( M \in \mathbb{R}^{d \times d} \) is matrix of the cofactors. The coefficient in position \((i, j)\) in \( \text{com}(M) \) is equal to \((-1)^{i+j} \) times the determinant of the \( d-1 \times d-1 \) sub-matrix obtained from after elimination of column \( j \) and line \( i \).

Therefore \( M^t \text{com}(M) = \det(M) I_d \) for all \( M \in \mathbb{R}^{d \times d} \). If \( M \) is non singular, it yields \( \text{com}(M) = \det(M) \times (M^t)^{-1} \). The transposition of the comatrix yields the adjugate of \( M \): \( \text{adj}(M) = \text{com}(M)^t \). The main result about change of coordinates in systems of balance laws can be now formulated with the comatrix of the inverse transformation \( \text{com}(\nabla \psi) \) and with the determinant of the Jacobian of the transformation
\[ J = \det(\nabla \psi). \]

**Theorem 1.** The system of balance laws with a source \( S : \mathbb{R}^d \to \mathbb{R}^n \)
\[ [\nabla \cdot f(U)](x) = S(x), \quad x \in \mathbb{R}^d, \tag{18} \]
is equivalent to the system of balance laws in the new frame \( X = \psi^{-1}(x) \in \mathbb{R}^d \)
\[ [\nabla \cdot (f(U(\psi))\text{com}(\nabla \psi))] (X) = [JS(\psi)](X), \quad X \in \mathbb{R}^d. \tag{19} \]

Moreover one has the Piola identity
\[ [\nabla \cdot \text{com}(\nabla \psi)](X) = 0, \quad X \in \mathbb{R}^d. \tag{20} \]

**Remark 1.** The divergence is taken with respect to \( x \) in (18) and with respect to \( X \) in (19). Note that the dimension of the product of matrices is correct, that is \( f(U(\psi))\text{com}(\nabla \psi) \in \mathbb{R}^{n \times d} \).

**Proof.** Let \( \Omega \subset \mathbb{R}^d \) be any smooth open subset of \( \mathbb{R}^d \). Its frontier is \( \partial \Omega \). The outgoing exterior normal vector define on the boundary is denoted as \( n \in \mathbb{R}^d \). The boundary measure is \( d\sigma \). One has the Stokes formula
\[ \int_{x \in \Omega} \nabla \cdot f(U(x))dx = \int_{x \in \partial \Omega} f(U(x))n d\sigma. \]

After integration \( \Omega \) the system of balance laws (18) is formally equivalent to
\[ \int_{x \in \partial \Omega} f(U(x))n d\sigma = \int_{x \in \Omega} S(x)dx, \quad \forall \Omega \subset \mathbb{R}^d \]
which is now written for all smooth open subset $\Omega \subset \mathbb{R}^d$. Perform the change of coordinates $x = \psi(X)$ and use the Nanson formula (21) for points $x \in \partial \Omega$ and $X = \varphi(x) \in \partial \varphi(\Omega)$. One gets

$$\int_{X \in \partial \varphi(\Omega)} f(U(\psi(X))) \text{com}(\nabla \psi)(X)n_X d\sigma_X = \int_{X \in \varphi(\Omega)} S(\psi(X))J(X)dX.$$  

A change of notation $\theta = \varphi(\Omega) \subset \mathbb{R}^d$ shows

$$\int_{X \in \partial \theta} f(U(\psi(X))) \text{com}(\nabla \psi)(X)n_X d\sigma_X = \int_{X \in \theta} S(\psi(X))J(X)dX, \quad \theta \subset \mathbb{R}^d.$$  

Rewrite the left hand side with the Stokes formula

$$\int_{X \in \theta} \nabla \cdot [f(U(\psi(X))) \text{com}(\nabla \psi)(X)] dX = \int_{X \in \theta} S(\psi(X))J(X)dX, \quad \theta \subset \mathbb{R}^d.$$  

Since this identity holds for any smooth $\theta$, it yields the first part of the claim (19). The Piola identity (20) is obtained taking $f = I_d \in \mathbb{R}^{d \times d}$ and vanishing source $S = 0$ in (19). The proof is ended.

Therefore the proof of the theorem relies essentially on the Nanson’s formula (21) which is of pure geometrical nature since it expresses the effect on differential elements of a general transformation of the space coordinates.

**Proposition 1** (Nanson’s formula). One has the identity between differential elements

$$n d\sigma = \text{com}(\nabla \psi)n_X d\sigma_X.$$  

(21)

**Proof.** The proof relies on two well know formulas. The first formula is a multidimensional generalization of (2). Let $V(t) \subset \mathbb{R}^n$ a moving smooth bounded domain in dimension $d$. Assume that the points on the boundary $x \in \partial V$ move with a velocity $u(x)$. One has

$$\frac{d}{dt} \int_{x \in V(t)} h(x, t)dx = \int_{x \in V(t)} \partial_t h(x, t)dx + \int_{x \in \partial V} h(x, t)(u(x), n(x))d\sigma.$$  

(22)

The second formula is the change of coordinates in integrals

$$|V| = \int_{x \in V} dx = \int_{x \in \psi(V)} \det(\nabla \psi)dX.$$  

(23)

Hence the variation of the volume of a moving domain computed in the reference frame $x$ with formula (22) (take $h = 1$) is

$$\frac{d}{dt} |V(t)| = \int_{x \in \partial V(t)} (u(x), n(x))d\sigma.$$  

It can also be computed in the reference frame $X$

$$\frac{d}{dt} |V(t)| = \frac{d}{dt} \int_{x \in \psi(V(t))} \det(\nabla \psi)dX = \int_{x \in \partial \psi(V(t))} \det(\nabla \psi)(u_X(X), n_X)d\sigma_X.$$  

12
In this expression \( u_X(X) \) is the velocity (measured in frame \( X \)) of points \( X \in \partial \varphi(V(t)) \). The chain rule yields
\[
\begin{align*}
u_X(\varphi(x)) &= \nabla \varphi(x) u(x), \quad x \in \partial V(t).
\end{align*}
\]
So
\[
\begin{align*}
\int_{x \in \partial V(t)} (u(x), n(x)) d\sigma &= \int_{X \in \partial \varphi(V(t))} \det(\nabla \psi)(X) (u(x), (\nabla \varphi(X)^t n_X(X))) d\sigma_X
\end{align*}
\]
The velocity field is arbitrary. Therefore
\[
\begin{align*}
n(x) d\sigma &= \det(\nabla \psi)(X) (\nabla \psi(X)^t n_X(X)) d\sigma_X
\end{align*}
\]
where \( x \in \partial V(t) \) and \( X = \varphi(x) \in \partial \varphi(V(t)) \). By definition of the comatrix
\[
\begin{align*}
\det(\nabla \psi)(X) \nabla \varphi(X)^t = \det(\nabla \psi)(X) \nabla \psi(X)^t = \text{com}(\nabla \psi).
\end{align*}
\]
The proof is ended.

1.8 Lagrangian gas dynamics in dimension \( d = 1 \)

Let us rewrite the system (14) as a time-space divergence
\[
\nabla_{tx} \cdot (U, f(U)) = 0 \in \mathbb{R}^n.
\]
Since \( U = (\rho, \rho u, \rho e)^t \) and \( f(U) = (\rho u, \rho u^2 + p, \rho u e + \rho u)^t \), one gets
\[
\nabla_{tx} \cdot \begin{pmatrix}
\rho & \rho u \\
\rho u & \rho u^2 + p \\
\rho e & \rho u e + \rho u
\end{pmatrix} = 0.
\]
Consider the Euler-to-Lagrange change of space-time coordinate from \((t, x)\) to \((t', X)\)
\[
t' = t \quad \text{and} \quad \frac{\partial x(t', X)}{\partial t'} = u(t', x(t', X)) \text{ with } x(0, X) = X.
\]
This transformation is regular if \( u \) is a smooth enough function. We apply the transformation (19) to the space-time free-divergence equation. It means that one must determine the comatrix of the transformation. One has
\[
\nabla_{(v, X)}(t, x) = \begin{pmatrix}
1 & 0 \\
u & J
\end{pmatrix}, \quad J = \frac{\partial x}{\partial X}.
\]
So \( \text{com} (\nabla_{(v, X)}(t, x)) = \begin{pmatrix}
J & -u \\
0 & 1
\end{pmatrix} \). Therefore equation (19) writes
\[
\nabla_{v, X} \left[(U, f(U)) \begin{pmatrix} J & -u \\
0 & 1 \end{pmatrix} \right] = 0.
\]
that is
\[
\begin{align*}
\partial_t (\rho J) &= 0, \\
\partial_t (\rho u J) + \partial_X p &= 0, \\
\partial_t (\rho e J) + \partial_X (\rho u) &= 0.
\end{align*}
\]

This system is not closed since the determinant \( J \) of the Jacobian matrix shows up. It shows the necessity of the Piola transform which is used to close the system. Using (20) the Piola identity is just one equation \( \partial_t J - \partial_X u = 0 \), which is also obtained by differentiation of \( \partial_t x(t', X) = u \) with respect to \( X \). We can now replace the notation \( t' \) with the usual notation \( t \) for time. It yields a system of 4 conservation laws
\[
\begin{align*}
\partial_t (\rho J) &= 0, \\
\partial_t (\rho u J) + \partial_X p &= 0, \\
\partial_t (\rho e J) + \partial_X (\rho u) &= 0, \\
\partial_t J - \partial_X u &= 0.
\end{align*}
\]

(24)

It is convenient to define the mass variable.

**Theorem 2** (Lagrangian formulation with the mass variable in 1D). Consider the Euler system (24) for compressible gas dynamics. Assume the density is positive \( \rho > 0 \). Define the specific volume \( \tau = \frac{1}{\rho} \) and the mass variable \( dm = \rho(0, X) dX \). Then the Euler system can be written in Lagrange coordinates under the form
\[
\begin{align*}
\partial_t \tau - \partial_m u &= 0, \\
\partial_t u + \partial_m p &= 0, \\
\partial_t e + \partial_m (\rho u) &= 0.
\end{align*}
\]

(25)

**Proof.** The mass variable is just one more change of coordinate \( m(X) = \int_0^X \rho(0, y) dy \), which can also be recast as \( dm = \rho_0 dX \). The first equation in (24) yields \( (\rho J)(t, X) = J(0, X) \rho(0, X) \). By definition \( J(0, X) = \frac{\partial X}{\partial X} = 1 \). So \( (\rho J)(t, X) = \rho(0, X) \) which does not depend on the time \( t \). One obtains for example
\[
\partial_t J - \partial_X u = 0 \Rightarrow \rho_0 \partial_t \tau - \partial_X u = 0 \Rightarrow \partial_t \tau - \frac{1}{\rho_0} \partial_X u = 0 \Rightarrow \partial_t \tau - \partial_m u = 0.
\]

The algebra is similar for the two last equations. One has
\[
\partial_t (\rho u J) + \partial_X p = 0 \Rightarrow \rho_0 \partial_t u + \partial_X p = 0 \Rightarrow \partial_t u + \frac{1}{\rho_0} \partial_X p = 0 \Rightarrow \partial_t u + \partial_m p = 0
\]

and
\[
\partial_t (\rho e J) + \partial_X (\rho u) = 0 \Rightarrow \rho_0 \partial_t e + \partial_X (\rho u) = 0
\]

\[
\Rightarrow \partial_t e + \frac{1}{\rho_0} \partial_X (\rho u) = 0 \Rightarrow \partial_t e + \partial_m (\rho u) = 0.
\]

The proof is ended. \( \square \)

**Proposition 2.** The differential operators \( \partial_t \) and \( \partial_m \) in the Lagrangian system (25) are Galilean invariant.
Remark 2. The understanding of the property is eased using a notation standard in mechanical sciences. We write $\partial_{a|b}$ the partial derivation with respect to the variable $a$ with respect to a frozen variable $b$. It clarifies the notations so that one makes a clear distinction between $\partial_{t|x}$ where the frozen variable is $x$ and $\partial_{t|m}$ where the frozen variable is $X$ or $m$. So the proposition 25 means more exactly: $\partial_{t|m}$ and $\partial_{m|t}$ are Galilean invariant.

Proof. One has

$$\begin{align*}
\partial_{t|m} &= \partial_{t|x} + u\partial_{x|t} = \partial_{t'|x'} + (u + v)\partial_{x'|t'} = \partial_{t'|x'} + u'\partial_{x'|t'} = \partial_{t'|m'}, \\
\partial_{m|t} &= \frac{1}{\rho}\partial_{x|t} = \frac{1}{\rho}\partial_{x'|t'} = \partial_{m'|t'}.
\end{align*}$$

So $\partial_t = \partial_{t'}$ and $\partial_m = \partial_{m'}$. The proof is ended.

Other notations for the time derivative with respect to the mass variable are

$$\partial_{t|m} = \partial_t + u\partial_x = \frac{d}{dt} = D_t.$$  

1.9 Lagrangian gas dynamics in dimension $d = 2$

Start from (15). Define

$$A = \partial_x X, \quad B = \partial_y y, \quad L = \partial_y x, \quad M = \partial_y y.$$  

(26)

The Jacobian matrix of the space-time transformation is

$$\frac{\partial (t, x, y)}{\partial (t, X, Y)} = \begin{pmatrix} 1 & 0 & 0 \\ u & A & L \\ v & B & M \end{pmatrix}$$

so

$$\text{com} \left( \frac{\partial (t, x, y)}{\partial (t, X, Y)} \right) = \begin{pmatrix} J & -uM + vL & uB - vA \\ 0 & M & -B \\ 0 & -L & A \end{pmatrix}, \quad J = AM - BL.$$  

The matrix-matrix product in (19) writes

$$\rho \begin{pmatrix} \rho J & \rho u & \rho v \\ \rho u & \rho u^2 + p & \rho u^2 + p \\ \rho v & \rho v^2 + p & \rho v^2 + p \\ \rho e & \rho u e + pu & \rho u e + pv \end{pmatrix} \begin{pmatrix} J & -uM + vL & uB - vA \\ 0 & M & -B \\ 0 & -L & A \end{pmatrix} = \begin{pmatrix} \rho J & 0 & 0 \\ \rho u J & pM & -pB \\ \rho v J & -pL & pA \\ \rho e J & puM - pvL & -puB + pvA \end{pmatrix}.$$  

15
One obtains

\[
\begin{align*}
\partial_t (\rho J) &= 0, \\
\partial_t (\rho J u) + \partial_X (pM) + \partial_Y (-pB) &= 0, \\
\partial_t (\rho J v) + \partial_X (-pL) + \partial_Y (pA) &= 0, \\
\partial_t (\rho J e) + \partial_X (puM - pvL) + \partial_Y (pvA - puB) &= 0.
\end{align*}
\]

(27)

The Piola identities write

\[
\begin{align*}
\partial_t J - \partial_X (uM - vL) - \partial_Y (vA - uB) &= 0, \\
\partial_X M - \partial_Y B &= 0, \\
-\partial_X L + \partial_Y A &= 0.
\end{align*}
\]

(28)

The two last compatibility relations are evident. It is not the case for the first one. We add the definition of the Lagrange-Euler transformation

\[
\begin{align*}
\partial_t x(t, X, Y) &= u, & x(0, X, Y) &= X, \\
\partial_t y(t, X, Y) &= v, & y(0, X, Y) &= Y.
\end{align*}
\]

(29)

1.10 Hui’s Formulation

The Hui formulation [8, 9, 10, 11] of Lagrangian gas dynamics is slightly different from (27-29) and has been studied in [14] and [5]. Its objective is to provide a closed system of conservation laws convenient for numerical discretization. Indeed one can notice that \( A, B, L \) and \( M \) are the gradient (26) of the transformation. But combining (26) and (29) one gets without difficulty the evolution equations

\[
\begin{align*}
\partial_t A &= \partial_X u, & \partial_t B &= \partial_X v, & \partial_t L &= \partial_Y u & \partial_t M &= \partial_Y v.
\end{align*}
\]

(30)

This system is closed. It will be used to discuss the weak hyperbolicity of Lagrangian gas dynamics.

1.11 Lagrangian gas dynamics in dimension \( d = 3 \)

The system of Lagrangian gas dynamics in dimension \( d = 3 \) shows increasing complexity, essentially because the comatrix becomes quadratic with respect to gradient of the deformation. It makes the interpretation of the equations less evident and it is probably the reason why it is delicate to use them directly for the design of numerical methods.
The deformation gradient is
\[
\mathcal{J} = \begin{pmatrix}
\frac{\partial X}{\partial x}, & \frac{\partial Y}{\partial x}, & \frac{\partial Z}{\partial x} \\
\frac{\partial X}{\partial y}, & \frac{\partial Y}{\partial y}, & \frac{\partial Z}{\partial y} \\
\frac{\partial X}{\partial z}, & \frac{\partial Y}{\partial z}, & \frac{\partial Z}{\partial z}
\end{pmatrix} = \begin{pmatrix}
A & L & P \\
B & M & Q \\
C & N & R
\end{pmatrix}.
\]
Define \( J = \det(\mathcal{J}) \) and
\[
\text{com}(\mathcal{J}) = \begin{pmatrix}
MR - NQ & -BR + CQ & BN - CM \\
-LR + NP & AR - CP & -AN + CL \\
LQ - MP & -AQ + BP & AM - BM
\end{pmatrix}.
\]
With compact notations, the gradient of the space-time transformation is
\[
\frac{\partial (t, x, y, z)}{\partial (t, X, Y, Z)} = \begin{pmatrix} 1 \\ 0 \\ u \end{pmatrix}, \quad u = \begin{pmatrix} u \\ v \\ w \end{pmatrix}.
\]
So
\[
\text{com} \left( \frac{\partial (t, x, y, z)}{\partial (t, X, Y, Z)} \right) = \begin{pmatrix} J \\ 0 \\ -u' \text{com}(\mathcal{J}) \end{pmatrix}.
\]
The calculation of (19) reduces to
\[
\begin{pmatrix}
\rho \\ \rho u \\ \rho u^t
\end{pmatrix} \begin{pmatrix}
\rho u \\ \rho u \otimes u + p I_d \\ \rho e u^t + p u^t
\end{pmatrix} \begin{pmatrix}
J \\ 0 \\ -u' \text{com}(\mathcal{J})
\end{pmatrix} = \begin{pmatrix}
\rho J \\ p u J \\ p u' \text{com}(\mathcal{J})
\end{pmatrix}.
\]
Therefore the Lagrangian system of gas dynamics in dimension \( d = 3 \) is made of 5 physical conservation laws
\[
\nabla_t \cdot \left( \begin{pmatrix}
\rho J \\ \rho u J \\ p \text{com}(\mathcal{J})
\end{pmatrix} \frac{0}{\rho e J} \right) = 0,
\]
with four scalar Piola identities
\[
\nabla_t \cdot \left( \begin{pmatrix} J \\ 0 \\ -u' \text{com}(\mathcal{J}) \end{pmatrix} \text{com}(\mathcal{J}) \right) = 0
\]
and with the definition of the transformation Lagrange-Euler
\[
\partial_t x(t, X) = u, \quad x(0, X) = X.
\]

2 Linear stability and hyperbolicity

Linear stability is a fundamental notion for dynamical systems. The idea is to add a small perturbation to a small constant initial data. After linearization of
the equations, the study of the time evolution of the small perturbation yields information about the linear stability (or instability) of the system.

Consider a system of conservation laws
\[ \partial_t U + \partial_x f(U) = 0, \quad U, f(U) \in \mathbb{R}^n. \] (37)
Assume the flux is differentiable and define the Jacobian matrix of the flux
\[ A(U_0) = \nabla_U f(U)(U_0) \in \mathbb{R}^{n \times n}, \quad U_0 \in \mathbb{R}^n. \] (38)
Let \( U_\varepsilon \) be a certain solution obtained from perturbation around the state \( U_0 \)
\[ U_\varepsilon(t, x) = U_0 + \varepsilon V(t, x) + o(\varepsilon). \] (39)
The expansion of all terms of the equation \( \partial_t U_\varepsilon + \partial_x f(U_\varepsilon) = 0 \) in ascending powers of \( \varepsilon \) yields
\[ (\partial_t U_0 + \partial_x f(U_0)) + \varepsilon (\partial_t V + \partial_x (A(U_0)V)) + o(\varepsilon) = 0. \]
Neglecting high order terms, the perturbation \( V \) is solution of the linear equation
\[ \partial_t V(t, x) + A\partial_x V(t, x) = 0, \quad V(t, x) \in \mathbb{R}^n, \quad A = A(U_0) = \nabla_U f(U_0) \in \mathbb{R}^{n \times n}. \] (40)
The study of the linear stability is based on the study of bounded solutions of this equation. The standard notion of hyperbolicity is as follows.

**Definition 2 (Hyperbolic non linear system of conservation laws).** A non linear system of conservation laws (37) will be said strongly hyperbolic in the domain \( \Omega \subset \mathbb{R}^n \) if and only if the companion linear system (40) is strongly linearly stable for all \( U_0 \in \Omega \).

It remains of course to define what is strong linear stability, this is performed below. We will see that, in the general case, Lagrangian gas dynamics is not strongly hyperbolic, but only weakly hyperbolic.

### 2.1 Classification in dimension \( d = 1 \)
The most general method to establish the stability property of the linear system is based on Fourier-Laplace modes
\[ V(t, x) = e^{i(kx - \omega t)} W, \quad W \in \mathbb{R}^n \text{ or } \mathbb{C}^n, \quad k \in \mathbb{R}. \]
Plugging such representation in (40), one obtains that \( W \) is solution of an eigenproblem
\[ -i\omega e^{i(kx - \omega t)} W + ikA(U_0)e^{i(kx - \omega t)} W = 0, \] that is
\[ A(U_0)W = \lambda W, \quad W \in \mathbb{R}^n \text{ or } \mathbb{C}^n, \quad \lambda = \frac{\omega}{k} \in \mathbb{C}. \] (41)
It is clear that the location of the eigenvalues $\lambda \in \mathbb{C}$ in the complex plane matters. Indeed let us assume for example that there exists an eigenvalue $\lambda \notin \mathbb{R}$. So $\lambda$ has a non zero imaginary part: this has an immediate consequence on the behavior of $V(t, x) = e^{ik(x-\lambda t)}W$ with respect to the time variable. If $\text{Im}(\lambda) > 0$, then, by increasing $k >> 1$, $\text{Re}(-ik\lambda) = k\text{Re}(-i\lambda) > 0$ can be made arbitrarily large. If $\text{Im}(\lambda) < 0$, $\text{Re}(-ik\lambda) > 0$ can also be made arbitrarily large by decreasing $k << -1$. Moreover the conjugate $\overline{\lambda}$ of a non real eigenvalue $\lambda \notin \mathbb{R}$ of a real matrix is also an eigenvalue. It yields additional exponentially growing solutions. In all these cases, one gets an exponential growth in time with an arbitrary large exponential growth factor: this behavior characterizes ill-posed problems.

**Definition 3 (Strong linear instability).** The linear system (40) is said to be strongly unstable at $U_0$ if and only if there exists non real eigenvalues $\lambda \notin \mathbb{R}$ to the eigenproblem (41).

**Definition 4 (Linear stability).** The linear system (40) is said to be stable at $U_0$ if and only if all eigenvalues $\lambda$ of the eigenproblem (41) are real. The eigenvalues are interpreted as the velocity of the Fourier modes.

The interpretation stems from the expression $V(t, x) = e^{ik(x-\lambda t)}W$ which indeed show that $\lambda$ has the dimension of a velocity. Next we detail the structure of the eigenvectors.

**Definition 5 (Linear stability, strong or weak).** Assume the linear system (40) is stable. It is said strongly stable if the eigenvectors span $\mathbb{R}^n$. If the eigenvectors do not span $\mathbb{R}^n$, the linear system is weakly stable.

