

# Equations Différentielles Ordinaires Equations aux Dérivées Partielles

Franck Boyer

M1 Enseignement Supérieur et Recherche

Université Paul Sabatier - Toulouse 3

26 février 2021

Ce document est mis à disposition selon les termes de la licence [Creative Commons](#)  
“Attribution - Pas d’utilisation commerciale - Partage dans les mêmes conditions 4.0  
International”



Ces notes sont en construction permanente. Ne pas hésiter à signaler des erreurs ou imprécisions à l'adresse  
[franck.boyer@math.univ-toulouse.fr](mailto:franck.boyer@math.univ-toulouse.fr)



# Table des matières

<b>I</b>	<b>Equations différentielles ordinaires</b>	<b>1</b>
I	Introduction . . . . .	1
II	Préliminaires . . . . .	2
II.1	Les fonctions lipschitziennes . . . . .	2
II.2	Equations et inéquations différentielles linéaires scalaires. Lemmes de Gronwall . . . . .	10
II.3	Notions de solution d'un problème de Cauchy . . . . .	13
III	Théorème de Cauchy-Lipschitz global et applications . . . . .	13
III.1	Enoncé et preuve du théorème principal . . . . .	14
III.2	Equations différentielles linéaires . . . . .	16
III.3	Flot d'un champ de vecteurs . . . . .	18
IV	Théorème de Cauchy-Lipschitz local . . . . .	24
IV.1	Enoncé et preuve du théorème principal . . . . .	24
IV.2	Critères de globalité . . . . .	26
V	Equilibres. Stabilité. Stabilité asymptotique . . . . .	28
V.1	Cas linéaire . . . . .	29
V.2	Cas nonlinéaire . . . . .	34
VI	Etude détaillée d'un exemple : un modèle de propagation d'épidémie . . . . .	45
VI.1	Existence et unicité d'une solution globale en temps positif . . . . .	45
VI.2	Etats d'équilibre . . . . .	46
VI.3	Quelques exemples de trajectoires . . . . .	48
<b>II</b>	<b>Equations de transport</b>	<b>53</b>
I	Modèles de transport en 1D . . . . .	53
I.1	Trafic routier . . . . .	53
I.2	Dynamique des gaz simplifiée . . . . .	56
II	Modèles de transport en dimension quelconque . . . . .	58
II.1	Théorème de Liouville . . . . .	58
II.2	Théorème de Reynolds . . . . .	59
II.3	Exemple d'application à l'établissement d'une loi de conservation . . . . .	60
III	Solutions classiques des équations de transport . . . . .	61
III.1	Cas général de l'équation de convection . . . . .	61
III.2	Cas particuliers importants . . . . .	63
III.3	Autres applications de la méthode des caractéristiques . . . . .	64
IV	Solutions faibles de l'équation de transport . . . . .	67
IV.1	Définition des solutions faibles . . . . .	67
IV.2	Validité de la formule de représentation par les caractéristiques . . . . .	70
IV.3	Unicité . . . . .	72
IV.4	Le problème des conditions aux limites / conditions au bord . . . . .	72
IV.5	Un exemple de modèle en dynamique des populations . . . . .	75
<b>III</b>	<b>Formulations variationnelles de problèmes aux limites elliptiques</b>	<b>81</b>
I	Le problème de la corde/membrane élastique à l'équilibre . . . . .	81
I.1	Présentation . . . . .	81
I.2	Les questions mathématiques que l'on veut résoudre . . . . .	83
I.3	Comment montrer l'existence d'un minimiseur ? . . . . .	87
II	Espaces de Sobolev en dimension 1 . . . . .	88
II.1	L'espace $H^1(\cdot, \cdot)$ . . . . .	88
II.2	L'espace $H_0^1(I)$ . . . . .	92
II.3	Résolution du problème variationnel pour la corde élastique . . . . .	93

III	Formulation variationnelle d'un problème aux limites linéaires. Théorème de Lax-Milgram. . . . .	94
III.1	Principe général . . . . .	95
III.2	Exemples en dimension 1 . . . . .	96
III.3	Preuve du théorème de Lax-Milgram . . . . .	103
IV	Espaces de Sobolev et problèmes elliptiques sur un domaine de $\mathbb{R}^d$ . . . . .	104
IV.1	Espaces de Sobolev sur un domaine de $\mathbb{R}^d$ . . . . .	104
IV.2	Problèmes aux limites elliptiques . . . . .	108
<b>A</b>	<b>Eléments de la théorie des distributions</b> . . . . .	<b>113</b>
I	Intégration par parties en dimension $d$ : le cas des fonctions à support compact . . . . .	113
II	Un lemme important de la théorie de l'intégration . . . . .	114
III	Espace des fonctions test. Espace des distributions. . . . .	117
III.1	Définitions, exemples . . . . .	117
III.2	Convergence au sens des distributions . . . . .	121
IV	Dérivation au sens des distributions. . . . .	122
<b>B</b>	<b>La formule de Stokes</b> . . . . .	<b>125</b>
I	Hypersurfaces de $\mathbb{R}^d$ . Intégrale de surface . . . . .	125
I.1	Courbes planes . . . . .	125
I.2	Intégrales sur des hypersurfaces de $\mathbb{R}^d$ . . . . .	127
II	Domaines réguliers de $\mathbb{R}^d$ . . . . .	129
III	Formule de Stokes . . . . .	129
III.1	Le cas du demi-espace $\mathbb{R}_+^d$ . . . . .	129
III.2	Le cas du demi-espace à frontière non plane . . . . .	130

# Chapitre I

## Equations différentielles ordinaires

### I Introduction

Le but de ce chapitre est l'étude des équations différentielles ordinaires (du premier ordre) de la forme

$$x' = f(t, x),$$

où  $x$  est une **fonction inconnue** de la variable réelle  $t$  et l'application  $f$  est une donnée du problème. A chaque instant  $t$ , la valeur de  $x(t)$  est appelée l'état du système à l'instant  $t$  et l'espace dans lequel la fonction inconnue  $x$  prend ses valeurs est appelé *espace d'états*. Comme la mécanique est principalement à l'origine de l'étude de ces équations, on parlera souvent de **variable de temps** pour la variable  $t$  et de **trajectoire du système** pour une solution  $t \mapsto x(t)$ .

Dans toute la suite, l'espace d'états considéré sera un (sous-ensemble d'un) espace vectoriel de dimension finie. De très nombreuses questions se posent sur l'équation ci-dessus :

- Si on se donne un état initial du système, c'est-à-dire un instant initial  $t_0$  et l'état  $x_0$  du système à cet instant  $t_0$ , existe-t'il une ou plusieurs solution(s) de l'équation qui satisfont à la condition initiale  $x(t_0) = x_0$  ?
- Si une telle solution existe, peut-on la calculer explicitement ?
- Si le calcul explicite est impossible, peut-on au moins les décrire qualitativement ?
- Si  $f$  est une fonction périodique du temps, existe-t'il des solutions périodiques en temps de l'équation ?
- Si  $x$  et  $y$  sont deux solutions de l'équation, associées à des données initiales  $x_0$  et  $y_0$  proches (en un sens à préciser), est-ce que les solutions  $x$  et  $y$  restent proches au cours du temps ?
- Etc ...

Nous ne répondrons que très partiellement à un certain nombre de ces questions dans la suite. Il se trouve que la réponse à beaucoup de ces questions repose d'une façon ou d'une autre sur la première d'entre elles. On appelle cela la théorie de Cauchy : elle consiste essentiellement à donner des hypothèses sur les données du problème (principalement sur l'application  $f$ ) pour assurer l'existence et l'unicité des solutions **du problème de Cauchy** suivant

$$\begin{cases} x'(t) = f(t, x(t)), \\ x(t_0) = x_0, \end{cases} \quad (\text{I.1})$$

pour des données initiales  $(t_0, x_0)$  fixées.

On peut rencontrer différents types de représentations graphiques de solutions d'équations différentielles et il est bon de comprendre tout de suite les différences.

- *Equations scalaires.*

Exemple de l'équation logistique  $x' = x(1 - x)$ .

On peut tracer plusieurs solutions sous la forme d'une fonction du temps sur le même graphe (Figure I.1).

- *Equations non scalaires.*

Exemple d'un modèle de Lotka-Volterra  $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} x_1(0.4 - x_2) \\ -x_2(1 - 3x_1) \end{pmatrix}$

On peut tracer une solution sous la forme de deux fonctions du temps sur le même graphe (Figure I.2) ou sous la forme d'une courbe paramétrée  $t \mapsto (x_1(t), x_2(t)) \in \mathbb{R}^2$  dans le plan  $\mathbb{R}^2$  (Figure I.3).

On rencontre aussi fréquemment le tracé d'un champ de vecteurs sous la forme d'un ensemble de flèches en un certain nombre de points du domaine (Figure I.4). Il en existe deux versions : à gauche la version usuelle (chaque flèche a une longueur proportionnelle à la norme euclidienne de  $f(t, x)$ ), à droite une version normalisée (chaque flèche est alors de longueur fixe). Dans ce deuxième cas on perd donc l'information sur le module de la vitesse d'évolution de la solution en tout point en ne conservant que sa direction.

