

Cours EDPs MACS3

Pascal Omnes

Septembre 2022

1 Modèles physiques

1.1 Écoulements

1.1.1 Équations d'Euler et de Navier-Stokes compressibles

Gaz ou liquide décrit macroscopiquement par $\rho(\mathbf{x}, t)$ sa masse volumique (souvent appelée densité) (kg.m^{-3}), $\mathbf{u}(\mathbf{x}, t)$ sa vitesse (m.s^{-1}), e la densité volumique d'énergie ($\text{J.m}^{-3} = \text{N.m}^{-2}$) et $p(\mathbf{x}, t)$ la pression ($\text{Pa} = \text{N.m}^{-2}$). Dans un premier temps sans phénomène de dissipation visqueuse ou thermique (équations d'Euler).

La conservation de la masse s'écrit:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (1)$$

Le bilan de quantité de mouvement s'écrit:

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p = \mathbf{f}, \quad (2)$$

où \mathbf{f} est une densité volumique de force (N.m^{-3}) (par exemple $\mathbf{f} = -\rho g \mathbf{e}_z$ est la densité de force liée à la gravité). Le bilan d'énergie s'écrit:

$$\frac{\partial e}{\partial t} + \nabla \cdot ((e + p)\mathbf{u}) = \mathbf{f} \cdot \mathbf{u}. \quad (3)$$

Il est aussi souvent écrit à l'aide de la variable $E = \frac{e}{\rho}$, densité massique d'énergie totale (J.kg^{-1}):

$$\frac{\partial \rho E}{\partial t} + \nabla \cdot ((\rho E + p)\mathbf{u}) = \mathbf{f} \cdot \mathbf{u}. \quad (4)$$

On introduit aussi la densité massique d'énergie interne ε , qui est liée à la densité massique d'énergie totale par la formule

$$E = \varepsilon + \frac{|\mathbf{u}|^2}{2}. \quad (5)$$

Il faut une relation supplémentaire pour fermer les équations: c'est en général une loi de pression de la forme

$$p = p(\rho, \varepsilon). \quad (6)$$

Par exemple pour un gaz parfait, il existe une variable T (la température) telle que $\varepsilon = C_v T$, où C_v est la capacité thermique massique à volume constant. Or puisque $pV = nRT$, on tire $p = \frac{\rho}{M} RT = \frac{R}{MC_v} \rho \varepsilon = (\gamma - 1) \rho \varepsilon$, avec M la masse molaire du gaz, $\gamma = \frac{C_p}{C_v}$ le coefficient adiabatique, C_p la capacité thermique massique à pression constante, liée à C_v par $C_p - C_v = \frac{R}{M}$.

Dans un second temps avec viscosité et dissipation thermique: équations de Navier-Stokes (sans viscosité les oiseaux ne peuvent pas voler... résultante des forces nulle ou quasi-nulle).

On note:

$$\underline{\mathbf{D}} = \frac{1}{2} [(\nabla \mathbf{u}) + (\nabla \mathbf{u})^T]$$

le tenseur des déformations, $\mu \geq 0$ la viscosité dynamique du milieu, $\lambda \geq 0$ le "coefficient de Lamé" et

$$\underline{\boldsymbol{\sigma}}_{\mathbf{v}} = 2\mu \underline{\mathbf{D}} + \lambda(\nabla \cdot \mathbf{u}) \mathbf{Id}$$

le tenseur des contraintes visqueuses. On note que $\underline{\boldsymbol{\sigma}}_{\mathbf{v}}$ est symétrique. L'équation de quantité de mouvement s'écrit alors

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p - \nabla \cdot (\underline{\boldsymbol{\sigma}}_{\mathbf{v}}) = \mathbf{f}. \quad (7)$$

(On a défini, en dimension deux $\nabla \cdot (\underline{\boldsymbol{\sigma}}_{\mathbf{v}}) = (\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y}, \frac{\partial \sigma_{yx}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y})^T$). On note ensuite $k \geq 0$ la conductivité thermique du fluide, on définit le flux de chaleur dans le milieu par $\mathbf{q} = -k \nabla T$ et le bilan d'énergie s'écrit alors

$$\frac{\partial \rho E}{\partial t} + \nabla \cdot ((\rho E + p) \mathbf{u}) - \nabla \cdot (k \nabla T) - \nabla \cdot (\underline{\boldsymbol{\sigma}}_{\mathbf{v}} \mathbf{u}) = \mathbf{f} \cdot \mathbf{u}. \quad (8)$$

1.1.2 Dérivation de l'équation de conservation de la masse

On prend un domaine Ω fixe. On fait un bilan de masse entre les instants t et $t + dt$: Au temps $t + dt$, la masse contenue dans le domaine est $\int_{\Omega} \rho(\mathbf{x}, t + dt) d\mathbf{x}$ alors qu'elle était de $\int_{\Omega} \rho(\mathbf{x}, t) d\mathbf{x}$ à l'instant t . La variation de masse est donc

$$\Delta m = \int_{\Omega} \rho(\mathbf{x}, t + dt) d\mathbf{x} - \int_{\Omega} \rho(\mathbf{x}, t) d\mathbf{x} = \int_t^{t+dt} \left(\int_{\Omega} \frac{\partial \rho}{\partial t}(\mathbf{x}, \tau) d\mathbf{x} \right) d\tau. \quad (9)$$

En l'absence de création de matière par injection ou au contraire de prélèvement de matière, la variation de masse dans le domaine est nécessairement due au flux de masse \mathbf{f}_m qui entre ou sort du domaine (le flux de masse est la masse qui traverse une unité de surface en une unité de temps: $\text{kg.m}^{-2}.\text{s}^{-1}$):

$$\Delta m = \int_t^{t+dt} \left(\int_{\partial \Omega} \mathbf{f}_m(s, \tau) \cdot \mathbf{n}_{\partial \Omega} ds \right) d\tau,$$

où $\mathbf{n}_{\partial\Omega}$ est le vecteur unitaire normal sortant à Ω au point considéré de la frontière. On remarque que la présence du produit scalaire de \mathbf{f}_m avec $\mathbf{n}_{\partial\Omega}$ est justifiée car la composante tangentielle de \mathbf{f}_m ne joue aucun rôle: la masse ne pénètre pas tangentiellement dans le domaine.

Maintenant ce flux de masse est proportionnel à la vitesse du fluide \mathbf{u} et à la masse volumique du fluide ρ . Par ailleurs si la vitesse est dirigée vers l'extérieur $\mathbf{u} \cdot \mathbf{n}_{\partial\Omega} \geq 0$ alors que dans ce cas on a une perte de masse, donc le coefficient de proportionnalité est -1 . Au final on a $\mathbf{f}_m = -\rho\mathbf{u}$ (on vérifie les unités: $\rho\mathbf{u}$ est bien $\text{kg.m}^{-2}.\text{s}^{-1}$) et donc

$$\Delta m = - \int_t^{t+dt} \left(\int_{\partial\Omega} \rho\mathbf{u}(s, \tau) \cdot \mathbf{n}_{\partial\Omega} ds \right) d\tau. \quad (10)$$

Avec l'aide de la formule de Green-Gauß et (9) et (10), on obtient, pour tout t , dt et domaine Ω la forme intégrale correspondant à (1):

$$\int_{\Omega \times [t, t+dt]} \left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho\mathbf{u}) \right) (\mathbf{x}, \tau) d\mathbf{x} d\tau = 0, \quad (11)$$

d'où se déduit (1), l'égalité ci-dessus étant vraie pour tout t , dt et domaine Ω .

1.1.3 Adimensionnement

C'est une étape particulièrement importante de l'étude des phénomènes physiques: elle permet de comparer entre eux les régimes physiques en se référant à des nombres sans dimensions. Cela a pour avantage de mettre en évidence des phénomènes importants par rapport à d'autres (régime dominé par la convection ou par la diffusion par ex.) et de réaliser des similitudes (expérience équivalente à échelle réduite par ex.).

Le principe est de passer de grandeurs dimensionnées (ρ en kg.m^{-3} , \mathbf{x} en m, ...) à des grandeurs sans dimensions par un changement de variables multiplicatif à l'aide d'une grandeur de référence.

L longueur du dispositif (de la pièce par ex.): $\mathbf{x} = L\mathbf{x}'$

\bar{t} temps d'observation des phénomènes: $t = \bar{t}t'$

\bar{U} vitesse de référence

$\bar{\rho}$ densité de référence (par exemple celle du fluide au repos) $\rho = \bar{\rho}\rho'$

\bar{p} pression de référence, liée à la densité de référence par $\bar{p} = \bar{\rho} \frac{\partial p}{\partial \rho}(\bar{\rho})$ (en raison du gradient de pression: $\nabla p = \frac{\partial p}{\partial \rho} \nabla \rho$).

Les variables avec un $'$ sont maintenant sans dimensions. Par ailleurs la dérivation composée donne $\frac{\partial \bullet}{\partial t} = \frac{1}{\bar{t}} \frac{\partial \bullet}{\partial t'}$, etc... Si on se contente d'Euler isentropique par simplicité (pas d'équation d'énergie et p ne dépend que de la densité, on obtient (en laissant tomber les $'$ pour plus de facilité d'écriture): La conservation de la masse adimensionnée:

$$\frac{\bar{\rho}}{\bar{t}} \frac{\partial \rho}{\partial t} + \frac{\bar{\rho}\bar{U}}{L} \nabla \cdot (\rho\mathbf{u}) = 0 \quad (12)$$

et le bilan de quantité de mouvement adimensionné:

$$\frac{\bar{\rho}\bar{U}}{\bar{t}} \frac{\partial \rho \mathbf{u}}{\partial t} + \frac{\bar{\rho}\bar{U}^2}{L} \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \frac{\bar{\rho}}{L} \frac{\partial p}{\partial \rho}(\bar{\rho}) \nabla p = \mathbf{f}, \quad (13)$$

soit aussi, si l'on choisit $\bar{U} = \frac{L}{\bar{t}}$ (i.e. le temps d'observation est choisi égal au temps nécessaire à une particule se déplaçant à la vitesse de référence pour traverser le dispositif):

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (14)$$

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \frac{1}{M^2} \nabla p = \frac{\bar{t}}{\bar{\rho}\bar{U}} \mathbf{f}, \quad (15)$$

où $M := \frac{\bar{U}}{(\frac{\partial p}{\partial \rho}(\bar{\rho}))^{\frac{1}{2}}}$ est appelé le nombre de Mach; c'est le rapport entre la vitesse de référence et une vitesse du son de référence dans le milieu considéré. Sachant que pour un gaz parfait isentropique la loi de pression est donnée par $p = C\rho^\gamma$, la vitesse du son caractéristique est donnée par

$$c_s = \sqrt{\frac{\partial p}{\partial \rho}(\bar{\rho})} = \sqrt{\gamma C \bar{\rho}^{\gamma-1}} = \sqrt{\gamma \frac{p(\bar{\rho})}{\bar{\rho}}}$$

soit encore, si l'on écrit la loi des gaz parfaits ($\bar{p} = \frac{R}{M_s} \bar{\rho} \bar{T}$)

$$c_s = \sqrt{\gamma \frac{R}{M_s} \bar{T}}.$$

Par exemple si la densité volumique de force est due à la gravité, la force adimensionnée est

$$\tilde{\mathbf{f}} = -\frac{g}{\frac{\bar{U}}{\bar{t}}} \rho \mathbf{e}_z = -\frac{gL}{\bar{U}^2} \rho \mathbf{e}_z = -\frac{1}{Fr^2} \rho \mathbf{e}_z,$$

elle sera donc (ou non) négligeable en fonction du nombre de Froude $Fr = \frac{\bar{U}}{\sqrt{gL}}$, qui est le rapport des deux accélérations g et $\frac{\bar{U}}{\bar{t}}$. Si on observe des temps "très courts", ou des phénomènes à vitesse "très rapide", alors la force due à la gravité n'aura pas d'effet notable et sera négligée. Fr est aussi la racine carrée du rapport de deux énergies (cinétique et potentielle).

Enfin, la loi de pression s'écrit

$$p = \tilde{p}(\rho), \quad (16)$$

avec $\tilde{p}(\rho) := \frac{1}{\bar{\rho} \frac{\partial p}{\partial \rho}(\bar{\rho})} p(\bar{\rho} \rho)$ ($= \frac{1}{\gamma} \rho^\gamma$ pour les gaz parfaits).

Concernant le nombre de Mach, sachant que la vitesse du son dans l'air est d'environ 340 m.s⁻¹ et que les phénomènes physiques "usuels" se font à quelques m.s⁻¹ ou quelques dizaines de m.s⁻¹, le nombre de Mach de référence est toujours petit devant 1 (entre 10⁻³ et 10⁻¹).

1.1.4 Passage à l'incompressible

Lorsque $M \ll 1$, on peut avoir l'idée d'effectuer un développement asymptotique des inconnues en puissances de M :

$$p = p_0 + Mp_1 + M^2p_2 + \dots \quad (17)$$

$$\rho = \rho_0 + M\rho_1 + M^2\rho_2 + \dots \quad (18)$$

$$\mathbf{u} = \mathbf{u}_0 + M\mathbf{u}_1 + M^2\mathbf{u}_2 + \dots \quad (19)$$

On injecte (17)–(19) dans le système (14)–(15) et l'on égalise les puissances de M , en commençant par M^{-2} . On obtient:

- ordre M^{-2} : $\nabla p_0 = 0$. Cela signifie que la pression est uniforme spatialement à l'ordre 0.

- ordre M^{-1} : $\nabla p_1 = 0$. Cela signifie que la pression est uniforme spatialement également à l'ordre 1.