These notions are justified with the Jordan representation of a square matrix $A \in \mathbb{R}^{n \times n}$

$$A = P(D + T)P^{-1}$$

where $P = P^*$ is a complex unitary change of basis matrix (it is an orthogonal change of basis in case it is real valued), $D$ is a diagonal matrix which diagonal coefficients are all the eigenvalues of $A$ with multiplicity, and $T$ is an upper triangular matrix with vanishing diagonal such that $DT = TD$. Since $T$ is upper triangular with vanishing diagonal, it is also a nilpotent matrix which order is the smallest $1 \leq r \leq n$ such that $T^r = 0$. A Fourier representation of a function $V(t, x) = e^{ikx}Z(t)$ solution of (40) yields the equation

$$Z'(t) + ikAZ(t) = 0.$$

The general solution is easily written with the matrix exponential formula

$$Z(t) = e^{-iktA}Z_0 = e^{-iktP(D+T)P^{-1}}Z_0 = Pe^{-ikt(D+T)}P^{-1}Z_0.$$

Since $DT = TD$, one gets $e^{-ikt(D+T)} = e^{-iktD}e^{-iktT} = e^{-iktD} \sum_{p=0}^{r-1} \frac{(-ikt)^p}{p!}$. Therefore one gets the general representation formula

$$V(t, x) = \sum_{p=0}^{r-1} \left(e^{ikx}Pe^{-iktD}T_p P^{-1}Z_0\right) \frac{(-ik)^p}{p!} t^p. \quad (43)$$
Since \( r \) is the nilpotent order of \( T \), there exists a vector \( Z_0 \in \mathbb{R}^n \) such that \( T^p P^{-1} Z_0 \neq 0 \), \( p = r - 1 \). Since the matrix \( D = \text{diag}(d_j)_{1 \leq j \leq n} \) is real diagonal, the matrix \( e^{-iktD} \) is also diagonal with coefficients

\[
(e^{-iktD})_{jj} = e^{-ikt d_j}, \quad d_j \in \mathbb{R}.
\]

At inspection of (43) the quadratic norm (in \( \mathbb{C}^n \)) of \( V(t, x) \) is of order \( t^{r-1} \) for \( k \neq 0 \)

\[
\|V(t, x)\| = O(t^{r-1}). \tag{44}
\]

Moreover \( r-1 \) is the maximal integer with such growth. So we can now interpret the definition 5.

- If \( r = 1 \), meaning \( T = 0 \) and \( A \) can be made diagonal with a change of basis matrix, then the norm of Fourier solutions is uniformly bounded in time. This is the notion of strong linear stability.
- Assume \( 1 < r \leq n \). There exists Fourier solutions with polynomial growth in time. Moreover the growth order in time \( t^{r-1} \) corresponds to a multiplication by \( k^{r-1} \). Since such a multiplication in Fourier corresponds to \( r-1 \) derivations with respect to \( x \), we observe that weak linear stability also implies loss of \( r-1 \) derivatives. More precisely one has the asymptotic behavior for large \( t \)

\[
\|V(t, x)\| \approx Ct^{(r-1)} \| \partial_x^{r-1} V(0, x) \|, \quad C > 0. \tag{45}
\]

**Definition 6** (Order of weak linear stability). The order of weak linear stability is the integer \( q = r - 1 \in \mathbb{N} \), where \( r \) is the nilpotent order of \( T \) in the Jordan representation (42) of the matrix \( A = A(U_0) \).

**Definition 7** (Well prepared data). A well prepared data for a weakly stable linear system (40) is any Fourier mode \( W(x) = e^{ikx} Z_0 \) with \( TP^{-1} Z_0 = 0 \).

Of course the definition makes sense only if \( T \neq 0 \), or, more precisely, if \( T = 0 \) all datas are well prepared. For well prepared datas, the general representation formula (43) is rewritten as

\[
V(t, x) = e^{ikx} P e^{-iktD} P^{-1} Z_0 \quad \text{for} \quad V(0, x) = W(x) \tag{46}
\]

which turns into the uniform in time estimate

\[
\|V(t, x)\| = O(1), \quad TP^{-1} V(0, x) = 0. \tag{47}
\]

Moreover we get from (46) that

\[
TP^{-1} V(t, x) = TP^{-1} (e^{ikx} P e^{-iktD} P^{-1} Z_0)
= T e^{-iktD} P^{-1} Z_0 = e^{-iktD} TP^{-1} Z_0 = 0,
\]

which implies that the norm of Fourier solutions is uniformly bounded in time. This is the notion of strong linear stability.

**Definition 7** (Well prepared data). A well prepared data for a weakly stable linear system (40) is any Fourier mode \( W(x) = e^{ikx} Z_0 \) with \( TP^{-1} Z_0 = 0 \).

Of course the definition makes sense only if \( T \neq 0 \), or, more precisely, if \( T = 0 \) all datas are well prepared. For well prepared datas, the general representation formula (43) is rewritten as

\[
V(t, x) = e^{ikx} P e^{-iktD} P^{-1} Z_0 \quad \text{for} \quad V(0, x) = W(x) \tag{46}
\]

which turns into the uniform in time estimate

\[
\|V(t, x)\| = O(1), \quad TP^{-1} V(0, x) = 0. \tag{47}
\]

Moreover we get from (46) that

\[
TP^{-1} V(t, x) = TP^{-1} (e^{ikx} P e^{-iktD} P^{-1} Z_0)
= T e^{-iktD} P^{-1} Z_0 = e^{-iktD} TP^{-1} Z_0 = 0,
\]

which implies that the norm of Fourier solutions is uniformly bounded in time. This is the notion of strong linear stability.
\[ \partial_t \begin{pmatrix} u \\ v \end{pmatrix} + \partial_x \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = 0 \quad \text{unstable} \]

\[ \partial_t \begin{pmatrix} u \\ v \end{pmatrix} + \partial_x \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = 0 \quad \text{strongly stable, that is hyperbolic} \]

\[ \partial_t \begin{pmatrix} u \\ v \end{pmatrix} + \partial_x \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = 0 \quad \text{weakly stable, with order 1} \]

\[ \partial_t \begin{pmatrix} u \\ v \end{pmatrix} + \partial_x \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = 0 \quad \text{well prepared data satisfy } \partial_x u = 0 \]

Table 1: Linear systems.

which means that if \( V(0,x) \) is a well prepared data at initial time, it remains a well prepared data at any time.

The technicalities attached to the discussion of the eigenvector structure of a given matrix \( A \) might be avoided using the strict hyperbolicity criterion which goes back to [12].

**Definition 8.** A matrix \( A = A(U_0) \) is said strictly hyperbolic if and only if all its eigenvalues are real and distinct.

Indeed \( n \) distinct real eigenvalues yield \( n \) linearly independent real eigenvectors, that is strong linear stability.

**Remark 3.** However we will show later that multiple eigenvalues is a general situation for many systems of conservation laws that come from continuum mechanics. It rules out the possibility to use the strict hyperbolicity for general purposes.

### 2.2 A useful property

The previous notions give a framework which permits the linear stability analysis of a general system of conservation laws. But it remains to compute the eigenstructure of the Jacobian matrix of the flux. In practice it can be quite tricky. A useful property is that the eigenstructure is not changed if one performs a change of unknowns. So, for a given system, it suggests to look for astute changes of variables.

Let us start with the system of conservation laws

\[ \partial_t U + \partial_x f(U) = 0. \tag{48} \]
We introduce a change of unknown under the form \( W = \phi(U) \in \mathbb{R}^n \) where \( \phi : \mathbb{R}^n \to \mathbb{R}^n \) is a differentiable transformation of the phase space. The transformation is non singular that is \( \det(\nabla \phi) \neq 0 \). Firstly we rewrite in a quasilinear form

\[
\partial_t U + A(U) \partial_x U = 0, \quad A = \nabla f(U).
\]

Secondly we perform the change of unknowns and get another quasilinear form

\[
\partial_t W + Q^{-1}A(U)Q \partial_x W = 0, \quad A = \nabla f(U) \in \mathbb{R}^{n \times n}, \; Q = \nabla \phi(U) \in \mathbb{R}^{n \times n}.
\]

**Proposition 3.** The linear stability/instability (strong or weak) properties of the initial system (48) are the same as linear instability/stability (strong or weak) properties of the equation (49) written with the unknown \( W \).

**Proof.** The matrix \( A(U) \) and \( Q^{-1}A(U)Q \) are similar, so have exactly the same eigenstructure. The same eigenvalues are equal with multiplicities. The eigenvectors coincide. Most of all the number of missing eigenvectors is the same, together with the order of weak stability. \( \square \)

In practice one can try to determine a convenient variable \( W \) inspired by the underlying physics so that the structure of \( Q^{-1}A(U)Q \) is simpler. This method is useful to avoid the brute force calculation of the characteristic polynomial of \( A \), which yields the eigenvalues. It is also quite useful for calculations of eigenvectors.

### 2.3 Generalization in dimension \( d \geq 2 \)

Consider \( \partial_t U + \partial_x f(U) + \partial_y g(U) = 0 \) where both fluxes \( U \mapsto f(U) \) and \( U \mapsto g(U) \) are differentiable. Since the problem is additive with respect to the directions, the corresponding linearized system at a given state \( U_0 \) writes

\[
\partial_t V + A \partial_x V + B \partial_y V = 0, \quad A = \nabla f(U_0) \text{ and } B = \nabla g(U_0).
\]

(50)

Rotate the coordinates by an arbitrary angle \( \theta \in \mathbb{R} \)

\[
\begin{cases} 
  x' = \cos \theta x + \sin \theta y, \\
  y' = -\sin \theta x + \cos \theta y.
\end{cases}
\]

\( \iff \)

\[
\begin{cases} 
  x = \cos \theta x' - \sin \theta y', \\
  y = \sin \theta x' + \cos \theta y'.
\end{cases}
\]

Studying the solutions \( V \) which are independent with respect to the direction \( y' \) is equivalent to add the constraint that \( \partial_{y'} = 0 \) that is using the chain rule

\[
\frac{\partial y'}{\partial x} \partial_x + \frac{\partial y'}{\partial y} \partial_y = -\sin \theta \partial_x + \cos \theta \partial_y = 0.
\]

One also has that

\[
\partial_x = \frac{\partial x'}{\partial x} \partial_{x'} + \frac{\partial y'}{\partial x} \partial_{y'}, \quad \partial_y = \frac{\partial x'}{\partial y} \partial_{x'} + \frac{\partial y'}{\partial y} \partial_{y'} = \sin \theta \partial_{x'}.
\]

In this case the equation (50) is written in one dimensional form

\[
\partial_t V + (\cos \theta A + \sin \theta B) \partial_{x'} V = 0.
\]
Define
\[
A(\theta) = \cos \theta A + \sin \theta B, \quad \forall \theta \in \mathbb{R}.
\]
We will say that the problem is hyperbolic at \( U_0 \) if and only if the matrix \( A(\theta) \) is hyperbolic for all real \( \theta \). The generalization in any dimension is as follows.

**Definition 9.** Consider the non linear system of conservation laws \( \partial_t U + \sum_{i=0}^d \partial_x f_i(U) = 0 \) where the fluxes are differentiable. This problem is hyperbolic (at \( U_0 \)) if and only if the matrix
\[
A(\alpha) = \sum_{i=0}^d \alpha_i \nabla f_i(U_0) \in \mathbb{R}^{n \times n}
\]
is hyperbolic for all \( \alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d \).

All previous definitions of strong and weak stability generalize as well, together with the notion of ill posedness.

### 2.4 Examples

We consider the different equations and systems that have been constructed, and show that they correspond to linearly well posed systems. We also determine the eigenvalues of the Jacobian matrix. These eigenvalues correspond to the velocity of small perturbations.

#### 2.4.1 Traffic flow

Consider the equation (6) for traffic flow. Linearize all quantities around a given density \( \rho_0 \) : that is \( \rho_\varepsilon(t,x) = \rho_0 + \varepsilon \mu(t,x) + o(\varepsilon) \). The linear equation for the perturbation \( \mu \) writes
\[
\partial_t \mu + a \partial_x \mu = 0, \quad a = u_{\text{max}} \left( 1 - 2 \frac{\rho_0}{\rho_{\text{max}}} \right).
\]
The explicit solution is \( \mu(t,x) = \mu(x - at) \). Hence we identify the velocity of perturbations: it is \( \lambda = a \).

Let \( \rho_c \) be the critical density defined by \( \rho_c = \frac{\rho_{\text{max}}}{2} \): we note that if \( \rho_0 < \rho_c \) then the velocity is positive \( a > 0 \). On the contrary if \( \rho_0 > \rho_c \) then the velocity is negative \( a < 0 \). This introduces a strong difference between the velocity of vehicles which is always non negative, and the velocity of small linear perturbations which can take both signs. This is illustrated in 6.

#### 2.4.2 Shallow water

**Lemma 1.** The flux of the shallow water system (10) is differentiable provided \( h \neq 0 \). If \( h > 0 \), the model is strictly hyperbolic. If \( h < 0 \), the model is linearly unstable.
Remark 4. A negative height $h < 0$ does not make sense on physical grounds. So it is satisfactory to attach instability to such non-physical data.

Proof. Set $a = h$ and $b = hu$. So the flux of the shallow water system (10) writes $f(a, b) = \left(\frac{b^2}{a} + \frac{g}{2a^2}\right)$. One has

$$A = \begin{pmatrix} 0 & 1 \\ -\frac{b^2}{a^2} + ga & \frac{b}{2} \end{pmatrix}, \quad \text{tr}(A) = \frac{2b}{a} = 2u, \quad \det(A) = \frac{b^2}{a^2} - ga = u^2 - gh.$$ 

The eigenvalue equation is $\lambda^2 - \text{tr}(A)\lambda + \det(A) = 0$. Therefore the two eigenvalues are

$$\lambda = \frac{2u \pm \sqrt{(2u)^2 - 4(u^2 - gh)}}{2} = u \pm c,$$

where $c$ is identified as the velocity of small perturbation in the local fluid frame.

- If $h > 0$ the eigenvalues are different which yields the strict hyperbolicity.
- If $h < 0$ then $c \in i\mathbb{R}^*$. In this case the eigenvalues are pure imaginary and complex conjugate, so the linear system is unstable. The proof is ended.

Moreover the flux is not differentiable for $h = 0$. Even if it does not exactly correspond to the previous definitions, one can nevertheless study $A_0 = \lim_{h \to 0^+} A$ which admits the eigenvalue $u$ with multiplicity two. But $A_0 \neq uI_d$. Therefore $A_0$ is not diagonalizable. So $A_0$ is weakly hyperbolic.

A numerical application is as follows. One considers that the mean height in oceans is approximately 4000 m. It can be used to calculate the velocity of tsunamis with $c = \sqrt{gh} \approx \sqrt{9.81 \times 4000} \approx 200\text{ms}^{-1} = 720\text{km}\text{h}^{-1}$. This is a reasonable value.
Lemma 2. Consider the Euler system (14) with the polytropic pressure law (13) and a positive density \( \rho > 0 \).

Then: a) the flux is differentiable; b) if \( \varepsilon > 0 \), the model is strictly hyperbolic with eigenvalues

\[
\lambda_1 = u - c, \quad \lambda_2 = u, \quad \lambda_3 = u + c, \quad c = \sqrt{\gamma(\gamma - 1)\varepsilon},
\]

where \( c \) is the speed of sound. The model is linearly unstable for \( \varepsilon < 0 \).

Proof. Set \( a = \rho \), \( b = \rho u \) and \( c = \rho e \). The flux writes

\[
f(a, b, c) = \begin{pmatrix}
\frac{\nu^2}{a} + (\gamma - 1) \left( c - \frac{\nu^2}{2a} \right) = \frac{3-\gamma}{2} \frac{b^2}{a} + (\gamma - 1)c \\
\frac{bc}{a} + (\gamma - 1) \left( \frac{bc}{a} - \frac{b^3}{2a^2} \right) = \frac{bc}{a} - \frac{\gamma(b^2)}{2a^2}
\end{pmatrix}.
\]

The Jacobian matrix of the flux is

\[
A = \begin{pmatrix}
0 & \frac{3-\gamma}{2} \frac{b^2}{a^2} & 1 \\
-\frac{bc}{a^2} + (\gamma - 1) \frac{b^3}{a^2} & (3 - \gamma) \frac{b}{a} & (\gamma - 1) \\
\frac{bc}{a} - \frac{\gamma(b^2)}{2a^2} & \frac{bc}{a} - \frac{\gamma(b^2)}{2a^2}
\end{pmatrix}.
\]

The three invariants of \( A \) are

\[
\text{tr}(A) = 3 \frac{b}{a} = 3u, \quad \Delta_2(A) = \frac{\gamma^2 - \gamma + 6 \frac{b^2}{a^2} - \gamma(\gamma - 1) \frac{c}{a}}{2} = 3u^2 - \gamma(\gamma - 1)\varepsilon u
\]

and the determinant

\[
\det(A) = \frac{\gamma^2 - \gamma + 2 \frac{b^3}{a^3} - \gamma(\gamma - 1) \frac{bc}{a^2}}{2} = u^3 - \gamma(\gamma - 1)\varepsilon.
\]

The eigenvalue equation \( \det(A - \lambda I_3) = 0 \) writes

\[
\lambda^3 - 3u\lambda^2 + (3u^2 - \gamma(\gamma - 1)\varepsilon)\lambda - u^3 + \gamma(\gamma - 1)\varepsilon = 0.
\]

An evident solution is \( \lambda = u \). Factorization yields

\[
(\lambda - u)(\lambda^2 - 2u\lambda + u^2 - \gamma(\gamma - 1)\varepsilon) = 0.
\]

Therefore the eigenvalues of \( A \) are real and equal to (51). If \( \varepsilon > 0 \) the system is strictly hyperbolic. If \( \varepsilon < 0 \) it is linearly unstable. The proof is ended. \( \square \)

The quantity \( c = \sqrt{\gamma(\gamma - 1)\varepsilon} = \sqrt{\frac{\gamma}{\rho}} \) is the speed of sound. In standard conditions one can measure \( \rho, p, c \). Therefore one has access to the value of \( \gamma \) which is actually related to the microscopic structure of the gas. A numerical application is as follows. The density of air is \( \rho = 1.28 \times 10^3 \text{ g/m}^3 \). The pressure
at the surface of the earth is \( p = 1 \) atm = 1.013 \( \times 10^5 \) gm\(^{-1}\)s\(^{-2}\).

Since \( \gamma_{\text{air}} = 1.4 \) one finds \( c = \sqrt{\frac{\gamma_{\text{air}} p}{\rho}} \approx 332.88 \text{ms}^{-1} \) which is in good accordance with the experiment values.

It is instructive to follow what Newton and Poisson did at their time by neglecting the influence of the temperature. It is called the Boyle pressure law \( p = \mu \rho \) where \( \mu > 0 \) is a constant. We would find \( c\sqrt{\mu} = \sqrt{\frac{\mu}{\rho}} \approx 281.31 \text{ms}^{-1} \).

This value does not correspond to measurements. This refutation of the Boyle is an indirect justification of a polytropic pressure law \( p = (\gamma - 1)\rho \varepsilon \) for air, with \( \gamma_{\text{air}} = 1.4 \).

### 2.4.4 Eulerian compressible gas dynamics in dimension \( d > 1 \)

We firstly consider the Euler system (15) in dimension \( d = 3 \).

**Proposition 4.** Consider the system (16) with a positive density and positive pressure described by a polytropic equation of state. Define \( c = \sqrt{\frac{\gamma p}{\rho}} \).

Then the system is hyperbolic. The 5 eigenvalues of the Jacobian matrix in direction \( \alpha \in \mathbb{R}^3 \), \( |\alpha| = 1 \), are

\[
\begin{align*}
    u_\alpha - c, u_\alpha, u_\alpha, u_\alpha, u_\alpha + c \text{ with } u_\alpha = \mathbf{u} \cdot \alpha.
\end{align*}
\]

**Proof.** We study the stability with the method of section 2.3 and with a preliminary rotation of the axis. It is therefore sufficient to study the solutions of the Euler system which are invariant with respect to \( y \) and \( z \), taking into account that the horizontal velocity stands for \( u_\alpha = \mathbf{u} \cdot \alpha \). We use a new variable

\[
W = (p, u, v, w, \varepsilon)^t
\]

and the notation for the material derivative \( \frac{d}{dt} = \partial_t + \mathbf{u} \partial_x \). One gets after convenient algebra

\[
\begin{align*}
    \frac{d}{dt} \rho + \rho \partial_x u &= 0, \\
    \rho \frac{d}{dt} u + \partial_x p &= 0, \\
    \rho \frac{d}{dt} v &= 0, \\
    \rho \frac{d}{dt} w &= 0, \\
    \rho \frac{d}{dt} \varepsilon + p \partial_x u &= 0.
\end{align*}
\]

which turns into

\[
\begin{align*}
    \partial_t p + u \partial_x p + \rho c^2 \partial_x u &= 0, \\
    \partial_t u + u \partial_x u + \frac{1}{\rho} \partial_x p &= 0, \\
    \partial_t v + u \partial_x v &= 0, \\
    \partial_t w + u \partial_x w &= 0, \\
    \partial_t \varepsilon + u \partial_x \varepsilon + p \partial_x u &= 0.
\end{align*}
\]

(52)

Therefore the quasi-linear system for the new variable \( W \) is \( \partial_t W + B \partial_x W = 0 \)
with
\[
B = uI + \begin{pmatrix}
0 & \rho c^2 & 0 & 0 & 0 \\
\frac{1}{\rho} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & p & 0 & 0 & 0
\end{pmatrix}.
\]

The eigenvectors of \(B\) are
\[
s_1 = \begin{pmatrix}
\rho c \\
-1 \\
0 \\
0 \\
0
\end{pmatrix},
\quad s_2 = \begin{pmatrix}
0 \\
0 \\
1 \\
0 \\
0
\end{pmatrix},
\quad s_3 = \begin{pmatrix}
0 \\
0 \\
0 \\
1 \\
0
\end{pmatrix},
\quad s_4 = \begin{pmatrix}
0 \\
0 \\
0 \\
0 \\
1
\end{pmatrix},
\quad s_5 = \begin{pmatrix}
\rho c \\
1 \\
0 \\
0 \\
0
\end{pmatrix},
\]
with eigenvalues \(u - c\), \(u\) with multiplicity 3 and \(u + c\). Since the eigenvectors are linear independent for \(\rho c \neq 0\), it shows this problem admits 5 real eigenvalues and 5 real eigenvectors. It ends the proof.

The system is not strictly hyperbolic.

### 2.4.5 Lagrangian compressible gas dynamics in dimension \(d = 1\)

**Proposition 5.** Consider a pressure law written as \((\tau, \varepsilon) \mapsto p(\tau, \varepsilon)\). Assume \(p_\tau - pp_\varepsilon > 0\), \(\tau > 0\) and define \(c = \frac{1}{\rho} \sqrt{p_\tau - pp_\varepsilon} > 0\). Then the Lagrangian gas dynamics system (25) in the mass variable dimension one is hyperbolic. The eigenvalues of the Jacobian matrix are \(-\rho c, 0, \rho c\).

**Proof.** Consider the system with a pressure law \(p = p(\rho, \varepsilon)\)

\[
\begin{align*}
\partial_t \tau - \partial_m u &= 0, \\
\partial_t u + \partial_m p &= 0, \\
\partial_t \varepsilon + \partial_m (pu) &= 0, \quad \varepsilon = \varepsilon + \frac{1}{2}u^2.
\end{align*}
\]

The Jacobian matrix of the flux calculated with respect to the variables \((\tau, u, \varepsilon)\) is

\[
A = \begin{pmatrix}
0 & -1 & 0 \\
p_\tau & -up_\tau & p_\varepsilon \\
up_\tau & p - u^2 p_\varepsilon & up_\varepsilon
\end{pmatrix}.
\]

The characteristic polynomial is \(\det(A - \lambda I) = -\lambda^3 - (p_\tau - pp_\varepsilon)\lambda\). Set \(c = \frac{1}{\rho} \sqrt{p_\tau - pp_\varepsilon}\). One can check that a polytropic pressure law yields \(c = \sqrt{\frac{\gamma p_\tau}{\rho}}\). The roots of the characteristic polynomial are \(\lambda_1 = -\rho c\), \(\lambda_2 = 0\) and \(\lambda_3 = \rho c\). Since the eigenvalues are real and different, the system is strictly hyperbolic. The proof is ended.