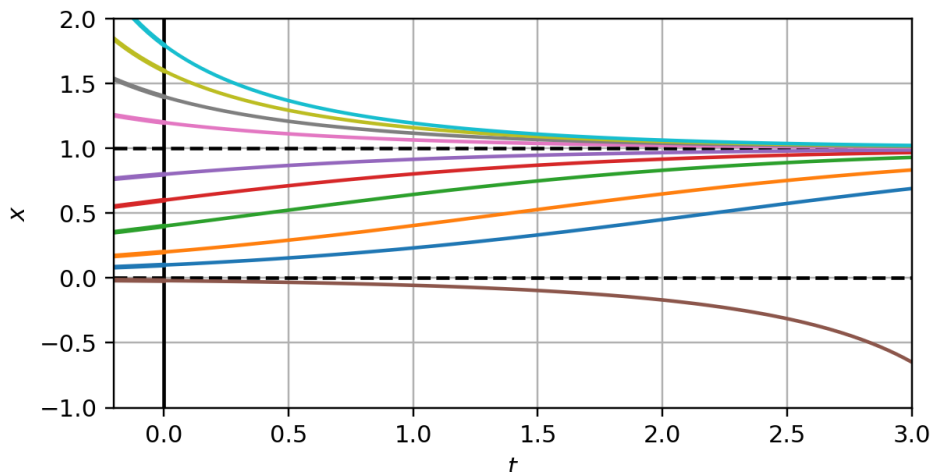


FIGURE I.1 – Plusieurs solutions de l'équation logistique

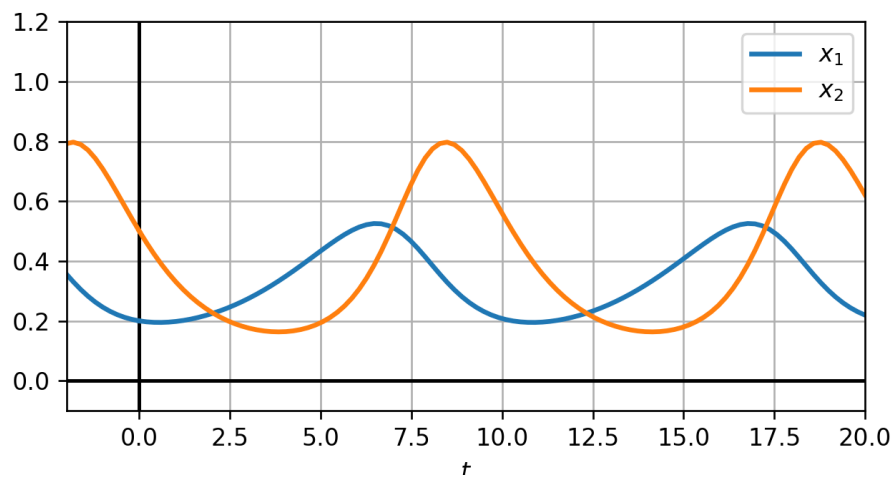


FIGURE I.2 – Une seule solution du système de Lotka-Volterra (2 composantes)

## II Préliminaires

### II.1 Les fonctions lipschitziennes

#### II.1.a Définitions et premiers exemples

On munit tous les espaces  $\mathbb{R}^d$  de la norme euclidienne usuelle.

##### Définition I.1

Soit  $A \subset \mathbb{R}^m$  et  $f : A \mapsto \mathbb{R}^n$ . On dit que  $f$  est (globalement) lipschitzienne s'il existe un nombre  $L > 0$  tel que

$$\forall x, y \in A, \quad \|f(x) - f(y)\| \leq L\|x - y\|.$$

Si on veut préciser la valeur de la constante, on dit que  $f$  est  $L$ -lipschitzienne. La plus petite valeur de  $L$  qui vérifie cette propriété est appelée la constante de Lipschitz de  $f$  et noté  $\text{Lip}(f)$ .

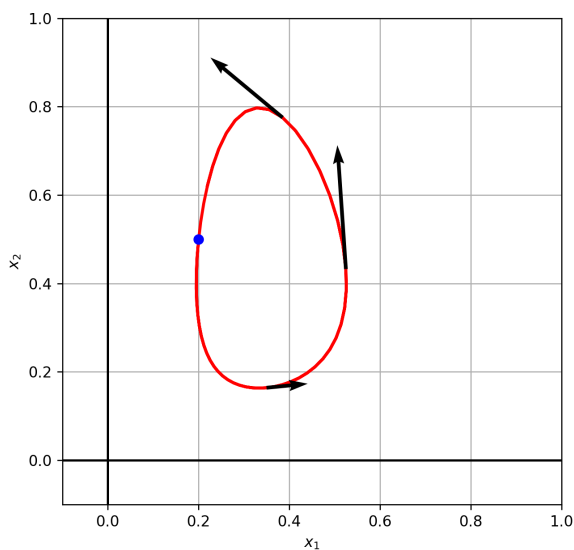


FIGURE I.3 – Trajectoire dans l'espace d'états

**Remarque I.2**

*Dans la définition ci-dessus les deux normes qui apparaissent sont différentes (la première est sur  $\mathbb{R}^m$  et la seconde est sur  $\mathbb{R}^n$ ).*

*Si on prend d'autres normes sur ces espaces, le caractère lipschitzien de la fonction  $f$  ne change pas mais la valeur de la constante de Lipschitz peut changer.*

**Proposition I.3**

*Toute fonction lipschitzienne est uniformément continue (et donc continue).*

Un exemple assez usuel de fonctions lipschitziennes est donné par la proposition suivante.

**Proposition I.4**

*Soit  $\Omega \subset \mathbb{R}^m$  un **ouvert convexe** de  $\mathbb{R}^m$  et  $f : \Omega \rightarrow \mathbb{R}^n$  une fonction de classe  $\mathcal{C}^1$ .*

*La fonction  $f$  est lipschitzienne si et seulement si sa différentielle est bornée sur  $\Omega$  et on a alors*

$$\text{Lip}(f) = \sup_{x \in \Omega} \|df(x)\|_{L(\mathbb{R}^m, \mathbb{R}^n)}. \quad (\text{I.2})$$

On peut généraliser ce résultat sur des ouverts un peu plus généraux que les ouverts convexes, mais cela demande des

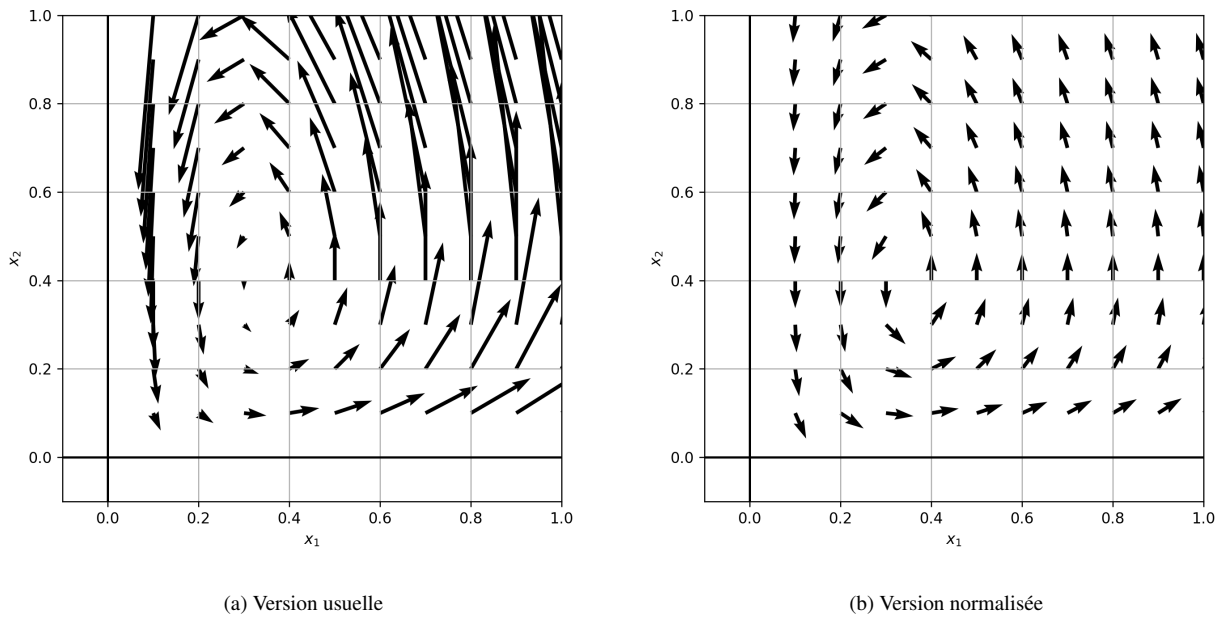


FIGURE I.4 – Tracé du champ de vecteurs de Lotka-Volterra

précautions et l'égalité (I.2) n'est plus forcément vraie.