À l'ordre M^0 :

$$\frac{\partial \rho_0}{\partial t} + \nabla \cdot (\rho_0 \mathbf{u}_0) = 0 \quad (20)$$

$$\frac{\partial \rho_0 \mathbf{u}_0}{\partial t} + \nabla \cdot (\rho_0 \mathbf{u}_0 \otimes \mathbf{u}_0) + \nabla p_2 = \tilde{\mathbf{f}}, \quad (21)$$

si l'on a supposé la force adimensionnée d'ordre 0 par rapport à M . Enfin, la loi d'état fournit

$$p_0 = \tilde{p}(\rho_0). \quad (22)$$

Puisque p_0 (et p_1) ne dépendent pas de l'espace, il en est de même de ρ_0 (et de ρ_1) par cette loi. En intégrant (22) sur le domaine de calcul et en supposant une intégrale de flux nulle (soit parce qu'il n'y a pas d'entrée / sortie de matière, soit parce que celles-ci se compensent, soit par application de conditions aux limites périodiques), on obtient

$$\frac{d}{dt} \int_{\Omega} \rho_0(\mathbf{x}, t) d\mathbf{x} + \int_{\partial\Omega} \rho_0 \mathbf{u}_0 \cdot \mathbf{n} = |\Omega| \frac{d}{dt} \rho_0(t) = 0,$$

ce qui signifie que ρ_0 est constant également en temps, et d'après la loi de pression, il en est de même de p_0 . À partir de cette propriété, un développement de (20) fournit

$$\nabla \cdot \mathbf{u}_0 = 0. \quad (23)$$

Ensuite, à partir de (21) et de l'égalité $\nabla \cdot (\mathbf{u}_0 \otimes \mathbf{u}_0) = (\nabla \cdot \mathbf{u}_0) \mathbf{u}_0 + \mathbf{u}_0 \cdot \nabla \mathbf{u}_0$, on obtient (pour une solution régulière),

$$\frac{\partial \mathbf{u}_0}{\partial t} + \mathbf{u}_0 \cdot \nabla \mathbf{u}_0 + \nabla \left(\frac{p_2}{\rho_0} \right) = \frac{1}{\rho_0} \tilde{\mathbf{f}}. \quad (24)$$

Le système (23)–(24) constitue le système 'd'Euler incompressible', dans lequel on renomme souvent $\frac{p_2}{\rho_0}$ "la pression" et on la note de nouveau par p . Toutefois, on voit que cette nouvelle variable est assez éloignée de la pression d'origine.

Certains considèrent cette nouvelle pression comme 'le multiplicateur de Lagrange' de la contrainte d'incompressibilité $\nabla \cdot \mathbf{u}_0 = 0$. On voit aussi qu'il n'y a plus besoin de loi d'état pour fermer le système, puisque la densité a disparu des équations. Certains disent que c'est la contrainte d'incompressibilité $\nabla \cdot \mathbf{u}_0 = 0$ qui joue le rôle de loi d'état.

Si maintenant on reste dans le cadre incompressible et que l'on regarde les équations de Navier-Stokes, le tenseur des contraintes visqueuses se simplifie en

$$\underline{\sigma}_v = 2\mu \underline{\mathbf{D}}.$$

De plus, si l'on fait l'hypothèse que μ est une constante, alors

$$\nabla \cdot (\underline{\sigma}_v) = \mu (\nabla(\nabla \cdot \mathbf{u}) + \Delta \mathbf{u}) = \mu \Delta \mathbf{u}.$$

L'équation de quantité de mouvement adimensionnée s'écrit alors

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p - \frac{1}{Re} \Delta \mathbf{u} = \mathbf{F}, \quad (25)$$

avec $Re^{-1} = \frac{\nu}{LU}$ le nombre de Reynolds et $\nu := \frac{\mu}{\rho}$ est la viscosité cinématique du fluide. Sachant que ν s'exprime en $\text{m}^2 \cdot \text{s}^{-1}$, le nombre de Reynolds est bien sans dimension. Si le nombre de Reynolds est petit devant 1, ce sont les effets visqueux qui l'emportent, s'il est grand devant 1, ce sont les effets inertiels qui l'emportent. Lorsque le nombre de Reynolds est élevé, l'écoulement peut devenir turbulent.

1.1.5 Convection naturelle et forcée - Approximation de Boussinesq

On considère ici les phénomènes de convection forcée et naturelle, qui sont la mise en mouvement du fluide et le transport de la chaleur par celui-ci. Ce transport de chaleur est décrit par l'équation portant sur l'énergie. Celle-ci peut être transformée en partant de (8) et en ôtant le produit scalaire de (7) avec \mathbf{u} , et en tenant compte du fait que ρ est (presque) constant, que $\nabla \cdot \mathbf{u} = 0$ et que $\underline{\sigma}_v$ est symétrique. On a:

$$\begin{aligned} \frac{\partial \rho \mathbf{u}}{\partial t} \cdot \mathbf{u} &= \rho \frac{1}{2} \frac{\partial |\mathbf{u}|^2}{\partial t} \\ [\nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u})] \cdot \mathbf{u} &= \rho (\mathbf{u} \cdot \nabla \mathbf{u}) \cdot \mathbf{u} = \frac{1}{2} \rho \nabla \cdot (|\mathbf{u}|^2 \mathbf{u}) \\ \nabla p \cdot \mathbf{u} &= \nabla \cdot (p \mathbf{u}) \\ \nabla \cdot (\underline{\sigma}_v) \cdot \mathbf{u} &= \nabla \cdot (\underline{\sigma}_v \cdot \mathbf{u}) - \underline{\sigma}_v : \nabla \mathbf{u} = \nabla \cdot (\underline{\sigma}_v \mathbf{u}) - 2\mu \underline{\mathbf{D}} : \underline{\mathbf{D}}. \end{aligned}$$

Étant donné la relation entre les énergies (5), on obtient:

$$\rho \frac{\partial \varepsilon}{\partial t} + \rho \mathbf{u} \cdot \nabla \varepsilon - k \Delta T = 2\mu \underline{\mathbf{D}} : \underline{\mathbf{D}}. \quad (26)$$

De plus, des éléments de thermodynamique indiquent qu'à pression constante, il existe une relation simple entre les variations d'énergie interne et les variations de températures:

$$d\varepsilon = C_p dT.$$

En remplaçant dans (26), et en négligeant le terme $2\mu\mathbf{D} : \mathbf{D}$ (variation de température due aux frottements des molécules d'air sur elles-mêmes), et en tenant compte d'un terme source (chauffage) on obtient:

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T - \frac{k}{\rho C_p} \Delta T = s, \quad (27)$$

qui, en adimensionné, donne

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T - \frac{1}{Pe} \Delta T = \tilde{s}, \quad (28)$$

où l'on a défini le nombre de Péclet $Pe = \frac{\bar{U} L \rho C_p}{k}$.

Dans la convection forcée, un élément extérieur exerce une force prescrite sur le fluide (ventilateur etc) et le met en mouvement; la force de gravité est alors négligeable devant cette force et n'est pas prise en compte dans le modèle. Les équations du fluide et de la thermique sont découplées: On résout (23) et (25), puis l'équation de l'énergie (28).

En régime de convection naturelle, un élément extérieur chauffe le fluide et ce sont les variations de la densité, dues à la dilatation qui mettent en mouvement le fluide: Dans (25), la force \mathbf{F} se résume à la force de gravité. Dans ce régime, celle-ci joue donc un rôle essentiel.

En particulier, on définit une pression d'équilibre (la pression hydrostatique, celle qui est obtenue lorsque la vitesse est nulle, c'est-à-dire lorsqu'il n'y a pas de mouvement) par

$$\nabla p_{eq} = -\frac{1}{Fr^2} \mathbf{e}_z.$$

Ensuite, on relie les variations de densité aux variations de température par la relation suivante, si l'on suppose que $\rho = \rho(T)$ puisque la pression est quasiment constante: $\rho = \rho_0(1 - \alpha \delta T)$ avec $\alpha = -\frac{1}{\rho_0} \frac{\partial \rho}{\partial T}(T_0)$ (gaz parfaits $\alpha = \frac{1}{T_0}$); on obtient, en éliminant la pression d'équilibre et en adimensionnant la perturbation de température par la température de référence T_0 :

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla \delta p - \frac{1}{Re} \Delta \mathbf{u} = \frac{\alpha T_0}{Fr^2} \delta T \mathbf{e}_z, \quad (29)$$

où δT est maintenant la perturbation de température adimensionnée. Puisque dans (28) la température n'intervient que par ses dérivées et que la température de référence est constante en espace et en temps, on obtient que la perturbation de température δT vérifie la même équation:

$$\frac{\partial \delta T}{\partial t} + \mathbf{u} \cdot \nabla \delta T - \frac{1}{Pe} \Delta \delta T = \tilde{s}. \quad (30)$$

On voit qu'en régime de convection naturelle, la vitesse du fluide et la variation de température sont liées par les équations (29) et (30).

1.2 Classification des EDPs

1.2.1 Rappel de la définition usuelle

On se place dans un cadre scalaire, avec une fonction ne dépendant que de deux variables (x, y) par ex. ou encore (x, t) lorsque la seconde variable est le temps. Une EDP linéaire du second ordre, écrite sous la forme

$$A\partial_{xx}u + B\partial_{xy}u + C\partial_{yy}u + \dots = 0$$

(où \dots représente les termes avec des dérivées premières ou pas de dérivées) est dite

- hyperbolique si $B^2 - 4AC > 0$ (exemple type: ondes $A = 1, B = 0, C = -1$),
- parabolique si $B^2 - 4AC = 0$ (exemple type: chaleur $A = -1, B = 0, C = 0$),
- elliptique si $B^2 - 4AC < 0$ (exemple type: Laplace $A = 1, B = 0, C = 1$).

1.2.2 EDPs hyperboliques: acoustique - Euler

Ici nous nous plaçons dans le cadre de systèmes, possiblement multidimensionnels, et d'ordre un. La notion d'hyperbolicité est donc à redéfinir, en vérifiant que dans le cas le plus simple on retombe sur la définition ci-dessus.

On mène en parallèle à titre d'exemple le système des ondes en dimension 2, après l'avoir obtenu des équations d'Euler isentropiques dont on considère de petites perturbations.

On reprend l'équation de conservation de la masse et de la quantité de mouvement non visqueuse, ainsi que la loi d'état pour un gaz parfait isentropique:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (31)$$

Le bilan de quantité de mouvement s'écrit:

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p = 0, \quad (32)$$

La loi d'état est la suivante:

$$p = C\rho^\gamma \quad (33)$$

Ici, on considère de très petites perturbations autour de l'état d'équilibre du fluide (typiquement l'air). Cet état d'équilibre est décrit par $\rho_0 = \text{cte}$ (et donc $p_0 = \text{cte}$) et $\mathbf{u}_0 = 0$. On écrira donc

$$\mathbf{u} = \varepsilon \mathbf{u}_1 \ ; \ \rho = \rho_0 + \varepsilon \rho_1 \quad (34)$$

On obtient alors

$$p(\rho) = p(\rho_0 + \varepsilon \rho_1) \approx p(\rho_0) + \varepsilon \rho_1 \frac{\partial p}{\partial \rho}(\rho_0)$$

et donc

$$\nabla p \approx \varepsilon \frac{\partial p}{\partial \rho}(\rho_0) \nabla \rho_1 \quad (35)$$

En injectant (34) dans (31), on obtient

$$\frac{\partial(\rho_0 + \varepsilon \rho_1)}{\partial t} + \nabla \cdot ((\rho_0 + \varepsilon \rho_1) \varepsilon \mathbf{u}_1) = 0$$

Soit, à l'ordre 1 en ε :

$$\frac{\partial \rho_1}{\partial t} + \rho_0 \nabla \cdot \mathbf{u}_1 = 0$$

En notant $r = \frac{\rho_1}{\rho_0}$, on a

$$\frac{\partial r}{\partial t} + \nabla \cdot \mathbf{u}_1 = 0 \quad (36)$$

En injectant (34) et (35) dans (32), on obtient

$$\frac{\partial[(\rho_0 + \varepsilon \rho_1) \varepsilon \mathbf{u}_1]}{\partial t} + \nabla \cdot ((\rho_0 + \varepsilon \rho_1) \varepsilon^2 \mathbf{u}_1 \otimes \mathbf{u}_1) + \varepsilon \frac{\partial p}{\partial \rho}(\rho_0) \nabla \rho_1 = 0.$$

Soit, en ne conservant que les termes d'ordre ε :

$$\rho_0 \frac{\partial \mathbf{u}_1}{\partial t} + 0 + \frac{\partial p}{\partial \rho}(\rho_0) \nabla \rho_1 = 0$$

En divisant par ρ_0 et en tenant compte du fait que la loi d'état fournit

$$\frac{\partial p}{\partial \rho}(\rho_0) = \gamma C \rho_0^{\gamma-1} = \gamma \frac{C \rho_0^\gamma}{\rho} = \gamma \frac{p_0}{\rho_0},$$

et en utilisant la loi des gaz parfaits $p_0 = \frac{R}{M} \rho_0 T_0$, on obtient finalement

$$\frac{\partial \mathbf{u}_1}{\partial t} + \gamma \frac{R}{M} T_0 \nabla r = 0.$$

En remarquant que $\gamma \frac{R}{M} T_0$ est homogène au carré d'une vitesse, on définit la vitesse du son dans le milieu par

$$c = \sqrt{\gamma \frac{R}{M} T_0}$$

et on obtient

$$\frac{\partial \mathbf{u}_1}{\partial t} + c^2 \nabla r = 0. \quad (37)$$

Le système (36)-(37) est l'équation des ondes sous forme de système du premier ordre. On peut en déduire une équation des ondes pour r en dérivant (36) par rapport à t et en remplaçant $\frac{\partial \mathbf{u}_1}{\partial t}$ par sa valeur tirée de (37):

$$\frac{\partial^2 r}{\partial t^2} + \nabla \cdot \left(\frac{\partial \mathbf{u}_1}{\partial t} \right) = 0,$$

$$\frac{\partial^2 r}{\partial t^2} + \nabla \cdot (-c^2 \nabla r) = 0,$$

$$\frac{\partial^2 r}{\partial t^2} - c^2 \Delta r = 0.$$

C'est bien une équation hyperbolique, et pour définir l'hyperbolicité des systèmes, nous nous plaçons dans le cadre suivant:

On considère un système de p équations à p inconnues $\mathbf{w} = (w_1, \dots, w_p)^T$, dépendant d'une variable temporelle t et de variables (x_1, \dots, x_d) dans \mathbb{R}^d , écrit sous la forme de lois de conservation:

$$\frac{\partial \mathbf{w}}{\partial t} + \sum_{j=1}^d \frac{\partial \mathbf{f}_j(\mathbf{w})}{\partial x_j} = 0, \quad (38)$$

où les $\mathbf{f}_j = (\mathbf{f}_{1j}, \dots, \mathbf{f}_{pj})^T$ sont d fonctions "de flux", supposées régulières et définies de \mathbb{R}^p (ou dans un sous-ensemble de \mathbb{R}^p) à valeurs dans \mathbb{R}^p . En posant $\mathbf{f} = (\mathbf{f}_1, \dots, \mathbf{f}_d)$, (\mathbf{f} est une matrice à p lignes et d colonnes), ce système s'écrit aussi souvent:

$$\frac{\partial \mathbf{w}}{\partial t} + \operatorname{div}_x(\mathbf{f}(\mathbf{w})) = 0. \quad (39)$$

Pour l'équation des ondes ci-dessus, dans \mathbb{R}^2 , nous avons $d = 2$ et $p = 3$ car $\mathbf{w} = (r, u_x, u_y)^T$ et on a

$$\begin{aligned} \frac{\partial r}{\partial t} + \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} &= 0 \\ \frac{\partial u_1}{\partial t} + c^2 \frac{\partial r}{\partial x} &= 0 \\ \frac{\partial u_2}{\partial t} + c^2 \frac{\partial r}{\partial y} &= 0 \end{aligned}$$

donc $\mathbf{f}_1(\mathbf{w}) = (u_1, c^2 r, 0)^T$ et $\mathbf{f}_2(\mathbf{w}) = (u_2, 0, c^2 r)^T$.