\(\square\)
2.4.6 Lagrangian compressible gas dynamics in dimension $d = 2$

It appears that the hyperbolicity of Eulerian gas dynamics does not imply the hyperbolicity of Lagrangian gas dynamics. The reason is the additional unknowns of the Lagrangian system which are endowed with a loss of eigenvectors in dimension $d > 1$. However the discussion depends on how one takes into account the Piola’s identities. To ease the discussion consider the Hu’s formulation (30) with invariance in direction $Y$

$$
\begin{align*}
\partial_t (\rho J) &= 0, \\
\partial_t (\rho J u) + \partial_x (p M) &= 0, \\
\partial_t (\rho J v) + \partial_x (-p L) &= 0, \\
\partial_t (\rho J e) + \partial_x (p u M - p v L) &= 0, \\
\partial_t A - \partial_x u &= 0, \\
\partial_t B - \partial_x v &= 0, \\
\partial_t L &= 0, \\
\partial_t M &= 0,
\end{align*}
$$

and add the identities $\partial_x L = \partial_x M = 0$. So, using the initial data $x(0) = X$, $L = 0$ and $M = 1$ are now constant coefficients. On gets the system

$$
\begin{align*}
\partial_t (\rho J) &= 0, \\
\partial_t (\rho J u) + \partial_x p &= 0, \\
\partial_t (\rho J v) &= 0, \\
\partial_t (\rho J e) + \partial_x (p u) &= 0, \\
\partial_t A - \partial_x u &= 0, \\
\partial_t B - \partial_x v &= 0.
\end{align*}
$$

(53)

The Piola identity is $\partial_x J = \partial_x (u M - v L) = \partial_x u$ is equivalent to $\partial_t A - \partial_x u = 0$ because $J = AM - BL = A$.

**Proposition 6.** The Lagrangian gas dynamics system (53) is weakly hyperbolic with wave velocities $-\rho c, 0, \rho c$. The multiplicity of the eigenvalue 0 is 3. The order of weak hyperbolicity is 1.

**Proof.** A full proof is in [4]. A simplified proof comes from the consideration that (53) yields $(\rho J)(t, X) = \rho_0(X)$ and the two decoupled sub-systems

$$
\begin{align*}
\partial_t (\rho J) &= 0, \\
\rho_0 \partial_t u + \partial_x p &= 0, \\
\partial_t (\rho J e) + \partial_x (p u) &= 0, \\
\partial_t A - \partial_x u &= 0, \\
\partial_t B - \partial_x v &= 0.
\end{align*}
$$

The first system is the hyperbolic 1D Lagrangian. For $\rho_0 > 0$, the second system is weakly hyperbolic, see the Table 1.

**Remark 5** (Physical interpretation of weak hyperbolicity). The geometrical variables are the cause of the loss of strong hyperbolicity, see also [11]. It is possible to show that shear velocities, which trigger shear between layers of flows, are archetypal of the phenomenon.
The point is the deformation gradient which is very sensitive to shear velocities. In Lagrangian formulations, the deformation gradient is an unknown. In Eulerian formulation, the deformation gradient is not an unknown. This is the reason why Lagrangian gas dynamics is only a weakly hyperbolic system, while Eulerian gas dynamics is strongly hyperbolic.

Consider the non linear equations. The results are the same with the linearized equations. One constructs a specific initial data for the Lagrangian system 30: the density and pressure are constants, \( \rho_0 = \rho_c, \ p_0 = p_c \); the vertical velocity vanishes identically \( v_0 = 0 \) and the horizontal velocity is function of the vertical variable, which is written with an abuse of notations as \( u_0 = u_0(Y) \). The other variables are initialized as \( A_0 = M_0 = 1, B_0 = L_0 = 0 \) and \( J_0 = A_0M_0 - B_0L_0 = 1 \).

Since the pressure is a constant, it is immediate to realize that such initial condition generates shears between layers of flows. The physical solution writes; \( \rho = \rho_c \) and \( p = p_c \); the velocity is constant on horizontal lines; the Euler to Lagrange transformation is

\[
x = X + tu_0(Y) \quad \text{and} \quad y = Y.
\]

Therefore \( A = \partial_X x = 1 \), \( B = \partial_X y = 0 \), \( M = \partial_Y y = 1 \) and most of all

\[
L = \partial_Y x = tu_0'(Y).
\]

The norm of \( L \) increases linearly in time and exhibits one loss of derivative with respect to the initial data. This behavior is typical of a weakly hyperbolic system, equation (45) and definition 7. In this case a well prepared initial data corresponds to \( u_0' = 0 \) which means that the horizontal velocity is independent of the vertical coordinate.

### 3 Conservation laws \( n = 1 \) in dimension \( d = 1 \)

For a general conservation law, we review the notions of strong solutions, weak solutions, entropy solutions, define a general entropic finite volume scheme. Then we apply it to the Eulerian and Lagrangian traffic flow equation.

#### 3.1 Strong solutions

Consider the Cauchy problem for a conservation law

\[
\begin{cases}
\partial_t u + \partial_x f(u) = 0, \\
u(0, x) = u_0(x)
\end{cases}
\]

and assume the initial data is differentiable \( u_0 \in C^1(\mathbb{R}) \). Assuming the solution \( u \) is also differentiable, one can write the quasi-linear non divergent form as well
\[ \partial_t u + a(u) \partial_x u = 0 \] where by definition \( a(u) = f'(u) \). Let us consider the change of coordinates

\[
\begin{align*}
\frac{\partial x(t', X)}{\partial t} &= a(u(t', x(t', X))), \\
(0, X) &= X,
\end{align*}
\]

(55)

with the chain rule

\[
\begin{align*}
\partial_t' &= \partial_t t \partial_t + \partial_t x \partial_x = \partial_t + a \partial_x, \\
\partial_X &= \partial_X t \partial_t + \partial_X x \partial_x = J \partial_x. 
\end{align*}
\]

(56)

By definition \( J = \partial X |_{t'=x} \). So \( \partial_t u + a(u) \partial_x u = 0 \) is equivalent to \( \partial_t' u(t', x(t', X)) = 0 \). It shows that \( u \) is constant along the characteristic curves \( u(t', x(t', X)) = u(0, x(t', 0)) = u_0(X) \). It implies that \( \frac{\partial x(t', X)}{\partial t'} = a(u_0(X)) \). The characteristic curves are actually straight lines

\[ x(t, X) = X + t a(u_0(X)), \quad a(u) = f'(u). \]

(57)

The construction of the solution with the method of characteristics consists to solve firstly the equation (55) and secondly to propagate the solution along the characteristics. Considering that \( t \) and \( x = x(t, X) \) are given, the characteristic equation is a non linear equation for \( X \)

\[ x(t, X) = X + t a_0(X), \quad a_0(X) = a(u_0(X)), \]

Proposition 7. Assume the function \( a_0 \in C^0(\mathbb{R}) \cap L^\infty(\mathbb{R}) \) is piecewise differentiable, with everywhere a derivative on the right and on the left

\[ C \leq a_0'(X^-), a_0'(X^+) \leq L, \quad \forall X \in \mathbb{R}, \ 0 < L. \]

(58)

Define the time \( T_* \)

\[ T_* = \frac{1}{-C} \text{ if } C < 0 \text{ and } T_* = \infty \text{ is } C \geq 0. \]

(59)

Then there exists a unique solution \((t, X) \mapsto x(t, X)\) of the characteristic equation for all time \( 0 \leq t < T_* \).

Proof. The proof is an easy consequence of the fixed point theorem for a contractive function. Rewrite the equation as

\[ X = g_a(X) \text{ with } g_a(X) = \frac{1}{1+\alpha} (x + \alpha X - t a(X)) \text{ and } \alpha \geq 0. \]

The derivatives of \( g_a \) are \( g'_a(X^\pm) = \frac{1}{1+\alpha} (\alpha - ta'(X^\pm)) \in \left[ \frac{\alpha-\alpha L}{1+\alpha}, \frac{\alpha-\alpha t C}{1+\alpha} \right] \). The contraction requirement \( \|g'_a\|_{L^\infty(\mathbb{R})} < 1 \) needed in the fixed point theorem writes \(-1 < \frac{\alpha-\alpha t C}{1+\alpha} < 1 \). It yields the conditions \(-1 - \alpha < \alpha - t L \) and \( \alpha - t C \leq 1 + \alpha \).

The second inequality is true if and only if \(-t C < 1 \). If \( C \geq 0 \) this is true for all time \( t \). If \( C < 0 \) it yields the condition \( t < T_* = \frac{1}{C} \) of the claim. The first inequality can always be made true by taking a parameter \( \alpha > 0 \) sufficiently large. The proof is ended. \( \square \)
Proposition 8 (Strong solutions). Assume the initial data \( x \mapsto u_0(x) \) is continuous with a piecewise continuous derivative that satisfies (58).

Then the continuous (with piecewise continuous derivative) function \( (t, x) \mapsto u(t, x) = u_0(X(t, x)) \) is solution of the Cauchy problem (54) for \( 0 \leq t < T \).

Proof. One has that
\[
\partial_t u(t, x) + \partial_x f(u(t, x)) = \partial_t u(t, x) + a(u(t, x))\partial_x u(t, x) = \partial_{t|X} u_0(X) = 0.
\]
This equality holds except at points where \( u \) is not \( C^1 \). The proof is ended. \( \square \)

3.2 Weak solutions

The construction of a strong solution with the method of characteristics fails for a time \( t = T_* \). Even worse the threshold time can vanish as well, that is \( T_* = 0 \). In this case the method of characteristics does not construct anything. It shows that the framework of strong solutions is too restrictive. It is therefore needed to extend the notion of what is a solution. This is done with the notion of weak solutions.

The question is to give a meaning to discontinuous solutions of the Cauchy problem
\[
\left\{ \begin{array}{l}
\partial_t u + \partial_x f(u) = 0, \quad t > 0, \quad x \in \mathbb{R} \\
u(0, x) = u_0(x), \quad x \in \mathbb{R}.
\end{array} \right. \tag{60}
\]

To this end assume that \( u \) is a strong solution. Consider an additional function \( (t, x) \mapsto \varphi(t, x) \) which is \( C^1 \) and with compact support in space-time: typcially \( \varphi(t, x) \equiv 0 \) if \( t > T \) or \( |x| > A \). The function \( \varphi \) is a priori non zero for \(-A < x < A\).

The set of such functions will be denoted as
\[
C^1_0 = \{ \varphi \in C^1(\mathbb{R}^+ \times \mathbb{R}); \varphi \text{ with compact support} \}, \quad \mathbb{R}^+ = [0, \infty).
\]

Multiply the equation (60) by \( \varphi \in C^1_0 \) and integrate in space and time
\[
\int_{\mathbb{R}} \int_{0 < t} (\partial_t u + \partial_x f(u)) \varphi \, dt \, dx = \int_{-A < x < A} \int_{0 < t} (\partial_t u + \partial_x f(u)) \varphi \, dt \, dx = 0.
\]
Integrate by parts
\[
- \int_{-A < x < A} \int_{0 < t} (u \partial_t \varphi + f(u) \partial_x \varphi) \, dx \, dt - \int_{-A < x < A} u_0(x) \varphi(0, x) \, dx = 0.
\]
We notice that the derivatives of \( u \) and \( f(u) \) do not show up anymore. It means that this expression allows to manipulate discontinuous functions. This is the weak formulation that was looked for.

Definition 10 (Weak solutions). Let \( u \in L^\infty(\mathbb{R} \times \mathbb{R}^+) \). It is a weak solution of the Cauchy problem (60) if and only if
\[
\int_{\mathbb{R}} \int_{0 < t} (u \partial_t \varphi + f(u) \partial_x \varphi) \, dx \, dt + \int_{\mathbb{R}} u_0(x) \varphi(0, x) \, dx = 0 \tag{61}
\]
holds for all \( \varphi \in C^1_0 \). This weak equation is also called a weak formulation.
Figure 7: Test function $\varphi \in C^1_0$ with compact support in $[0, T[ \times -A, A]$.

**Proposition 9.** Strong solutions are also weak solutions. A weak solution which is continuous and piecewise differentiable is also a strong solution.

**Theorem 3** (Characterization of discontinuous weak solutions). Let $u \in L^\infty(\mathbb{R} \times \mathbb{R}^+)$ and assume there exists a smooth curve $\Gamma : t \mapsto x(t)$ such that $u$ is piecewise differentiable on both sides of the curve. Then $u$ is a weak solution of (61) if and only if

a) $u$ is a strong solution separately on both sides of $\Gamma$;

b) one has the jump relation for all $x(t) \in \Gamma$:

$$-x'(t)[u] + [f(u)] = 0.$$  \hspace{1cm} (62)

**Remark 6.** By convention, $[g] = g^+ - g^-$ is the jump of $g$ across the curve.

**Proof.** It is sufficient to check (62). Let $\varphi \in C^1_0$ with a support that contains a piece of the curve $\Gamma = \{t \mapsto x(t)\}$. Split the space in two parts $\Omega^- = \{(t, x) : x < x(t)\}$ and $\Omega^+ = \{(t, x) : x > x(t)\}$.

For simplicity we assume that $\varphi(0, x) \equiv 0$ for all $x \in \mathbb{R}$ since it plays no role. So (61) becomes

$$\int \int_{\Omega^-} (u \partial_t \varphi + f(u) \partial_x \varphi) \, dx \, dt + \int \int_{\Omega^+} (u \partial_t \varphi + f(u) \partial_x \varphi) \, dx \, dt = 0.$$  

Using the Stokes formula on both sides, one gets

$$-\int \int_{\Omega^-} (\partial_t u + \partial_x f(u)) \varphi + \int_{\Gamma} ((f(u), u)^-, n^-) \varphi \, d\sigma$$

$$-\int \int_{\Omega^+} (\partial_t u + \partial_x f(u)) \varphi + \int_{\Gamma} ((f(u), u)^+, n^+) \varphi \, d\sigma = 0.$$  

32
Figure 8: The outgoing normals from $\Omega^\pm$ are denoted as $n^\pm \in \mathbb{R}^2$.

The function $u$ is a strong solution locally in $\Omega^-$ and $\Omega^+$, so $\int_{\Omega^-} (\partial_t u + \partial_x f(u)) \varphi = \int_{\Omega^+} (\partial_t u + \partial_x f(u)) \varphi = 0$. Therefore

$$\int_\Gamma ((f(u), u^-), n^-) \varphi d\sigma + \int_\Gamma ((f(u), u^+), n^+) \varphi d\sigma = 0.$$

Since $\varphi$ is arbitrary one gets the necessary and sufficient condition

$$((f(u), u^-), n^-) + ((f(u), u^+), n^+) = 0 \text{ sur } \Gamma.$$

The tangent and normals vectors are

$$t = \frac{1}{a} \begin{pmatrix} x'(t) \\ 1 \end{pmatrix}, \quad n^+ = \frac{1}{a} \begin{pmatrix} 1 \\ -x'(t) \end{pmatrix} = -n^-, \quad a = \sqrt{1 + x'(t)^2}.$$

It yields the claim $-x'(t)[u] + [f(u)] = 0$. The proof is ended. $\square$

It is convenient to consider a discontinuity that moves at a velocity $\sigma$

$$u(t, x) = u_L \text{ for } x < \sigma t, \quad u(t, x) = u_R \text{ for } x > \sigma t.$$

**Definition 11** (Rankine Hugoniot relation). The triplet $(\sigma, u_L, u_R)$ satisfies the Rankine-Hugoniot relation if and only if

$$-\sigma(u_R - u_L) + (f(u_R) - f(u_L)) = 0. \quad (63)$$

The Rankine-Hugoniot relation is a rephrasing of (62) for $\sigma = x'(t)$. One also notes $-\sigma[u] + [f(u)] = 0$. For example discontinuous solutions of the Burgers equation satisfy

$$\sigma = \frac{\left[u^2\right]}{[u]} = \frac{u_R + u_L}{2}. \quad (64)$$
This is a compatibility relation between the right and left states and the velocity of the discontinuity.

3.3 Entropy weak solutions

The use of weak solutions may generate a paradox in certain cases. Indeed the space of weak solutions is by definition much greater than the space of strong solutions. Therefore it is possible that the weak formulation allows additional solutions, even in the case where only one single solution is constructed by the method of characteristics. To have better insight into this paradox, consider the following situation.

**Proposition 10.** The Burgers equation with initial data

\[ u_0(x) = 0, \ x < 0; \quad u_0(x) = 1, \ x > 0, \]

admits two different weak solutions. One is a discontinuity with velocity \( \sigma = \frac{1}{T} \). The other one is

\[
0 \leq t < T = \infty \quad \begin{cases} u(t,x) = 0, & x < 0, \\ u(t,x) = \frac{x}{T}, & 0 < x < t, \\ u(t,x) = 1, & x - t > 0. \end{cases} \tag{65}
\]

**Proof.** The first function satisfies the criterion (64), so is indeed a weak solution. The second function is bounded, \( 0 \leq u(x,t) \leq 1 \) for all \( x \in \mathbb{R} \) and \( t \in \mathbb{R}^+ \). To show it is a weak solution we study the sum (61) of the two integrals in the definition of the weak solution. Taking care of the three different zones (65), it writes for \( \varphi \in C_0^1 \)

\[
I(\varphi) = \int_0^T \int_{0<x<t} \left( \frac{x}{t} \partial_t \varphi(x,t) + \frac{x^2}{2t^2} \partial_x \varphi(x,t) \right) dx dt
\]
\[
\int \int_{0<\tau<x} \left( \partial_t \varphi(x,t) + \frac{1}{2} \partial_x \varphi(x,t) \right) dx dt + \int_{0<x} \varphi(x,0) dx.
\]

We need to show it is exactly zero for all admissible \( \varphi \).

Consider the vector field \( \mathbf{G}(x,t) = \left( \frac{x^2}{2\tau^2} \varphi(x,t), \frac{x}{t} \varphi(x,t) \right) \) defined in the zone \( 0 < x < t \). One has
\[
\nabla_{xt} \cdot \mathbf{G} = \partial_x \left( \frac{x^2}{2\tau^2} \varphi \right) + \frac{x}{t} \partial_t \varphi = \frac{x}{t} \partial_x \varphi + \frac{x}{t} \partial_t \varphi.
\]

So by the Stokes formula the first term in \( I(\varphi) \) is
\[
I_1(\varphi) \equiv \int \int_{0<x<\tau} \left( \frac{x^2}{2\tau^2} \partial_x \varphi(x,t) + \frac{x}{t} \partial_t \varphi(x,t) \right) dx dt
= \int \int_{0<x<\tau} \nabla_{xt} \cdot \mathbf{G} dx dt = \int_{\partial(0<x<\tau)} \mathbf{G} \cdot \mathbf{n} d\sigma,
\]
where the exterior normal is \( \mathbf{n} = \frac{1}{\sqrt{2}} (1,-1) \) and the measure is \( d\sigma = \sqrt{2} dx \).

That is
\[
I_1(\varphi) = \int_0^\infty \left( \frac{x^2}{2t^2} \partial_x \varphi(x,t) - \frac{x}{t} \partial_t \varphi(x,t) \right) dx = -\frac{1}{2} \int_0^\infty \varphi(x,x) dx.
\]

A simple use of the Stokes formula shows that the second part in \( I(\varphi) \) is
\[
I_2(\varphi) \equiv \int \int_{0<x<\tau} \left( \partial_t \varphi(x,t) + \frac{1}{2} \partial_x \varphi(x,t) \right) dx dt
= \frac{1}{2} \int_0^\infty \varphi(x,x) dx - \int_{0<x} \varphi(x,0) dx.
\]

So \( I(\varphi) = 0 \) for all \( \varphi \in C^1_0 \) which shows this function is also a weak solution. The proof is ended.

Since it is not possible to distinguish between these two solutions solely on the basis of weak formulations, a natural question is to find an additional criterion, if any, that would select only one solution between the two. This criterion exists and is formulated with the notion of entropies. It is introduced by considering that admissible functions \( u \) should be the limit of viscous functions \( u_\varepsilon \) with a dissipative regularizing operator in the right hand side
\[
\begin{align*}
\left\{ \begin{array}{ll}
\partial_t u_\varepsilon + \partial_x f(u_\varepsilon) = \varepsilon \partial_{xx} u_\varepsilon, & t > 0 \\
u_\varepsilon(0,x) = u_{0,\varepsilon}(x).
\end{array} \right.
\end{align*}
\]

The dissipative regularizing operator is \( \varepsilon \partial_{xx} \) with a vanishing parameter \( \varepsilon \rightarrow 0^+ \) and we admit that \( u_\varepsilon \) is naturally a very smooth function (it can be justified \([6, 15, 3]\)).

35
Let us make an additional hypothesis which can also be justified: we assume that \( u_\varepsilon \) is a regular differentiable uniformly bounded function which tends to \( u_\varepsilon \in L^1_{\text{loc}} \) in the sense

\[
\lim_{\varepsilon \to 0^+} \|u - u_\varepsilon\|_{L^1_{\text{loc}}([0,T] \times \mathbb{R})} = 0, \quad \|u_\varepsilon\|_{L^\infty([0,T] \times \mathbb{R})} \leq C \text{ for all } \varepsilon
\]  

(67)

and that

\[
\lim_{\varepsilon \to 0^+} \|u_0 - u_{0,\varepsilon}\|_{L^1_{\text{loc}}(\mathbb{R})} = 0.
\]  

(68)

The functions \( u_\varepsilon \) are called viscous solutions. To identify all the limit equalities and inequalities satisfied by \( u \) limit of \( u_\varepsilon \), we need entropies.

**Definition 12** (Entropy and entropy flux). A function \( \eta : \mathbb{R} \to \mathbb{R} \) which is two times differentiable and convex is called an entropy. The corresponding entropy flux function \( \xi : \mathbb{R} \to \mathbb{R} \) is defined up to a constant by

\[
\eta'(u)f'(u) = \xi'(u), \quad \xi(u) = \int \eta'(v)f'(v)dv.
\]

Consider for example the Burgers equation with flux \( f(u) = \frac{u^2}{2} \). Any function \( \eta_p(u) = \frac{u^p}{2p} + \alpha \frac{u^2}{2} \) with \( p \in \mathbb{N}, p \geq 2 \) and \( \alpha > 0 \) is an entropy. The entropy flux is \( \xi_p(u) = \frac{u^{2p+1}}{2p+1} + \alpha \frac{u^3}{3} \) since

\[
\eta_p'(u)f'(u) = (u^{2p-1} + u)u = u^{2p} + u^2 = \xi_p'(u).
\]

Let us define the space \( C^1_{0,+} = C^1_0 \cap \{ \varphi \geq 0 \} \).

**Theorem 4.** Let \( u \) be at the limit of \( u_\varepsilon \) with (66-67,68). Then the limit function \( u \) satisfies two properties

a) It is a weak solution (61).

b) For all entropy-entropy flux pair \((\eta, \xi)\) and all \( \varphi \in C^1_{0,+} \), one has the weak entropy inequality

\[
- \int_{\mathbb{R}} \int_{0<t} (\eta(u)\partial_t \varphi + \xi(u)\partial_x \varphi) dxdt - \int_{\mathbb{R}} \eta(u_0)(x)\varphi(0,x)dx \leq 0.
\]  

(69)

**Proof.** From (66) one has \( \partial_t u_\varepsilon + f'(u_\varepsilon)\partial_x u_\varepsilon = \varepsilon \partial_{xx} u_\varepsilon \). Multiply by \( \eta'(u_\varepsilon) \)

\[
\partial_t \eta(u_\varepsilon) + \eta'(u_\varepsilon)f'(u_\varepsilon)\partial_x u_\varepsilon = \varepsilon \eta'(u_\varepsilon)\partial_{xx} u_\varepsilon.
\]

The entropy-entropy flux relation yields that

\[
\partial_t \eta(u_\varepsilon) + \partial_x \xi(u_\varepsilon) = \varepsilon \partial_{xx} \eta(u_\varepsilon) - \varepsilon \eta''(u_\varepsilon)(\partial_x u_\varepsilon)^2 \leq \varepsilon \partial_{xx} \eta(u_\varepsilon).
\]

Multiply by \( \varphi \in C^1_{0,+} \)

\[
\int \int (\partial_t \eta(u_\varepsilon) + \partial_x \xi(u_\varepsilon))\varphi dxdt \leq \varepsilon \int \int (\partial_{xx} \eta(u_\varepsilon))\varphi dxdt.
\]

36
Integrate by parts

\[- \int_{\mathbb{R}} \int_{0<t} (\eta(u_\varepsilon) \partial_t \varphi + \xi(u_\varepsilon) \partial_x \varphi) \, dx dt - \int_{\mathbb{R}} \eta(u_{0,\varepsilon}) \varphi(0,\cdot) \, dx \]

\[ \leq \varepsilon \int_{\mathbb{R}} \int_{0<t} \eta(u_\varepsilon) \partial_x \varphi \, dx dt. \]

It remains to show the convergence of the integrals with respect to \( \varepsilon \). Notice the test function \( \varphi \) is constant. One has

\[ \left| \int_{\mathbb{R}} \int_{0<t} \eta(u_\varepsilon) \partial_t \varphi \, dx dt - \int_{\mathbb{R}} \int_{0<t} \eta(u) \partial_t \varphi \, dx dt \right| \]

\[ \leq \left( \max \left| \partial_t \varphi \right| \max_{|v| \leq C} |\eta'(v)| \right) ||u_\varepsilon - u||_{L^1(\text{supp}(\varphi))}. \]

where the support is \( \text{supp}(\varphi) = \{(t, x) \in \mathbb{R}^+ \times \mathbb{R}; \varphi(t, x) > 0\} \). Therefore

\[ \int_{\mathbb{R}} \int_{0<t} \eta(u_\varepsilon) \partial_t \varphi \, dx dt \to \int_{\mathbb{R}} \int_{0<t} \eta(u) \partial_t \varphi \, dx dt. \]

The two other integrals converge for the same reasons. Finally the right hand side tends to zero since it depends linearly on \( \varepsilon \). The proof is ended. \( \square \)

**Definition 13** (Entropy weak solutions). A weak solution \( u \in L^\infty(\mathbb{R} \times \mathbb{R}^+) \) which satisfies the weak entropy inequalities (69) for all entropy-entropy flux pairs \((\eta, \xi)\) is called an entropy weak solution.