### Exemple I.5

*Quelques exemples typiques de fonctions lipschitziennes qui ne soient pas de classe  $\mathcal{C}^1$  :*

- La valeur absolue :  $x \in \mathbb{R} \mapsto |x|$ , ou plus généralement la norme  $x \in \mathbb{R}^m \mapsto \|x\| \in \mathbb{R}$ .
- L'application distance à un ensemble  $B$  :

$$x \mapsto d(x, B) = \inf_{y \in B} \|x - y\|.$$

- La projection sur un convexe fermé  $K \subset \mathbb{R}^n$ ,

$$x \in \mathbb{R}^n \mapsto P_K x \in \mathbb{R}^n,$$

où  $P_K x$  est l'unique point de  $K$  qui réalise  $\inf_{y \in K} \|x - y\|$ .

On laisse en exercice la preuve de la proposition suivante. On vérifiera également que le résultat peut être mis en défaut si l'une des deux fonctions n'est pas supposée bornée.

### Proposition I.6 (Produit de fonctions lipschitziennes bornées)

Soit  $A \subset \mathbb{R}^m$  et  $f : A \rightarrow \mathbb{R}^n$ ,  $g : A \rightarrow \mathbb{R}$  deux fonctions **lipschitziennes et bornées**, alors le produit  $fg : A \rightarrow \mathbb{R}^n$  est également une fonction lipschitzienne bornée.



### II.1.b Pour aller plus loin

On utilisera par la suite, les propriétés suivantes.

#### Théorème I.7 (de prolongement lipschitzien)

Soit  $A \subset \mathbb{R}^m$  et  $f : A \rightarrow \mathbb{R}^n$  une fonction  $L$ -lipschitzienne, alors il existe un prolongement  $(\sqrt{n}L)$ -lipschitzien de  $f$  à  $\mathbb{R}^m$  tout entier, c'est-à-dire une fonction  $F : \mathbb{R}^m \rightarrow \mathbb{R}^n$  qui vérifie :

- $F|_A = f$ .
- $F$  est  $(\sqrt{n}L)$ -lipschitzienne sur  $\mathbb{R}^m$ .

Si de plus  $f$  est bornée sur  $A$ , on peut choisir  $F$  bornée sur  $\mathbb{R}^m$ .

#### Remarque I.8

- On peut en fait montrer (mais c'est significativement plus difficile) que l'on peut trouver un prolongement lipschitzien de  $f$  qui soit exactement  $L$ -lipschitzien (théorème de Kirszbraun), et même si les espaces de départ et d'arrivée sont des Hilbert de dimension finie.

Le résultat ci-dessus sera très suffisant pour nos besoins.

- Si on rajoute l'hypothèse que  $A$  est un convexe borné alors la démonstration du théorème de Kirszbraun est facile :

- On commence par poser  $K = \overline{A}$  qui est un compact convexe.
- On prolonge  $f$  à  $K$ , par prolongement uniformément continu : pour tout point  $x \in K$  on prend une suite  $(x_n)_n$  de points de  $A$  qui converge vers  $x$  et on définit

$$\tilde{f}(x) = \lim_{n \rightarrow \infty} f(x_n),$$

en montrant que la limite existe et ne dépend pas de la suite.

La fonction  $\tilde{f}$  ainsi construite est bien  $L$ -lipschitzienne.

- On introduit enfin la projection orthogonale  $P_K$  sur  $K$  et on pose

$$F(x) := \tilde{f}(P_K(x)), \quad \forall x \in \mathbb{R}^d.$$

Il est clair que  $F$  prolonge  $f$  et par ailleurs, comme  $P_K$  est 1-lipschitzienne, on voit que  $F$  est  $L$ -lipschitzienne.

#### Preuve :

- Commençons par le cas  $n = 1$ . On peut alors donner une formule explicite qui résout le problème

$$F(x) := \inf_{y \in A} \left( f(y) + L\|x - y\| \right).$$

En effet :

- Si  $x \in A$ , le caractère  $L$ -lipschitzien de  $f$  donne

$$\forall y \in A, \quad f(x) \leq f(y) + L\|x - y\|,$$

ce qui prouve que

$$f(x) \leq F(x).$$

Comme par ailleurs  $F(x)$  est plus petit que  $f(x)$  (il suffit d'évaluer la quantité dans l'infimum en  $y = x$ ), on a bien montré que  $f(x) = F(x)$  pour  $x \in A$ .

- Montrons maintenant que  $F$  est  $L$ -lipschitzienne sur  $\mathbb{R}^m$  tout entier. On se donne  $x_1, x_2 \in \mathbb{R}^m$  et  $y \in A$  pour écrire par inégalité triangulaire

$$F(x_1) \leq f(y) + L\|x_1 - y\| \leq f(y) + L\|x_2 - y\| + L\|x_1 - x_2\|.$$

Si on prend l'infimum en  $y$ , on obtient

$$F(x_1) \leq F(x_2) + L\|x_1 - x_2\|.$$

On peut échanger les rôles de  $x_1$  et  $x_2$  et finalement obtenir

$$|F(x_1) - F(x_2)| \leq L\|x_1 - x_2\|,$$

ce qu'il fallait démontrer

- Dans le cas  $n > 1$ , on applique la formule précédente à chaque composante de la fonction (vectorielle)  $f = (f_1, \dots, f_n)$ . On obtient un prolongement lipschitzien  $F = (F_1, \dots, F_n)$  de  $f$  dont chaque composante est  $L$ -lipschitzienne.

Si maintenant  $x_1, x_2 \in \mathbb{R}^m$ , on a

$$\|F(x_1) - F(x_2)\|^2 = \sum_{i=1}^n |F_i(x_1) - F_i(x_2)|^2 \leq \sum_{i=1}^n L^2 \|x_1 - x_2\|^2 = nL^2 \|x_1 - x_2\|^2,$$

et le résultat suit. ■

On a vu plus haut que les fonctions  $\mathcal{C}^1$  à différentielle bornée étaient lipschitziennes et on a également vu (exemple I.5) que certaines fonctions lipschitziennes n'étaient pas de classe  $\mathcal{C}^1$ . Le résultat suivant (que nous n'utiliserons pas et ne mentionnons que pour la culture mathématique du lecteur ...) montre qu'il n'y a finalement pas tant de différences entre les deux notions.

### Théorème I.9 (Rademacher)

Si  $\Omega$  est un ouvert de  $\mathbb{R}^m$  et  $f : \Omega \rightarrow \mathbb{R}^n$  est une fonction  $L$ -lipschitzienne, alors elle est différentiable **presque partout** et, en dehors d'un ensemble de mesure nulle on a la borne  $\|df(x)\| \leq L$ .

Il y a en revanche une sorte de cas particulier de ce résultat utile dans l'étude des équations différentielles. Ce résultat dit essentiellement que, si  $x$  est une fonction régulière, alors on peut (presque) dériver  $|x|$  comme si de rien n'était.

### Proposition I.10

Soit  $x : \mathbb{R} \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^1$ , alors la fonction  $|x|$  vérifie l'identité intégrale suivante

$$|x(t)| = |x(s)| + \int_s^t x'(\tau) \operatorname{sgn}(x(\tau)) d\tau, \quad \forall t, s \in \mathbb{R}.$$

où  $\operatorname{sgn}$  est la fonction suivante

$$\operatorname{sgn}(u) = \begin{cases} 1 & \text{si } u > 0, \\ 0 & \text{si } u = 0, \\ -1 & \text{si } u < 0. \end{cases}$$

Donnons juste une idée de la preuve, les détails sont laissés en exercice :

#### Preuve :

Pour tout  $\varepsilon > 0$  on définit la fonction  $\beta_\varepsilon : \mathbb{R} \rightarrow \mathbb{R}$  par

$$\beta_\varepsilon(s) = \frac{s^2}{\sqrt{s^2 + \varepsilon}}, \quad \forall s \in \mathbb{R}.$$

On établit les propriétés suivantes :

- $\beta_\varepsilon$  est de classe  $\mathcal{C}^1$  sur  $\mathbb{R}$  et  $\|\beta'_\varepsilon\|_\infty \leq 2$  pour tout  $\varepsilon > 0$ .
- Pour tout  $s \in \mathbb{R}$ ,  $\beta_\varepsilon(s) \rightarrow |s|$  quand  $\varepsilon \rightarrow 0$ .
- Pour tout  $s \in \mathbb{R}$ ,  $\beta'_\varepsilon(s) \rightarrow \operatorname{sgn}(s)$  quand  $\varepsilon \rightarrow 0$ .