Ce système est écrit ici sous sa forme conservative. En appliquant la composition des dérivées:

$$\frac{\partial \mathbf{f}_j(\mathbf{w})}{\partial x_j} = \sum_{k=1}^p \frac{\partial \mathbf{f}_j}{\partial w_k} \frac{\partial w_k}{\partial x_j},$$

on obtient le système sous forme non-conservative:

$$\frac{\partial \mathbf{w}}{\partial t} + \sum_{j=1}^d A_j(\mathbf{w}) \frac{\partial \mathbf{w}}{\partial x_j} = 0, \quad (40)$$

où la matrice jacobienne $A_j(\mathbf{w})$ est dans $\mathbb{R}^{p \times p}$ et où ses éléments sont définis par

$$(A_j(\mathbf{w}))_{i,k} = \left(\frac{\partial f_{ij}}{\partial w_k}(\mathbf{w}) \right).$$

Chaque ligne de (40) s'écrit aussi

$$\frac{\partial w_i}{\partial t} + \sum_{j=1}^d \sum_{k=1}^p \left(\frac{\partial f_{ij}}{\partial w_k}(\mathbf{w}) \right) \frac{\partial \mathbf{w}}{\partial x_j} = 0. \quad (41)$$

On remarque que dans le cas où les fonctions de flux sont linéaires, alors les matrices sont constantes.

Dans le cas de l'équation des ondes, on a

$$\begin{aligned} \frac{\partial \mathbf{f}_1}{\partial x} &= \begin{pmatrix} \frac{\partial u_1}{\partial x} \\ c^2 \frac{\partial r}{\partial x} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ c^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial r}{\partial x} \\ \frac{\partial u_1}{\partial x} \\ \frac{\partial u_2}{\partial x} \end{pmatrix} \\ \frac{\partial \mathbf{f}_2}{\partial y} &= \begin{pmatrix} \frac{\partial u_2}{\partial y} \\ 0 \\ c^2 \frac{\partial r}{\partial y} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ c^2 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial r}{\partial y} \\ \frac{\partial u_1}{\partial y} \\ \frac{\partial u_2}{\partial y} \end{pmatrix} \end{aligned}$$

Et donc

$$A_1 = \begin{pmatrix} 0 & 1 & 0 \\ c^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ c^2 & 0 & 0 \end{pmatrix}$$

Le système est dit hyperbolique si pour tout \mathbf{w} (éventuellement sur un sous-ensemble de \mathbb{R}^p) et tout $\omega = (\omega_1, \dots, \omega_d) \in \mathbb{R}^d$, $\omega \neq 0$, la matrice $A(\mathbf{w}, \omega) = \sum_{j=1}^d \omega_j A_j(\mathbf{w})$ est diagonalisable dans \mathbb{R} , c'est-à-dire qu'elle possède p valeurs propres réelles et une base de p vecteurs propres associés. Le système est dit strictement hyperbolique si les valeurs propres sont toutes distinctes.

Vérifications sur l'équation des ondes:

$$A(\mathbf{w}, \omega) = w_1 A_1 + w_2 A_2 = \begin{pmatrix} 0 & \omega_1 & \omega_2 \\ c^2 \omega_1 & 0 & 0 \\ c^2 \omega_2 & 0 & 0 \end{pmatrix} \quad (42)$$

On a (développement par rapport à la troisième ligne)

$$\begin{aligned} \begin{vmatrix} -\lambda & \omega_1 & \omega_2 \\ c^2 \omega_1 & -\lambda & 0 \\ c^2 \omega_2 & 0 & -\lambda \end{vmatrix} &= c^2 \omega_2 \begin{vmatrix} \omega_1 & \omega_2 \\ -\lambda & 0 \end{vmatrix} - \lambda \begin{vmatrix} -\lambda & \omega_1 \\ c^2 \omega_1 & -\lambda \end{vmatrix} \\ &= c^2 \omega_2 (\lambda \omega_2) - \lambda (\lambda^2 - c^2 \omega_1^2) \\ &= -\lambda [\lambda^2 - c^2 (\omega_1^2 + \omega_2^2)]. \end{aligned}$$

On a donc trois valeurs propres réelles distinctes, le système est strictement hyperbolique.

Importance de l'hyperbolicité pour avoir la stabilité de la solution par rapport aux données en tout temps:

Considérons le système suivant

$$\begin{cases} \frac{\partial w_1}{\partial t} + c \frac{\partial w_2}{\partial x} = 0 \\ \frac{\partial w_2}{\partial t} - c \frac{\partial w_1}{\partial x} = 0 \end{cases}.$$

On est dans le cas $d = 1$ et $p = 2$, il n'y a donc qu'une seule matrice $A_j = A$

$$A = \begin{pmatrix} 0 & c \\ -c & 0 \end{pmatrix}.$$

On a

$$\begin{vmatrix} -\lambda & c \\ -c & -\lambda \end{vmatrix} = \lambda^2 + c^2$$

dont les valeurs propres $\pm ic$ ne sont pas réelles, mais complexes. La matrice est diagonalisable mais dans \mathbb{C} . Le système n'est pas hyperbolique (il ne vérifie pas la définition).

En additionnant la première ligne et i fois la seconde d'une part et en additionnant la première ligne et $-i$ fois la seconde d'autre part, on trouve

$$\begin{cases} \frac{\partial(w_1 + iw_2)}{\partial t} + c \frac{\partial(w_2 - iw_1)}{\partial x} = 0 \\ \frac{\partial(w_1 - iw_2)}{\partial t} + c \frac{\partial(w_2 + iw_1)}{\partial x} = 0 \end{cases}.$$

Soit, en posant $r_1 = (w_1 + iw_2)$ et $r_2 = (w_1 - iw_2)$:

$$\begin{cases} \frac{\partial r_1}{\partial t} - ic \frac{\partial r_1}{\partial x} = 0 \\ \frac{\partial r_2}{\partial t} + ic \frac{\partial r_2}{\partial x} = 0 \end{cases},$$

ce qui montre que si on choisit les conditions initiales telles que

$$w_1(t = 0, x) = A \cos(kx) \ ; \ w_2(t = 0, x) = -A \sin(kx),$$

alors cela implique

$$r_1(t = 0, x) = (w_1 + iw_2)(t = 0, x) = A \cos(kx) - iA \sin(kx) = A \exp(-ikx)$$

$$r_2(t = 0, x) = (w_1 - iw_2)(t = 0, x) = A \cos(kx) + iA \sin(kx) = A \exp(ikx),$$

et alors la solution est donnée par

$$r_1(t, x) = A \exp(ckt - ikx) \ ; \ r_2(t, x) = A \exp(ckt + ikx).$$

En effet, les conditions initiales sont vérifiées, et l'on calcule

$$\frac{\partial r_1}{\partial t} - ic \frac{\partial r_1}{\partial x} = [ck - ic(-ik)] r_1 = 0$$

$$\frac{\partial r_2}{\partial t} + ic \frac{\partial r_2}{\partial x} = [ck - ic(ik)] r_2 = 0.$$

Et donc, comme $w_1 = \frac{r_1+r_2}{2}$ et $w_2 = i \frac{(r_2-r_1)}{2}$:

$$w_1(x, t) = A \exp(ckt) \cos(kx) \ ; \ w_2(x, t) = -A \exp(ckt) \sin(kx)$$

ce qui montre une instabilité dans le système. On voit donc bien que le fait que la matrice soit diagonalisable dans \mathbb{C} pose problème.

Considérons maintenant le système suivant

$$\begin{cases} \frac{\partial w_1}{\partial t} + \lambda \frac{\partial w_1}{\partial x} + \frac{\partial w_2}{\partial x} = 0 \\ \frac{\partial w_2}{\partial t} + \lambda \frac{\partial w_2}{\partial x} = 0 \end{cases}$$

avec pour conditions initiales

$$w_1(x, t=0) = w_1^0(x) \ , \ w_2(x, t=0) = w_2^0(x).$$

Le flux s'écrit $\mathbf{f} = (\lambda w_1 + w_2, \lambda w_2)^T$ et la matrice 2×2 associée est

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Elle est bien à coefficients réels, à valeurs propres réelles mais non-diagonalisable. Le système n'est toujours pas hyperbolique.

On peut vérifier que la solution est

$$\begin{cases} w_1(x, t) = w_1^0(x - \lambda t) - t (w_2^0)'(x - \lambda t) \\ w_2(x, t) = w_2^0(x - \lambda t) \end{cases}$$

En effet, en $t = 0$, on retrouve bien les conditions initiales et de plus

$$\frac{\partial w_1}{\partial t} = -\lambda (w_1^0)' - (w_2^0)' + \lambda t (w_2^0)''$$

$$\frac{\partial w_1}{\partial x} = (w_1^0)' - t (w_2^0)''$$

et

$$\frac{\partial w_2}{\partial t} = -\lambda (w_2^0)'$$

$$\frac{\partial w_2}{\partial x} = (w_2^0)'$$

D'où

$$\begin{cases} \frac{\partial w_1}{\partial t} + \lambda \frac{\partial w_1}{\partial x} + \frac{\partial w_2}{\partial x} &= -\lambda(w_1^0)' - (w_2^0)' + \lambda t(w_2^0)'' + \lambda[(w_1^0)' - t(w_2^0)''] + (w_2^0)' \\ \frac{\partial w_2}{\partial t} + \lambda \frac{\partial w_2}{\partial x} &= -\lambda(w_2^0)' + \lambda(w_2^0)' \end{cases}$$

Propriété importante: l'hyperbolicité est conservée en cas de changement de variables.

1.2.3 Autre exemple hyperbolique: Maxwell

Les équations de Maxwell décrivent l'évolution d'un champ électromagnétique; les inconnues sont le champ électrique $\mathbf{E} = (E_x, E_y, E_z)$ et le champ magnétique $\mathbf{B} = (B_x, B_y, B_z)$. On distingue la loi de Faraday

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0, \quad (43)$$

et la loi d'Ampère

$$\frac{\partial \mathbf{E}}{\partial t} - c^2 \nabla \times \mathbf{B} = -\frac{\mathbf{j}}{\varepsilon_0}, \quad (44)$$

dans laquelle c est la vitesse de la lumière dans le milieu de propagation, \mathbf{j} une densité de courant électrique (considérée dans ce qui suit comme un terme source indépendant des inconnues; toutefois, dans les milieux conducteurs on peut avoir une relation du type $\mathbf{j} = \sigma \mathbf{E}$, avec σ la conductivité électrique du milieu) et ε_0 la permittivité diélectrique du milieu. On rappelle que

$$\nabla \times \mathbf{E} = \begin{pmatrix} \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} \\ \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} \\ \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} \end{pmatrix}.$$

Dans un cadre bidimensionnel, on suppose que les inconnues sont indépendantes de la variable verticale z . Par conséquent, si on considère que $\frac{\partial}{\partial z} = 0$ dans les équations précédentes, on s'aperçoit que (43)-(44) se découpent en deux sous-systèmes indépendants, l'un ne faisant intervenir que le jeu d'inconnues (B_z, E_x, E_y) et l'autre que le jeu d'inconnues (E_z, B_x, B_y) . Le traitement de ces deux sous-systèmes étant quasiment identique, nous nous contentons d'étudier le premier d'entre eux. Si on pose $\mathbf{w} = (B_z, E_x, E_y)^T$ on a

$$\begin{aligned} \frac{\partial B_z}{\partial t} + \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} &= 0 \\ \frac{\partial E_x}{\partial t} - c^2 \frac{\partial B_z}{\partial y} &= -\frac{j_x}{\varepsilon_0} \\ \frac{\partial E_y}{\partial t} + c^2 \frac{\partial B_z}{\partial x} &= -\frac{j_y}{\varepsilon_0} \end{aligned}$$

donc $\mathbf{f}_1(\mathbf{w}) = (E_y, 0, c^2 B_z)^T$ et $\mathbf{f}_2(\mathbf{w}) = (-E_x, -c^2 B_z, 0)^T$. On a

$$\begin{aligned}\frac{\partial \mathbf{f}_1}{\partial x} &= \begin{pmatrix} \frac{\partial E_y}{\partial x} \\ 0 \\ c^2 \frac{\partial B_z}{\partial x} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ c^2 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial B_z}{\partial x} \\ \frac{\partial E_x}{\partial x} \\ \frac{\partial E_y}{\partial x} \end{pmatrix} \\ \frac{\partial \mathbf{f}_2}{\partial y} &= \begin{pmatrix} -\frac{\partial E_x}{\partial y} \\ -c^2 \frac{\partial B_z}{\partial y} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ -c^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial B_z}{\partial y} \\ \frac{\partial E_x}{\partial y} \\ \frac{\partial E_y}{\partial y} \end{pmatrix}\end{aligned}$$

Et donc

$$A_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ c^2 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -1 & 0 \\ -c^2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

On a alors

$$A(\mathbf{w}, \omega) = w_1 A_1 + w_2 A_2 = \begin{pmatrix} 0 & -\omega_2 & \omega_1 \\ -c^2 \omega_2 & 0 & 0 \\ c^2 \omega_1 & 0 & 0 \end{pmatrix}$$

Cette matrice est très proche de la matrice associée à l'équation des ondes (42); elle en est obtenue en remplaçant ω_1 par $-\omega_2$ et ω_2 par ω_1 . On obtient donc directement

$$\begin{vmatrix} -\lambda & -\omega_2 & \omega_1 \\ -c^2 \omega_2 & -\lambda & 0 \\ c^2 \omega_1 & 0 & -\lambda \end{vmatrix} = -\lambda[\lambda^2 - c^2(\omega_1^2 + \omega_2^2)].$$

On a donc trois valeurs propres réelles distinctes, le système est strictement hyperbolique.