**Theorem 5** (Characterization of discontinuous entropy weak solutions). Let \( u \in L^\infty(\mathbb{R} \times \mathbb{R}^+) \) and assume there exits a smooth curve \( \Gamma : t \mapsto x(t) \) such that \( u \) is piecewise differentiable on both sides of the curve. Then \( u \) is an entropy weak solution (61) if and only if

**a)** \( u \) is a strong solution locally on both sides of \( \Gamma \),

**b)** one has the jump relation

\[-x'(t) [u] + [f(u)] = 0, \quad t > 0, \quad (70)\]

**c)** one has the jump relation for all entropy-entropy flux pairs

\[-x'(t) [\eta(u)] + [\xi(u)] \leq 0, \quad t > 0. \quad (71)\]

Consider a triplet \( (\sigma, u_L, u_R) \in \mathbb{R}^3 \) solution of the Rankine Hugoniot relation, with \( \sigma \) the velocity of the discontinuity, \( u_L \) the left state and \( u_R \) the right state. Since an entropy weak solution satisfies the jump inequality for all entropies (71) one obtains a system of one equality plus an infinite number of inequalities

\[
\begin{align*}
-\sigma(u_R - u_L) + f(u_R) - f(u_L) &= 0, \\
-\sigma(\eta(u_R) - \eta(u_L)) + \xi(u_R) - \xi(u_L) &\leq 0, \quad \forall (\eta, \xi) \quad \eta'' \geq 0 \text{ and } \xi' = \eta' f'.
\end{align*}

(72)
3.4 Peculiarities of Lagrangian traffic flow

The traffic flow equation is a paradigm for the comparison of Eulerian and Lagrangian formulations. We start from the Eulerian equation

$$\partial_t \rho + \partial_x f(\rho) = 0, \quad f(\rho) = \rho u(\rho)$$  \hspace{1cm} (73)

and add the assumption that $u'(\rho) \leq 0$ which corresponds to the fact that the velocity of a driver is adapted to the density: the denser the flow the slower the velocity. Using the Euler-to-Lagrange method that allows to rewrite the one dimensional system of gas dynamics as a the purely Lagrangian system (25), one obtains the Lagrangian form of the traffic flow equation

$$\partial_t \tau - \partial_m v(\tau) = 0, \quad v(\tau) = u(\rho), \quad \rho = \frac{1}{\tau}.$$  \hspace{1cm} (74)

The mass variable is $dm = \rho_0(x)dx$.

- A first remark is, of course, that a zero initial density is a problem for the Lagrangian formulation which becomes nonsense. However one may considers a well modified initial density $\rho_0^\varepsilon = \max(\rho_0(x), \varepsilon)$ for a small value of the parameter $\varepsilon > 0$. One expects from basic physical considerations some continuity properties of the solution with respect to the initial data (actually this is what we observe on the highway).

- A second remark is about the speed of sound.

**Proposition 11.** Assuming $u' \leq 0$, the speed of sound $a(\tau)$ of the Lagrangian system (74) has a constant sign

$$a(\tau) = -\rho^2 u'(\rho) \geq 0.$$

**Proof.** Indeed $a(\tau) = -v'(\tau) = \frac{d}{d\tau} u\left(\frac{1}{\tau}\right) = -\frac{1}{\tau^2} u'\left(\frac{1}{\tau}\right) = -\rho^2 u'(\rho).$ \[\square\]

This property has the important consequence that it allows to use numerical methods with simpler structure. It will be generalized to systems in the next chapter.

- The third remark is about a comparison of Rankine-Hugoniot relation of the Eulerian system

$$-\sigma (\rho_R - \rho_L) + (f(\rho_R) - f(\rho_L)) = 0$$  \hspace{1cm} (75)

with the Rankine-Hugoniot relation of the Lagrangian system

$$-D (\tau_R - \tau_L) - (v(\tau_R) - v(\tau_L)) = 0.$$  \hspace{1cm} (76)

**Proposition 12.** The Eulerian and Lagrangian Rankine-Hugoniot relations are equivalent. Moreover $\tau_{R,L}^1 = \rho_{R,L}^1$ and one has the correspondence between the shock velocities $D$ and $\sigma$

$$D = \rho_R(\sigma - u_R) = \rho_L(\sigma - u_L).$$  \hspace{1cm} (77)

Moreover the Eulerian and Lagrangian entropy criterions are equivalent.
Proof. The first part of the claim is easy. Indeed the equation (75) with $f = \rho u$ rewrites as $\rho_R(\sigma - u_R) = \rho_L(\sigma - u_L)$. The interpretation is that the mass flux across the shock, the dimension is the product of a density times a velocity, is constant. Set $D = \rho_R(\sigma - u_R) = \rho_L(\sigma - u_L)$. One has

$$-D(\tau_R - \tau_L) = -\rho_R(\sigma - u_R)\tau_R + \rho_L(\sigma - u_L)\tau_L = (u_R - \sigma) - (u_L - \sigma) = u_R - u_L.$$ 

The second part is shown as follows. First we consider the entropy criterion for the Eulerian equation

$$-\sigma(\eta(\rho_R) - \eta(\rho_L)) + (\xi(\rho_R) - \xi(\rho_L)) \leq 0$$  (78)

for all entropy $\eta'' > 0$ and entropy flux $\xi' = f'\eta'$. Second we define for $0 < \tau < \infty$

$$\hat{\eta}(\tau) = \tau\eta\left(\frac{1}{\tau}\right) \quad \text{and} \quad \hat{\xi}(\tau) = \xi\left(\frac{1}{\tau}\right) - u\left(\frac{1}{\tau}\right)\eta\left(\frac{1}{\tau}\right)$$

and show that $(\hat{\eta}, \hat{\xi})$ is an entropy-entropy flux pair for the Lagrangian system. The concavity criterion is checked as follows

$$\frac{d^2}{d\tau^2}\hat{\eta}(\tau) = \frac{d^2}{d\tau^2}\left(\tau\eta\left(\frac{1}{\tau}\right)\right) = \frac{d}{d\tau}\left(\eta\left(\frac{1}{\tau}\right) - \frac{1}{\tau}\eta'\left(\frac{1}{\tau}\right)\right)$$

$$= -\frac{1}{\tau^2}\eta'\left(\frac{1}{\tau}\right) + \frac{1}{\tau^2}\eta'\left(\frac{1}{\tau}\right) + \frac{1}{\tau^3}\eta''\left(\frac{1}{\tau}\right) = \frac{1}{\tau^3}\eta''\left(\frac{1}{\tau}\right) \geq 0$$

so the function $\hat{\eta}$ is indeed convex. The relation $\xi'(\rho) = (u(\rho) + \rho u'(\rho))\eta'(\rho)$ can be used to checked the identity

$$\frac{d}{d\tau}\hat{\xi}(\tau) = \frac{d}{d\tau}\left[\xi\left(\frac{1}{\tau}\right) - u\left(\frac{1}{\tau}\right)\eta\left(\frac{1}{\tau}\right)\right]$$

$$= -\frac{1}{\tau^2}\left[\xi'\left(\frac{1}{\tau}\right) - u'\left(\frac{1}{\tau}\right)\eta\left(\frac{1}{\tau}\right) - u\left(\frac{1}{\tau}\right)\eta'\left(\frac{1}{\tau}\right)\right]$$

$$= -\frac{1}{\tau^2}\left[\frac{1}{\tau}u'\left(\frac{1}{\tau}\right)\eta'\left(\frac{1}{\tau}\right) - u'\left(\frac{1}{\tau}\right)\eta\left(\frac{1}{\tau}\right)\right] = -\frac{1}{\tau^2}u'\left(\frac{1}{\tau}\right)\frac{1}{\tau}\eta'\left(\frac{1}{\tau}\right) - \eta\left(\frac{1}{\tau}\right)\right].$$

But by definition $\hat{\eta}(\tau) = \tau\eta\left(\frac{1}{\tau}\right)$ which yields $\hat{\xi}'(\tau) = \eta\left(\frac{1}{\tau}\right) - \frac{1}{\tau}\eta'\left(\frac{1}{\tau}\right)$. Therefore

$$\frac{d}{d\tau}\hat{\xi}(\tau) = \frac{1}{\tau^2}u'\left(\frac{1}{\tau}\right)\eta'\left(\frac{1}{\tau}\right) = \frac{d}{d\tau}v(\tau) \frac{d}{d\tau}\hat{\eta}(\tau)$$

with $v$ the Lagrangian flux (74). It shows $\hat{\xi}$ is indeed the Lagrangian entropy flux.

We can now transform (78) with the help of these identities. One rewrite (78) as

$$-(\sigma - u_R)\eta(\rho_R) - (\sigma - u_L)\eta(\rho_L) + \hat{\xi}(\tau_R) - \hat{\xi}(\tau_L) \leq 0$$

39
and as
\[-(\sigma - u_R)\rho_R\tilde{\eta}(\tau_R) - (\sigma - u_L)\rho_L\tilde{\eta}(\tau_L) + \tilde{\xi}(\tau_R) - \tilde{\xi}(\tau_L) \leq 0.\]

Use the formula (77) for the Lagrangian shock velocity and get
\[-D(\tilde{\eta}(\tau_R) - \tilde{\eta}(\tau_L)) + \tilde{\xi}(\tau_R) - \tilde{\xi}(\tau_L) \leq 0.\]  \hspace{1cm} (79)

In summary all Eulerian entropy relations (78) are equivalent to all Lagrangian entropy relations (79). It shows the last part of the claim. The proof is ended.

\[\Box\]

### 3.5 Numerical computation of entropy weak solutions

The calculation of correct entropic weak solutions generate important constraints on numerical methods. We concentrate on construction issues.

#### 3.5.1 Notion of a conservative finite volume scheme

It is known in the literature that Finite Volume (FV) schemes are among the most efficient ones for the calculation of discontinuous weak solutions. To justify this fact, we start from three versions of a basic Finite Difference scheme and show that if the exact solution is discontinuous, only one version of the scheme is correct. The correct version is the one which can be reinterpreted as a FV scheme. It enlightens the notion of a conservative scheme.

The example is based on the Burgers equation
\[\partial_t u + \partial_x u^2 = 0\]
with the initial data
\[
\begin{align*}
  &u(t = 0, x) = u_0(x) = 1 & \text{for } x < 0.5 \\
  &u(t = 0, x) = u_0(x) = 1 + (0.5 - x) & \text{for } 0.5 < x < 1.5 \\
  &u(t = 0, x) = u_0(x) = 0 & \text{for } 1.5 < x.
\end{align*}
\]

For small time \( t < 1 \), the entropic weak solution contains a ramp between \( x = 1 + t \) and \( x = 2 \). At \( t = 1 \) an entropic shock is formed. After that time, the shock propagates at velocity \( \sigma = \frac{1}{2} \).

Let us now discretize the equation, with a general numerical method in Finite Difference form
\[
\frac{u_j^{n+1} - u_j^n}{\Delta t} + a_j^n \frac{u_j^n - u_{j-1}^n}{\Delta x} = 0, \hspace{1cm} (80)
\]
but with three different evaluations of the velocity
\[
\begin{align*}
  &\text{Choice 1} \quad a_j^n = \frac{u_j^n + u_{j-1}^n}{2}, \\
  &\text{Choice 2} \quad a_j^n = u_{j-1}^n, \\
  &\text{Choice 3} \quad a_j^n = u_j^n.
\end{align*} \hspace{1cm} (81)
\]

In the classical Finite Difference sense and for smooth reference solutions, the discrete velocity is clearly consistent with the exact velocity. This verification
is left to the reader. Let us now present at figure 10 the numerical results at
time $t = 0.5$ and at figure 11 the numerical results at time $t = 1.5$.

The explanation why only choice 1 yields a correct numerical approximation
of the shock is the related to the notion of conservation. Indeed the numerical
scheme with first velocity (81) is recast as

$$
\text{Choice 1 : } \frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{(u_j^n)^2}{2} - \frac{(u_{j-1}^n)^2}{2} \frac{\Delta x}{\Delta x} = 0
$$

One says the scheme is conservative since the formal conservation of the total
mass holds

$$
\text{Choice 1 : } \sum_j u_j^{n+1} = \sum_j u_j^n. \quad (82)
$$

It is clear at inspection of figure 11 that the conservation of the total mass is
strongly correlated to an accurate numerical prediction of the shock velocity. In
that sense preservation of the total mass seems to be a necessary condition for
accurate shock calculations. A scheme which preserves a property like (82) is
called a conservative scheme.

### 3.5.2 Finite Volume scheme

The previous example shows the importance of having a procedure which guaran-
tees by construction the preservation of the total mass (82). To achieve this
end, we consider a space-time predefined grid with mesh size $\Delta x > 0$ and
time step $\Delta t > 0$. The so-called volumes of the FV scheme are the intervals
$(x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}})$ where $x_{j+\frac{1}{2}} = (j + \frac{1}{2})\Delta x$ for all $j \in \mathbb{Z}$. The scheme writes under
the general form

$$
\frac{u_{j}^{n+1} - u_{j}^{n}}{\Delta t} + \frac{f_{j+\frac{1}{2}}^{n} - f_{j-\frac{1}{2}}^{n}}{\Delta x} = 0, \quad \forall j \in \mathbb{Z}, \forall n \in \mathbb{N}, \quad (83)
$$
The numerical solution at $t = 0.5$ for the 3 numerical velocities. The numerical solutions are very different. Only choice 1 yields a reasonable prediction of the theoretical shock position: $x_{\text{shock}} = 1,5 + \frac{1}{2}(1,5 - 1) = 1,75$.

with the mean value initial data

$$u_j^0 = \frac{1}{\Delta x} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u_0(x)dx, \quad \forall j \in \mathbb{Z}. \quad (84)$$

The numerical solution in cell $j$ and at time step $n$ is $u_j^n$. This scheme is explicit, that is one can determine explicitly the new value in the cell in function of the previous one and of the fluxes $f_{j+\frac{1}{2}}^n$

$$u_j^{n+1} = u_j^n + \frac{\Delta t}{\Delta x} f_{j-\frac{1}{2}}^n - \frac{\Delta t}{\Delta x} f_{j+\frac{1}{2}}^n, \quad j \in \mathbb{Z}. \quad (85)$$

Therefore the construction of the scheme ultimately relies on the determination of the numerical flux $f_{j+\frac{1}{2}}^n$ for $j \in \mathbb{Z}$.

**Proposition 13.** Whatever are the fluxes the scheme (83) is formally conservative

$$\sum_{j \in \mathbb{Z}} u_j^{n+1} = \sum_{j \in \mathbb{Z}} u_j^n.$$ 

**Proof.** Indeed

$$\sum_{j \in \mathbb{Z}} u_j^{n+1} - \sum_{j \in \mathbb{Z}} u_j^n = \sum_{j \in \mathbb{Z}} (u_j^{n+1} - u_j^n) = \frac{\Delta t}{\Delta x} \sum_{j \in \mathbb{Z}} (f_{j-\frac{1}{2}}^n - f_{j+\frac{1}{2}}^n) = 0$$

since it is a telescopic sum.

### 3.5.3 Definition of a generic flux

We consider

$$f_{j+\frac{1}{2}}^n = \frac{1}{2} (f(u_{j+1}^n) + f(u_j^n)) + \frac{c}{2}(u_j^n - u_{j+1}^n), \quad j \in \mathbb{Z}, \quad n \in \mathbb{N}. \quad (85)$$
This flux is the half sum of the fluxes on both sides plus a term \( \frac{c}{2} (u^n_j - u^{n+1}_j) \) which is called a viscous term. This formula is referred to in the literature as the Rusanov flux.

**Proposition 14.** Consider the scheme (83-84). Assume

\[
m = \inf_j (u^n_j) \quad \text{and} \quad M = \sup_j (u^n_j).
\]

and take the generic flux (85). Assume the parameter \( c \) is sufficiently large

\[
\max_{m \leq u \leq M} |f'(u)| \leq c.
\]

Assume the CFL constraint \( c \Delta t \leq 1 \).

Then the discrete solution satisfies the maximum principle

\[
m \leq u^{n+1}_j \leq M, \quad j \in \mathbb{Z}, \ n \in \mathbb{N}.
\]

**Remark 7.** Notice that \( c \) does not vanish in (87), except in the trivial case where \( f \) is constant.

**Proof.** Note \( \nu = \frac{\Delta t}{\Delta x} \). One has

\[
u c = \frac{\nu c}{2} (u^{n+1}_j - 2u^n_j + u^{n-1}_j) - \frac{\nu}{2} [f(u^{n+1}_j) - f(u^{n}_j)]
\]

\[= u^n_j + \frac{\nu c}{2} (u^{n+1}_j - 2u^n_j + u^{n-1}_j) - \frac{\nu}{2} [a^n_j (u^{n+1}_j - u^{n-1}_j)].
\]

By construction one has

\[a^n_j = \frac{f(u^{n+1}_j) - f(u^{n}_j)}{u^{n+1}_j - u^{n-1}_j} = f'(z^n_j), \quad \min(u^{n+1}_j, u^{n}_j) \leq z^n_j \leq \max(u^{n+1}_j, u^{n}_j).
\]

So

\[u^{n+1}_j = [1 - \nu c] u^n_j + \frac{\nu}{2} [c - a^n_j] u^{n+1}_j + \frac{\nu}{2} [c + a^n_j] u^{n-1}_j.
\]

Assume the coefficients in front of \( u^n_j, \ u^{n+1}_j \) and \( u^{n-1}_j \) are all non negative. Since the sum is equal to one, \( u^{n+1}_j \) is a convex combination of \( u^n_j, \ u^{n+1}_j \) and \( u^{n-1}_j \). The CFL constraint at \( n = 0 \) insures the non negativity at first step. By iteration this property propagates from one time to the other. It proves the maximum principles and ends the proof.

**Proposition 15.** Make the assumptions of proposition 14. Define the numerical entropy flux

\[
\xi^\frac{n}{2}_j = \xi(u^{n+1}_j) + \xi(u^n_j) + \frac{c}{2} (\eta(u^n_j) - \eta(u^{n+1}_j)).
\]

Then the numerical solution satisfies the discrete entropy inequality

\[
\frac{\eta(u^{n+1}_j) - \eta(u^n_j)}{\Delta t} + \frac{\xi^\frac{n}{2}_j - \xi^\frac{n-1}{2}_j}{\Delta x} \leq 0, \quad j \in \mathbb{Z}, \ n \in \mathbb{N}.
\]
Proof. The proof is split in three steps.

**Step 1.** Rewrite the scheme as

\[
  u_j^{n+1} = \frac{1}{2} (u_j^n + \nu c (-u_j^n + u_{j-1}^n) - \nu \left[ f(u_j^n) - f(u_{j-1}^n) \right] )
\]

\[
+ \frac{1}{2} (u_j^n + \nu c (u_{j+1}^n - u_j^n) - \nu \left[ f(u_{j+1}^n) - f(u_j^n) \right] )
\]

The convexity of \( \eta \) yields

\[
\eta(u_j^{n+1}) \leq \frac{1}{2} \eta \left( u_j^n + \nu c (-u_j^n + u_{j-1}^n) - \nu \left[ f(u_j^n) - f(u_{j-1}^n) \right] \right)
\]

\[
+ \frac{1}{2} \eta \left( u_j^n + \nu c (u_{j+1}^n - u_j^n) - \nu \left[ f(u_{j+1}^n) - f(u_j^n) \right] \right).
\]

So one can write

\[
\eta(u_j^{n+1}) - \eta(u_j^n) + \nu \left( \xi_j^n + \xi_{j-1}^n \right) \leq \frac{1}{2} \varphi(u_{j+1}^n) + \frac{1}{2} \phi(u_{j-1}^n)
\]

where

\[
\varphi(w) = \eta \left( u_j^n + \nu c (w - u_j^n) - \nu \left[ f(w) - f(u_j^n) \right] \right)
\]

\[-\eta(u_j^n) - \nu c (\eta(w) - \eta(u_j^n)) + \nu \left( (\xi(w) - \xi(u_j^n)) \right) \]

and

\[
\phi(z) = \eta \left( u_j^n + \nu c (-u_j^n + z) - \nu \left[ f(u_j^n) - f(z) \right] \right)
\]

\[-\eta(u_j^n) - \nu c (\eta(u_j^n) + \eta(z)) + \nu \left( (\xi(u_j^n) - \xi(z)) \right) \]

So the proof of the entropy inequality can be done by proving separately that \( \varphi(w) \leq 0 \) and \( \phi(z) \leq 0 \).

**Step 2.** One has \( \varphi(u_j^n) = 0 \) and

\[
\varphi'(w) = \nu (c - f'(w)) \times \left( \eta' \left( u_j^n + \nu c (w - u_j^n) - \nu \left[ f(w) - f(u_j^n) \right] \right) \right) \]

The first term is non negative. The second term is the product of a non negative term by \( (u_j^n + \nu c (w - u_j^n) - \nu \left[ f(w) - f(u_j^n) \right] ) \) - \( w \). Using the CFL condition, this term is itself the product of a non negative term by \( u_j^n - w \). Therefore one has

\[
\varphi'(w) = k(w)(u_j^n - w), \quad k(w) \geq 0 \quad \forall w.
\]

It shows that \( \varphi(w) \leq 0 \).

**Step 3.** One has \( \phi(u_j^n) = 0 \). Derivation yields

\[
\phi'(z) = \nu (c + f'(z)) \times \left( \eta' \left( u_j^n + \nu c (-u_j^n + z) - \nu \left[ f(u_j^n) - f(z) \right] \right) \right) \]

The first terms non negative and the second term is the product of \( u_j^n - z \) by a non negative term under CFL. So one has \( \phi(z) \leq 0 \). The proof is ended. \( \square \)
**Proposition 16.** Make the assumptions of proposition 14 and a more restrictive CFL condition \( c \frac{\Delta t}{\Delta x} \leq \frac{1}{2} \). Then the numerical solution satisfies the inequalities

\[
\sum_{j \in \mathbb{Z}} |u_{j+1}^{n+1} - u_j^{n+1}| \leq \sum_{j \in \mathbb{Z}} |u_{j+1}^n - u_j^n|, \quad n \in \mathbb{N},
\]

and

\[
\sum_{0 \leq n \Delta t \leq T} \left( \Delta t \sum_{j \in \mathbb{Z}} |u_{j+1}^{n+1} - u_j^{n+1}| + \Delta x \sum_{j \in \mathbb{Z}} |u_{j+1}^{n+1} - u_j^n| \right) \leq T \left( 1 + 2c \frac{\Delta t}{\Delta x} \right) \sum_{j \in \mathbb{Z}} |u_{j+1}^0 - u_j^0|.
\]

**Proof.** The recurrence relation (88) is rewritten with Harten’s notation

\[
u^n_{j+1} = u_j^n + C^n_j (u_{j-1}^n - u_j^n) + D^n_j (u_{j+1}^n - u_j^n)
\]

with \( C^n_j = \frac{(c + a^n_j)}{2} \) and \( D^n_j = \frac{(c - a^n_j)}{2} \). So

\[
u^n_{j+1} - u^n_{j+1} = (1 - C^n_j - D^n_j) (u_j^n - u_{j-1}^n) + C^n_{j-1} (u_{j-1}^n - u_{j-2}^n) + D^n_j (u_{j+1}^n - u_j^n) + C^n_{j-1} (u_{j-1}^n - u_{j-2}^n) + D^n_j (u_{j+1}^n - u_j^n)
\]

Under the restricted CFL condition of the proposition, one has

\[
0 \leq \min \left( C^n_{j-1}, D^n_j, 1 - C^n_j - D^n_{j-1} \right).
\]

So

\[
\sum_j |u_{j+1}^{n+1} - u_{j-1}^n| \leq \sum_j (1 - C^n_j - D^n_{j-1}) |u_j^n - u_{j-1}^n| + \sum_j C^n_{j-1} |u_{j-1}^n - u_{j-2}^n| + \sum_j D^n_j |u_{j+1}^n - u_j^n| = \sum_j |u_j^n - u_{j-1}^n|.
\]

The second inequality comes from the fact that (88) yields

\[
|u_{j+1}^{n+1} - u_j^n| \leq \frac{\nu}{2} (c - a^n_j) |u_{j+1}^n - u_j^n| + \frac{\nu}{2} (c + a^n_j) |u_j^n - u_{j-1}^n| \leq c |u_{j+1}^n - u_j^n| + c |u_j^n - u_{j-1}^n|.
\]

The end of the proof is evident. \( \square \)

**3.5.4 Convergence**

The question addressed now is the convergence of the discrete solution to a certain limit, as the mesh parameters \( \Delta x \) and \( \Delta t \) tend to zero. We can split the answer to this question in two subquestions. A first one is the identification of the limit. That is if we assume that the discrete solution admits a limit, can we identify the equations and in equations satisfied by the limit: the answer is not
trivial, in particular because non linear equations may have multiple solutions. The second subquestion concerns the reason why the discrete solution admits a limit. For scalar conservation laws, it can be proved. This the reason why concentrate on the properties of the limit.