On écrit ensuite la formule

$$\beta_\varepsilon(x(t)) = \beta_\varepsilon(x(s)) + \int_s^t \beta'_\varepsilon(x(\tau)) x'(\tau) d\tau,$$

et on justifie le passage à la limite  $\varepsilon \rightarrow 0$  dans cette égalité pour obtenir le résultat. ■

### II.1.c Fonctions localement lipschitziennes

Dans tout ce paragraphe, on utilise la topologie induite sur une partie de  $\mathbb{R}^m$ .

#### Définition I.11

Soit  $A \subset \mathbb{R}^m$  et  $f : A \rightarrow \mathbb{R}^n$  une fonction. On dit que  $f$  est **localement lipschitzienne** sur  $A$  si :  
Pour tout point  $x_0 \in A$ , il existe un ouvert  $U$  de  $A$  contenant  $x_0$  tel que  $f$  est lipschitzienne sur  $U$ .

#### Proposition I.12

1. Toute fonction localement lipschitzienne sur un ensemble  $A$  est continue dans  $A$ .
2. Si  $A$  est ouvert, toute fonction de classe  $C^1$  sur  $A$  est localement lipschitzienne sur  $A$ .

#### Preuve :

1. C'est immédiat à partir des définitions.
2. C'est une application de l'inégalité des accroissements finis.

On se donne un  $x_0 \in A$  puis un  $r > 0$  tel que  $K = \bar{B}(x_0, r) \subset A$  (ceci est possible car  $A$  est ouvert). Comme  $f$  est de classe  $C^1$  et que  $K$  est compact, il existe  $L > 0$  tel que  $\|df(x)\| \leq L$  pour tout  $x \in K$ .

L'ouvert  $U = B(x_0, r)$  est convexe et l'inégalité des accroissements finis permet d'établir que pour tous  $x, y \in U$ , on a

$$\|f(x) - f(y)\| \leq L\|x - y\|,$$

ce qui conclut la preuve. ■

En pratique, la caractérisation suivante est souvent très utile dans les preuves.

#### Proposition I.13

On reprend les notations de la définition et on suppose que  $A$  est un ouvert de  $\mathbb{R}^m$ .  
Les deux assertions suivantes sont équivalentes :

1.  $f$  est localement lipschitzienne sur  $A$ .
2. Pour tout compact  $K$  dans  $A$ ,  $f$  est (globalement) lipschitzienne sur  $K$ .

Notez que le compact  $K$  peut être aussi gros que voulu (à condition qu'il soit inclus dans  $A$ ) mais que bien entendu, la constante de lipschitz de  $f$  sur  $K$ , dépend de  $K$ , et peut tout à fait tendre vers l'infini quand  $K$  grossit jusqu'à remplir  $A$ .

#### Preuve :

1.  $\Rightarrow$  2. Soit  $K$  un compact de  $A$ . On va raisonner par l'absurde.

On suppose que  $f$  n'est pas lipschitzienne sur  $K$ . Cela signifie que, pour tout  $n \geq 1$ , il existe des points  $x_n, y_n \in K$  tels que

$$\|f(x_n) - f(y_n)\| > n\|x_n - y_n\|, \quad \forall n \geq 1. \quad (\text{I.3})$$

Comme  $K$  est un compact d'un espace métrique, et donc que  $K \times K$  est également un compact de l'espace produit, on peut extraire des sous-suites communes  $(x_{\varphi(n)})_n, (y_{\varphi(n)})_n$  qui convergent vers  $x^* \in K$  et  $y^* \in K$  respectivement. Comme  $f$  est continue, nous savons que

$$f(x_{\varphi(n)}) \xrightarrow{n \rightarrow \infty} f(x^*),$$

$$f(y_{\varphi(n)}) \xrightarrow{n \rightarrow \infty} f(y^*).$$

Il s'en suit que  $f(x_{\varphi(n)}) - f(y_{\varphi(n)})$  est bornée et donc, en utilisant (I.3), on obtient que

$$\|x_{\varphi(n)} - y_{\varphi(n)}\| \xrightarrow{n \rightarrow \infty} 0.$$

On déduit de cela que les deux limites  $x^*$  et  $y^*$  sont égales. On note ce point  $x_0$ .

On utilise maintenant l'hypothèse sur  $f$ , qui nous dit en particulier qu'il existe  $\delta > 0$  et  $L > 0$  tel que

$$\|f(x) - f(y)\| \leq L\|x - y\|, \quad \forall x, y \in A \cap B(x_0, \delta). \quad (\text{I.4})$$

Comme les suites  $(x_{\varphi(n)})_n$  et  $(y_{\varphi(n)})_n$  convergent vers  $x_0$ , on sait qu'il existe un entier suffisamment grand  $N$  tel que

$$x_{\varphi(n)} \in A \cap B(x_0, \delta), \text{ et } y_{\varphi(n)} \in A \cap B(x_0, \delta), \text{ pour tout } n \geq N.$$

On peut donc appliquer (I.4) et obtenir

$$\|f(x_{\varphi(n)}) - f(y_{\varphi(n)})\| \leq L\|x_{\varphi(n)} - y_{\varphi(n)}\|, \quad \forall n \geq N.$$

Comparant ceci à (I.3), on obtient bien une contradiction pour  $n$  assez grand.

2.  $\Rightarrow$  1. Soit  $x_0 \in A$ . Comme  $A$  est ouvert, il existe  $\delta > 0$  tel que  $\bar{B}(x_0, \delta) \subset A$ . Comme cette boule fermée est compacte (nous sommes dans un espace vectoriel normé de dimension finie !) on peut appliquer l'hypothèse et conclure directement que  $f$  est lipschitzienne sur la boule ouverte  $B(x_0, \delta)$ . ■

En combinant le Théorème I.7 et la proposition I.13, on obtient le corollaire utile suivant.

#### Corollaire I.14

Soit  $A$  un ouvert de  $\mathbb{R}^m$  et  $f : A \rightarrow \mathbb{R}^n$  une fonction localement lipschitzienne. Pour tout compact  $K$  dans  $A$ , il existe une fonction globalement lipschitzienne  $f_K : \mathbb{R}^m \rightarrow \mathbb{R}^n$  telle que

$$f_K = f, \text{ sur } K.$$

### II.1.d Champs de vecteurs. Variable de temps. Variable d'état

Comme on l'a vu en introduction, une équation différentielle est définie par une fonction  $f$  dont on va maintenant discuter les propriétés.

#### Définition I.15

Soit  $I \subset \mathbb{R}$  un intervalle ouvert et  $\Omega \subset \mathbb{R}^n$  un ouvert. Une application  $f : (t, x) \in I \times \Omega \mapsto \mathbb{R}^n$  est appelée un **champ de vecteurs** sur  $I \times \Omega$ .

On dit que ce champ de vecteurs est **autonome** si  $I = \mathbb{R}$  et si  $f$  ne dépend pas explicitement du temps (auquel cas on identifiera  $f$  à une fonction de  $\Omega$  dans  $\mathbb{R}^n$  ...).

Dans ce cours on ne considérera que des champs de vecteurs au moins continus.

#### Définition I.16

On dit qu'un champ de vecteurs continu  $f : I \times \Omega \rightarrow \mathbb{R}^n$  est **globalement lipschitzien par rapport à la variable d'état** s'il existe une fonction continue  $L : t \in I \rightarrow L(t) \in \mathbb{R}^+$  telle que

$$\text{Pour tout } t \in I, \text{ la fonction } x \in \Omega \mapsto f(t, x) \in \mathbb{R}^n \text{ est } L(t)\text{-lipschitzienne sur } \Omega.$$

Autrement dit,

$$\forall t \in I, \forall x_1, x_2 \in \Omega, \quad \|f(t, x_1) - f(t, x_2)\| \leq L(t)\|x_1 - x_2\|.$$

#### Définition I.17

On dit qu'un champ de vecteurs continu  $f : I \times \Omega \rightarrow \mathbb{R}^n$  est **localement lipschitzien par rapport à la variable d'état** si :

— Pour tout  $(t_0, x_0) \in I \times \Omega$ , il existe  $L > 0$ ,  $\delta > 0$ , un voisinage ouvert  $U$  de  $x_0$  tels que

$$\text{Pour tout } t \in I \cap [t_0 - \delta, t_0 + \delta], \text{ la fonction } x \in U \mapsto f(t, x) \in \mathbb{R}^n \text{ est } L\text{-lipschitzienne sur } U.$$

Attention à cette dernière définition qui ne signifie pas simplement que  $f(t, \cdot)$  est localement lipschitzienne pour tout  $t$ . C'est un peu plus fort que cela car on demande l'uniformité locale en temps de la constante de Lipschitz. Par exemple la fonction suivante

$$f(t, x) = \begin{cases} t \sin\left(\frac{x}{t^2}\right), & \text{si } t \neq 0, \\ 0, & \text{si } t = 0, \end{cases}$$

est continue, lipschitzienne par rapport à  $x$  pour tout  $t$ , mais n'est pas localement lipschitzienne par rapport à la variable  $x$  au sens de la définition précédente.