Le cas vraiment tridimensionnel est plus compliqué car d'une part de plus grande dimension ($p = 6, d = 3$) et d'autre part l'analyse montre que la matrice associée possède trois valeurs propres doubles ($c, 0$ et $-c$), il faut donc pour chacune de ces trois valeurs propres montrer que le sous-espace propre associé est de dimension deux (laissé en exercice au lecteur).

2 Hyperbolique scalaire 1D

On se place ici dans le cas $p = d = 1$, c'est le cas le plus simple, et on va pourtant voir que des difficultés importantes apparaissent en raison de la non-linéarité de l'équation.

On va étudier l'équation suivante, d'inconnue la fonction u dépendant des variables (x, t) (x est la position dans l'espace à une dimension et t est le temps)

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0 \quad , \quad \forall (x, t) \in \mathbb{R} \times \mathbb{R}^+ \quad (45)$$

avec la condition initiale

$$u(x, t = 0) = u_0(x) \quad , \quad \forall x \in \mathbb{R}. \quad (46)$$

La fonction f est supposée **convexe** et suffisamment régulière. On notera

$$a = f', \quad (47)$$

et on notera que a est une fonction croissante, par convexité de f .

Tant que la fonction u est régulière, on peut appliquer la loi de composition des dérivées, et on obtient la forme non-conservative de l'équation (45) (qui était, elle, sous forme conservative)

$$\frac{\partial u}{\partial t} + a(u) \frac{\partial u}{\partial x} = 0. \quad (48)$$

On tombe donc sur une équation de transport dont la vitesse $a(u)$ dépend de la valeur de la solution.

Les deux cas typiques sont $f(u) = cu$, avec $c \in \mathbb{R}$, on a donc $a(u) = c$, qui ne dépend pas de u , c'est l'advection **linéaire** bien connue, et d'autre part $f(u) = \frac{u^2}{2}$, qui se nomme alors équation de Burgers (sans viscosité). L'équation de Burgers est très fortement liée à l'équation d'Euler unidimensionnelle dans laquelle on considère que densité (et donc pression) sont constantes.

2.1 Caractéristiques

2.1.1 Cas du transport linéaire

Dans le cas linéaire, c'est-à-dire lorsque $f(u) = cu$, avec $c \in \mathbb{R}$, la solution de (45)-(46) est donnée par

$$u(x, t) = u_0(x - ct). \quad (49)$$

Il est en effet immédiat de vérifier que la relation donnée par (49) évaluée en $t = 0$ coïncide avec (46) d'une part, et que d'autre part la solution proposée par (49) a pour dérivées partielles (en supposant u_0 dérivable)

$$\frac{\partial u}{\partial t}(x, t) = -cu'_0(x - ct) \quad \text{et} \quad \frac{\partial u}{\partial x}(x, t) = u'_0(x - ct),$$

ce qui permet d'obtenir (45).

Or, ce que l'on peut conclure de (49), c'est que la valeur de la solution $u(x, t)$ pour tous les points (x, t) situés sur une même droite d'équation

$$x = \xi + ct \quad \text{avec} \quad \xi \in \mathbb{R} \quad (50)$$

est identique (et vaut $u_0(\xi)$). Autrement dit, la valeur de u est constante le long de chacune des droites d'équation donnée par (50). Ces droites sont paramétrées par $\xi \in \mathbb{R}$, c'est-à-dire que pour chaque $\xi \in \mathbb{R}$ on peut tracer une droite différente, le long de laquelle $u(x, t)$ vaut la même valeur, à savoir $u_0(\xi)$. Ces droites sont appelées **droites caractéristiques**, ou plus simplement **caractéristiques**. Le réel ξ est appelé **le pied de la caractéristique**. Lorsqu'on les trace dans le plan (x, t) , elles sont toutes parallèles, de coefficient directeur $\frac{1}{c}$.

2.1.2 Cas non-linéaire

Dans le cas non-linéaire, nous allons essayer de déterminer des courbes dans le plan (x, t) le long desquelles $u(x, t)$ serait constant. "Déterminer des courbes" revient à trouver des fonctions $t \mapsto \tilde{x}(t)$. En effet, si l'on connaissait de telles courbes, et si par chaque point du plan (x, t) ne passait qu'une et une seule de ces courbes, notée $s \mapsto (\tilde{x}_{(x,t)}(s), s)$, alors on aurait

$$u(x, t) = u(\tilde{x}_{(x,t)}(s), s) \quad , \quad \forall s \geq 0 \quad (51)$$

et donc en particulier

$$u(x, t) = u(\tilde{x}_{(x,t)}(0), 0) = u_0(\tilde{x}_{(x,t)}(0)) \quad , \quad (52)$$

ce qui fournit la valeur de u au point (x, t) en fonction de la valeur de u_0 au pied $\tilde{x}_{(x,t)}(0)$ de la caractéristique.

Cherchons donc si de telles courbes $t \mapsto (\tilde{x}(t), t)$ existent, notons $\xi = \tilde{x}(0)$ et raisonnons par condition nécessaire. Considérons la fonction $v : t \mapsto u(\tilde{x}(t), t)$. Si elle ne dépend pas de t , alors sa dérivée par rapport à t est nulle et on a

$$\frac{dv}{dt}(t) = \frac{\partial u}{\partial x}(\tilde{x}(t), t) \frac{d\tilde{x}}{dt}(t) + \frac{\partial u}{\partial t}(\tilde{x}(t), t) = 0. \quad (53)$$

Si l'on compare (53) et (48), on se rend compte que choisir

$$\frac{d\tilde{x}}{dt}(t) = a(u(\tilde{x}(t), t)) \quad (54)$$

nous permettrait d'arriver à nos fins. De plus, puisque la valeur de u est constante le long de ces courbes, elle vaut $u(\tilde{x}(t), t) = u(\tilde{x}(0), 0) = u_0(\xi)$, et la relation (54) devient alors

$$\frac{d\tilde{x}}{dt}(t) = a(u_0(\xi)) \quad \text{avec} \quad \tilde{x}(t=0) = \xi. \quad (55)$$

À nouveau, comme dans le cas linéaire, les courbes définies par (55) sont des droites, paramétrées par $\xi \in \mathbb{R}$, d'équation

$$\tilde{x}(t) = \xi + a(u_0(\xi)) t. \quad (56)$$

Inversement, vérifions à présent que, si pour tout (x, t) (avec $0 \leq t \leq t^*$ pour un $t^* > 0$ que nous déterminerons) il existe un unique $\xi(x, t)$ vérifiant

$$\xi + a(u_0(\xi)) t = x, \quad (57)$$

alors $u : (x, t) \mapsto u(x, t) = u_0(\xi(x, t))$ est solution de (45)–(46), dès que $a \circ u_0$ est dérivable. En effet, cherchons tout d’abord à partir de (57) une relation liant les dérivées partielles de ξ par rapport à x et à t . En dérivant (57) par rapport à x et à t on obtient

$$\begin{cases} (1 + (a \circ u_0)'(\xi) t) \frac{\partial \xi}{\partial x} &= 1 \\ (1 + (a \circ u_0)'(\xi) t) \frac{\partial \xi}{\partial t} + (a \circ u_0)(\xi) &= 0 \end{cases}$$

En multipliant la première de ces deux égalités par $(a \circ u_0)(\xi)$ et en ajoutant le résultat à la seconde, on obtient après simplification:

$$(1 + (a \circ u_0)'(\xi) t) \left(\frac{\partial \xi}{\partial t} + (a \circ u_0)(\xi) \frac{\partial \xi}{\partial x} \right) = 0.$$

Tant que $t < t^* := (\inf - \frac{1}{(a \circ u_0)'})$ si u_0 est strictement décroissante au moins sur un intervalle ou alors pour tout t si u_0 est croissante au sens large sur \mathbb{R} , on peut alors conclure que

$$\left(\frac{\partial \xi}{\partial t} + (a \circ u_0)(\xi) \frac{\partial \xi}{\partial x} \right) = 0. \quad (58)$$

Étant donné que pour $u : (x, t) \mapsto u(x, t) = u_0(\xi(x, t))$ on a

$$\left(\frac{\partial u}{\partial t} + (a \circ u) \frac{\partial u}{\partial x} \right) (x, t) = u_0'(\xi(x, t)) \left(\frac{\partial \xi}{\partial t} + (a \circ u_0)(\xi) \frac{\partial \xi}{\partial x} \right) (x, t)$$

on peut conclure de (58) que $u : (x, t) \mapsto u(x, t) = u_0(\xi(x, t))$ satisfait (45); de plus (46) est également vérifiée car $\xi(x, t = 0) = x$.

À présent, une question légitime est de savoir si par un point donné (x, t) du plan, il passe une et une seule droite caractéristique. En effet, si c’est le cas et que l’on sait trouver la valeur de ξ résolvant (57), alors on pourra calculer la solution $u(x, t)$, elle vaudra $u_0(\xi(x, t))$. Il faut bien avoir à l’esprit que (57) est une équation sur ξ , dont la solution dépend de x et de t (c’est pour cela qu’on a noté $\xi(x, t)$). Nous énonçons une hypothèse suffisante dans laquelle cela sera le cas:

Proposition: La fonction a étant supposée continue, **si la fonction u_0 est continue et croissante**, alors il existe une unique solution à l’équation (57). En effet, f étant convexe, sa dérivée a est une fonction croissante; elle est de plus supposée continue. Si u_0 est supposée croissante continue, alors par composition $a \circ u_0$ est aussi croissante et continue. D’autre part, $t \geq 0$ et $\xi \mapsto \xi$ est strictement croissante, continue et tend vers $\pm\infty$ en $\pm\infty$. Tout ceci implique que la fonction $\xi \mapsto \xi + a(u_0(\xi)) t - x$ est une fonction continue, strictement croissante et qui tend vers $\pm\infty$ en $\pm\infty$. Elle s’annule donc une et une seule fois sur \mathbb{R} , en un point $\xi(x, t)$.

Nous donnons maintenant un exemple pour lequel ces conditions sont vérifiées, et construisons également ensuite deux contre-exemples, dans lesquels soit u_0

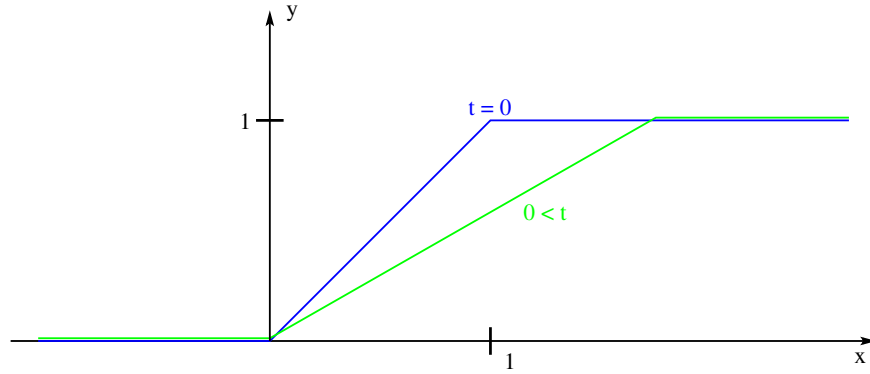


Figure 1: Cas où la donnée initiale est croissante et continue: courbe $y = u(x, t)$ pour $t = 0$ (condition initiale en bleu) et pour $t > 0$ (en vert)

est croissante mais discontinue, soit u_0 est continue mais décroissante. Dans tous les cas nous considérons l'équation de Burgers pour laquelle $a(u) = u$.

Cas de u_0 croissante continue

Nous choisissons u_0 sous la forme

$$u_0(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x \geq 1 \end{cases} \quad (59)$$

qui est bien croissante et continue. Cette condition initiale est représentée en bleu sur la figure 1.

Calculons les équations des caractéristiques. D'après (56), on a

$$x(t) = \xi + u_0(\xi)t.$$

Il faut donc discuter selon les valeurs de ξ . Selon (59), cela mène à

$$\begin{cases} x(t) = \xi & \text{si } \xi \leq 0 \\ x(t) = \xi + \xi t & \text{si } 0 \leq \xi \leq 1 \\ x(t) = \xi + t & \text{si } \xi \geq 1 \end{cases} \quad (60)$$

Les droites correspondant à (60) sont tracées sur la figure 2.