Define a function $u_{\Delta x, \Delta t} : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ by

$$u_{\Delta x, \Delta t}(t, x) = u^n_j$$

where $(j - \frac{1}{2})\Delta x < x < (j + \frac{1}{2})\Delta x$ and $n\Delta t \leq t < (n + 1)\Delta t$. Some of these strict inequalities can be replaced by large inequalities.

**Theorem 6.** Consider $u_{\Delta x, \Delta t}$ provided by the scheme (83-85) with an initial data $u_0 \in L^\infty(\mathbb{R})$. Assume the CFL condition of proposition 14. Assume there exists $u \in L^1_{\text{loc}}(\mathbb{R}^+ \times \mathbb{R})$ such that

$$\lim_{(\Delta x, \Delta t) \to 0} \|u_{\Delta x, \Delta t} - u\|_{L^1_{\text{loc}}(\mathbb{R}^+ \times \mathbb{R})} = 0$$

and

$$\lim_{(\Delta x, \Delta t) \to 0} \|u_{\Delta x, \Delta t}(0, \cdot) - u_0\|_{L^1_{\text{loc}}(\mathbb{R})} = 0.$$ 

Assume the parameter $c$ is such that

$$c\Delta x \to 0 \text{ for } \Delta x \to 0.$$ (93)

Then $u \in L^\infty(\mathbb{R}^+ \times \mathbb{R})$ is an entropy weak solution.

### 3.6 Schemes for traffic flow equation

#### 3.6.1 Eulerian LWR traffic flow

Consider the Eulerian LWR traffic flow equation $\partial_t \rho + \partial_x (\rho - \rho^2) = 0$ with the initial data

$$\rho_0 = 0.4 \text{ for } x < 0.3,$$

$$\rho_0 = 1. \text{ for } 0.3 < x < 0.7,$$

$$\rho_0 = 0. \text{ for } 0.7 < x.$$ (94)

The numerical solution is computed with the scheme (83)-(84)-(85). We take three meshes, 100, 200 and 400 cells on the interval $[0, 1]$. The constant of the flux is $c = 2$. The CFL condition writes $c\Delta t = \frac{1}{10}\Delta x$.

The results are displayed in figure 12 at the final time $t = 0.2$. Accordingly to the theory, the solution is made of a rarefaction for the rightmost part, which characterizes the exit from the traffic jam. The leftmost part is a shock, typical of the entry in a traffic jam. We observe the numerical solution satisfies the entropy condition.

However it must be noticed that wrong solutions show up easily if the discrete entropy inequalities are not satisfied. To illustrate this phenomenon consider the following initial data

$$u_0 = 1 \text{ for } 0.4 < x < 0.6, \quad u_0 = 0 \text{ elsewhere.}$$
Moreover instead of computing correctly \( A = \max_{0 \leq u \leq 1} |f'(u)| \) which takes the value \( A = 1 \), we replace \( A \) with the lazy prediction is

\[
B = \max_{u_j \neq u_{j+1}} \frac{|f(u_{j+1}) - f(u_j)|}{u_{j+1} - u_j}.
\]

Even worse we privilege implementation considerations and use

\[
B_\varepsilon = \max_{|u_j - u_{j+1}| \geq \varepsilon > 0} \frac{|f(u_{j+1}) - f(u_j)|}{u_{j+1} - u_j},
\]

where \( \varepsilon > 0 \) is a purely artificial number, so as to avoid division by zero. Let us take for example \( \varepsilon = 10^{-6} \). If one uses this value, it is easy to check that the numerical valuation gives the numerical value \( B_\varepsilon = 0 \).

Let us now assume that one uses the scheme with a CFL constant equal to 1 and the numerical parameter in the flux \( c = B_\varepsilon = 0 \). With this parameters one obtains the result of figure 13 where the numerical solution is stable in \( L^\infty \).

Notice that the rightmost discontinuity is not admissible from the entropy perspective since \( 0 = u_R < u_L = 1 \). Since the numerical solution is independent of the mesh parameters, the Lax-Wendroff theorem can be invoked to prove the convergence to a weak solution as \( \Delta x \to 0 \). This weak limit solution is not an entropic weak solution.

### 3.6.2 Lagrangian LWR traffic flow

Start the discussion from the Lagrangian scheme

\[
\frac{\tau_j^{n+1} - \tau_j^n}{\Delta t} - \frac{\dot{u}_j^n + \dot{u}_j^n}{\Delta m_j} = 0
\]

Figure 12: Entry in a traffic jam and exit. Numerical convergence for 100, 200 and 400 cells.
with a definition of the mass
\[ \Delta m_j = \rho_j^0 \Delta X_j, \quad \Delta X_j = X_{j+\frac{1}{2}} - X_{j-\frac{1}{2}} = \Delta x_j^0. \]

The flux velocity can be either upwinded
\[ u_{j+\frac{1}{2}}^n = u_{j+1}^n \]
because the sound velocity is always negative, or given by the more general procedure (83)-(84)-(85). A natural definition of the mesh displacement is
\[ x_{j+\frac{1}{2}}^{n+1} = x_{j+\frac{1}{2}}^n + \Delta t u_{j+\frac{1}{2}}^n, \quad x_{j+\frac{1}{2}}^0 = X_{j+\frac{1}{2}}. \tag{96} \]

The discrete Jacobian of the transformation is
\[ J_j^n = \frac{x_{j+\frac{1}{2}}^n - x_{j-\frac{1}{2}}^n}{x_{j+\frac{1}{2}}^0 - x_{j-\frac{1}{2}}^0}. \]

**Proposition 17.** The Lagrangian scheme in mass variable is equivalent to the discretization of the system
\[
\begin{cases}
\frac{\rho_{j+1}^{n+1}}{\rho_{j+1}^n} J_{j+\frac{1}{2}}^{n+1} - \frac{\rho_{j}^{n}}{\rho_{j}^{n}} J_{j-\frac{1}{2}}^{n} = 0, \\
\frac{J_{j+1}^{n+1} - J_{j}^{n}}{\Delta t} - \frac{u_{j+\frac{1}{2}}^n - u_{j-\frac{1}{2}}^n}{\Delta X_j} = 0.
\end{cases} \tag{97} \]

**Proof.** The mesh displacement is rewritten as
\[ x_{j+\frac{1}{2}}^{n+1} - x_{j-\frac{1}{2}}^{n+1} = x_{j+\frac{1}{2}}^n - x_{j-\frac{1}{2}}^n + \Delta t \left( u_{j+\frac{1}{2}}^n - u_{j-\frac{1}{2}}^n \right) \]
Divison by \[ x_{j+\frac{1}{2}}^0 - x_{j-\frac{1}{2}}^0 = \Delta X_j \] yields the second equation of (97).
Consider the first time step \( n = 0 \). Rewrite is as
\[
\rho_0^j \Delta x_j^0 \tau_j^1 - (\Delta x_j^0 + \Delta t (u_j^0 + \frac{1}{2} - u_j^0 - \frac{1}{2})) = 0
\]
that is \( \rho_j^1 \Delta x_j^0 - \Delta x_j^1 = 0 \). It yields the first equation of (97). The proof is ended after iteration on \( n \).

The analogy between Lagrangian discretization and particular methods is easy by accepting the natural idea that a vehicle or a group of vehicles is like a particle of a group of particles in kinetic equations. This is illustrated below

Figure 14: Interpretation of the Lagrangian scheme as a particle discretization.

A Lagrangian simulation on a mobile mesh is presented at the figure 15. The initial data is a slight perturbation of the data (94)
\[
\begin{align*}
\rho_0 & = 0.4 \text{ for } x < 0.3, \\
\rho_0 & = 1. \text{ for } 0.3 < x < 0.7, \\
\rho_0 & = 0.001 \text{ for } 0.7 < x.
\end{align*}
\]
The modification is done to allow the computation of the specific volume \( \tau = \frac{1}{\rho} \) for small but non zero density. This modification has no impact on the time step for this calculation.

An interpretation of the result is easy considering the visible mesh displacement. Indeed the cells are stretched in the rarefaction fan and are compressed at the shock. This feature can be compared with the Eulerian result of the figure (12) where no such mesh modification is possible of course. The numerical solution is quite similar even if the numerical methods are extremely different.

Note however the important dilatation of the mesh near the foot of the fan.

4 Compressible gas dynamics and the entropy

4.1 The entropy

Proposition 18. Let \((h, u)\) be a solution of the shallow water system (10) of class \( C^1 \) in space-time. Then an additional conservation law is satisfied
\[
\partial_t \left( gh^2 + hu^2 \right) + \partial_x \left( (2gh^2 + hu^2) u \right) = 0.
\]

(98)
Moreover, for \( g, h > 0 \), the function \( \eta(h, hu) = gh^2 + hu^2 \) is strictly convex with respect to \( h \) and \( hu \).

**Proof.** To obtain this relation we rewrite (10) as

\[
\begin{align*}
\partial_t h + u \partial_x h + h \partial_x u &= 0, \\
(h \partial_t u + u \partial_x u) + \partial_x \left( \frac{h}{2} h^2 \right) &= 0
\end{align*}
\]

with a use of the identity \( \partial_t hu + \partial_x hu^2 = h (\partial_t u + u \partial_x u) \). Multiply the first equation by \( gh \) and the second one by \( 2u \)

\[
\begin{align*}
gh (\partial_t h + u \partial_x h) + gh^2 \partial_x u &= 0, \\
h (\partial_t u^2 + u \partial_x u^2) + u \partial_x (gh^2) &= 0.
\end{align*}
\]

Use \( h (\partial_t u^2 + u \partial_x u^2) = \partial_t (hu^2) + \partial_x (hu^3) \) and \( gh(\partial_t h + u \partial_x h) = \partial_t gh^2 + \partial_x ghu^2 \), then sum up. It yields (98). Notice that the fact the solution \((h, hu)\) is \( C^1 \) with \( h > 0 \) is crucial.

Define \( \eta(a, b) = ga^2 + \frac{b^2}{a} \) such that \( \eta(h, hu) = gh^2 + hu^2 \). One has

\[
\nabla^2 \eta = 2 \begin{pmatrix} a^2 + \frac{b^2}{a^2} & -\frac{b}{a} \\
-\frac{b}{a} & -\frac{b}{a^2}
\end{pmatrix}.
\]

This is symmetric matrix with a positive trace and a positive determinant \( D = \frac{4b}{a} > 0 \). So \( \nabla^2 \eta > 0 \) which ends the proof of the strict convexity.

One says that \( \eta \) is an entropy for the shallow water equations. The Euler system of compressible gas dynamics satisfies a similar property.

**Proposition 19.** Smooth solutions of the Euler equations

\[
\begin{align*}
\partial_t \rho + \partial_x (\rho u) &= 0, \\
\partial_t (\rho u) + \partial_x (\rho u^2 + p) &= 0, \\
\partial_t (\rho e) + \partial_x (\rho e u + pu) &= 0,
\end{align*}
\]
with a perfect gas pressure law satisfy an additional conservation law
\[ \partial_t(\rho S) + \partial_x(\rho u S) = 0, \quad S = \log \varepsilon \tau^{\gamma-1} \quad \text{and} \quad \varepsilon = e - \frac{1}{2}u^2. \quad (100) \]

Proof. The proof is given for $\rho > 0$. Rewrite the Eulerian system as
\[
\begin{cases}
\rho D_t \tau - \partial_x u = 0, \\
\rho D_t u + \partial_x p = 0, \\
\rho D_t e + \partial_x pu = 0,
\end{cases}
\]
with $D_t = \partial_t + u \partial_x$ the material derivative. Let $(\tau, u, e)$ be a $C^1$-space-time solution with a perfect gas pressure law $p = (\gamma - 1)\rho \varepsilon = (\gamma - 1)\varepsilon$. One easily checks the differential identity
\[ TdS = d\varepsilon + pd\tau, \quad T = \varepsilon, \quad S = \log(\varepsilon \tau^{\gamma-1}). \quad (101) \]
Notice that $S$ is, physically speaking, a thermodynamical entropy per unit mass for a polytropic gas. The differential identity is clearly related to the fundamental principle of thermodynamics. The function $T$ is the temperature. Assume $T > 0$ which is physically relevant. One has also
\[ TdS = de - udu + pd\tau. \quad (102) \]
It yields for smooth functions $TD_t S = D_t e - uD_t u + pD_t \tau = 0$ which turns into
\[ D_t S = 0. \quad (103) \]
This expression implies $\partial_t(\rho S) + \partial_x(\rho u S) = \rho D_t S + S (\partial_t \rho + \partial_x (\rho u)) = 0$. The proof is ended. \hfill \Box

It can be proved that the function
\[ \langle \rho, \rho u, \rho e \rangle \mapsto \eta(\rho, \rho u, \rho e) \equiv -\rho S(\rho, \rho u, \rho e) \quad (104) \]
is strictly convex for $\rho > 0$ and $\varepsilon > 0$.
As suggested by the two previous examples, it is possible for systems like
\[ \partial_t U + \partial_x f(U) = 0, \quad U, f(U) \in \mathbb{R}^n \quad (105) \]
that a certain non linear function of the unknown satisfies an additional conservation law for all $C^1$ solutions. This additional equation writes
\[ \partial_t \eta(U) + \partial_x \xi(U) = 0, \quad \eta(U), \xi(U) \in \mathbb{R}. \quad (106) \]

**Definition 14 (Entropy of a system).** The pair $(\eta(U), \xi(U))$ is an entropy-entropy flux pair for the system of conservation laws (106) if and only if they satisfy the two properties

a) The function $\eta$ is strictly convex,
b) The compatibility relation between first derivatives holds
\[ \nabla \eta(U) \nabla f(U) = \nabla \xi(U). \quad (107) \]

**Definition 15 (Entropy variable).** Let \( (\eta(U), \xi(U)) \) be an entropy-entropy flux pair for the system of conservation laws (106). The column vector
\[ V = (\nabla_U \eta(U))^t \in \mathbb{R}^n \]
will be called the entropy variable.

Godunov made the extremely important remark that the sole existence of an entropy-entropy flux pair guarantees the linear stability, that is stability around constant states, as enunced in definition 2. This is of paramount importance since many systems coming from the physics are endowed with such a pair. This principle establishes a connection between the existence of an entropy-entropy flux pair which comes form the underlying physical context and the mathematical well-posedness of the linearized system.

**Theorem 7 (Godunov theorem).** Assume the system of conservation laws (106) is endowed with an entropy-entropy flux pair. Then it is hyperbolic.

Consider the system of conservation laws
\[ \partial_t U + \partial_x f(U) = 0, \quad U, f(U) \in \mathbb{R}^n, \quad (108) \]
endowed with an entropy-entropy flux pair. Assume the initial data is
\[ U(0, x) = U_0(x), \quad x \in \mathbb{R} \quad (109) \]
Consider the space-time domain \( \Omega \)
\[ \Omega = (-A, A) \times [0, T] \quad T, A > 0. \]
If \( U \in C^1(\Omega) \) and satisfies the initial condition (109), then one says it is a strong solution of the system (108) in \( \Omega \).

It is possible to relax the smoothness requirement: for example \( U \in C^1(\Omega) \) and piecewise \( C^1 \) is possible.

**Definition 16 (Weak entropic solution).** The bounded function \( U \in L^\infty(\Omega) \) is a weak solution of (108) with initial data (109) if and only if
\[ \int_{\mathbb{R}} \int_{0 < t} \left( (U, \partial_t \varphi) + (f(U), \partial_x \varphi) \right) dx \, dt + \int_{\mathbb{R}} (U_0(x), \varphi(0, x)) dx = 0 \quad (110) \]
for all functions \( \varphi \in (C^1_0(\Omega))^n \).

Of course, as it is the case for scalar conservation laws, a strong solution is also a weak solution, and a smooth weak solution is also a strong solution.
We start from a general definition of a viscous solution, so with a generic viscous tensor proportional to the identity matrix. This will be completed with the study of more physical viscous tensors at the end of the section. Consider \( U_\varepsilon \), a sequence of smooth \( C^2(\Omega) \) strong solutions of the system with evanescent viscosity

\[
\partial_t U_\varepsilon + \partial_x f(U_\varepsilon) = \varepsilon \partial_{xx} U_\varepsilon, \quad \varepsilon \to 0^+.
\]

Rewrite as \( \partial_t U_\varepsilon + (\nabla U \cdot f(U_\varepsilon)) \partial_x U_\varepsilon = \varepsilon \partial_{xx} U_\varepsilon, \varepsilon \to 0^+ \). One is interested in the calculation of the variation of entropy \( \partial_t \eta(U_\varepsilon) \). Take the scalar product against \( \nabla U \eta(U_\varepsilon) \)

\[
\partial_t \eta(U_\varepsilon) + \partial_x \xi(U_\varepsilon) = \varepsilon ((\nabla U \eta(U_\varepsilon)), \partial_x U_\varepsilon),
\]

that is

\[
\partial_t \eta(U_\varepsilon) + \partial_x \xi(U_\varepsilon) = \varepsilon \partial_{xx} \eta(U_\varepsilon) - \varepsilon (\partial_x U_\varepsilon, (\nabla^2 U \eta(U_\varepsilon)) \partial_x U_\varepsilon)
\]

and

\[
\partial_t \eta(U_\varepsilon) + \partial_x \xi(U_\varepsilon) \leq \varepsilon \partial_{xx} \eta(U_\varepsilon).
\]

Notice the compatibility relation (107) is used to assemble \( \partial_x \xi(U_\varepsilon) \), and the convexity of the entropy is used to obtain the inequality \( (\partial_x U_\varepsilon, (\nabla^2 U \eta(U_\varepsilon)) \partial_x U_\varepsilon) \geq 0 \). It remains to multiply by a smooth non negative function with compact support \( \varphi \in C^1_{0,+} \) and to integrate by parts as in the algebra (69). Standard assumptions (67-68) are used to formalize the fact that \( U_\varepsilon \) tends to the limit \( U \).

**Definition 17** (Weak entropy solutions). Let \( U \) be a weak solution \((110)\). One says it is an entropy weak solution if and only if

\[
- \int_{\mathbb{R}} \int_{t<0} (\eta(U) \partial_t \varphi + \xi(U) \partial_x \varphi) \, dx \, dt - \int_{\mathbb{R}} \eta(U_0)(x) \varphi(0, x) \, dx \leq 0, \quad \forall \varphi \in C^1_{0,+}.
\]

The abstract viscosity tensor \((111)\) is mathematically convenient but has no physical foundations. It is necessary to complement the previous analysis of viscous solutions by at least one example which shows the generality of the definition 14. The example is based on the compressible Euler system with viscosity and thermic dissipation

\[
\begin{align*}
\partial_t \rho_{\nu,\kappa} + \partial_x (\rho_{\nu,\kappa} u_{\nu,\kappa}) &= 0, \\
\partial_t (\rho_{\nu,\kappa} u_{\nu,\kappa}) + \partial_x (\rho_{\nu,\kappa} u_{\nu,\kappa}^2 + p_{\nu,\kappa}) &= \nu \partial_{xx} u_{\nu,\kappa}, \\
\partial_t (\rho_{\nu,\kappa} e_{\nu,\kappa}) + \partial_x (\rho_{\nu,\kappa} u_{\nu,\kappa} e_{\nu,\kappa} + p_{\nu,\kappa} u_{\nu,\kappa}) &= \nu \partial_x (u_{\nu,\kappa} \partial_x u_{\nu,\kappa}) + \kappa \partial_{xx} T_{\nu,\kappa}.
\end{align*}
\]

The term \( \nu \partial_{xx} u_{\nu,\kappa} \) is the viscosity which is scaled by a viscous parameter \( \nu > 0 \). Note the viscosity as a counterpart in the right hand side of the energy equation. On the other hand the thermic dissipation \( \partial_{xx} T_{\nu,\kappa} \) is scaled by \( \kappa > 0 \) and shows in just one equation and is scaled by the same factor. Both therm is evanescent for \( \nu, \kappa \to 0^+ \). Rewrite the system with the material derivative \( D_t = \partial_t + u_{\nu} \partial_x \)

\[
\begin{align*}
\rho_{\nu} D_t \tau_{\nu,\kappa} - \partial_x u_{\nu,\kappa} &= 0, \\
\rho_{\nu,\kappa} D_t (u_{\nu,\kappa} + \partial_x p_{\nu,\kappa}) &= \nu \partial_{xx} u_{\nu,\kappa}, \\
\rho_{\nu,\kappa} D_t e_{\nu,\kappa} + \partial_x (p_{\nu,\kappa} u_{\nu,\kappa}) &= \nu \partial_x (u_{\nu,\kappa} \partial_x u_{\nu,\kappa}) + \kappa \partial_{xx} T_{\nu,\kappa}.
\end{align*}
\]

53
Let us admit for simplicity a perfect gas pressure law for which the entropy is (101). It yields
\[
\rho \nu,\kappa \frac{\partial}{\partial t} S_{\nu,\kappa} = \rho \nu,\kappa p_{\nu,\kappa} \frac{\partial}{\partial x} u_{\nu,\kappa} - \rho \nu,\kappa u_{\nu,\kappa} \frac{\partial}{\partial x} p_{\nu,\kappa} + \rho \nu,\kappa \frac{\partial}{\partial x} \kappa
\]
\[
= p_{\nu,\kappa} \frac{\partial}{\partial x} u_{\nu,\kappa} - u_{\nu,\kappa} (\nu \partial_{xx} u_{\nu,\kappa} - \partial_x p_{\nu,\kappa}) + \nu \partial_x (u_{\nu,\kappa} \partial_x u_{\nu,\kappa}) - \partial_x (p_{\nu,\kappa} u_{\nu,\kappa})
\]
\[
= \nu (\partial_x u_{\nu,\kappa})^2 + \kappa \partial_{xx} T_{\nu,\kappa}.
\]
One obtains after dividing by the temperature and rearrangement of the right hand side
\[
\rho \nu,\kappa \frac{\partial}{\partial t} S_{\nu,\kappa} = \left( \frac{\nu}{T_{\nu,\kappa}} (\partial_x u_{\nu,\kappa})^2 + \kappa \frac{\partial_x T^2_{\nu,\kappa}}{T_{\nu,\kappa}} \right) + \kappa \partial_{xx} \ln T_{\nu,\kappa} \tag{114}
\]
which can be rewritten in a eulerian form
\[
\frac{\partial}{\partial t} (\rho S_{\nu,\kappa}) + \frac{\partial}{\partial x} (\rho u S_{\nu,\kappa}) \geq \kappa \partial_{xx} \ln T_{\nu,\kappa}.
\]
Noticing that \(\kappa \partial_{xx} \ln T_{\nu,\kappa}\) tends to zero in the sense of distributions, after passing to the limit \((\nu, \kappa) \to 0\) one obtains \(\partial_t (\rho S) + \partial_x (\rho u S) \geq 0\) in the sense of distribution. This is exactly (112) taking \(\eta = -\rho S\) and \(\xi = -\rho u S\).