### Proposition I.18

*Un champ de vecteurs continu  $f : I \times \Omega \rightarrow \mathbb{R}^n$  est localement lipschitzien par rapport à la variable d'état si et seulement si : pour tout compact  $\mathcal{K} \subset I \times \Omega$ , il existe  $L > 0$  telle que*

$$\forall t \in I, x_1, x_2 \in \Omega, \text{ tels que } (t, x_1) \in \mathcal{K} \text{ et } (t, x_2) \in \mathcal{K} \text{ on a } \|f(t, x_1) - f(t, x_2)\| \leq L\|x_1 - x_2\|.$$

### Corollaire I.19

*Soit  $f : I \times \Omega \rightarrow \mathbb{R}^n$  un champ de vecteurs continu et localement lipschitzien par rapport à la variable d'état,  $K$  un compact de  $\Omega$  et  $[a, b] \subset I$  un sous-intervalle compact de  $I$ .*

*Il existe un champ de vecteur  $\tilde{f} : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  continu et globalement lipschitzien par rapport à la variable d'état tel que*

$$\tilde{f} = f, \text{ sur } [a, b] \times K.$$

### Preuve :

En raisonnant composante par composante, on voit qu'il suffit de montrer ce résultat de restriction/prolongement pour toute fonction scalaire  $g : I \times \Omega \rightarrow \mathbb{R}$  continue et localement lipschitzienne par rapport à la variable d'état. D'après la proposition précédente appliquée au compact  $\mathcal{K} = [a, b] \times K$ , il existe  $L > 0$  tel que

$$\forall t \in [a, b], \forall x_1, x_2 \in K, |g(t, x_1) - g(t, x_2)| \leq L\|x_1 - x_2\|.$$

On va commencer par fabriquer la fonction  $\tilde{g}$  recherchée l'ensemble  $[a, b] \times \mathbb{R}^n$  en utilisant, pour tout  $t \in [a, b]$  fixé qui joue ici le rôle d'un paramètre, la même formule quand dans la preuve du Théorème I.7 :

$$\tilde{g}(t, x) = \inf_{y \in K} \left( g(t, y) + L\|x - y\| \right).$$

On a déjà vu que cette fonction est bien globalement  $L$ -lipschitzienne par rapport à la variable  $x$ . Il reste à montrer que  $\tilde{g}$  est continue. Pour cela, on utilise le fait que  $g$  est uniformément continue sur le compact  $[a, b] \times K$  (théorème de Heine). Ainsi, pour  $\varepsilon > 0$  fixé, il existe  $\delta > 0$  tel que

$$\forall t_1, t_2 \in [a, b], \forall x_1, x_2 \in K, |t_1 - t_2| + \|x_1 - x_2\| \leq \delta \implies |g(t_1, x_1) - g(t_2, x_2)| \leq \varepsilon. \quad (\text{I.5})$$

Prenons maintenant  $t_1, t_2 \in [a, b], \forall x_1, x_2 \in \mathbb{R}^n$  vérifiant  $|t_1 - t_2| + \|x_1 - x_2\| \leq \delta$  et prenons un  $y \in K$  quelconque. Par (I.5) et l'inégalité triangulaire il vient

$$g(t_1, y) + L\|x_1 - y\| \leq g(t_2, y) + \varepsilon + L\|x_1 - x_2\| + L\|x_2 - y\| \leq g(t_2, y) + L\|x_2 - y\| + \varepsilon + L\delta,$$

et en prenant l'infimum sur  $y$ , on obtient

$$\tilde{g}(t_1, x_1) \leq \tilde{g}(t_2, x_2) + \varepsilon + L\delta.$$

Comme on peut échanger  $t_1, x_1$  avec  $t_2, x_2$  dans cette formule, on a finalement obtenu que

$$|\tilde{g}(t_1, x_1) - \tilde{g}(t_2, x_2)| \leq \varepsilon + L\delta.$$

Ceci montre que la fonction  $\tilde{g}$  ainsi contruite est uniformément continue sur  $[a, b] \times \mathbb{R}^n$  et donc en particulier continue.

Il reste à étendre  $\tilde{g}$  à  $\mathbb{R} \times \mathbb{R}^n$  tout entier, ce que l'on fait aisément en posant

$$\tilde{g}(t, x) = \tilde{g}(a, x), \quad \forall t \leq a,$$

$$\tilde{g}(t, x) = \tilde{g}(b, x), \quad \forall t \geq b.$$

On vérifie aisément que cette extension préserve les propriétés attendues de  $\tilde{g}$ . ■

## II.2 Equations et inéquations différentielles linéaires scalaires. Lemmes de Gronwall

Avant d'attaquer l'étude générale des équations différentielles, on peut déjà constater que la résolution des équations linéaires (homogènes) scalaires est facile. Il s'agit du cas où  $d = 1$  et où le « champ de vecteurs » (ici c'est un champ scalaire ...) est de la forme

$$f(t, x) = a(t)x, \quad \forall t \in I, \forall x \in \mathbb{R}.$$

### Proposition I.20

Soit  $I \subset \mathbb{R}$  un intervalle non vide et  $t \in I \mapsto a(t) \in \mathbb{R}$  une fonction continue.  
Si  $x : I \rightarrow \mathbb{R}$  est une fonction de classe  $C^1$  vérifiant l'équation

$$x'(t) = a(t)x(t), \quad \forall t \in I,$$

alors on a

$$x(t) = x(s) \exp \left( \int_s^t a(\tau) d\tau \right), \quad \forall t, s \in I.$$

### Preuve :

Fixons la valeur de  $s \in I$  quelconque et posons

$$z_s(t) = x(t) \exp \left( - \int_s^t a(\tau) d\tau \right), \quad \forall t \in I.$$

Par construction, cette fonction est de classe  $C^1$  et vérifie

$$z'_s(t) = (x'(t) - a(t)x(t)) \exp \left( - \int_s^t a(\tau) d\tau \right),$$

ce qui donne, avec l'équation vérifiée par  $x$ ,

$$z'_s(t) = 0, \quad \forall t \in I.$$

Autrement dit,  $z_s$  est une fonction constante, et en particulier

$$z_s(t) = z_s(s) = x(s), \quad \forall t \in I.$$

■

### Corollaire I.21

Pour tout  $t_0 \in I$  et toute valeur  $x_0 \in \mathbb{R}$ , il existe une unique solution  $x$ , définie sur  $I$  tout entier, du problème suivant

$$\begin{cases} x'(t) = a(t)x(t) \\ x(t_0) = x_0. \end{cases}$$

Elle est donnée par la formule suivante

$$x(t) = x_0 \exp \left( \int_{t_0}^t a(\tau) d\tau \right), \quad \forall t \in I.$$

Il se trouve que le calcul précédent permet également de donner des informations importantes sur la fonction  $x$  dans

le cas où l'équation différentielle est remplacée par une inéquation différentielle. C'est l'objet du résultat suivant.

### Proposition I.22

Soit  $I \subset \mathbb{R}$  un intervalle non vide et  $t \in I \mapsto a(t) \in \mathbb{R}$  une fonction continue. Soit  $x : I \rightarrow \mathbb{R}$  est une application de classe  $C^1$ .

Si  $x$  vérifie l'inégalité différentielle suivante

$$x'(t) \leq a(t)x(t), \quad \forall t \in I,$$

alors on a

$$x(t) \leq x(s) \exp \left( \int_s^t a(\tau) d\tau \right), \quad \forall t, s \in I, \text{ t.q. } t \geq s,$$

et

$$x(t) \geq x(s) \exp \left( \int_s^t a(\tau) d\tau \right), \quad \forall t, s \in I, \text{ t.q. } t \leq s.$$

### Remarque I.23

Si l'inégalité différentielle vérifiée par  $x$  est inversée

$$x'(t) \geq a(t)x(t), \quad \forall t \in I,$$

alors la conclusion demeure en inversant les sens des inégalités. Il suffit pour s'en convaincre de changer  $x$  en  $-x$ .

### Preuve :

On reprend mot pour mot la preuve de la proposition précédente. Cette fois, on obtient que

$$z'_s(t) = (x'(t) - a(t)x(t)) \exp(\dots) \leq 0,$$

Autrement dit, on a obtenu que  $z_s$  était décroissante (et non plus constante). Cela donne

$$z_s(t) \leq z_s(s) = x(s), \quad \forall t \in I, t \geq s,$$

$$z_s(t) \geq z_s(s) = x(s), \quad \forall t \in I, t \leq s.$$

■

Un des outils **fondamentaux** dans la théorie des équations différentielles est le résultat suivant, qui stipule essentiellement que le résultat de la proposition précédente persiste si l'inégalité différentielle est satisfaite sous forme intégrale à condition toutefois que la fonction  $a$  soit positive. Bien que son nom usuel soit *Lemme de Gronwall*, il mérite d'être présenté comme un théorème dans ce cours.