On peut à présent vérifier que par un point du plan (x, t) il passe une et unique caractéristique, dont nous allons pouvoir déterminer le pied, ce qui permettra de calculer $u(x, t)$. Sur la figure 2, on remarque qu'un point (x, t) peut se situer dans trois zones distinctes du plan:

1. Soit $x \leq 0$ (caractéristiques verticales)
2. Soit $0 \leq x \leq 1 + t$ (caractéristiques de pentes de plus en plus faibles)
3. Soit $1 + t \leq x$ (caractéristiques toutes parallèles, de pente 1)

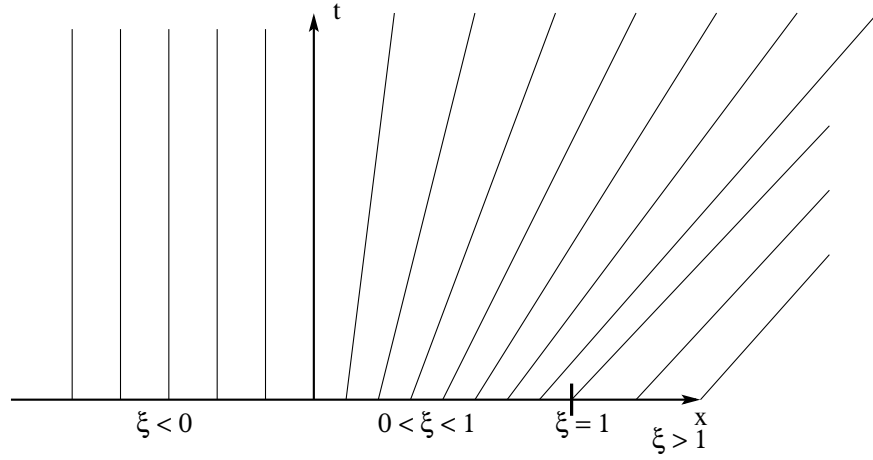


Figure 2: Droites caractéristiques dans le cas où u_0 est croissante et continue

Dans le cas 1., l'unique caractéristique passant par (x, t) a pour équation $x = \xi$, donc $\xi = x$ et on a $u(x, t) = u_0(\xi) = u_0(x) = 0$.

Dans le cas 2., l'unique caractéristique passant par (x, t) a pour équation $x = \xi + \xi t$, donc $\xi = \frac{x}{(1+t)}$ et on a $u(x, t) = u_0(\xi) = \xi = \frac{x}{(1+t)}$ (en effet $0 \leq x \leq 1+t$ implique $0 \leq \xi \leq 1$).

Dans le cas 3., l'unique caractéristique passant par (x, t) a pour équation $x = \xi + t$, donc $\xi = x - t$ et on a $u(x, t) = u_0(\xi) = 1$ (en effet $1+t \leq x$ implique $1 \leq x - t = \xi$).

Pour résumer, la solution est donnée pour tout temps t par

$$u(x, t) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x}{(1+t)} & \text{si } 0 \leq x \leq 1+t \\ 1 & \text{si } x \geq 1+t \end{cases} . \quad (61)$$

Cette solution est représentée en vert sur la figure 1.

Cas de u_0 croissante discontinue

Nous choisissons u_0 sous la forme

$$u_0(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{cases} \quad (62)$$

qui est bien croissante et discontinue. Cette condition initiale est représentée en bleu sur la figure 3.

Calculons les équations des caractéristiques. D'après (56), on a

$$x(t) = \xi + u_0(\xi)t.$$

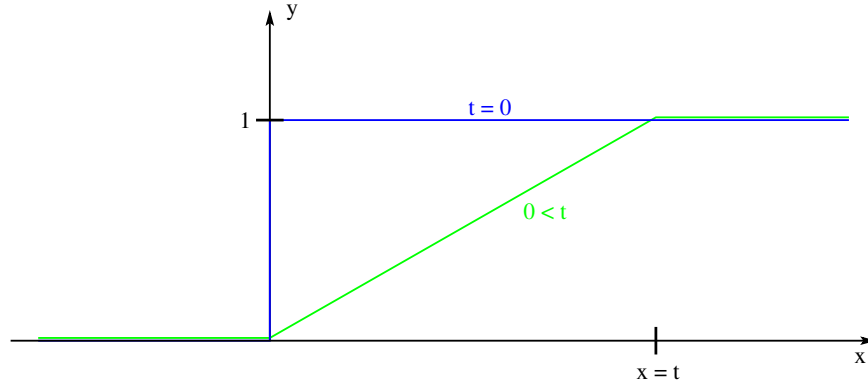


Figure 3: Cas où la donnée initiale est croissante et discontinue: courbe $y = u(x, t)$ pour $t = 0$ (condition initiale en bleu) et pour $t > 0$ (en vert)

Il faut donc discuter selon les valeurs de ξ . Selon (62), cela mène à

$$\begin{cases} x(t) = \xi & \text{si } \xi \leq 0 \\ x(t) = \xi + t & \text{si } \xi > 0 \end{cases} \quad (63)$$

Les droites correspondant à (63) sont tracées sur la figure 4 en traits pleins.

On peut à présent vérifier qu'il n'y a que deux zones du plan (x, t) pour lesquelles par un point (x, t) donné, il passe une et unique caractéristique, dont nous allons pouvoir déterminer le pied, ce qui permettra de calculer $u(x, t)$. Sur la figure 4, on remarque qu'un point (x, t) peut se situer dans trois zones distinctes du plan:

1. Soit $x \leq 0$ (caractéristiques verticales)
2. Soit $0 < x \leq t$ (pas de caractéristiques)
3. Soit $t < x$ (caractéristiques toutes parallèles, de pente 1)

Dans le cas 1., l'unique caractéristique passant par (x, t) a pour équation $x = \xi$, donc $\xi = x$ et on a $u(x, t) = u_0(\xi) = u_0(x) = 0$.

Dans le cas 2., on peut vérifier que $(x, t) \mapsto u(x, t) = \frac{x}{t}$ est solution de l'équation. En effet, on a alors

$$\frac{\partial u}{\partial t}(x, t) = -\frac{x}{t^2} \text{ et } \frac{\partial u}{\partial x}(x, t) = \frac{1}{t}$$

ce qui assure que

$$\frac{\partial u}{\partial t}(x, t) + u(x, t) \frac{\partial u}{\partial x}(x, t) = -\frac{x}{t^2} + \frac{x}{t} \frac{1}{t} = 0.$$

Les droites d'équation $\frac{x}{t} = c$ (i.e. $x = ct$), avec $c \in [0, 1]$ sont tracées en traits pointillés sur la figure 4, et par extension on les appelle aussi droites

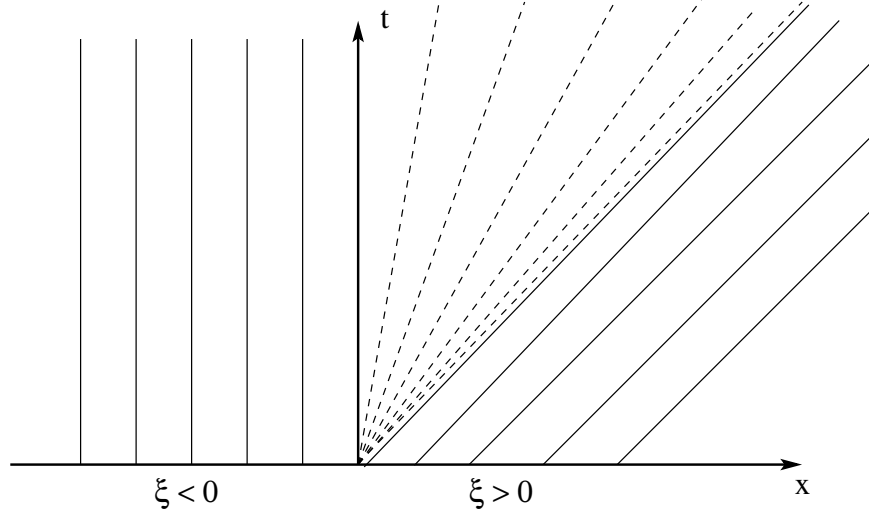


Figure 4: Droites caractéristiques dans le cas où u_0 est croissante et discontinue

caractéristiques, car u est effectivement constante sur ces courbes (et vaut $\frac{x}{t} = c$). Ceci est valable pour l'équation de Burgers. Dans le cas plus général où l'on connaît a et que cette fonction est continue et strictement croissante et donc inversible, d'inverse notée a^{-1} , on cherche $u(x, t)$ sous la forme $u(x, t) = v(\frac{x}{t})$. On a alors

$$\frac{\partial u}{\partial t}(x, t) + a(u(x, t)) \frac{\partial u}{\partial x}(x, t) = -\frac{x}{t^2} v' \left(\frac{x}{t} \right) + a \circ v \left(\frac{x}{t} \right) \frac{1}{t} v' \left(\frac{x}{t} \right).$$

On a donc, après factorisation

$$\frac{1}{t} v' \left(\frac{x}{t} \right) \left[-\frac{x}{t} + a \circ v \left(\frac{x}{t} \right) \right] = 0.$$

Ceci conduit à $v' = 0$ et donc u constante sur tout (x, t) , soit $a \circ v \left(\frac{x}{t} \right) = \frac{x}{t}$, et donc $u(x, t) = a^{-1}(\frac{x}{t})$ (on retrouve bien pour Burgers $u(x, t) = \frac{x}{t}$).

Dans le cas 3., l'unique caractéristique passant par (x, t) a pour équation $x = \xi + t$, donc $\xi = x - t$ et on a $u(x, t) = u_0(\xi) = 1$ (en effet $t < x$ implique $0 < x - t = \xi$).

Pour résumer, la solution est donnée pour tout temps $t > 0$ par

$$u(x, t) = \begin{cases} 0 & \text{si } x \leq 0 \\ \frac{x}{t} & \text{si } 0 < x \leq t \\ 1 & \text{si } x > t \end{cases} . \quad (64)$$

Cette solution est représentée en vert sur la figure 3. On remarque que cette solution est continue pour $t > 0$.

Cas de u_0 décroissante continue

Nous choisissons u_0 sous la forme

$$u_0(x) = \begin{cases} 1 & \text{si } x \leq 0 \\ 1 - x & \text{si } 0 \leq x \leq 1 \\ 0 & \text{si } 1 \leq x \end{cases} \quad (65)$$

qui est bien décroissante et continue. Cette condition initiale est représentée en bleu sur la figure 5.

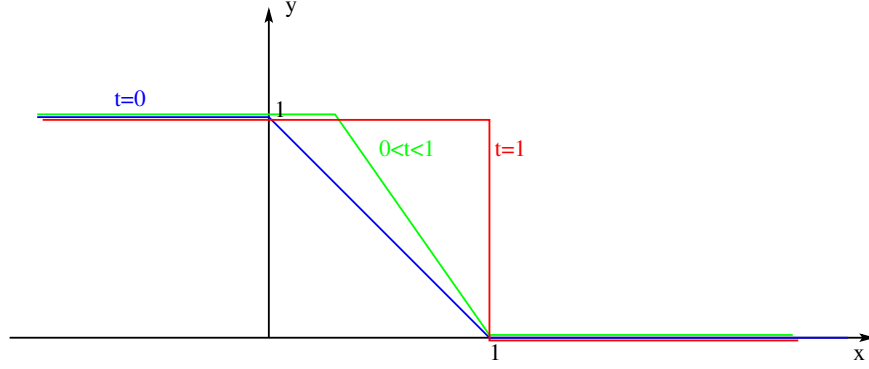


Figure 5: Cas où la donnée initiale est décroissante et continue: courbe $y = u(x, t)$ pour $t = 0$ (condition initiale en bleu), pour $1 > t > 0$ (en vert) et pour $t = 1$ (naissance d'un choc en rouge)

Calculons les équations des caractéristiques. D'après (56), on a

$$x(t) = \xi + u_0(\xi)t.$$

Il faut donc discuter selon les valeurs de ξ . Selon (65), cela mène à

$$\begin{cases} x(t) = \xi + t & \text{si } \xi \leq 0 \\ x(t) = \xi + (1 - \xi)t & \text{si } 0 \leq \xi \leq 1 \\ x(t) = \xi & \text{si } 1 \leq \xi \end{cases} \quad (66)$$

Les droites correspondant à (66) sont tracées sur la figure 6.

On peut à présent vérifier que pour $t < 1$, il passe une et unique caractéristique par un point (x, t) donné, dont nous allons pouvoir déterminer le pied, ce qui permettra de calculer $u(x, t)$. Sur la figure 6, on remarque que pour $t < 1$ un point (x, t) peut se situer dans trois zones distinctes du plan:

1. Soit $x \leq t$ (caractéristiques toutes parallèles, de pente 1)
2. Soit $t < x \leq 1$ (caractéristiques de pente croissante)
3. Soit $1 < x$ (caractéristiques verticales)

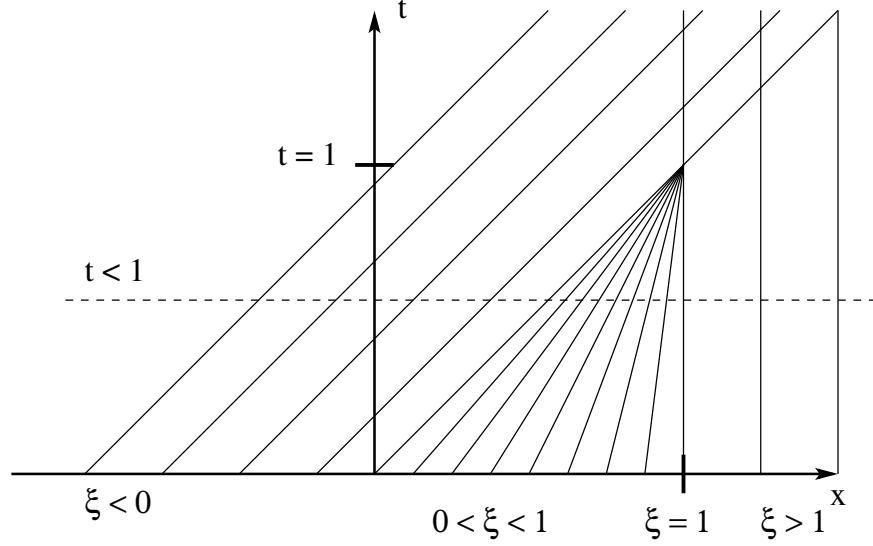


Figure 6: Droites caractéristiques dans le cas où u_0 est décroissante et continue

Dans le cas 1., l'unique caractéristique passant par (x, t) a pour équation $x = \xi + t$, donc $\xi = x - t$ et on a $u(x, t) = u_0(\xi) = u_0(x - t) = 1$ (car $x \leq t$ implique $\xi \leq 0$).