### 4.2 Entropy and Lagrangian systems

Some features of Eulerian and Lagrangian systems of conservation laws in continuum mechanics are common. For example both satisfy Galilean invariance principles. Both obey to the general theory of hyperbolic systems. In particular discontinuous solutions such as shocks or contact discontinuities must be included in the theory. There is nevertheless an important property which is different in Euler or Lagrange coordinates, which is related to the entropy law. Indeed many of the systems of conservation laws in continuum mechanics are endowed with a density equation and an entropy equation (for smooth solutions) which are very similar
\[
\partial_t \rho + \partial_x (\rho u) = 0 \quad \text{and} \quad \partial_t (\rho S) + \partial_x (\rho u S) = 0.
\]
In Lagrange coordinates, that is in mass variable in dimension one, a system like
\[
\partial_t U + \partial_m f(U) = 0 \tag{115}
\]
is often endowed with a physical entropy relation
\[
\partial_t S = 0 \text{ with a zero entropy flux.}
\]
That is the mathematical entropy is \(\eta(U) = -S\) and the entropy flux is \(\xi(U) = 0\). This property is, when considering examples, so general that it can even become a definition.
Hypothesis 1 (Lagrangian systems and zero entropy flux). One property of Lagrangian systems of conservation laws is the zero entropy flux: $\xi(U) \equiv 0$. It is characterized by

$$\nabla \eta(U) \nabla f(U) = 0, \quad \forall U. \quad (116)$$

This relation admits a very strong physical meaning if one considers that the physical entropy measures some degree of a irreversibility in the physical process sustained by the matter. Irreversible processes correspond to a strict increase of the local entropy, while the entropy is constant for reversible processes. In the case of compressible gas dynamics, one can compare (100) which expresses a reversibility for smooth solutions to (114) which expresses the possibility of an irreversible process.

Figure 16: Irreversible process: $S_1 < S_2 < S_3 < S_4$. Reversible process $S_1 = S_2 = S_3 = S_4$.

We examine some consequences of the hypotheses of a zero entropy flux.

**Proposition 20.** Assume the entropy flux is zero (116). Then the flux $f(U)$ can be expressed as an homogenous function of degree zero with respect to the entropy variable $V$.

**Proof.** The definition of a homogeneous function of degree $p$ is that $g(\lambda x) = \lambda^p g(x)$ for all $\lambda \in \mathbb{R}$. If $g$ is moreover differentiable, one has the equivalent characterization with the Euler relation

$$\nabla_x g(x) \cdot x = pg(x). \quad (117)$$
To prove the claim, note that $V = \nabla \eta(U)^t = -\nabla S(U)^t$. Using the chain rule
\[ \nabla_U(f(U(V))) = \nabla_V(f(U(V))) \nabla_U V, \]
the relation (116) is equivalent to
\[ \nabla_U \eta(U) \nabla_V(f(U(V))) \nabla_U V = 0 \]
that is $\nabla_U \eta(U) \nabla_V(f(U(V))) = 0$ after simplification by the non singular matrix $\nabla_U V$. A transposition yields
\[ [\nabla_V(f(U(V))]^t V = 0. \]
But $f(U(V)) = \nabla_V \xi^*(V)$ where $\xi^* = (V, f(U)) - \xi(U) = (V, f(U))$ is the polar transform of the entropy flux, the matrix $\nabla_V(f(U(V)))$ is symmetric. One obtains the Euler relation
\[ [\nabla_V(f(U(V))] V = 0. \]
Therefore the flux $f$ is homogeneous of degree $p = 0$ with respect to $V$. 

One can make the additional assumption that the last component of the entropy vector $V$ is non zero, that is $V_n \neq 0$. It will be shown later on that it is a generic assumption. In this case one can express the flux as $f(U) = g(\Psi)$ where $g : \mathbb{R}^{n-1} \rightarrow \mathbb{R}^n$ is some function of the variable
\[ \Psi = \left( \frac{V_1}{V_n}, \frac{V_2}{V_n}, \ldots, \frac{V_{n-1}}{V_n} \right)^t \in \mathbb{R}^{n-1}. \]
(118)
To be even more specific let us consider the system of Lagrangian gas dynamics
\[
\begin{align*}
\frac{\partial t}{\tau} - \partial_m u &= 0, \\
\partial_t u + \partial_m p &= 0, \\
\partial_t e + \partial_m (pu) &= 0.
\end{align*}
\]
(119)
for which the differential relation (102) shows that $V = -\frac{1}{\tau}(p, -u, 1)^t$. The sign $-$ is because $\eta = -S$. Inserting in (118) one gets $\Psi = (p, -u)^t$. The flux is a function of two independent variables, which are the pressure $p$ and the opposite of the velocity $-u$.

Remark 8. At inspection of (119) the structure of the flux is extremely simple, in the sense that it is only a linear-quadratic function of $\Psi = (p, -u)^t$.

In order to show that the linear-quadratic structure of the flux is a general property, one can rely on certain Galilean invariance properties. To do so, we assume that the main unknown $U$ can be decomposed in three parts as in
\[ U = \begin{pmatrix} \mathbf{v} \in \mathbb{R}^{n-d-1} \\ \mathbf{u} \in \mathbb{R}^d \\ \mathbf{e} \in \mathbb{R} \end{pmatrix} \in \mathbb{R}^n \]
(120)
where the vector $\mathbf{v}$ regroups density-like variables, $\mathbf{u}$ regroups velocity-like variable and $\mathbf{e}$ is the total energy. The internal energy is
\[ \varepsilon = \mathbf{e} - \frac{1}{2} |\mathbf{u}|^2. \]
Notice that the necessity of having a vectorial velocity variable is easily understood considering the one dimensional system
\[
\begin{align*}
\frac{\partial \tau}{\partial t} - \partial_m u &= 0, \\
\frac{\partial u}{\partial t} + \partial_m p &= 0, \\
\frac{\partial v}{\partial t} &= 0, \\
\frac{\partial e}{\partial t} + \partial_m (pu) &= 0.
\end{align*}
\]

(121)

for which \( u = (u, v) \).

**Hypothesis 2** (Lagrangian Galilean invariance). Let \( U \) be any smooth solution of the system of conservation laws (115) having the structure (120). Galilean invariance expresses that for all \( u_0 \in \mathbb{R}^d \)
\[ U_{u_0} = (v, u + u_0, \varepsilon + \frac{1}{2} |u + u_0|^2)^t \]
is also smooth solution of (115).

**Hypothesis 3** (Lagrangian reversibility). Let \( U \) be a smooth solution of the system of conservation laws (115) having the structure (120). Reversibility means that \( U_- = (v, -u, e)^t \) is also solution of
\[ -\partial_t U_- + \partial_m f(-U_-) = 0. \]

**Hypothesis 4** (Entropy is function of \( v \) and \( \varepsilon \)). The physical entropy \( S \) can be written as a function of the variables \( v \) and \( \varepsilon \), and so the mathematical entropy can be written as
\[ \eta(v, u, e) = -S(v, e - \frac{1}{2} u^2). \]
Moreover we assume that \( V_n < 0 \) which generalizes the case of compressible gas dynamics since in this case \( V_n = -\frac{1}{T} \).

**Definition 18** (Precise definition of \( \Psi \)). Assume 4 and define
\[ w(v, \varepsilon) = \nabla_v S(v, \varepsilon) \nabla_\varepsilon S(v, \varepsilon)^{-1} \]
(122)

It yields the representation of the reduced entropy variable (118)
\[ \Psi = (w(v, \varepsilon), -u)^t \in \mathbb{R}^{n-1}. \]

**Theorem 8** (Representation formula of the flux). Consider the system of conservation laws (115) and make the assumptions 1, 2, 3 and 4.

Then the flux of the system is linear-quadratic in \( \Psi \). More precisely there exists a symmetric matrix \( M = M^t \in \mathbb{R}^{n-1 \times n-1} \) such that
\[ f(U) = \left( \begin{array}{c} M \Psi \\ -\frac{1}{2} (\Psi, M \Psi) \end{array} \right). \]
(123)

Moreover there exists a matrix \( N \in \mathbb{R}^{n-1-d \times d} \) such that
\[ M = \left( \begin{array}{cc} 0 & N^t \\ N & 0 \end{array} \right). \]
(124)

By identification the Lagrangian system of gas dynamics (119) corresponds to
\[ M = \left( \begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right) \] and \( \Psi = \left( \begin{array}{c} p \\ -u \end{array} \right). \)

That is \( N = 1 \). Therefore the theorem means that the structure of many Lagrangian systems coming from continuum mechanics have the same structure.
Proof of the representation theorem 8. The proof is done in three steps.

Step 1. Using hypothesis 4, we begin to write \( f(U) = \begin{pmatrix} h(\Psi) \in \mathbb{R}^{n-1} \\ g(\Psi) \in \mathbb{R} \end{pmatrix} \) where \( \Psi = (w(v, \varepsilon), -u)^t \in \mathbb{R}^{n-1} \). So the \( n-1 \) first equations of the system of conservation laws are

\[
\partial_t \begin{pmatrix} v \\ u \end{pmatrix} + [\nabla \Psi h(w, -u)] \partial_m \begin{pmatrix} w \\ -u \end{pmatrix} = 0.
\]

Step 2. The hypothesis 2 allows to write

\[
\partial_t \begin{pmatrix} v \\ u + u_0 \end{pmatrix} + [\nabla \Psi h(w, -u - u_0)] \partial_m \begin{pmatrix} w \\ -u - u_0 \end{pmatrix} = 0, \quad \forall u_0 \in \mathbb{R}^d.
\]

That is

\[
\partial_t \begin{pmatrix} v \\ u \end{pmatrix} + [\nabla \Psi h(w, -u)] \partial_m \begin{pmatrix} w \\ -u \end{pmatrix} = 0, \quad \forall u_0 \in \mathbb{R}^d.
\]

So the matrix \([\nabla \Psi h(w, -u)]\) is independent of \( u_0 \)

\[
\nabla_{u|w} [\nabla \Psi h(w, -u)] = \nabla \Psi [\nabla_{u|w} h(w, -u)] = 0.
\]

Therefore the matrix \( \nabla_{u|w} h(w, -u) \) is independent of \( \Psi \): one can write \([\nabla_{u|w} h(w, -u)] = B \) for a given matrix \( B \in \mathbb{R}^{n-1 \times d} \). Integration yields \( h(w, -u) = -Bu + l(w) \) where the function \( w \mapsto l(w) \) is unknown at this stage of the analysis. A technique of separation of variables yields

\[
h(w, -u) = \begin{pmatrix} -B_1 u + l_1(w) \in \mathbb{R}^{n-1-d} \\ -B_2 u + l_2(w) \in \mathbb{R}^d \end{pmatrix}.
\]

Step 3. The hypothesis 3 implies

\[
\pm \partial_t \begin{pmatrix} v \\ u \end{pmatrix} + \partial_m \begin{pmatrix} \pm B_1 u + l_1(w) \\ \pm B_2 u + l_2(w) \end{pmatrix} = 0.
\]

Therefore \( l_1 \equiv 0 \) and \( B_2 \equiv 0 \). More precisely \( \partial_m l_1 = 0 \) so one can eliminate \( l_1 \). A similar relation holds for \( B_2 \). So one can write

\[
f(U) = \begin{pmatrix} -B_1 u \\ l_2(w) \\ g(\Psi) \end{pmatrix}.
\]

Step 4. The polar transform of the entropy flux is

\[
\xi^*(V) = (V, f(U)) = V_n [-\langle w, B_1 u \rangle - \langle u, l_2(w) \rangle + g(w, -u)].
\]

One has that \( \nabla_V \xi^* = f(U) \). Let us differentiate with respect to \((V_1, \cdots, V_{n-1-d}) = V_n \) \( w \) and after that with respect to \((V_{n-d}, \cdots, V_{n-1}) = -V_n u \). One finds

\[
\begin{cases} 
- B_1 u - (\nabla w l_2(w))^t u + \nabla w g(w, -u) = -B_1 u, \\
B_1^t w + l_2(w) + \nabla u g(w, -u) = l_2(w).
\end{cases}
\]

So

\[
\begin{cases} 
\nabla w g(w, -u) = (\nabla w l_2(w))^t u \\
\nabla u g(w, -u) = -B_1^t w.
\end{cases}
\]

The Maxwell relation for cross derivatives yields

\[
\nabla u (\nabla w l_2(w))^t u = (\nabla w (-B_1^t w))^t
\]

that is \( \nabla w l_2(w) = B_1^t \) which solution writes \( l_2(w) = B_1^t w \). Furthermore \( g(w, -u) = (u, B_1^t w) \). One can characterize this result for the flux using the notation \( N = B_1 \)

\[
f(U) = \begin{pmatrix} M \Psi \\ -\frac{1}{2} [\Psi, M \Psi] \end{pmatrix} \quad \text{with} \quad M = \begin{pmatrix} 0 & N^t \\ N & 0 \end{pmatrix}.
\]

The proof is ended. \( \square \)
Next we study a first consequence of the structure (123-124) on the eigenstructure of the Jacobian matrix. Consider the Lagrangian system with the flux (125)

\[
\begin{align*}
\partial_t v - N \partial_m u & = 0, \\
\partial_t u + N^t \partial_m w & = 0, \\
\partial_t e + \partial_m (w, Nu) & = 0.
\end{align*}
\]

(126)

In view of the computation of some properties of the wave velocities, one can think of using proposition 3 with the variable \(W = (w, u, S)^t\). Since \(\partial_t S = 0\) one obtains the quasi-linear system

\[
\begin{align*}
\partial_t w - [\nabla_{w[S]} v]^{-1} N \partial_m u & = 0, \\
\partial_t u + N^t \partial_m w & = 0, \\
\partial_t S & = 0.
\end{align*}
\]

(127)

So the wave velocities are the eigenvectors of the matrix

\[
B = \begin{pmatrix} 0 & -[\nabla_{w[S]} v]^{-1} N & 0 \\ N^t & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{n \times n}.
\]

Let us write a generic eigenvector as \((a, b, c) \in \mathbb{R}^{n-1-d} \times \mathbb{R}^d \times \mathbb{R}\)

\[
\begin{align*}
- [\nabla_{w[S]} v]^{-1} N b & = \mu a, \\
N^t a & = \mu b, \\
0 & = \mu c.
\end{align*}
\]

(128)

One obtains a first trivial eigenvector \((a, b, c) = (0, 0, 1)\) for the eigenvalue \(\mu = 0\). Other null eigenvectors are described in the following proposition.

**Proposition 21** (Zero spectrum and non strict hyperbolicity). If \(n - 1 - d \neq d\) then the system (126) is not strictly hyperbolic and the null eigenvalue has multiplicity \(p \geq 1 + \max(n - 1 - 2d, 2d + 1 - n) \geq 2\).

**Proof.** The size of \(N\) is \(n - 1 - d \times d\). So the hypothesis of the proposition implies that \(N\) is in general a rectangular matrix. The rank of \(N^t\) is \(\text{rk}(N) \leq \min(n - 1 - d, d)\): therefore \(N\) and \(N^t\) admit a non trivial eigenvector for the null eigenvalue. It means that there exists a non trivial eigenvector of \(B\) for the null eigenvalue. With the addition of the trivial eigenvector \((0, 0, 1)\), it generates an eigenspace for the null eigenvector of dimension \(\geq 2\). The proof is ended. \(\Box\)

Assume for example the size of the velocity variable is \(d = 1\). It corresponds to a system which can be considered as truly one-dimensional. Then the system cannot be strictly hyperbolic for \(n \geq 4\). This is already the case for the example (121) for which \(n = 4\) and \(d = 1\).

In certain physical cases, it is necessary to introduce a modification of the general structure (123) and (124). Indeed a careful examination often shows that
some assumptions of theorem 8 may miss, which results in subtle modifications of the structure of the matrix $M$. This is the reason to focus on systems like

$$\partial_t U + \partial_m \left( \begin{pmatrix} M \Psi \\ - \frac{1}{2} (\Psi, M \Psi) \end{pmatrix} \right) = 0. \quad (129)$$

The only assumption on $M \in \mathbb{R}^{n-1}$ is it is a symmetric matrix $M = M^t$. The vector $\Psi \in \mathbb{R}^{n-1}$ is deduced from the entropy via the entropy variable $V = \nabla_U S$

$$\Psi_i = \frac{V_i}{V_n}, \quad 1 \leq i \neq n-1.$$

We consider the three dimensional Eulerian MHD model written under conservative form

$$\begin{cases}
\partial_t \rho + \nabla \cdot \rho u = 0, \\
\partial_t (\rho u) + \nabla \cdot (\rho u \otimes u) + \nabla P - \nabla \cdot \frac{B \otimes B}{\mu} = 0, \\
\partial_t B + \nabla \cdot (u \otimes B) = 0, \\
\partial_t e + \nabla \cdot (\rho e + P u - \frac{1}{\mu} B(B, u)) = 0.
\end{cases} \quad (130)$$

The magnetic field is $B \in \mathbb{R}^3$ and $u \otimes B$ designs the vectorial production of $u$ by $B$. The operator $\nabla \cdot$ is the curl operator in dimension $d = 3$. By definition $\mu = 4\pi$ and

$$P = p + \frac{1}{2\mu} |B|^2.$$

The total energy is the sum of the internal thermal energy, the kinetic energy and the magnetic energy

$$\rho e = \rho \varepsilon_{th} + \frac{1}{2} \rho u |u|^2 + \frac{1}{2\mu} |B|^2. \quad (131)$$

Correct solutions must satisfy the free divergence constraint on the magnetic field $\nabla \cdot B = 0$. Furthermore the magnetic equation implies that

$$\partial_t \nabla \cdot B = - \nabla \cdot (\nabla \cdot (u \otimes B)) = 0,$$

so, if the magnetic field is divergence free at $t = 0$, it remains free divergence. The natural physical entropy relation writes for smooth solutions

$$\partial_t \rho S + \nabla \cdot (\rho u S) = 0. \quad (132)$$

Make the hypothesis that the flow is invariant in directions $y$ and $z$

$$\partial_y = \partial_z = 0.$$

Let us decompose the magnetic field as $B = (B_x, B_y, B_z)$ with

$$\partial_t \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} + \begin{pmatrix} \partial_x \\ 0 \\ 0 \end{pmatrix} \wedge \begin{pmatrix} u_y B_z - u_z B_y \\ u_z B_x - u_x B_z \\ u_x B_y - u_y B_x \end{pmatrix} = 0$$

60
not forgetting the free divergence constraint

$$\partial_x B_x + \partial_y B_y + \partial_z B_z = \partial_x B_x = 0.$$  

So $B_x$ is constant in time and space $\partial_x B_x = \partial_t B_x = 0$. One can simplify

$$\begin{cases} 
\partial_t \rho + \partial_x (\rho u_x) = 0, \\
\partial_t \rho u_x + \partial_x (\rho u_x^2) - \frac{\mu}{\mu} B_x = 0, \\
\partial_t \rho u_y + \partial_y (\rho u_y B_y - \frac{\mu}{\mu} B_y) = 0, \\
\partial_t \rho u_z + \partial_z (\rho u_z B_z - \frac{\mu}{\mu} B_z) = 0, \\
\partial_t \rho e + \partial_x (\rho u_x + \partial_x \rho u_x B_x) = 0,
\end{cases}$$  

(133)

with

$$\tau = P - \frac{1}{\mu} B_x^2 p = \frac{1}{\mu} (-B_x^2 + B_y^2 + B_z^2).$$

It is striking to observe that the components of the magnetic field do play the same role in the total pressure $\tau$. The associated Lagrangian formulation is

$$\begin{cases} 
\partial_t \tau - \partial_m u_x = 0, \\
\partial_t \tau B_y - \partial_m u_y B_x = 0, \\
\partial_t \tau B_z - \partial_m u_z B_x = 0, \\
\partial_t u_x + \partial_m \tau = 0, \\
\partial_t u_y - \partial_m B_x B_y = 0, \\
\partial_t u_z - \partial_m B_x B_z = 0, \\
\partial_t e + \partial_m (\tau u_x - \frac{B_y}{\mu} (B_x u_x + B_z u_z)) = 0.
\end{cases}$$  

(134)

The Lagrangian entropy law for smooth solutions writes $\partial_t S = 0$ therefore one can use the general method to identify the entropic variable $V$ and the reduced entropic variable $\Psi$.

Starting from the fundamental law of thermodynamics $TdS = d\varepsilon + pd\tau$ and the definition $e$, one gets

$$TdS = de - u du + \tau d\tau - \frac{B_y}{\mu} dB_y - \frac{B_z}{\mu} dB_z.$$  

One obtains immediately

$$U = \begin{pmatrix} 
\tau \\
\tau B_y \\
\tau B_z \\
u_x \\
u_y \\
u_z \\
e
\end{pmatrix}, \quad V = \frac{1}{\tau} \begin{pmatrix} 
\tau \\
B_y \\
B_z \\
u_x \\
u_y \\
u_z \\
e
\end{pmatrix}, \quad \Psi = \begin{pmatrix} 
\frac{B_x}{\mu} \\
\frac{B_y}{\mu} \\
\frac{B_z}{\mu} \\
u_x \\
u_y \\
u_z
\end{pmatrix}.$$
Rewrite the Lagrangian system (134) as

$$
\partial_t U + \partial_m \left( -\frac{1}{2} (\Psi, M \Psi) \right) = 0,
$$

where the matrix $M$ is by identification

$$
M = M' = \begin{pmatrix}
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & B_x & 0 \\
0 & 0 & 0 & 0 & 0 & B_x \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & B_x & 0 & 0 & 0 & 0 \\
0 & 0 & B_x & 0 & 0 & 0
\end{pmatrix}.
$$

This matrix is indeed constant because the horizontal component of the magnetic field is constant $B_x$. The tangential components $B_y$ and $B_z$ are unknowns.

The wave velocity of slow waves, Alfvén waves and fast waves are

$$
\begin{align*}
\omega_s^2 &= \frac{1}{2} \left( a^2 - \sqrt{a^4 - 4c^2 B_x^2 \mu \rho} \right), \\
\omega_a^2 &= \frac{B_x^2 \mu \rho}{\mu \tau}, \\
\omega_f^2 &= \frac{1}{2} \left( a^2 + \sqrt{a^4 - 4c^2 B_x^2 \mu \rho} \right)
\end{align*}
$$

with the inequalities $\omega_s^2 \leq \omega_a^2 \leq \omega_f^2$. The slow and fast eigenvectors are

$$
s_s = \begin{pmatrix}
\varphi_s - \varphi_a \\
B_x B_z \\
B_x B_z
\end{pmatrix} \quad \text{and} \quad
s_f = \begin{pmatrix}
\varphi_f - \varphi_a \\
-\frac{B_x B_y}{\mu \tau} \\
-\frac{B_x B_z}{\mu \tau}
\end{pmatrix}.
$$

The waves velocities of the Eulerian system (133) are

$$
-c_f + u \leq -c_a + u \leq -c_s + u \leq u \leq c_s + u \leq c_a + u \leq c_s.
$$

A natural question is to extend in dimensions $d > 1$ some of the results for one-dimensional Lagrangian systems. It is enough to express the entropy principle for a large variety of multidimensional systems.

**Definition 19** (Multidimensional Lagrangian systems). Consider a system written in dimension $d \geq 1$ as

$$
\partial_t U + \sum_{1 \leq i \leq d} \partial_{x_i} \left( -\frac{1}{2} (\Psi, M_i \Psi) \right) = \begin{pmatrix}
M_0 \Psi \\
0
\end{pmatrix},
$$

where $U \in \mathbb{R}^n$ is the unknown, and $\Psi \in \mathbb{R}^{n-1}$ is a reduced entropy variable computed with the help of an entropy function $S : \mathbb{R}^n \to \mathbb{R}$ as follows: $V = \nabla U S$ and $\Psi_i = \frac{V_i}{V_n}$ for $1 \leq i \neq n - 1$ (assuming $V_n \neq 0$). Assume the matrices
$M_i \in \mathbb{R}^{n-1}$ are symmetric $M_i = M_i^t$ for $1 \leq i \leq d$. Assume the matrices, which are not necessarily constant, satisfy the compatibility relation

$$
\sum_{1 \leq i \leq d} \partial_X M_i = M_0 + M_0^t.
$$

Notice that the function $\eta(U) = -S(U)$ is not a mathematical entropy in the strict sense because we never stated the matrices $M_i$ are functions of $U$. The system (136) is not necessarily a closed one. The main asset of the structure (136) is the compatibility with the entropy law in the following sense.