### Théorème I.24 (Lemme de Gronwall)

Soit  $I$  un intervalle non vide,  $t \in I \mapsto a(t) \in \mathbb{R}$  une fonction **positive** et  $C \in \mathbb{R}$  une constante.

Si  $x : I \rightarrow \mathbb{R}$  est une fonction **continue** qui vérifie, pour un certain  $s \in I$ , la propriété suivante :

$$x(t) \leq C + \int_s^t a(\tau)x(\tau)d\tau, \quad \forall t \in I, t \geq s, \tag{I.6}$$

alors nous avons l'inégalité

$$x(t) \leq C \exp \left( \int_s^t a(\tau)d\tau \right), \quad \forall t \in I, t \geq s. \tag{I.7}$$

Dans certains ouvrages, les auteurs ajoutent l'hypothèse «  $x$  est une fonction positive ». Il est bon de remarquer que c'est strictement inutile (même s'il est vrai que le lemme de Gronwall est souvent employé pour des fonctions  $x$  positives

...)

### Remarque I.25

Un résultat analogue dans le cas  $t \leq s$  est le suivant : si  $x$  vérifie

$$x(t) \leq C + \int_t^s a(\tau)x(\tau)d\tau, \quad \forall t \in I, t \leq s,$$

alors on a

$$x(t) \leq C \exp\left(\int_t^s a(\tau)d\tau\right), \quad \forall t \in I, t \leq s.$$

Notez bien que les bornes dans les intégrales ont été échangées.

La démonstration de cette propriété est un excellent exercice.

### Preuve :

Bien évidemment s'il était possible de « dériver une inégalité »<sup>1</sup>, et si  $x$  était supposée dérivable, alors on pourrait directement obtenir l'inégalité différentielle  $x' \leq ax$  et conclure par la proposition précédente. Toute l'idée de la preuve consiste à contourner ces difficultés et *faire comme si*.

Pour cela, on va donner un nom au membre de droite de l'inégalité (I.6)

$$y(t) := C + \int_s^t a(\tau)x(\tau)d\tau, \quad \forall t \in I.$$

Comme  $x$  et  $a$  sont continues, cette fonction  $y$  est bien de classe  $C^1$  et vérifie

$$y'(t) = a(t)x(t).$$

Mais par hypothèse (I.6) nous avons  $x \leq y$  et la fonction  $a$  est positive<sup>2</sup> donc nous avons l'inégalité

$$y'(t) \leq a(t)y(t), \quad \forall t \in I, t \geq s.$$

On peut donc appliquer à  $y$  la proposition I.22 et conclure que

$$y(t) \leq y(s) \exp\left(\int_s^t a(\tau)d\tau\right), \quad \forall t \in I, t \geq s.$$

Par construction, nous avons  $y(s) = C$  et comme  $x \leq y$  par hypothèse, on a bien prouvé que

$$x(t) \leq C \exp\left(\int_s^t a(\tau)d\tau\right), \quad \forall t \in I, t \geq s.$$

■

Ce lemme fondamental admet de nombreuses variantes et généralisations. On peut par exemple énoncer le résultat suivant dont la preuve élémentaire est laissée en exercice.

### Proposition I.26 (lemme de Gronwall généralisé)

Soit  $I$  un intervalle non vide,  $t \in I \mapsto a(t) \in \mathbb{R}$  une fonction **positive** et  $b : I \rightarrow \mathbb{R}$  une fonction **croissante**. Si  $x : I \rightarrow \mathbb{R}$  est une fonction **continue** qui vérifie, pour un certain  $s \in I$ , la propriété suivante :

$$x(t) \leq b(t) + \int_s^t a(\tau)x(\tau)d\tau, \quad \forall t \in I, t \geq s,$$

alors nous avons l'inégalité

$$x(t) \leq b(t) \exp\left(\int_s^t a(\tau)d\tau\right), \quad \forall t \in I, t \geq s.$$

1. Se convaincre que ça n'est pas licite !!

2. Notez bien à quel endroit cette hypothèse cruciale intervient



## II.3 Notions de solution d'un problème de Cauchy

### II.3.a Définitions

On se donne un champ de vecteurs continu  $f : I \times \Omega \rightarrow \mathbb{R}^d$ , ainsi qu'un instant initial  $t_0 \in I$  et un état initial  $x_0 \in \Omega$ .

La définition suivante est fondamentale car elle formalise le fait que l'intervalle de temps sur lequel une solution est définie est inconnu *a priori* et sa détermination fait partie intégrante du problème.

#### Définition I.27

Une solution du problème de Cauchy (I.1) est un couple  $(J, x)$  où  $J \subset I$  est un intervalle contenant  $t_0$ , et  $x : J \rightarrow \Omega$  une application de classe  $\mathcal{C}^1$  vérifiant  $x(t_0) = x_0$  ainsi que l'équation différentielle

$$x'(t) = f(t, x(t)), \quad \forall t \in J.$$

Comme l'intervalle de définition des solutions n'est pas connu *a priori*, on a besoin de préciser les définitions suivantes.

#### Définition I.28

Une solution  $(J, x)$  du problème de Cauchy (I.1) est dite :

- **Maximale** : s'il n'existe pas de solution  $(\tilde{J}, \tilde{x})$  qui prolonge strictement  $(J, x)$ , autrement dit telle que  $J \subsetneq \tilde{J}$  et  $\tilde{x} = x$  sur  $J$ . Voir la figure I.5.
- **Globale** : si  $J = I$  (dans ce cas elle est bien évidemment maximale).

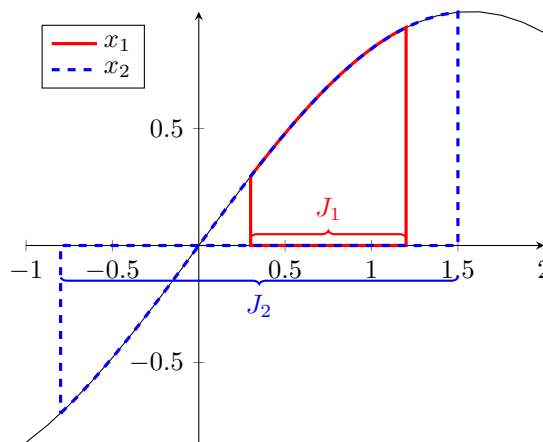


FIGURE I.5 – Notion de prolongement de solution.  $(J_1, x_1)$  n'est pas une solution maximale.

## III Théorème de Cauchy-Lipschitz global et applications

On suppose dans toute cette section que l'ensemble des états possibles du système est l'espace entier :  $\Omega = \mathbb{R}^d$ .

Toute la théorie exposée ici repose sur la remarque fondamentale mais immédiate suivante : pour  $J \subset I$  contenant  $t_0$  et  $x : J \rightarrow \mathbb{R}^d$ , les propriétés suivantes sont équivalentes

1.  $x \in \mathcal{C}^1(J, \mathbb{R}^d)$ , vérifie  $x(t_0) = x_0$  et  $x'(t) = f(t, x(t))$  pour tout  $t \in J$ .
2.  $x \in \mathcal{C}^0(J, \mathbb{R}^d)$  et vérifie

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds.$$

### III.1 Enoncé et preuve du théorème principal

Le but de cette sous-section est de démontrer le théorème suivant.

#### Théorème I.29 (Cauchy-Lipschitz global)

On suppose que le champ de vecteurs  $f : I \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  est continu et **globalement lipschitzien** par rapport à la variable d'état.

Alors, pour toute donnée de Cauchy  $(t_0, x_0)$ , il **existe une unique solution globale**  $(I, x)$  du problème de Cauchy (I.1).

De plus, toute autre solution  $(J, \tilde{x})$  du problème de Cauchy est la restriction à  $J$  de  $x$ .

#### Preuve :

La preuve est structurée en plusieurs parties.

##### — Unicité :

On montre la propriété suivante qui inclut l'unicité d'une éventuelle solution globale et la dernière assertion du théorème : soient  $(J_1, x_1)$  et  $(J_2, x_2)$  deux solutions du problème de Cauchy (I.1) et soit  $J = J_1 \cap J_2$ . On va montrer que  $x_1 = x_2$  sur  $J$ .

La différence  $z = x_1 - x_2$  vérifie, par soustraction

$$z'(t) = f(t, x_1(t)) - f(t, x_2(t)), \quad \forall t \in J,$$

ainsi que  $z(t_0) = 0$ .