Dans le cas 2., l'unique caractéristique passant par (x, t) a pour équation $x = \xi + (1 - \xi)t$, donc $\xi = \frac{x-t}{1-t}$ et on a $u(x, t) = u_0(\xi) = u_0(\frac{x-t}{1-t}) = 1 - \frac{x-t}{1-t} = \frac{1-x}{1-t}$ (car $t \leq x \leq 1$ implique $0 \leq \xi \leq 1$).

Dans le cas 3., l'unique caractéristique passant par (x, t) a pour équation $x = \xi$, donc $\xi = x$ et on a $u(x, t) = u_0(\xi) = 0$ (en effet $1 < x$ implique $1 < \xi$).

Pour résumer, la solution est donnée pour tout temps $0 \leq t < 1$ par

$$u(x, t) = \begin{cases} 1 & \text{si } x \leq t \\ \frac{1-x}{1-t} & \text{si } t < x \leq 1 \\ 0 & \text{si } 1 < x \end{cases} . \quad (67)$$

Cette solution est représentée en vert sur la figure 5 pour $0 < t < 1$. On remarque que cette solution reste continue pour $0 < t < 1$ mais devient discontinue en $t = 1$ (solution en rouge sur la figure 5). L'intervalle $t < x \leq 1$ s'est réduit à un point et la valeur $\frac{1-x}{1-t}$ n'est plus définie. La solution en $t = 1$ est donc donnée par

$$u(x, t) = \begin{cases} 1 & \text{si } x \leq 1 \\ 0 & \text{si } 1 < x \end{cases} . \quad (68)$$

De manière plus générale, supposons que u_0 est strictement décroissante sur $[\xi_1, \xi_2]$ avec $\xi_1 < \xi_2$ et que a est strictement croissante. Les deux caractéristiques

issues de ξ_1 et de ξ_2 ont pour équations

$$\begin{aligned} x &= \xi_1 + a(u_0(\xi_1))t \\ x &= \xi_2 + a(u_0(\xi_2))t \end{aligned}$$

Ces deux droites se coupent lorsque

$$\xi_1 + a(u_0(\xi_1))t = \xi_2 + a(u_0(\xi_2))t$$

c'est-à-dire pour

$$t = \frac{(\xi_2 - \xi_1)}{[a(u_0(\xi_1)) - a(u_0(\xi_2))]} \quad (69)$$

qui est bien défini et strictement positif car $\xi_1 < \xi_2$ implique $a(u_0(\xi_1)) - a(u_0(\xi_2)) > 0$ puisque u_0 est strictement décroissante et a strictement croissante.

Au point d'intersection des deux droites, on aurait alors

$$u(x, t) = u_0(\xi_1) \quad \text{et} \quad u(x, t) = u_0(\xi_2),$$

ce qui est impossible, puisque $\xi_1 < \xi_2$ et u_0 est strictement décroissante. On a donc une discontinuité.

On remarque que l'instant d'intersection donné par (69) s'écrit aussi

$$t = -\frac{1}{\frac{a(u_0(\xi_2)) - a(u_0(\xi_1))}{\xi_2 - \xi_1}}$$

qui tend vers $-\frac{1}{(a \circ u_0)'(\xi)}$ lorsque $\xi_2 \rightarrow \xi_1 \rightarrow \xi$. On peut démontrer, et on admettra, que le temps maximal d'existence d'une solution continue lorsque u_0 est continue et strictement décroissante sur au moins un intervalle de \mathbb{R} est

$$t^* = \inf_{\xi \in \mathbb{R}} \left(-\frac{1}{(a \circ u_0)'(\xi)} \right),$$

qui est le temps que nous avons défini

2.2 Solutions faibles

Nous venons de voir que des discontinuités peuvent apparaître en un temps fini $t^* < +\infty$ dans la solution, même si la condition initiale est continue (même si elle est \mathcal{C}^∞). L'équation

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0 \quad (70)$$

ne peut plus être considérée au sens classique (la fonction u étant discontinue, elle n'est plus dérivable). Il faut alors considérer cette équation au sens des distributions, c'est-à-dire que l'on considère des solutions faibles. Par un procédé classique, choisissons $T > 0$ et considérons ϕ suffisamment régulière et à support

compact dans $\mathbb{R} \times [0, T[$ (ϕ est nulle en $t = T$ mais pas nécessairement nulle en $t = 0$, afin de garder la mémoire de la condition initiale sur u). Si u est une solution régulière de (70), alors en prenant le produit des deux membres de (70) par ϕ et en intégrant sur $\mathbb{R} \times [0, T]$ on obtient

$$\int \int_{\mathbb{R} \times [0, T]} \frac{\partial u}{\partial t} \phi + \int \int_{\mathbb{R} \times [0, T]} \frac{\partial f(u)}{\partial x} \phi = 0.$$

Dans la première intégrale, on effectue l'intégrale sur t , et dans la seconde l'intégrale sur x , qui se ramène à une intégrale sur $[-L, L]$ avec $L \geq 0$, tel que le support (en x) de ϕ soit inclus dans $[-L, L]$. On obtient

$$\int_{\mathbb{R}} [u \phi]_0^T dx - \int \int_{\mathbb{R} \times [0, T]} u \frac{\partial \phi}{\partial t} + \int_{[0, T]} [f(u) \phi]_{-L}^L - \int \int_{\mathbb{R} \times [0, T]} f(u) \frac{\partial \phi}{\partial x} = 0.$$

Soit, puisque $\phi(T) = 0$ et que $\phi(-L) = \phi(L) = 0$

$$\int \int_{\mathbb{R} \times [0, T]} \left[u \frac{\partial \phi}{\partial t} + f(u) \frac{\partial \phi}{\partial x} \right] dx dt = - \int_{\mathbb{R}} u_0(x) \phi(x, t=0) dx. \quad (71)$$

On note bien ici que si l'on avait choisi ϕ nulle en $t = 0$, la condition initiale u_0 aurait complètement disparu de la formulation.

Ainsi, toute solution régulière de (70) munie de la condition initiale u_0 vérifie (71). Inversement, toute solution régulière de (71) vérifie (70) et la condition initiale.

Toutefois, on se rend compte que (71) ne nécessite pas de condition de continuité ni de dérivabilité pour u ($u \in L_{\text{loc}}^\infty$ suffit) et que cette expression est bien définie pour $\phi \in \mathcal{C}^1$ à support compact dans $\mathbb{R} \times [0, T[$.

Par définition, on dira que u est une solution faible de (70) si u est dans L_{loc}^∞ et vérifie (71) pour tout $\phi \in \mathcal{C}^1$ à support compact dans $\mathbb{R} \times [0, T[$.

2.2.1 Relations de Rankine-Hugoniot

Considérons la figure 7. Nous supposons que u est solution faible de l'équation de transport non-linéaire, c'est-à-dire qu'elle vérifie (71) pour tout $\phi \in \mathcal{C}^1$ à support compact dans $\mathbb{R} \times [0, T[$ et qu'elle est régulière par morceaux: le plan est divisé en deux sous-parties ouvertes notées Σ^- et Σ^+ séparées par une courbe Σ de discontinuité de la solution faible, courbe d'équation $x = \sigma(t)$, et la solution u est au moins C^1 dans Σ^- et y est notée u^- et au moins C^1 dans Σ^+ et y est notée u^+ . Notons que le raisonnement se généralise à un nombre quelconque de courbes de discontinuité, car le raisonnement qui suit est local.

Alors, en considérant (71) pour une fonction ϕ à support compact dans la zone notée A , les intégrales se ramènent à des intégrales sur A dans laquelle u vaut u^- et est régulière; on peut donc calculer les intégrales par intégration par partie, et on obtient (les termes de bord étant nuls)

$$\int \int_A \left(\frac{\partial u^-}{\partial t} + \frac{\partial f(u^-)}{\partial x} \right) \phi = 0. \quad (72)$$

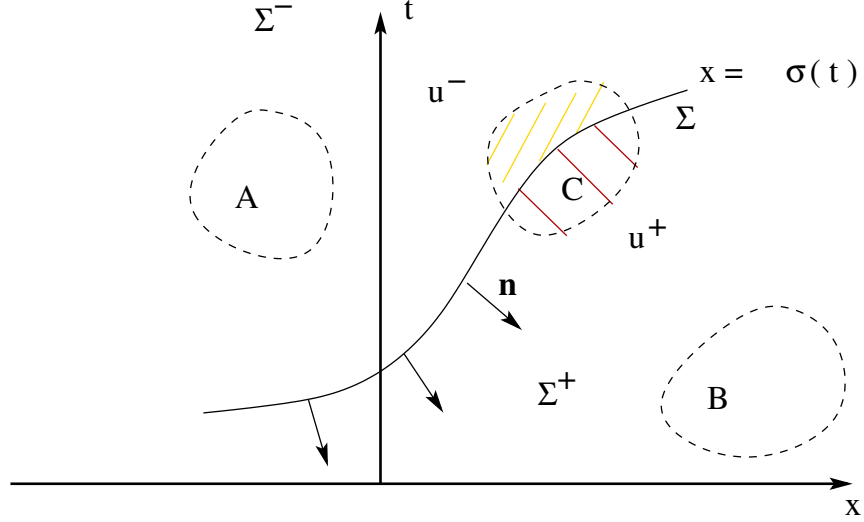


Figure 7: Courbe de discontinuité dans le plan (x, t)

Puisque (72) est valable pour tout ϕ régulier, on en déduit que

$$\frac{\partial u^-}{\partial t} + \frac{\partial f(u^-)}{\partial x} = 0 \quad (73)$$

sur A ; puisque A est arbitraire dans Σ^- , on en déduit que u^- est solution au sens classique de l'équation de transport: elle vérifie (73) sur Σ^- .

On raisonne de même pour u^+ en considérant des fonctions ϕ à support compact dans B , un ouvert quelconque de Σ^+ , et on déduit que u^+ est solution au sens classique de l'équation de transport sur Σ^+ :

$$\frac{\partial u^+}{\partial t} + \frac{\partial f(u^+)}{\partial x} = 0. \quad (74)$$

À présent, considérons une fonction ϕ à support dans la zone C qui intersecte la courbe $x = \sigma(t)$ et notons $\mathbf{n} = (n_x, n_t)$ le vecteur unitaire normal à la courbe de discontinuité orienté vers les x croissants

$$\mathbf{n} = \frac{1}{\sqrt{1 + \sigma'(t)^2}} (1, -\sigma'(t))^T.$$

Notons que cette expression pour \mathbf{n} est trouvée à partir des coordonnées d'un vecteur tangent à la courbe et dirigé vers les x croissant, qui a pour expression $(\sigma'(t), 1)$ (quand t augmente de dt , alors x augmente de $dx = \left(\frac{dx}{dt}\right) dt = \sigma'(t)dt$).

Appliquons alors la définition de la solution faible. On a

$$\begin{aligned} \int \int_{\mathbb{R} \times [0, T]} \left[u \frac{\partial \phi}{\partial t} + f(u) \frac{\partial \phi}{\partial x} \right] dx dt &= \\ \int \int_{C \cap \Sigma^-} \left[u \frac{\partial \phi}{\partial t} + f(u) \frac{\partial \phi}{\partial x} \right] dx dt &+ \int \int_{C \cap \Sigma^+} \left[u \frac{\partial \phi}{\partial t} + f(u) \frac{\partial \phi}{\partial x} \right] dx dt \end{aligned}$$

Et d'autre part

$$- \int_{\mathbb{R}} u_0(x) \phi(x, t=0) dx = 0.$$

Puisque u vaut u^- (respectivement u^+) dans Σ^- (resp. Σ^+), on obtient

$$\int \int_{C \cap \Sigma^-} \left[u^- \frac{\partial \phi}{\partial t} + f(u^-) \frac{\partial \phi}{\partial x} \right] dx dt + \int \int_{C \cap \Sigma^+} \left[u^+ \frac{\partial \phi}{\partial t} + f(u^+) \frac{\partial \phi}{\partial x} \right] dx dt = 0. \quad (75)$$

Calculons chacun de ces termes en utilisant la formule de Green-Ostrogradski et en se souvenant que ϕ est nulle sur ∂C

$$\begin{aligned} \int \int_{C \cap \Sigma^-} u^- \frac{\partial \phi}{\partial t} &= - \int \int_{C \cap \Sigma^-} \frac{\partial u^-}{\partial t} \phi + \int_{\partial(C \cap \Sigma^-)} u^- \phi n_t \\ &= - \int \int_{C \cap \Sigma^-} \frac{\partial u^-}{\partial t} \phi + \int_{\Sigma} u^- \phi n_t. \end{aligned} \quad (76)$$

De même:

$$\int \int_{C \cap \Sigma^-} f(u^-) \frac{\partial \phi}{\partial x} = - \int \int_{C \cap \Sigma^-} \frac{\partial f(u^-)}{\partial x} \phi + \int_{\Sigma} f(u^-) \phi n_x \quad (77)$$

$$\int \int_{C \cap \Sigma^+} u^+ \frac{\partial \phi}{\partial t} = - \int \int_{C \cap \Sigma^+} \frac{\partial u^+}{\partial t} \phi + \int_{\Sigma} u^+ \phi (-n_t) \quad (78)$$

$$\int \int_{C \cap \Sigma^+} f(u^+) \frac{\partial \phi}{\partial x} = - \int \int_{C \cap \Sigma^+} \frac{\partial f(u^+)}{\partial x} \phi + \int_{\Sigma} f(u^+) \phi (-n_x) \quad (79)$$