**Theorem 9.** Assume the matrices $M_0, M_1, M_2, \cdots$ are smooth. Then smooth solutions of (136) satisfy

$$
\partial_t S = 0.
$$

**Proof.** Since the matrices are differentiable by hypothesis, smooth solutions can be written as

$$
\partial_t U + \sum_{1 \leq i \leq d} \left( -\frac{1}{2} (\Psi, M_i \partial X_i (\Psi)) - (\Psi, M_i \partial X_i (\Psi)) \right) = \begin{pmatrix} M_0 \Psi \\ 0 \end{pmatrix}.
$$

One has that

$$
\partial_t S = V_n \left( \begin{pmatrix} \Psi \\ 1 \end{pmatrix}, U \right),
$$

so

$$
\partial_t S + \frac{V_n}{2} \sum_i (\Psi, \partial X_i M_i \Psi) = V_n (\Psi, M_0 \Psi) = \frac{V_n}{2} (\Psi, [M_0 + M_0^t] \Psi).
$$

Therefore

$$
\partial_t S = -\frac{V_n}{2} (\Psi, [M_0 + M_0^t - \sum_i \partial X_i M_i] \Psi) = 0
$$

using the compatibility relation (137). The proof is ended. \qed

### 4.3 Numerical methods

Our goal is to analyze some numerical methods which apply to the discretization of

$$
\begin{align*}
\partial_t \rho + \partial_x (\rho u) + \partial_y (\rho v) + \partial_z (\rho w) &= 0, \\
\partial_t (\rho u) + \partial_x (\rho u^2 + p) + \partial_y (\rho uv) + \partial_z (\rho uw) &= 0, \\
\partial_t (\rho v) + \partial_x (\rho pv) + \partial_y (\rho v^2 + p) + \partial_z (\rho vw) &= 0, \\
\partial_t (\rho w) + \partial_x (\rho pw) + \partial_y (\rho vw) + \partial_z (\rho w^2 + p) &= 0, \\
\partial_t (\rho e) + \partial_x (\rho we + pu) + \partial_y (\rho ve + pv) + \partial_z (pwe + pw) &= 0.
\end{align*}
$$

The total energy is the sum of the internal energy and kinetic energy $e = \varepsilon + \frac{1}{2}(u^2 + v^2 + w^2)$. The pressure is assumed to be a function of the density and internal energy $p = p(\rho, \varepsilon)$. One fundamental assumption is that there exists a thermodynamical entropy $S$ with two properties

**a)** It is strictly concave with respect to $\varepsilon$ and $\tau = \frac{1}{\rho}$.

**b)** One has the fundamental principle of thermodynamics and there exists a function $T$, called temperature, such that

$$
TdS = d\varepsilon + p\,d\tau, \quad T > 0.
$$

63
The system (138) is invariant by rotation of the frame. Actually this principle is widely used for the development of numerical methods. For example consider the discretization of

\[
\begin{align*}
\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) &= 0, \\
\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u}^2 + p) + \nabla \cdot (\rho \mathbf{u} \mathbf{v}) &= 0, \\
\frac{\partial (\rho \mathbf{v})}{\partial t} + \nabla \cdot (\rho \mathbf{u} \mathbf{v}) + \nabla \cdot (\rho \mathbf{v}^2 + p) &= 0, \\
\frac{\partial (\rho \mathbf{e})}{\partial t} + \nabla \cdot (\rho \mathbf{u} \mathbf{e} + p \mathbf{u}) + \nabla \cdot (\rho \mathbf{v} \mathbf{e} + p \mathbf{v}) &= 0.
\end{align*}
\]

(140)

on the Cartesian mesh of figure 17. All cells of the mesh can be referenced with two scalar indices \(j, k \in \mathbb{Z}\). We will consider a directional splitting technique. That is we discretize firstly the system

\[
\begin{align*}
\frac{\partial \rho}{\partial t} + \nabla_x (\rho \mathbf{u}) &= 0, \\
\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla_x (\rho \mathbf{u}^2 + p) &= 0, \\
\frac{\partial (\rho \mathbf{v})}{\partial t} + \nabla_x (\rho \mathbf{u} \mathbf{v}) &= 0, \\
\frac{\partial (\rho \mathbf{e})}{\partial t} + \nabla_x (\rho \mathbf{u} \mathbf{e} + p \mathbf{u}) &= 0,
\end{align*}
\]

(141)

in the horizontal direction, that is on every horizontal line of the cartesian mesh. Secondly we solve the system

\[
\begin{align*}
\frac{\partial \rho}{\partial t} + \nabla_y (\rho \mathbf{v}) &= 0, \\
\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla_y (\rho \mathbf{u} \mathbf{v}) &= 0, \\
\frac{\partial (\rho \mathbf{v})}{\partial t} + \nabla_y (\rho \mathbf{v}^2 + p) &= 0, \\
\frac{\partial (\rho \mathbf{e})}{\partial t} + \nabla_y (\rho \mathbf{v} \mathbf{e} + p \mathbf{v}) &= 0,
\end{align*}
\]

(142)

in the vertical direction.
on all vertical line of the cartesian mesh. This splitting is done at every time step which is decomposed in two intermediate steps. The extension in dimension three is obvious.

4.4 Principle of a Lagrange+remap scheme in dimension one

The basic principle starts from

\[
\begin{align*}
\partial_t \tau - \partial_m u &= 0, \\
\partial_t u + \partial_m p &= 0, \\
\partial_t e + \partial_m (pu) &= 0,
\end{align*}
\]

(143)

which is a Lagrangian form of (141), with the mass variable defined by \(dm = \rho_0 dX = \rho dx\). One adds

\[
\partial_t x = u \quad \rho J = \rho(0, X)
\]

(144)

which express the displacement of the frame. Let us consider an initial mesh

\[\Delta x\]

\[\Delta t\]

Mesh at time step \(t_n = n\Delta t\)

Figure 18: Principle of a Lagrange+remap scheme in dimension \(d = 1\). Notice the different scenarios during the first Lagrange stage: cell number 1 translates to the left; cell number 2 expands; cells number 3 and 4 translate to the right; cell number 6 is compressed. A time step restriction is needed to guarantee that the cells do not cross. At the end of this Lagrangian stage, one gets cells number \(j'\) for \(j = 1, \ldots, 6\). The second stage consists in remapping the Lagrangian mesh at the end of the first stage on the initial eulerian mesh.

with mesh size \(\Delta x > 0\). In a first stage one solves the lagrangian system (143). In a second stage one uses the discrete equivalent of (144). It yields finally a consistent discretization of the eulerian system. The principle of the mesh displacement is depicted in figure 18.
4.5 Principle of an entropy Lagrangian solver

Consider a continuous-in-time finite volume Lagrangian scheme

\[
\begin{align*}
& m_j \tau_j^\prime(t) - u_{j+\frac{1}{2}}^* + u_{j-\frac{1}{2}}^* = 0, \\
& m_j u_j^\prime(t) + p_{j+\frac{1}{2}} - p_{j-\frac{1}{2}} = 0, \\
& m_j c_j^\prime(t) + p_{j+\frac{1}{2}} u_{j+\frac{1}{2}}^* - p_{j-\frac{1}{2}} u_{j-\frac{1}{2}}^* = 0,
\end{align*}
\] (145)

and let us assume the fluxes are chosen as

\[
\begin{align*}
& p_{j+\frac{1}{2}}^* = \frac{1}{2} \left( p_j + p_{j+1} \right) + \frac{(\rho c)_{j+\frac{1}{2}}}{2} (u_j - u_{j+1}), \\
& u_{j+\frac{1}{2}}^* = \frac{1}{2} (u_j + u_{j+1}) + \frac{1}{2(\rho c)_{j+\frac{1}{2}}} (p_j - p_{j+1}).
\end{align*}
\] (146)

These formulas will be justified later, however it is already clear that they are an extension of similar formulas for the scalar case (85). The system is in semi-discrete form, that is continuous in time. The important property is the following.

**Proposition 22.** The continuous-in-time scheme (145) satisfies the identity

\[
m_j T_j S_j^\prime(t) = \left( p_{j+\frac{1}{2}}^* - p_j \right) \left( u_{j+\frac{1}{2}}^* - u_j \right) + \left( p_{j-\frac{1}{2}}^* - p_j \right) \left( u_{j-\frac{1}{2}}^* - u_j \right).
\] (147)

With the fluxes (147) it satisfies the entropy inequality \( S_j^\prime(t) \geq 0 \).

**Proof.** The proof is only a matter of basic algebra using the fundamental principle of thermodynamics (139). Indeed one has

\[
m_j T_j S_j^\prime = m_j \left( e_j^\prime + p_j \tau_j^\prime \right) = m_j \left( e_j^\prime - u_j u_j^\prime + p_j \tau_j^\prime \right)
\]

\[
\begin{align*}
= & - \left( p_{j+\frac{1}{2}}^* u_{j+\frac{1}{2}}^* - p_{j-\frac{1}{2}}^* u_{j-\frac{1}{2}}^* \right) + u_j \left( p_{j+\frac{1}{2}}^* - p_{j-\frac{1}{2}}^* \right) + p_j \left( u_{j+\frac{1}{2}}^* - u_{j-\frac{1}{2}}^* \right) \\
= & \left[ -p_{j+\frac{1}{2}}^* u_{j+\frac{1}{2}}^* + u_j p_{j+\frac{1}{2}}^* + p_j u_{j+\frac{1}{2}}^* - p_j u_j \right] \\
+ & \left[ p_{j-\frac{1}{2}}^* u_{j-\frac{1}{2}}^* - u_j p_{j-\frac{1}{2}}^* - p_j u_{j-\frac{1}{2}}^* + p_j u_j \right] \\
= & \left[ \left( p_{j+\frac{1}{2}}^* - p_j \right) \left( u_{j+\frac{1}{2}}^* - u_j \right) \right] + \left[ \left( p_{j-\frac{1}{2}}^* - p_j \right) \left( u_{j-\frac{1}{2}}^* - u_j \right) \right].
\end{align*}
\]

It ends the proof of the first part of the claim.

The proof of the second part easily comes from the remark that the fluxes can be rewritten as the solution of the specific linear system

\[
\begin{align*}
& p_{j+\frac{1}{2}}^* - p_j + (\rho c)_{j+\frac{1}{2}} \left( u_{j+\frac{1}{2}}^* - u_j \right) = 0, \\
& p_{j-\frac{1}{2}}^* - p_{j+1} + (\rho c)_{j+\frac{1}{2}} \left( u_{j+\frac{1}{2}}^* - u_{j+1} \right) = 0.
\end{align*}
\] (148)

Inserting in (147) one obtains after elimination of the pressures

\[
m_j T_j S_j^\prime(t) = \left( (\rho c)_{j+\frac{1}{2}} \left( u_{j+\frac{1}{2}}^* - u_j \right) \right)^2 + \left( (\rho c)_{j-\frac{1}{2}} \left( u_{j-\frac{1}{2}}^* - u_j \right) \right)^2 \geq 0.
\]

The velocities can be eliminated as well with the same conclusion. The proof is ended. \(\square\)
One has also Proposition 23. Provided a local CFL condition is satisfied under the form \( \frac{\Delta t}{\Delta x} \leq \text{CFL} \), the fully discrete scheme is entropic \( S^L_j \geq S^e_j \).

Set the function \( g(s) = S(U^n + s(U^L - U^n)) \) which satisfies \( g(0) = S(U^n) \), \( g(1) = S(U^L) \) and

\[
\nabla S(U^L) = \frac{1}{T^L} (p^L - u^L, 1)^t.
\]

With a Taylor expansion, there exists \( 0 < s_* < 1 \) such that

\[
S(U^L) = S(U^n) - \nabla S(U^L) \cdot (U^L - U^n) - \frac{1}{2} \nabla^2 S(U^L + s_*(U^L - U^n)) \cdot (U^L - U^n).
\]

For time steps bounded \( \Delta t \leq \Delta t_* \), one has

\[
\nabla^2 S(U^L + s_*(U^L - U^n)) \cdot (U^L - U^n) \geq c_* |U^L - U^n|^2, \quad c_* > 0.
\]

One has also

\[
-\frac{T^L m^2}{\Delta t} \nabla S(U^L) \cdot (U^L - U^n) = u_j^L \left( p^L_{j+\frac{1}{2}} - p^L_{j-\frac{1}{2}} \right) + p^L_j \left( u^*_{j+\frac{1}{2}} - u^*_{j-\frac{1}{2}} \right)
\]

\[
- \left( p^*_{j+\frac{1}{2}} u^*_{j+\frac{1}{2}} - p^*_{j-\frac{1}{2}} u^*_{j-\frac{1}{2}} \right) - \left( p^L_j u^L_j - p^L_j u^L_j \right)
\]

\[
= - \left( p^*_j - p^L_j \right) \left( u^*_{j+\frac{1}{2}} - u^L_j \right) + \left( p^*_j - p^L_j \right) \left( u^*_{j-\frac{1}{2}} - u^L_j \right)
\]

An expansion yields

\[
\begin{align*}
p^*_j + \frac{1}{2} p^*_{j+1} - p^L_j - (pc)_j\cdot \left( u^*_{j+1} - u^L_j \right) \\
u^*_{j+\frac{1}{2}} - u^L_j = \frac{1}{2} \left( u^*_{j+1} - u^L_j - (pc)_j \cdot \left( u^n_{j+1} - u^L_j \right) \right) \\
+ \frac{1}{2} \left( u^n_j - u^L_j + (pc)_j \cdot \left( p^L_j - p^L_j \right) \right) \\
- \frac{1}{2(p^L_j - p^L_j)} \left( p^L_{j+1} - p^L_j \right) \left( u^n_{j+1} - u^n_j \right)
\end{align*}
\]

Using the identity \((a + b)(a - b) = a^2 - b^2\), one gets

\[
- \left( p^*_j + \frac{1}{2} p^L_j \right) \left( u^*_{j+\frac{1}{2}} - u^L_j \right) \geq - \frac{1}{2(p^L_j - p^L_j)} \left( p^L_{j+1} - p^L_j \right) \left( u^n_{j+1} - u^n_j \right).
\]

67
Similarly
\[
(p_j^* - p^{L}) (u_j^* - u_j^L) \geq - \frac{1}{2} (\rho c_{j - \frac{1}{2}} - (\rho c)_{j + \frac{1}{2}} (u_j^* - u_j^L))^2.
\]

For time steps bounded \( \Delta t \leq \Delta t_* \), one has
\[
\frac{T^L \tau^L_j}{\Delta t} \nabla S(U^L) \cdot (U^L - U^n) \leq d_* |U^L - U^n|^2, \quad d_* > 0.
\]

So
\[
S(U^L) - S(U^n) \geq \left( c_* - \frac{\Delta t}{m_j} d_* \right) |U^L - U^n|^2.
\]

The number between parenthesis is \( c_* - \frac{\Delta t}{\rho_j \Delta x_j} d_* \geq 0 \) provided the local CFL is satisfied \( \frac{\Delta t}{\rho_j \Delta x_j} d_* \leq c_* \).

### 4.6 Moving grid

Since the Lagrangian phase is a discretization of (143), it is natural to ask the Lagrangian scheme to be compatible with a discrete version of (144). It is indeed the case as shown below.

Define \( x^n_{j + \frac{1}{2}} = (j + \frac{1}{2}) \Delta x \) the edge between cell \( j \) and cell \( j + 1 \) at the beginning of the time step. It is natural to define
\[
x^L_{j + \frac{1}{2}} = x^n_{j + \frac{1}{2}} + \Delta t u^*_j + \frac{1}{2}.
\]

**Proposition 24.** The first equation of the Lagrangian scheme and the grid displacement are compatible in the sense that the local mass is constant
\[
\rho^L_j \left( x^L_{j + \frac{1}{2}} - x^L_{j - \frac{1}{2}} \right) = \rho^n_j \Delta x.
\]

**Proof.** By substitution this relation is equivalent to
\[
\frac{1}{\Delta x} \left( \Delta x + \Delta t (u^*_j + \frac{1}{2} - u^*_{j - \frac{1}{2}}) \right) = \frac{\rho^n_j}{\rho^L_j},
\]
that is \( \frac{1}{\rho^L_j} - \frac{1}{\rho^L_j} = \frac{1}{\rho^n_j \Delta x} (\Delta t(u^*_j + \frac{1}{2} - u^*_{j - \frac{1}{2}})) \). The proof is ended. \( \square \)

Define the length
\[
\Delta x^L_j = x^L_{j + \frac{1}{2}} - x^L_{j - \frac{1}{2}}, \quad j \in \mathbb{Z}.
\]

Since the mass is constant \( \Delta x^L_j \rho^L_j = \Delta x \rho^n_j = \Delta m_j \), the Lagrangian scheme is conservative in the sense that four formally conservative relations hold
\[
\begin{align*}
\sum_{j \in \mathbb{Z}} \Delta m_j \tau^L_j &= \sum_{j \in \mathbb{Z}} \Delta m_j \tau^n_j, \\
\sum_{j \in \mathbb{Z}} \Delta m_j u^L_j &= \sum_{j \in \mathbb{Z}} \Delta m_j u^n_j, \\
\sum_{j \in \mathbb{Z}} \Delta m_j v^L_j &= \sum_{j \in \mathbb{Z}} \Delta m_j v^n_j, \\
\sum_{j \in \mathbb{Z}} \Delta m_j e^L_j &= \sum_{j \in \mathbb{Z}} \Delta m_j e^n_j.
\end{align*}
\]
4.7 Remap

The principles of remapping are simple: they essentially consist to write the discrete equations that correspond to the second stage of what is depicted in the figure 18. The main constraints stem from natural conservativity and stability requirements. This is achieved without difficulties as show below. Note however that remapping can be presented as a splitting of operators, see below in section 4.10.

Consider one more time figure 18. The projection corresponds to the computation of mean values inside the cells of the fixed mesh. The only information to take care about is the sign of the velocities $u^*_{j+\frac{1}{2}}$ and $u^*_{j-\frac{1}{2}}$ which determine the evolution of the cell number $j$. A summary of the main cases is

\begin{align*}
  j = 1 & : \quad \Delta x \rho_j^{n+1} = (\Delta x + \Delta tu^*_{j+\frac{1}{2}}) \rho_j^L - \Delta tu^*_{j+\frac{1}{2}} \rho_j^{L+1}, \\
  j = 2 & : \quad \Delta x \rho_j^{n+1} = \Delta x \rho_j^L, \\
  j = 3, 4, 5 & : \quad \Delta x \rho_j^{n+1} = (\Delta x - \Delta tu^*_{j-\frac{1}{2}}) \rho_j^L + \Delta tu^*_{j-\frac{1}{2}} \rho_j^{L-1}, \\
  j = 6 & : \quad \Delta x \rho_j^{n+1} = (\Delta x - \Delta tu^*_{j+\frac{1}{2}} + \Delta tu^*_{j-\frac{1}{2}}) \rho_j^L - \Delta tu^*_{j+\frac{1}{2}} \rho_j^{L+1} + \Delta tu^*_{j-\frac{1}{2}} \rho_j^{L-1}.
\end{align*}

(150)

Replacing the density $\rho$ by their analogues $\rho u$, $\rho v$ and $\rho e$ one obtains the other relations. More precisely $\rho_j^{n+1}$ and $\rho_j^L$ is replace by $\rho_j^{n+1} u_j^{n+1}$ and $\rho_j^L u_j^L$ and so forth. For example in cell $j = 1$ one gets

\begin{align*}
  \Delta x \rho_j^{n+1} = (\Delta x - \Delta tu^*_{j+\frac{1}{2}} + \Delta tu^*_{j-\frac{1}{2}}) \rho_j^L - \Delta tu^*_{j+\frac{1}{2}} \rho_j^{L+1} + \Delta tu^*_{j-\frac{1}{2}} \rho_j^{L-1}.
\end{align*}

(151)

where $\rho_j^{L+1}$ (resp. $\rho_j^{L-1}$, $u_j^{n+1}$ or $\rho_j^L e_j^L$) is the density (resp. momentum or total energy) which is upwinded accordingly to the sign of the edge velocity $u^*_{j+\frac{1}{2}}$. That is

\begin{align*}
  \begin{cases}
    \text{if } u^*_{j+\frac{1}{2}} > 0 & \rho_j^{L+1} = \rho_j^L, \\
    \text{if } u^*_{j+\frac{1}{2}} < 0 & \rho_j^{L+1} = \rho_j^{L+1}, \\
    \text{if } u^*_{j+\frac{1}{2}} = 0 & \text{arbitrary since the product vanishes}.
  \end{cases}
\end{align*}

Notice that (151) is equivalent to

\begin{align*}
  \Delta x \rho_j^{n+1} = \Delta x \rho_j^L - \Delta tu^*_{j+\frac{1}{2}} \rho_j^{L+1} + \Delta tu^*_{j-\frac{1}{2}} \rho_j^{L-1}.
\end{align*}

(152)

At inspection of the design principle in figure 18, a criterion is needed to guarantee that $x_{j+\frac{1}{2}}$ does not cross more than one cell. This condition writes

\begin{align*}
  \left( \max_j \left| u^*_{j+\frac{1}{2}} \right| \right) \frac{\Delta t}{\Delta x} \leq 1.
\end{align*}

(153)

A slightly more severe constraint can be used as an extra guarantee that no Lagrangian cell has a zero or negative length. Considering cell number 6 in
the figure one gets the more severe condition $\left( \max_j |u_{j+\frac{1}{2}}^*| \right) \frac{\Delta t}{\Delta x} \leq \frac{1}{2}$ where the maximal time step is divided by a factor 2.

**Proposition 25.** The remapping is formally conservative up to boundaries

$$\begin{cases}
\sum_{j \in Z} \Delta x \rho_{j+1} = \sum_{j \in Z} \Delta x \rho_j^L, \\
\sum_{j \in Z} \Delta x \rho_{j+1} v_{j+1} = \sum_{j \in Z} \Delta x \rho_j^L v_j, \\
\sum_{j \in Z} \Delta x \rho_{j+1} e_{j+1} = \sum_{j \in Z} \Delta x \rho_j^L e_j.
\end{cases} \tag{154}$$

**Proof.** Geometrically evident on figure 18. \hfill \Box

### 4.8 Eulerian formulation of a Lagrange+remap scheme

A Lagrange+remap scheme is a particular discretization of the Eulerian formulation of compressible gas dynamics. To make this statement more evident, one can rewrite the two stages Lagrange+remap scheme as a one step eulerian scheme under the form

$$\begin{cases}
\rho_{j+1}^n + \frac{u_j^* + \rho_{j+\frac{1}{2}}^L u_{j+\frac{1}{2}} - u_{j-\frac{1}{2}}^* \rho_{j-\frac{1}{2}}^L}{\Delta t} = 0, \\
\frac{\rho^n_{j+1} u^n_{j+1} - \rho^n_j u^n_j}{\Delta t} + \frac{u_{j+\frac{1}{2}}^* \rho_{j+\frac{1}{2}}^L u_{j+\frac{1}{2}} - u_{j-\frac{1}{2}}^* \rho_{j-\frac{1}{2}}^L u_{j-\frac{1}{2}}}{\Delta x} = 0, \\
\frac{\rho_{j+1}^n e_{j+1} - \rho^n_j e^n_j}{\Delta t} + \frac{u_{j+\frac{1}{2}}^* \rho_{j+\frac{1}{2}}^L e_{j+\frac{1}{2}} - u_{j-\frac{1}{2}}^* \rho_{j-\frac{1}{2}}^L e_{j-\frac{1}{2}}}{\Delta x} = 0.
\end{cases} \tag{155}$$

Indeed the first equation comes from (152) after elimination of $\Delta x \rho^L_j = \Delta x \rho^n_j$. The remaining three ones can be retrieved from the generalization of (152) to the variables $\rho u$, $\rho v$ and $\rho e$. For example the equation for $u$ is obtained from

$$\Delta x \rho_{j+1}^n u_{j+1} = \Delta x \rho_j^L u_j - \Delta t u_{j+\frac{1}{2}}^* \rho_{j+\frac{1}{2}}^L u_{j+\frac{1}{2}} + \Delta t u_{j-\frac{1}{2}}^* \rho_{j-\frac{1}{2}}^L u_{j-\frac{1}{2}}.$$

The second equation of the Lagrangian scheme rewrites

$$\frac{1}{\Delta t} (\Delta x \rho_j^L u_j - \rho_j^e u_j) + p_{j+\frac{1}{2}} - p_{j-\frac{1}{2}} = 0.$$

An evident combination yields the second equation of (155). A same algebra yields the two last equation.