Pour tout  $t \in J$ ,  $t \geq t_0$ , on intègre l'équation ci-dessus entre  $t_0$  et  $t$  pour obtenir

$$z(t) = \int_{t_0}^t (f(s, x_1(s)) - f(s, x_2(s))) ds.$$

Prenons la norme de cette identité (on utilise ici que  $t \geq t_0$  !)

$$\|z(t)\| \leq \int_{t_0}^t \|f(s, x_1(s)) - f(s, x_2(s))\| ds,$$

et utilisons le caractère globalement lipschitzien de  $f$  par rapport à la variable d'état (remarquez que  $f$  est évaluée au même temps  $s$  dans les deux termes), pour obtenir

$$\|z(t)\| \leq \int_{t_0}^t L(s) \|x_1(s) - x_2(s)\| ds = \int_{t_0}^t L(s) \|z(s)\| ds.$$

Comme ceci est vrai pour tout  $t \geq t_0$ , que  $t \mapsto \|z(t)\|$  est continue et que  $L \geq 0$  (c'est la constante de Lipschitz), on peut appliquer le lemme de Gronwall (avec ici  $C = 0$ ) et en déduire que

$$\|z(t)\| \leq 0 \times \exp\left(\int_{t_0}^t L(s) ds\right), \quad \forall t \geq t_0.$$

Ceci montre bien que  $z$  est nulle pour  $t \geq t_0$ , autrement dit que  $x_1 \equiv x_2$  pour  $t \in J \cap [t_0, +\infty[$ . Un raisonnement similaire en tout point<sup>3</sup> montre que  $x_1 \equiv x_2$  pour  $t \in J \cap ]-\infty, t_0]$ .

##### — Borne a priori :

Il s'agit ici de supposer que la solution globale  $(I, x)$  existe et d'établir une borne qu'elle devra nécessairement vérifier. Pour cela, on intègre à nouveau entre  $t_0$  et  $t \geq t_0$ , l'équation différentielle pour obtenir (en utilisant que  $x(t_0) = x_0$  qui est une valeur fixée)

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds, \quad \forall t \geq t_0.$$

On aimerait pouvoir utiliser le lemme de Gronwall pour estimer  $x$  mais il faut pour cela arriver à majorer le terme en  $f$  et la seule hypothèse à notre disposition pour cela, c'est le caractère lipschitzien par rapport à la variable d'état. Il nous faut donc faire apparaître une différence de valeurs de  $f$ , par exemple de la façon suivante

$$x(t) - x_0 = \int_{t_0}^t f(s, x_0) ds + \int_{t_0}^t (f(s, x(s)) - f(s, x_0)) ds, \quad \forall t \geq t_0.$$

3. Vous êtes invités à le faire en détail

La première intégrale ne dépend que des données (et pas de la solution  $x$ !), et dans la seconde on peut utiliser l'hypothèse. Posons donc  $z(t) = x(t) - x_0$  et écrivons

$$\|z(t)\| \leq \int_{t_0}^t \|f(s, x_0)\| ds + \int_{t_0}^t L(s) \|z(s)\| ds, \quad \forall t \geq t_0.$$

Le lemme de Gronwall (généralisé, Prop. I.26) nous donne l'estimation

$$\|z(t)\| \leq \left( \int_{t_0}^t \|f(s, x_0)\| ds \right) \exp \left( \int_{t_0}^t L(s) ds \right), \quad \forall t \geq t_0.$$

Revenant à la variable  $x$ , ceci s'écrit

$$\|x(t) - x_0\| \leq \left( \int_{t_0}^t \|f(s, x_0)\| ds \right) \exp \left( \int_{t_0}^t L(s) ds \right), \quad \forall t \geq t_0.$$

Une estimation similaire se démontre pour les  $t \leq t_0$  (en mettant des valeurs absolues sur les intégrales ...) de sorte que, si on pose dorénavant

$$\varphi(t) := 1 + \|x_0\| + \left| \int_{t_0}^t \|f(s, x_0)\| ds \right|, \quad \forall t \in I,$$

$$\psi(t) := \left| \int_{t_0}^t L(s) ds \right|, \quad \forall t \in I,$$

qui ne dépendent que des données  $t_0, x_0$  et  $f$ , alors on a obtenu, l'estimation *a priori* suivante

$$\|x(t)\| \leq \varphi(t) e^{\psi(t)}, \quad \forall t \in I. \quad (\text{I.8})$$

On notera que l'on a arbitrairement ajouté 1 à la fonction  $\varphi$  de sorte que  $\varphi(t) \geq 1$  et  $\psi(t) \geq 0$ , pour tout  $t \in I$ . Par ailleurs, on observe que  $\varphi$  est décroissante sur  $I \cap ]-\infty, t_0]$  et croissante sur  $I \cap [t_0, +\infty[$ .

— *Introduction d'un bon espace fonctionnel :*

Fort des calculs précédents, on sait maintenant qu'il suffit de chercher  $x$  dans l'ensemble des fonctions qui vérifient l'estimation (I.8). En réalité il va falloir relâcher un peu cette contrainte en introduisant l'espace suivant

$$E := \left\{ z \in \mathcal{C}^0(I, \mathbb{R}^d), \sup_{t \in I} \left( \varphi(t)^{-1} e^{-2\psi(t)} \|z(t)\| \right) < +\infty \right\},$$

que l'on munit de la norme <sup>4</sup>

$$\|z\|_E := \sup_{t \in I} \left( \varphi(t)^{-1} e^{-2\psi(t)} \|z(t)\| \right).$$

Notez bien le facteur 2 que l'on a mis dans l'exponentielle par rapport à la quantité obtenue dans la borne *a priori*. La présence de ce facteur est fondamentale dans la suite (n'importe quel facteur  $> 1$  ferait l'affaire). Par ailleurs, la division par  $\varphi$  car on a pris soin de s'assurer que  $\varphi$  restait plus grande que 1.

On laisse en exercice le résultat suivant (la preuve est très similaire à celle de la complétude de l'espace des fonctions continues bornées muni de la norme infinie).

### Lemme I.30

L'espace  $(E, \|\cdot\|_E)$  est un Banach.

— *Introduction d'un problème de point fixe :*

Pour toute fonction  $z \in E$ , on définit une nouvelle fonction  $\theta(z) \in \mathcal{C}^0(I, \mathbb{R}^d)$  par la formule

$$(\theta(z))(t) := x_0 + \int_{t_0}^t f(s, z(s)) ds, \quad \forall t \in I.$$

Étudions les propriétés de cette application :

— On constate que les fonctions constantes sont dans  $E$ . Calculons  $\theta(x_0)$  (c'est un très léger abus de notation, on devrait écrire  $\theta(t \mapsto x_0)$ ) :

$$(\theta(x_0))(t) = x_0 + \int_{t_0}^t f(s, x_0) ds,$$

4. Si nécessaire, vérifiez que c'est bien une norme !

ce qui donne

$$\|(\theta(x_0))(t)\| \leq \|x_0\| + \left| \int_{t_0}^t \|f(s, x_0)\| ds \right| \leq \varphi(t),$$

et donc  $\theta(x_0) \in E$ .

— Montrons que  $\theta$  envoie  $E$  dans lui-même : pour cela on prend  $z, \tilde{z} \in E$  et on écrit

$$(\theta(z))(t) - (\theta(\tilde{z}))(t) = \int_{t_0}^t (f(s, z(s)) - f(s, \tilde{z}(s))) ds,$$

que l'on majore de la façon suivante, en utilisant le caractère lipschitzien de  $f$  et les propriétés de monotonie de  $\varphi$ ,

$$\begin{aligned} \|(\theta(z))(t) - (\theta(\tilde{z}))(t)\| &\leq \left| \int_{t_0}^t L(s) \|z(s) - \tilde{z}(s)\| ds \right| \\ &= \left| \int_{t_0}^t \varphi(s) L(s) e^{2\psi(s)} \varphi(s)^{-1} e^{-2\psi(s)} \|z(s) - \tilde{z}(s)\| ds \right| \\ &\leq \|z - \tilde{z}\|_E \varphi(t) \left| \int_{t_0}^t L(s) e^{2\psi(s)} ds \right| \\ &\leq \|z - \tilde{z}\|_E \varphi(t) \left[ \frac{e^{2\psi(s)}}{2} \right]_{t_0}^t \\ &\leq \frac{1}{2} \|z - \tilde{z}\|_E \varphi(t) e^{2\psi(t)}. \end{aligned}$$

Cela donne

$$\sup_{t \in I} \left( \varphi(t)^{-1} e^{-2\psi(t)} \|(\theta(z))(t) - (\theta(\tilde{z}))(t)\| \right) \leq \frac{1}{2} \|z - \tilde{z}\|_E. \quad (\text{I.9})$$

On en déduit que  $\theta(z) - \theta(\tilde{z})$  est dans  $E$ . En prenant  $\tilde{z} = x_0$ , comme on a vu que  $\theta(x_0) \in E$ , on en déduit que  $\theta(z) \in E$  pour tout  $z \in E$ , ce qui montre bien que  $\theta$  envoie  $E$  dans lui-même.

— Maintenant que l'on a établi que  $\theta(E) \subset E$ , on peut revenir à (I.9) qui s'écrit

$$\|\theta(z) - \theta(\tilde{z})\|_E \leq \frac{1}{2} \|z - \tilde{z}\|_E, \quad \forall z, \tilde{z} \in E.$$

On a bien montré que  $\theta$  est contractante sur  $E$ . Observez ici l'importance d'avoir mis le facteur 2 dans l'exponentielle dans la définition de  $E$ , c'est lui qui fournit le taux de contraction 1/2 dans cette inégalité. Pour s'en convaincre, on peut reprendre les calculs précédents, en remplaçant ce coefficient dans la définition de  $E$  par un nombre  $\alpha > 0$  arbitraire et observer ce que cela donne.