En rassemblant (75)–(79), on obtient

$$\begin{aligned} &\int \int_{C \cap \Sigma^-} \left[\frac{\partial u^-}{\partial t} + \frac{\partial f(u^-)}{\partial x} \right] \phi + \int \int_{C \cap \Sigma^+} \left[\frac{\partial u^+}{\partial t} + \frac{\partial f(u^+)}{\partial x} \right] \phi \\ &+ \int_{\Sigma} \{ (u^+ - u^-) n_t + [f(u^+) - f(u^-)] n_x \} \phi = 0. \end{aligned}$$

De plus, u^- et u^+ sont solutions classiques de l'équation de transport respectivement dans Σ^- et Σ^+ , cf (73) et (74), ce qui permet de conclure que

$$\int_{\Sigma} \{ (u^+ - u^-) n_t + [f(u^+) - f(u^-)] n_x \} \phi = 0.$$

Ceci étant vrai pour tout ϕ , on en déduit que sur Σ

$$(u^+ - u^-) n_t + [f(u^+) - f(u^-)] n_x = 0.$$

En tenant compte de l'expression du vecteur \mathbf{n} , ceci fournit la relation de Rankine-Hugoniot qui relie le saut de u , celui de $f(u)$ et la vitesse du choc:

$$\sigma'(t)[u^+ - u^-] = [f(u^+) - f(u^-)]. \quad (80)$$

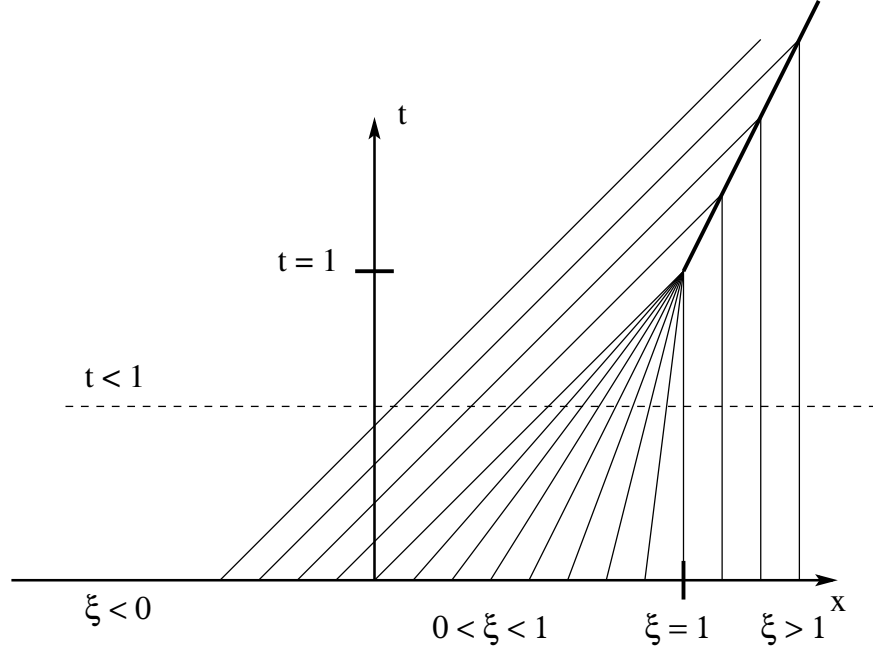


Figure 8: Caractéristiques et courbe de discontinuité ($x = \sigma(t) = \frac{1+t}{2}$ pour $t \geq 1$)

Dans le cas Burgers, on a

$$\sigma'(t)[u^+ - u^-] = \frac{1}{2}[(u^+)^2 - (u^-)^2]$$

soit

$$\sigma'(t) = \frac{1}{2}[(u^+) + (u^-)] \quad (81)$$

Le choc se déplace donc à la vitesse $\frac{1}{2}[(u^+) + (u^-)]$.

Si l'on reprend l'exemple présenté page 22, à la naissance du choc à $t = 1$, on a $u^- = 1$ et $u^+ = 0$. Le choc a pour vitesse $\frac{1}{2}$ et a donc pour équation $\sigma(t) = \frac{1+t}{2}$, pour $t \geq 1$. La figure 8 présente cette courbe de discontinuité et on peut s'apercevoir que les caractéristiques convergent vers la ligne de choc. La solution s'exprime alors, pour $t \geq 1$, par

$$u(x, t) = \begin{cases} 1 & \text{si } x \leq \frac{1+t}{2} \\ 0 & \text{si } \frac{1+t}{2} < x \end{cases} \quad (82)$$

De plus, la figure 9 représente la solution en des temps $t > 1$.

Pour résumer, nous retiendrons **la proposition suivante**: Une fonction $u(x, t)$, régulière de part et d'autre de toute courbe de discontinuité est solution faible de l'équation de transport si et seulement si, en plus de satisfaire à la condition initiale, on a les deux propriétés suivantes:

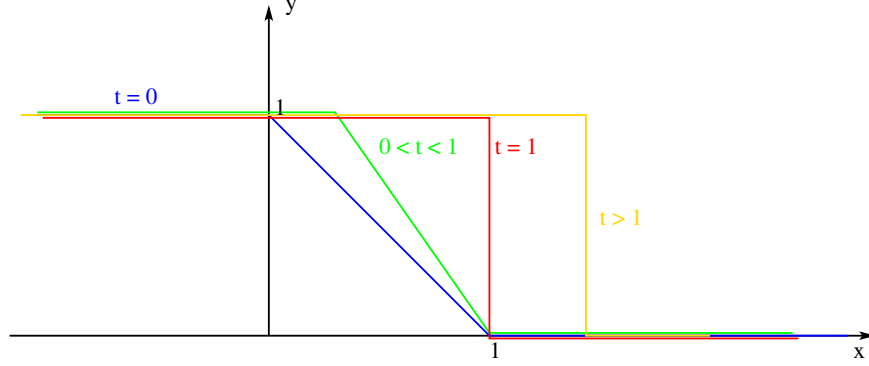


Figure 9: Évolution de la solution au cours du temps, en particulier après le choc (courbe jaune pour $t > 1$)

- u est solution de $\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0$ au sens classique partout où elle est régulière,
- au travers de toute courbe de discontinuité d'équation $x = \sigma(t)$, elle vérifie $\sigma'(t)[u^+ - u^-] = [f(u^+) - f(u^-)]$, avec la définition

$$u^\pm = \lim_{\varepsilon \rightarrow 0^+} u((x, t) \pm \varepsilon \mathbf{n}).$$

2.2.2 Non-unicité

Dans la section précédente, nous avons donc défini ce qu'est une solution faible de l'équation de transport non-linéaire. Malheureusement, nous allons constater qu'il n'y a pas unicité de la solution faible; il y en a même une infinité. Considérons par exemple l'équation de Burgers, munie de la donnée initiale suivante:

$$u_0 = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 & \text{si } 0 < x \end{cases}, \quad (83)$$

et fixons un réel $\gamma < 1$. Alors on peut vérifier que la fonction v_γ donnée par

$$v_\gamma(x, t) = \begin{cases} 0 & \text{si } x \leq \frac{\gamma}{2}t \\ \gamma & \text{si } \frac{\gamma}{2}t < x \leq \frac{t}{2} \\ 1 - \gamma & \text{si } \frac{t}{2} < x \leq (1 - \frac{\gamma}{2})t \\ 1 & \text{si } (1 - \frac{\gamma}{2})t < x \end{cases} \quad (84)$$

est solution faible de l'équation de Burgers. En effet, en $t = 0$ la fonction donnée par (84) vérifie bien (presque partout) $v_\gamma(x, t = 0) = u_0(x)$ et d'autre part:

- v_γ est solution au sens classique partout où elle est régulière, car elle est constante par morceaux, et donc $\frac{\partial u}{\partial t} = 0$ et $\frac{\partial u^2}{\partial x} = 0$,

- au travers de la discontinuité d'équation $x = \frac{\gamma}{2}t$, la vitesse du choc est $\frac{\gamma}{2}$ et on a $u^- = 0$ et $u^+ = \gamma$, et donc $\frac{u^- + u^+}{2} = \frac{\gamma}{2}$.
- au travers de la discontinuité d'équation $x = \frac{t}{2}$, la vitesse du choc est $\frac{1}{2}$ et on a $u^- = \gamma$ et $u^+ = 1 - \gamma$, et donc $\frac{u^- + u^+}{2} = \frac{1}{2}$,
- au travers de la discontinuité d'équation $x = (1 - \frac{\gamma}{2})t$, la vitesse du choc est $(1 - \frac{\gamma}{2})$ et on a $u^- = 1 - \gamma$ et $u^+ = 1$, et donc $\frac{u^- + u^+}{2} = (1 - \frac{\gamma}{2})$.

Il va donc falloir un critère pour trouver quelle est "la bonne" solution parmi toutes les solutions faibles. C'est ce que nous allons voir dans la section qui suit.

2.2.3 Entropie

Commençons par une définition: on appelle "couple entropie - flux d'entropie", ou plus simplement "entropie" pour l'équation

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0 \quad (85)$$

un couple de fonctions $(U, F) \mathcal{C}^1(\mathbb{R})$ telles que

- (i) U est strictement convexe,
- (ii) $F'(x) = U'(x)f'(x)$, $\forall x \in \mathbb{R}$.

La définition est plus complexe en dimension supérieure, mais en dimension un, on peut construire autant d'entropies que l'on dispose de fonctions U strictement convexes et appartenant à $\mathcal{C}^1(\mathbb{R})$, car pour U donnée, il suffit ensuite de choisir pour F une primitive quelconque de $U'f'$. Un exemple d'entropie pour l'équation de Burgers est $(U : x \mapsto \frac{x^2}{2}, F : x \mapsto \frac{x^3}{3})$.

Notons que la relation entre F et U permet d'écrire une loi de conservation supplémentaire pour u solution de (85), partout où u est régulière; en effet, dans ce cas, on a

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0 &\implies \frac{\partial u}{\partial t} + f'(u) \frac{\partial u}{\partial x} = 0 \\ &\implies U'(u) \frac{\partial u}{\partial t} + U'(u) f'(u) \frac{\partial u}{\partial x} = 0 \\ &\implies U'(u) \frac{\partial u}{\partial t} + F'(u) \frac{\partial u}{\partial x} = 0 \\ &\implies \frac{\partial U(u)}{\partial t} + \frac{\partial F(u)}{\partial x} = 0. \end{aligned} \quad (86)$$

En revanche, si u est solution faible discontinue, alors l'égalité (86) n'est pas vérifiée au sens des distributions; il faudrait pour cela (par un calcul analogue à celui fait précédemment lors de l'établissement de la relation de Rankine-Hugoniot pour les solutions faibles) que le long de la courbe de discontinuité d'équation $x = \sigma(t)$, on ait

$$\sigma'(t)[U(u^+) - U(u^-)] = [F(u^+) - F(u^-)], \quad (87)$$

ce qui est en général impossible à satisfaire simultanément avec la relation de Rankine-Hugoniot

$$\sigma'(t)[u^+ - u^-] = [f(u^+) - f(u^-)]. \quad (88)$$

Par exemple, pour l'équation de Burgers et l'entropie ($U : x \mapsto \frac{x^2}{2}, F : x \mapsto \frac{x^3}{3}$), l'égalité (87), fournit

$$\frac{1}{2}\sigma'(t)((u^+)^2 - (u^-)^2) = \frac{1}{3}((u^+)^3 - (u^-)^3)$$

soit, après simplification par $(u^+ - u^-)$:

$$\frac{1}{2}\sigma'(t)(u^+ + u^-) = \frac{1}{3}((u^+)^2 + (u^+)(u^-) + (u^-)^2). \quad (89)$$

Or, la relation (88) fournit pour l'équation de Burgers

$$\sigma'(t) = \frac{1}{2}(u^+ + u^-) \quad (90)$$

Les relations (89) et (90) ne sont compatibles que si

$$\frac{1}{4}(u^+ + u^-)^2 - \frac{1}{3}((u^+)^2 + (u^+)(u^-) + (u^-)^2) = 0,$$

soit, après développement

$$\frac{1}{12} [3((u^+)^2 + 2(u^+)(u^-) + (u^-)^2) - 4((u^+)^2 + (u^+)(u^-) + (u^-)^2)] = 0,$$

soit

$$\frac{1}{12} [-(u^+)^2 + 2(u^+)(u^-) - (u^-)^2] = 0,$$

c'est-à-dire seulement si $u^+ = u^-$, c'est-à-dire seulement s'il n'y a pas de discontinuité!