One notices that the compact Eulerian formulation (155) is clearly conservative up to boundary conditions. It yields another proof of the eulerian part of the conservative relations (154). An important property in term of stability is the following.
Theorem 10. Assume the Lagrangian CFL constraint and the remapping CFL constraint. Then the Lagrange+remap scheme (155) satisfies the entropy inequality

\[
\frac{\rho_j^{n+1} S_j^{n+1} - \rho_j^n S_j^n}{\Delta t} + \frac{u_j^* \rho_j^L S_j^L + u_j^* \rho_j^L S_j^L - u_j^* \rho_j^L S_j^L}{\Delta x} \geq 0
\]  

(156)

which is a discrete counterpart of \( \partial_t S + \partial_x p u \geq 0 \).

Proof. Rewrite (151) as

\[
\begin{align*}
\rho_j^{n+1} &= \alpha \rho_j^n L + \beta \rho_j^{n+1} u_j^{n+1} + \gamma \rho_j^n L, \\
\rho_j^{n+1} u_j^{n+1} &= \alpha \rho_j^n L u_j^n L + \beta \rho_j^{n+1} u_j^{n+1} v_j^{n+1} + \gamma \rho_j^n L v_j^n L, \\
\rho_j^{n+1} v_j^{n+1} &= \alpha \rho_j^n L v_j^n L + \beta \rho_j^{n+1} v_j^{n+1} e_j^{n+1} + \gamma \rho_j^n L e_j^n L,
\end{align*}
\]

taking care that \( \beta \geq 0 \) and \( \gamma \geq 0 \). Indeed if the corresponding coefficients in (151) are negative, it is always possible to incorporate them in \( \alpha \) which is in factor of the central terms. Thanks to the CFL condition, \( \alpha \geq 0 \). So the coefficients \( (\alpha, \beta, \gamma) \) define a convex combination since \( \alpha, \beta, \gamma \geq 0 \) and \( \alpha + \beta + \gamma = 1 \), (notice that \( \gamma \) has nothing to do with the constant of a perfect gas). So \( U_j^{n+1} \) at time step \( n + 1 \) is a convex combination of \( U_j^{L-1}, U_j^L, U_j^{L+1} \) at the end of the Lagrangian time step \( L \). Since the function \( \rho S \) is concave with respect to its arguments, it yields \( \rho_j^{n+1} S_j^{n+1} \geq \alpha \rho_j^n L S_j^n + \beta \rho_j^{n+1} L S_j^{n+1} + \gamma \rho_j^n L S_j^{n+1} \). Using the Lagrangian entropy inequality \( S_j^L \geq S_j^n \) one gets

\[
\rho_j^{n+1} S_j^{n+1} \geq \alpha \rho_j^n L S_j^n + \beta \rho_j^{n+1} L S_j^{n+1} + \gamma \rho_j^n L S_j^{n+1}.
\]  

(157)

Multiply by \( \Delta x \)

\[
\Delta x \rho_j^{n+1} S_j^{n+1} \geq \Delta x \rho_j^n S_j^n - \Delta x u_j^* \rho_j^L S_j^L + \Delta x u_j^* \rho_j^L S_j^L
\]

which is the claim. The proof is ended. \( \square \)

Proposition 26. Under the assumptions of the theorem, one has another discrete entropy inequality

\[
S_j^{n+1} \geq \min \left( S_j^{n-1}, S_j^n, S_j^{n+1} \right).
\]  

(158)

Proof. This is an easy consequence of (157) and the identity \( \rho_j^{n+1} = \alpha \rho_j^n L + \beta \rho_j^{n+1} L + \gamma \rho_j^n L \). The proof is ended. \( \square \)

This inequality yields some non linear stability. Indeed consider a perfect gas for which \( S = \log(\varepsilon \tau^{\gamma-1}) \), \( \gamma > 1 \). Therefore the inequality implies for example

\[
\frac{\varepsilon_j^n}{(\rho_j^n)^{\gamma-1}} \geq C^0 = \min_j \left( \frac{\varepsilon_j^0}{(\rho_j^0)^{\gamma-1}} \right) > 0.
\]

71
It means the ratio of certain quantities which must remain non negative for physical correctness remains positive. This is almost an a priori estimate on the non negativity of $\rho$ and $\varepsilon$ separately.

**Remark 9.** An remarkable feature of the Lagrange+remap scheme is the discrete entropy inequalities which have been proved for any general pressure law, provided they admit a strictly concave entropy.

### 4.9 A simple numerical result

We display in figure 19 numerical results for the Sod shock tube test problem. Apart the smearing effect of numerical methods, the numerical solution is a correct approximation of the true solution. With an initial condition $u_L = u_R = a = 1$ this is called the Harten test problem: in figure 20 one observes a similar solution, but translated to the right at a uniform velocity $a = 1$. The Harten problem has been defined to illustrate the entropy defect of some schemes. The Lagrange+remap is not polluted by any such non entropic defect. This is direct consequence of the discrete entropy inequalities (156) and (158).

![Graphs showing numerical results for the Sod shock tube test problem](image)

**Figure 19:** Sod test problem computed with a Lagrange+remap scheme and 200 cells. Final time is $t = 0.14$.

### 4.10 Pure Lagrange and ALE in dimension one

ALE methods consider a grid velocity, denoted as $v$ in the following, which can be different from the fluid velocity. In dimension $d = 1$ the main interest in
ALE methods is pedagogical. In dimension $d > 1$, ALE is needed to regularize Lagrangian grids which can deform up to unacceptable proportions. This is discussed at the end of the chapter.

We use the notations of section 1.8. Start from the Euler system in dimension $d = 1$

$$
\begin{align*}
\partial_t \rho + \partial_x (\rho u) &= 0, \\
\partial_t (\rho u) + \partial_x (\rho u^2 + p) &= 0, \\
\partial_t (\rho e) + \partial_x (\rho e u + pu) &= 0.
\end{align*}
$$

Consider the change of coordinates

$$
t' = t, \quad \frac{\partial x(t',X)}{\partial t'} = v(t', x(t', X))
$$

where $(t, x) \mapsto v(t, x)$ is an arbitrary velocity called the grid velocity. Writing the equations in the set of coordinates $(t', X)$ amounts to compute the Jacobian matrix

$$
\nabla_{(t',X)}(t, x) = \begin{pmatrix}
1 & 0 \\
v & J
\end{pmatrix}, \quad J = \frac{\partial x}{\partial X}.
$$

The comatrix (definition 1) is $\text{com} \left( \nabla_{(t',X)}(t, x) \right) = \begin{pmatrix}
J & -v \\
0 & 1
\end{pmatrix}$. One obtains
from (19)
\[
\begin{aligned}
\partial_t' (\rho J) + \partial_X ((\rho (u - v))) &= 0, \\
\partial_t' (\rho u J) + \partial_X ((\rho u (u - v) + p)) &= 0, \\
\partial_t' (\rho e J) + \partial_X ((\rho e (u - v) + pu)) &= 0.
\end{aligned}
\]

The Piola identity (20) writes \( \partial_t' J - \partial_X v = 0 \). To simplify we use from now on the same notation for the time variable, that is \( t' = t \). One gets the closed system
\[
\begin{aligned}
\partial_t (\rho J) + \partial_X ((\rho (u - v))) &= 0, \\
\partial_t (\rho u J) + \partial_X ((\rho u (u - v) + p)) &= 0, \\
\partial_t (\rho e J) + \partial_X ((\rho e (u - v) + pu)) &= 0, \\
\partial_t J - \partial_X v &= 0.
\end{aligned}
\]

Since we wish to present ALE in the contact of splitting of operators, we prefer to write it like
\[
\begin{aligned}
\partial_t \begin{pmatrix} \rho J \\ \rho u J \\ \rho e J \end{pmatrix} + \partial_X \begin{pmatrix} 0 \\ -u \\ p \\ pu \end{pmatrix} + \partial_X \begin{pmatrix} \rho (u - v) \\ u - v \\ pu (u - v) \\ pe (u - v) \end{pmatrix} &= 0. 
\end{aligned}
\]

This formulation enlightens that the total flux is the sum of a first Lagrangian contribution and of a second contribution of convective nature. The initialization of \( J \) is naturally \( J(t = 0) = 1 \). It is now clear that three main cases occur.

**Eulerian formulation:** \( v = 0 \). Taking \( v = 0 \) is of course equivalent to the original system (159). The second equation which is eliminated.

**Lagrangian formulation:** \( v = u \). The first equation in (159) becomes an ODE and is therefore eliminated after exact integration, \( \rho J = \rho_0 \). One can define naturally the mass variable \( dm = \rho_0 dx \).

**ALE formulation:** \( v \neq 0 \) and \( v \neq u \). This third case is neither pure Eulerian nor pure Lagrangian. This is the Arbitrary Euler Lagrange formulation.

**Energy Lagrange system:** \( v = \left(1 + \frac{p}{pe}\right) u \). It corresponds to the grid velocity \( v = \left(1 + \frac{p}{pe}\right) u \) which zeroes the total flux of energy equation. For a perfect gas in usual condition \( 1 \leq \frac{p}{pe} \leq \gamma - 1 \) so this definition makes sense. Using \( u - v = -\frac{p}{pe} u \) one obtains the system
\[
\begin{aligned}
\partial_t (\rho J) - \partial_X \left( \frac{p}{pe} u \right) &= 0, \\
\partial_t (\rho u J) + \partial_X (p - \frac{p}{pe} u^2) &= 0, \\
\partial_t J - \partial_X \left( 1 + \frac{p}{pe} \right) u &= 0.
\end{aligned}
\]
The convenient mass variable writes $dm = \rho_0 e_0 dX$. One obtains

$$
\begin{align*}
\partial_t \frac{\rho}{\rho} - \partial_m \left( \frac{p}{\rho} u \right) &= 0, \\
\partial_t \frac{\varphi}{\rho} + \partial_m \left( p - \frac{\varphi}{\rho} u^2 \right) &= 0, \\
\partial_t \frac{\varphi}{\rho} - \partial_m \left( \left( 1 + \frac{\varphi}{\rho} \right) u \right) &= 0.
\end{align*}
$$

The use of this energy-Lagrange system for practical computations is an open problem and will not be considered.

4.10.1 Numerical discretization

The idea pursued below is to discretize the system (159) with a splitting strategy. One discretizes firstly

$$
\partial_t \left( \begin{array}{c} \rho J \\ \frac{J}{\rho u J} \\ \frac{\rho e J}{\rho e J} \end{array} \right) + \partial_X \left( \begin{array}{c} 0 \\ -u \\ \frac{p}{p u} \end{array} \right) = 0
$$

during a time step $\Delta t$. Secondly one discretizes

$$
\partial_t \left( \begin{array}{c} \rho J \\ \frac{J}{\rho u J} \\ \frac{\rho e J}{\rho e J} \end{array} \right) + \partial_X \left( \begin{array}{c} \rho (u - v) \\ u - v \\ \frac{p u (u - v)}{p e (u - v)} \end{array} \right) = 0
$$

during the same time step $\Delta t$.

4.10.2 Discretization of the first part (160)

Consider an infinite ($j \in \mathbb{Z}$) initial grid with intermediate points $X_{j+\frac{1}{2}}$, $\Delta X_j = X_{j+\frac{1}{2}} - X_{j-\frac{1}{2}}$. The lagrangian fluxes are

$$
\begin{align*}
&u_{j+\frac{1}{2}}^* = \frac{1}{2}(u_j^n + u_{j+1}^n) + \frac{1}{2\rho_j} \left( p_j^n - p_{j+1}^n \right), \\
p_{j+\frac{1}{2}}^* = \frac{1}{2}(p_j^n + p_{j+1}^n) + \frac{\rho_j^n}{2} (u_j^n - u_{j+1}^n).
\end{align*}
$$

The scheme is

$$
\begin{align*}
\rho_j^{n+\frac{1}{2}} J_j^{n+\frac{1}{2}} - \rho_j^n J_j^n &= 0, \\
\Delta X_j \left( J_j^{n+\frac{1}{2}} - J_j^n \right) - \Delta t \left( u_{j+\frac{1}{2}}^* - u_{j-\frac{1}{2}}^* \right) &= 0, \\
\Delta X_j \left( \rho_j^{n+\frac{1}{2}} J_j^{n+\frac{1}{2}} u_j^{n+\frac{1}{2}} - \rho_j^n J_j^n u_j^n \right) + \Delta t \left( p_{j+\frac{1}{2}}^* - p_{j-\frac{1}{2}}^* \right) &= 0, \\
\Delta X_j \left( \rho_j^{n+\frac{1}{2}} J_j^{n+\frac{1}{2}} e_j^{n+\frac{1}{2}} - \rho_j^n J_j^n e_j^n \right) + \Delta t \left( p_{j+\frac{1}{2}}^* u_{j+\frac{1}{2}}^* - p_{j-\frac{1}{2}}^* u_{j-\frac{1}{2}}^* \right) &= 0.
\end{align*}
$$
4.10.3 Discretization of the second part (161)

Define

\[ w^n_{j + \frac{1}{2}} = u^n_{j + \frac{1}{2}} - v^n_{j + \frac{1}{2}} \]

being understood that \( u^n_{j + \frac{1}{2}} \) is an explicit value at time step \( n \) and \( v^n_{j + \frac{1}{2}} \) is the velocity of the grid point. The discretization of (161) with a correct definition of the flux accordingly to the sign of the differential velocity \( w = u - v \) yields the scheme

\[
\begin{align*}
\Delta X_j \left( \rho_j^{n+1} J_j^{n+1} - \rho_j^{n+\frac{1}{2}} J_j^{n+\frac{1}{2}} \right) + \Delta t \left( \rho_j^{n+\frac{1}{2}} w^n_{j+\frac{1}{2}} - \rho_j^{n+\frac{1}{2}} w^n_{j-\frac{1}{2}} \right) &= 0, \\
\Delta X_j \left( J_j^{n+1} - J_j^{n+\frac{1}{2}} \right) + \Delta t \left( w^n_{j+\frac{1}{2}} - w^n_{j-\frac{1}{2}} \right) &= 0, \\
\Delta X_j \left( \rho_j^{n+1} J_j^{n+1} u^n_{j+1} - \rho_j^{n+\frac{1}{2}} J_j^{n+\frac{1}{2}} u^n_{j+\frac{1}{2}} \right) \\
&\quad + \Delta t \left( \rho_j^{n+\frac{1}{2}} u^n_{j+\frac{1}{2}} w^n_{j+\frac{1}{2}} - \rho_j^{n+\frac{1}{2}} u^n_{j-\frac{1}{2}} w^n_{j-\frac{1}{2}} \right) = 0, \\
\Delta X_j \left( \rho_j^{n+1} J_j^{n+1} e^n_{j+1} - \rho_j^{n+\frac{1}{2}} J_j^{n+\frac{1}{2}} e^n_{j+\frac{1}{2}} \right) \\
&\quad + \Delta t \left( \rho_j^{n+\frac{1}{2}} e^n_{j+\frac{1}{2}} w^n_{j+\frac{1}{2}} - \rho_j^{n+\frac{1}{2}} e^n_{j-\frac{1}{2}} w^n_{j-\frac{1}{2}} \right) = 0.
\end{align*}
\]

The convention is that \( f^n_{j+\frac{1}{2}} = f^n_{j+\frac{1}{2}} \) for \( w^n_{j+\frac{1}{2}} \geq 0 \) and \( f^n_{j+\frac{1}{2}} = f^n_{j+\frac{1}{2}} \) for \( w^n_{j+\frac{1}{2}} < 0 \).

4.10.4 Reformulation on a moving grid

The grid displacement is naturally defined by

\[
x^n_{j+\frac{1}{2}} = x^n_{j+\frac{1}{2}} + \Delta t u^n_{j+\frac{1}{2}}
\]

for the first stage, and by

\[
x^{n+1}_{j+\frac{1}{2}} = x^n_{j+\frac{1}{2}} + \Delta t \left( v^n_{j+\frac{1}{2}} - u^n_{j+\frac{1}{2}} \right)
\]

for the second stage. One gets the total displacement

\[
x^{n+1}_{j+\frac{1}{2}} = x^n_{j+\frac{1}{2}} + \Delta t v^n_{j+\frac{1}{2}}.
\]

Notice the initial condition \( x^0_{j+\frac{1}{2}} = X_{j+\frac{1}{2}} \). Define \( \Delta x_j^n = x^n_{j+\frac{1}{2}} - x^n_{j-\frac{1}{2}} \). One checks that by construction the variation in time of \( \Delta X_j J_j^n \) is equal to the variation in time of \( \Delta x_j^n \). So

\[
J_j^n = \frac{\Delta x_j^n}{\Delta X_j^n}.
\]

Define the mass in the cell by \( \Delta M_j^n = \Delta x_j^n \rho_j^n \). It is easy to rewrite the first stage (163) under the form

\[
\begin{align*}
\frac{\Delta M_j^n}{\Delta t} \left( \tau_j^{n+\frac{1}{2}} - \tau_j^{n} \right) - u^n_{j+\frac{1}{2}} + u^n_{j-\frac{1}{2}} &= 0, \\
\frac{\Delta M_j^n}{\Delta t} \left( u^n_{j+\frac{1}{2}} - u^n_{j-\frac{1}{2}} \right) + \rho^n_{j+\frac{1}{2}} - \rho^n_{j-\frac{1}{2}} &= 0, \\
\frac{\Delta M_j^n}{\Delta t} \left( e^n_{j+\frac{1}{2}} - e^n_{j} \right) + p^n_{j+\frac{1}{2}} u^n_{j+\frac{1}{2}} - p^n_{j-\frac{1}{2}} u^n_{j-\frac{1}{2}} &= 0./
\end{align*}
\]
So the CFL condition of this stage can be written as

$$\max_j \left( \frac{c_n^j}{\Delta x_j^n} \right) \Delta t \leq CFL < 1. \quad (169)$$

A similar entropy inequality is true under CFL

$$S_{j}^{n+\frac{1}{2}} \geq S_{j}^{n}. \quad (170)$$

The second stage (164) is analyzed as follows. Define $\Delta x_{j}^{n+\frac{1}{2}}$ which is the length of the Lagrangian cell at the end of the Lagrangian time step. The first equation of (164) writes

$$\Delta x_{j}^{n+1} \rho_{j}^{n+1} - \Delta x_{j}^{n+\frac{1}{2}} \rho_{j}^{n+\frac{1}{2}} + \Delta t \left( \rho_{j+\frac{1}{2}}^{n+\frac{1}{2}} w_{j+\frac{1}{2}}^{n} - \rho_{j-\frac{1}{2}}^{n+\frac{1}{2}} w_{j-\frac{1}{2}}^{n} \right) = 0.$$

Assume for simplicity that $w_{j+\frac{1}{2}}^{n} \geq 0$ and $w_{j-\frac{1}{2}}^{n} \geq 0$. The compatibility between $\rho_{j+\frac{1}{2}}^{n+1}$ and $\rho_{j-\frac{1}{2}}^{n+1}$ implies that

$$\rho_{j+\frac{1}{2}}^{n+1} = \rho_{j}^{n+\frac{1}{2}} \quad \text{and} \quad \rho_{j-\frac{1}{2}}^{n+1} = \rho_{j-1}^{n+\frac{1}{2}}.$$

Therefore

$$\Delta x_{j}^{n+1} \rho_{j}^{n+1} = \left( \Delta x_{j}^{n+\frac{1}{2}} - \Delta tw_{j+\frac{1}{2}}^{n} \right) \rho_{j}^{n+\frac{1}{2}} + \Delta tw_{j-\frac{1}{2}}^{n} \rho_{j-\frac{1}{2}}^{n+\frac{1}{2}}. \quad (171)$$

One obtains similarly for the second equation

$$\Delta x_{j}^{n+1} = \left( \Delta x_{j}^{n+\frac{1}{2}} - \Delta tw_{j+\frac{1}{2}}^{n} \right) + \Delta tw_{j-\frac{1}{2}}^{n}.$$

$$\Rightarrow \Delta x_{j}^{n+\frac{1}{2}} - \Delta tw_{j+\frac{1}{2}}^{n} = \Delta x_{j}^{n+1} - \Delta tw_{j-\frac{1}{2}}^{n}.$$

Substitution in (171) yields

$$\Delta x_{j}^{n+1} \rho_{j}^{n+1} = \left( \Delta x_{j}^{n+1} - \Delta tw_{j-\frac{1}{2}}^{n} \right) \rho_{j}^{n+\frac{1}{2}} + \Delta tw_{j+\frac{1}{2}}^{n} \rho_{j+\frac{1}{2}}^{n+\frac{1}{2}}. \quad (172)$$

By comparison with (150), it is clear that, for the example for the cells indexed by 3, 4 and 5, $u_{j+\frac{1}{2}}^{n} \geq 0$ has been replaced by $w_{j+\frac{1}{2}}^{n} \geq 0$. Up to this only difference, the general situation is unchanged. The other cases for $w_{j+\frac{1}{2}}^{n} \leq 0$ can be analyzed with the same method.

**Proposition 27.** The scheme (164) is equivalent to a geometric projection on the mesh of the figure 21.

**Proof.** Detail the different cases on the figure 21. □
Figure 21: The grid velocity is $u^*_{j+\frac{1}{2}}$ in the lagrangian first stage. It is equal to $w_{j+\frac{1}{2}} = v_{j+\frac{1}{2}} - u^*_{j+\frac{1}{2}}$ in the second remapping stage.

A natural stability constraint for the second stage writes

$$
\left( \max_j |w^*_{j+\frac{1}{2}}| \right) \frac{\Delta t}{\Delta x} \leq \frac{1}{2}.
$$

(173)

It prevents from any crossings. The following result is the generalization of theorem 10 for the ALE configuration.

**Theorem 11.** Assume the two CFL conditions (169) and (173). Then the ALE scheme (163-164) satisfies the discrete entropy inequality

$$
\Delta x_j \rho_j^{n+1} S_j^{n+1} - \Delta x_j \rho_j^n S_j^n + \Delta t \left( w_{j+\frac{1}{2}}^{n+\frac{1}{2}} \rho_{j+\frac{1}{2}}^{n+\frac{1}{2}} S_{j+\frac{1}{2}}^{n+\frac{1}{2}} - w_{j-\frac{1}{2}}^{n+\frac{1}{2}} \rho_{j-\frac{1}{2}}^{n+\frac{1}{2}} S_{j-\frac{1}{2}}^{n+\frac{1}{2}} \right) \geq 0.
$$

(174)

### 4.10.5 Numerical illustrations

The result of the Lagrange+remap scheme already presented for the Sod tube test problem at figure 19 are complemented by results for the same test problem, but in pure Lagrangian mode in figure 22 and in ALE mode in figure 23. In pure Lagrangian mode, there is no numerical smearing at the contact discontinuity. But on the other hand the numerical defaults at the interface are preserved by the scheme. This is specially visible on the density profile which exhibits a small pike denoted as the wall-heating phenomenon. The pike is more evident on the entropy profile. Velocity and pressure profiles are without noticeable oscillations. The results of an ALE calculation are given in figure 23 with an arbitrary grid velocity $v_{j+\frac{1}{2}} = 0.5 \times \sin(4\pi x) \times \sin(2\pi t)$. The results are comparable to those of the pure Lagrange computation and of the Eulerian computation in Lagrange+remap mode.

We also present the result of a Sod tube test problem in dimension $d = 2$ computed with the Lagrange+remap strategy combined with a directional splitting technique. That is the two dimensional problem is solved with a series of one dimensional numerical methods. Notice that the entropy inequality still holds in multi dimension.
Figure 22: Sod tube test problem. Pure Lagrangian simulation with 200 cells. Final time \( t = 0.14 \). Notice the compression of the mesh behind the shock and the numerical discrepancy (called a wall-heating) at the contact discontinuity.

The initial data is a Sod test problem from both sides of the interface defined by \( \sqrt{x^2 + y^2} = 0.5 \). An important numerical smearing is visible on the density profile at the contact discontinuity.

References


Figure 23: Same set-up as in figure 22, except that the calculation is ALE with grid velocity \( v_{j+\frac{1}{2}} = 0.5 \sin(4\pi x) \sin(2\pi t) \).


Figure 24: Density $\rho$ and modulus of the velocity $\sqrt{u^2 + v^2}$ at final time $T = 0.2$.


Figure 25: Pressure $p$ and entropy $S$ at final time $T = 0.2$. 

82