Ne pouvant pas assurer que (86) soit vérifiée au sens des distributions, on pose la définition suivante:

On appelle **solution faible entropique** de (85) toute solution faible de (85) vérifiant de plus

$$\frac{\partial U(u)}{\partial t} + \frac{\partial F(u)}{\partial x} \leq 0 \quad (91)$$

pour toute entropie (U, F) associée à l'équation (85). La relation (91) est à prendre au sens des distributions: nous rappelons qu'une distribution S est dite négative si pour toute fonction test positive ϕ on a $\langle S, \phi \rangle \leq 0$. Donc (91) signifie que pour toute fonction test ϕ positive on a

$$\langle U(u), \frac{\partial \phi}{\partial t} \rangle + \langle F(u), \frac{\partial \phi}{\partial x} \rangle \geq 0$$

Puisque nous avons vu que partout où u est régulière alors (86) est vérifiée, la relation (91) ne fournit en fait qu'une relation supplémentaire au travers des

discontinuités: de nouveau par un calcul analogue au précédent, on trouve en effet que

$$\int_{\Sigma} \{ (U(u^+) - U(u^-))n_t + [F(u^+) - F(u^-)]n_x \} \phi \leq 0$$

pour toute fonction test ϕ positive. En rappelant que le vecteur \mathbf{n} est proportionnel à $(1, -\sigma'(t))^T$ (avec une constante de proportionnalité positive), on en déduit que sur Σ

$$-\sigma'(t)[U(u^+) - U(u^-)] + [F(u^+) - F(u^-)] \leq 0$$

soit:

$$[F(u^+) - F(u^-)] \leq \sigma'(t)[U(u^+) - U(u^-)] \quad (92)$$

le long de la courbe de discontinuité Σ , pour toute entropie (U, F) associée à (85). On peut démontrer, grâce à la convexité de U et de f , que ceci implique que, le long de toute discontinuité de u , on a:

$$u^- > u^+. \quad (93)$$

Nous pouvons vérifier ceci pour l'équation de Burgers avec l'entropie $(U : x \mapsto \frac{x^2}{2}, F : x \mapsto \frac{x^3}{3})$. En effet, (92) fournit

$$\frac{1}{3}[(u^+)^3 - (u^-)^3] - \left(\frac{u^+ + u^-}{2} \right) \frac{1}{2}[(u^+)^2 - (u^-)^2] \leq 0,$$

soit, après factorisation

$$\frac{1}{12}(u^+ - u^-)\{4[(u^+)^2 + u^+u^- + (u^-)^2] - 3(u^+ + u^-)^2\} \leq 0$$

et donc

$$\frac{1}{12}(u^+ - u^-)[(u^+)^2 - 2u^+u^- + (u^-)^2] \leq 0$$

et donc nécessairement, lorsque $u^+ \neq u^-$ (sinon il n'y a pas de discontinuité):

$$u^+ - u^- < 0,$$

ce qui est bien (93).

Remarque: Justification de la condition (91). L'équation de transport (85) peut-être vue comme une approximation d'une équation plus générale comportant une faible viscosité ε :

$$\frac{\partial u_\varepsilon}{\partial t} + \frac{\partial f(u_\varepsilon)}{\partial x} - \varepsilon \frac{\partial^2 u_\varepsilon}{\partial x^2} = 0 \quad (94)$$

Si l'on admet que le terme de diffusion régularise la solution (c'est-à-dire que l'on admet que u_ε est au moins \mathcal{C}^2), alors on peut utiliser la forme non-conservative de (94) et en multipliant par $U'(u_\varepsilon)$, on obtient

$$U'(u_\varepsilon) \frac{\partial u_\varepsilon}{\partial t} + U'(u_\varepsilon) f'(u_\varepsilon) \frac{\partial u_\varepsilon}{\partial x} - U'(u_\varepsilon) \varepsilon \frac{\partial^2 u_\varepsilon}{\partial x^2} = 0. \quad (95)$$

En tenant compte des égalités

$$\frac{\partial^2 U(u_\varepsilon)}{\partial x^2} = \frac{\partial}{\partial x} \left[\frac{\partial u_\varepsilon}{\partial x} U'(u_\varepsilon) \right] = \frac{\partial^2 u_\varepsilon}{\partial x^2} U'(u_\varepsilon) + \frac{\partial u_\varepsilon}{\partial x} \frac{\partial u_\varepsilon}{\partial x} U''(u_\varepsilon)$$

on a

$$-U'(u_\varepsilon) \frac{\partial^2 u_\varepsilon}{\partial x^2} = \left(\frac{\partial u_\varepsilon}{\partial x} \right)^2 U''(u_\varepsilon) - \frac{\partial^2 U(u_\varepsilon)}{\partial x^2}. \quad (96)$$

D'autre part nous rappelons que

$$U'(u_\varepsilon) f'(u_\varepsilon) = F'(u_\varepsilon);$$

alors (96) et (95) entraînent

$$\frac{\partial U(u_\varepsilon)}{\partial t} + \frac{\partial F(u_\varepsilon)}{\partial x} - \varepsilon \frac{\partial^2 U(u_\varepsilon)}{\partial x^2} + \varepsilon \left(\frac{\partial u_\varepsilon}{\partial x} \right)^2 U''(u_\varepsilon) = 0. \quad (97)$$

Si on admet que $u_\varepsilon \rightarrow u$ au sens des distributions avec u solution de (85), alors

$$\begin{aligned} \left\langle \frac{\partial U(u_\varepsilon)}{\partial t}, \phi \right\rangle &= - \left\langle U(u_\varepsilon), \frac{\partial \phi}{\partial t} \right\rangle = - \int \int_{\mathbb{R} \times \mathbb{R}^+} U(u_\varepsilon) \frac{\partial \phi}{\partial t} \\ &\rightarrow - \int \int_{\mathbb{R} \times \mathbb{R}^+} U(u) \frac{\partial \phi}{\partial t} \\ &\rightarrow \left\langle \frac{\partial U(u)}{\partial t}, \phi \right\rangle \end{aligned} \quad (98)$$

par le théorème de convergence dominée de Lebesgue.

De même,

$$\frac{\partial F(u_\varepsilon)}{\partial x} \rightarrow \frac{\partial F(u)}{\partial x} \quad \text{et} \quad \varepsilon \frac{\partial^2 U(u_\varepsilon)}{\partial x^2} \rightarrow 0. \quad (99)$$

Comme par ailleurs U est convexe, sa dérivée seconde est positive et donc

$$\left\langle \varepsilon \left(\frac{\partial u_\varepsilon}{\partial x} \right)^2 U''(u_\varepsilon), \phi \right\rangle \geq 0 \quad (100)$$

pour toute fonction test $\phi \geq 0$.

On obtient donc (91) en rassemblant (97) (98), (99) et (100).

Par ailleurs, on admet **le résultat d'existence et d'unicité de la solution faible entropique**, en dimension d'espace quelconque pour $p = 1$ (équation scalaire):

Soit $u_0 \in L^\infty(\mathbb{R})$, alors l'équation (85) admet une unique solution entropique appartenant à $L^\infty(\mathbb{R} \times [0, +\infty[)$. De plus,

$$\|u(\cdot, t)\|_{L^\infty(\mathbb{R})} \leq \|u_0\|_{L^\infty(\mathbb{R})}, \quad \text{p.p. } t \geq 0.$$

De plus, si u et v sont deux solutions associées aux conditions initiales u_0 et v_0 respectivement, alors

$$u_0(x) \geq v_0(x) \quad \forall x \in \mathbb{R} \implies u(x, t) \geq v(x, t) \quad \forall x \in \mathbb{R}, \quad \text{p.p. } t \geq 0.$$

et, pour presque tout $t \geq 0$ et tout (x_1, x_2) avec $x_1 < x_2$

$$\int_{x_1}^{x_2} |u(x, t) - v(x, t)| dx \leq \int_{x_1 - At}^{x_2 + At} |u_0(x) - v_0(x)| dx$$

avec

$$A = \max \{ |f'(\xi)|; |\xi| \leq \max(\|u_0\|_{L^\infty(\mathbb{R})}, \|v_0\|_{L^\infty(\mathbb{R})}) \}.$$

Cette dernière propriété exprime en particulier la propagation à vitesse finie de l'information.

Pour résumer: lorsque f est strictement convexe, la solution faible entropique est caractérisée par

1. u est solution classique de l'équation de transport dans tout domaine où elle est \mathcal{C}^1
2. Le long de toute courbe de discontinuité éventuelle d'équation $x = \sigma(t)$, on a

$$\sigma'(t)(u^+ - u^-) = f(u^+) - f(u^-)$$

3. Le long de toute courbe de discontinuité éventuelle, on a $u^- > u^+$.

En particulier, si on trouve une fonction **continue** et \mathcal{C}^1 par morceaux qui est solution classique dans toute partie du plan où elle est \mathcal{C}^1 , alors c'est la solution entropique!

De plus, dans le plan (x, t) si l'on trace les caractéristiques et une éventuelle ligne de discontinuité, une façon de reconnaître graphiquement la solution entropique est de vérifier si les caractéristiques convergent vers la ligne de discontinuité. En effet, lorsque l'on a $u^- > u^+$ et f strictement convexe, alors pour $u \in]u^+; u^-[$ la fonction $u \mapsto \frac{f(u) - f(u^+)}{u - u^+}$ est strictement croissante. On en déduit

$$\sigma'(t) = \frac{f(u^-) - f(u^+)}{u^- - u^+} > \frac{f(u) - f(u^+)}{u - u^+} \geq \lim_{u \rightarrow u^+} \frac{f(u) - f(u^+)}{u - u^+} = a(u^+)$$

De même, du caractère strictement croissant de la fonction $u \mapsto \frac{f(u^-) - f(u)}{u^- - u}$ on déduit

$$\sigma'(t) = \frac{f(u^-) - f(u^+)}{u^- - u^+} < \frac{f(u^-) - f(u)}{u^- - u} \leq \lim_{u \rightarrow u^-} \frac{f(u^-) - f(u)}{u^- - u} = a(u^-)$$

Donc

$$a(u^+) < \sigma'(t) < a(u^-) \tag{101}$$

Or les caractéristiques ont pour équation

$$x = \xi + a(u_0(\xi))t$$

et u est constant le long de ces droites. Donc à gauche du choc les pentes des droites sont égales à $1/a(u^-)$ et à droite du choc les pentes sont égales à $1/a(u^+)$. Un simple dessin montre alors que la relation (101) indique que les caractéristiques convergent vers la ligne de choc, quelque soit le signe des diverses quantités dans (101).

3 Schémas volumes finis pour l'hyperbolique 1D

Le but de cette section est de présenter des méthodes d'approximation de la solution faible entropique de (45)–(46). On découpe pour cela l'espace et le temps en sous-intervalle sur lesquels nous écrivons un bilan issu de la loi de conservation (45); ce bilan fera apparaître des flux aux interfaces des intervalles en espace, qui devront être approchés.

3.1 Principe de base

On commence par discrétiser en espace. Nous choisissons des points $(x_{i+1/2})_{i \in \mathbb{Z}}$ et considérons les segments $T_i = [x_{i-1/2}; x_{i+1/2}]$; nous notons $|T_i| = x_{i+1/2} - x_{i-1/2}$ la longueur de T_i . Le pas du maillage est noté $h := \sup_{i \in \mathbb{Z}} |T_i|$.

Dans la direction du temps, nous découpons l'intervalle de calcul $[0, T]$ en N pas de temps $[t_n, t_{n+1}]$, pour $n \in [0, N-1]$. Pour simplifier, nous supposons que ces pas de temps sont tous de même durée $\Delta t = \frac{T}{N}$, mais ce qui suit peut être aisément généralisé.

Notons \hat{u} la solution faible entropique de (45)–(46). On effectue alors un bilan en intégrant (45) sur $T_i \times [t_n, t_{n+1}]$:

$$\int_{T_i \times [t_n, t_{n+1}]} \frac{\partial \hat{u}}{\partial t} + \frac{\partial f(\hat{u})}{\partial x} = 0.$$

En séparant la double intégrale en deux parties et en effectuant l'intégration de $\frac{\partial \hat{u}}{\partial t}$ sur $[t_n, t_{n+1}]$ et de $\frac{\partial f(\hat{u})}{\partial x}$ sur T_i , on obtient

$$\begin{aligned} & \int_{T_i} \hat{u}(x, t_{n+1}) dx - \int_{T_i} \hat{u}(x, t_n) dx \\ & + \int_{[t_n, t_{n+1}]} f(\hat{u}(x_{i+1/2}, t)) dt - \int_{[t_n, t_{n+1}]} f(\hat{u}(x_{i-1/2}, t)) dt = 0. \end{aligned} \quad (102)$$

Considérons la valeur moyenne de \hat{u} sur T_i au temps t_n :

$$\langle \hat{u} \rangle_i^n := \frac{1}{|T_i|} \int_{T_i} \hat{u}(x, t_n) dx$$

et la valeur moyenne du flux au point $x_{i+1/2}$ entre les instants t_n et t_{n+1} :

$$G_{i+1/2}^{n+1/2} = \frac{1}{\Delta t} \int_{[t_n, t_{n+1}]} f(\hat{u}(x_{i+1/2}, t)) dt.$$

De (102), on tire alors la relation exacte suivante:

$$\langle \hat{u} \rangle_i^{n+1} = \langle \hat{u} \rangle_i^n - \frac{\Delta t}{|T_i|} \left(G_{i+1/2}^{n+1/2} - G_{i-1/2}^{n+1/2} \right).$$

On construit alors un schéma numérique en s'inspirant de cette égalité et en considérant les inconnues $(u_i^n)_{i \in \mathbb{Z}, n \in [0, N]}$ destinées à être des approximations

des valeurs exactes $(\langle \hat{u} \rangle_i^n)_{i \in \mathbb{Z}, n \in [0, N]}$. La grande question à traiter est de savoir comment approcher les flux exacts $(G_{i+1/2}^{n+1/2})_{i \in \mathbb{Z}, n \in [0, N]}$ en fonctions des inconnues $(u_j^m)_{j \in \mathbb{Z}, m \in [0, N]}$, sachant que celles-ci sont des approximations des $\langle \hat{u} \rangle_j^m$.

On va écrire

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{|T_i|} \left(g_{i+1/2}^{n+1/2} - g_{i-1/2}^{n+1/2} \right),$$

où $g_{i+1/2}^{n+1/2}$ est appelé "flux numérique" à l'interface entre T_i et T_{i+1} sur le pas de temps $[t_n, t_{n+1}]$. Nous étudierons en particulier un grande classe de schémas, celle des schémas explicites à trois points pour lesquelles $g_{i+1/2}^{n+1/2}$ est de la forme

$$g_{i+1/2}^{n+1/2} = g(u_i^n, u_{i+1}^n).$$

On appelle alors la fonction g "fonction de flux" qui est un flux explicite "à deux points", puisqu'elle ne dépend que des deux valeurs u_i^n et u_{i+1}^n à l'interface des deux cellules. Nous allons voir dans les sections qui suivent des propriétés que g doit vérifier pour que le schéma fournisse une approximation satisfaisante de la solution exacte.