— *Résolution du problème de point fixe :*

Nous sommes dans le cadre d'application du théorème de point fixe de Banach, qui nous dit que l'application  $\theta$  admet un unique point fixe dans  $E$  que l'on note  $x$ . Cette fonction est continue et vérifie donc pour tout  $t \in I$

$$x(t) = (\theta(x))(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds.$$

En prenant  $t = t_0$ , on obtient que  $x(t_0) = x_0$ . D'autre part, cette égalité prouve que  $x$  est la primitive de la fonction continue  $s \mapsto f(s, x(s))$  et qu'elle est donc de classe  $\mathcal{C}^1$  et vérifie

$$x'(t) = f(t, x(t)), \quad \forall t \in I,$$

ce qui est bien l'équation différentielle initiale. Le théorème est démontré. ■

## III.2 Equations différentielles linéaires

Une équation différentielle linéaire est une équation associée à un champ de vecteurs de la forme

$$f(t, x) = A(t)x + b(t), \quad \forall t \in I, \forall x \in \mathbb{R}^d,$$

où  $t \in I \mapsto A(t) \in M_d(\mathbb{R})$  est une application continue à valeurs matricielles (ou endomorphismes si on remplace  $\mathbb{R}^d$  par un espace vectoriel de dimension finie) et  $b : t \in I \mapsto b(t) \in \mathbb{R}^d$  est une application à valeurs vectorielles. Si  $b \equiv 0$  (resp.  $b \not\equiv 0$ ) on dit que l'équation est homogène (resp. non-homogène).

Notez que, pour tout  $t \in I$  fixé, l'application  $f(t, \cdot)$  est affine et non pas linéaire ... ce qui donne lieu à une petite incohérence de vocabulaire qui est pourtant rentrée dans les mœurs ...

Pour un tel champ  $f$ , on a immédiatement

$$\forall t \in I, \forall x_1, x_2 \in \mathbb{R}^d, \|f(t, x_1) - f(t, x_2)\| = \|A(t)(x_1 - x_2)\| \leq \|A(t)\| \|x_1 - x_2\|,$$

ce qui montre que  $f$  est continu et globalement lipschitzien par rapport à la variable d'état. En conséquence de quoi, le théorème de Cauchy-Lipschitz global s'applique et devient

### **Théorème I.31 (Cauchy-Lipschitz linéaire)**

*Sous les hypothèses ci-dessus, pour toute donnée initiale  $(t_0, x_0) \in I \times \mathbb{R}^d$ , il existe une unique solution globale  $(I, x)$  au problème de Cauchy*

$$\begin{cases} x'(t) = A(t)x(t) + b(t), \\ x(t_0) = x_0, \end{cases} \quad (\text{I.10})$$

*Toute solution de ce problème est une restriction de cette solution globale.*

### **Proposition I.32**

*Soient  $x_1, \dots, x_m : I \rightarrow \mathbb{R}^d$  des solutions de l'équation différentielle linéaire homogène  $x' = A(t)x$ . Les propositions suivantes sont équivalentes :*

1. *Pour tout  $t \in I$ ,  $(x_1(t), \dots, x_m(t))$  est une famille libre de  $\mathbb{R}^d$ .*
2. *Il existe un  $t^* \in I$  tel que  $(x_1(t^*), \dots, x_m(t^*))$  est une famille libre de  $\mathbb{R}^d$ .*
3.  *$(x_1, \dots, x_m)$  est une famille libre de  $\mathcal{C}^0(I, \mathbb{R}^d)$ .*

On en déduit le corollaire immédiat mais fondamental suivant.

### **Corollaire I.33 (Structure des solutions d'une équation linéaire)**

*L'ensemble  $S_0$  de toutes les solutions de l'équation homogène  $x' = A(t)x$  est un espace vectoriel de dimension  $d$  et, pour tout  $b$ , l'ensemble  $S_b$  de toutes les solutions de l'équation générale  $x' = A(t)x + b(t)$  est un espace affine de dimension  $d$  dirigé par  $S_0$ .*

### **Preuve (de la proposition I.32):**

Les implications 1.  $\Rightarrow$  2. et 2.  $\Rightarrow$  3. sont immédiates, et sont en fait valables pour n'importe quelle famille de fonctions continues. On les laisse en exercice.

La dernière implication 3.  $\Rightarrow$  1., elle, n'est valable que parce que les  $x_i$  sont solutions de l'équation différentielle. En effet, en raisonnant par contraposition, supposons qu'il existe un  $t_0 \in I$  tel que  $(x_1(t_0), \dots, x_m(t_0))$  soit une famille liée de  $\mathbb{R}^d$ . Cela signifie qu'il existe des coefficients  $\alpha_1, \dots, \alpha_m \in \mathbb{R}$  non tous nuls tels que

$$\alpha_1 x_1(t_0) + \dots + \alpha_m x_m(t_0) = 0. \quad (\text{I.11})$$

Définissons la fonction  $y = \sum_{i=1}^m \alpha_i x_i(t)$ . Comme chaque  $x_i$  est solution de l'équation linéaire homogène,  $x' = A(t)x$ , on obtient immédiatement que  $y$  est également solution de cette équation. Mais par ailleurs, d'après (I.11), nous avons  $y(t_0) = 0$ . Ainsi  $y$  est l'unique solution globale de  $x' = A(t)x$  associée à la donnée de Cauchy  $(t_0, 0)$ , mais cette solution ne saurait être que la solution identiquement nulle.

On a ainsi prouvé que  $y(t) = 0$  pour tout  $t \in I$ , ce qui signifie que  $\sum_{i=1}^m \alpha_i x_i \equiv 0$ . Ceci est exclu car, par hypothèse, la famille  $(x_i)_i$  est libre dans  $\mathcal{C}^0(I, \mathbb{R}^d)$  et les  $\alpha_i$  sont non tous nuls. ■

— Dans le cas autonome, on dispose d'une formule de calcul des solutions similaire au cas scalaire : si  $A \in M_d(\mathbb{R})$  et  $x_0 \in \mathbb{R}^d$ , alors l'unique solution du problème de Cauchy

$$\begin{cases} x'(t) = Ax(t), \\ x(t_0) = x_0, \end{cases}$$

est donnée par la formule

$$x(t) = e^{(t-t_0)A} x_0.$$

- Dans le cas non-homogène, la formule ci-dessus se généralise sous la forme suivante (dite formule de Duhamel, ou de variation de la constante) : si  $A \in M_d(\mathbb{R})$ ,  $b : I \rightarrow \mathbb{R}^d$ , et  $x_0 \in \mathbb{R}^d$ , alors l'unique solution du problème de Cauchy

$$\begin{cases} x'(t) = Ax(t) + b(t), \\ x(t_0) = x_0, \end{cases}$$

est donnée par la formule

$$x(t) = e^{(t-t_0)A}x_0 + \int_{t_0}^t e^{(t-s)A}b(s) ds.$$

- Dans le cas non autonome (où  $A$  dépend du temps) et si la dimension est  $d = 1$ ,  $A(t)$  est alors un nombre réel noté  $a(t)$  pour éviter les confusions et on a alors les formules suivantes

$$x(t) = \exp\left(\int_{t_0}^t a(\tau)d\tau\right)x_0,$$

pour la solution du problème de Cauchy homogène, et

$$x(t) = \exp\left(\int_{t_0}^t a(\tau)d\tau\right)x_0 + \int_{t_0}^t \left[\exp\left(\int_s^t a(\tau)d\tau\right)b(s)\right] ds,$$

pour la solution du problème de Cauchy non-homogène.

- Enfin, dans le cas non autonome et si la dimension  $d$  est au moins égale à 2, alors on ne peut plus exprimer la solution du problème de Cauchy à l'aide des exponentielles de matrices. En particulier la fonction  $t \mapsto \exp\left(\int_{t_0}^t A(\tau)d\tau\right)x_0$  ne résout pas l'équation différentielle  $x' = A(t)x$  en général.

### III.3 Flot d'un champ de vecteurs

Au lieu de regarder chaque solution de l'équation différentielle indépendamment, on propose ici une vision plus globale de l'action de l'équation différentielle sur les différents états initiaux possibles.

#### III.3.a Définitions

Plaçons nous dans les conditions du théorème de Cauchy-Lipschitz global :  $f$  est un champ de vecteurs continu et globalement lipschitzien par rapport à la variable d'état. Le théorème 1.29 nous dit que pour tout  $t_0 \in I$ ,  $x_0 \in \mathbb{R}^d$ , il existe une unique solution globale au problème de Cauchy (1.1) que l'on notera ici provisoirement  $x_{t_0, x_0} \in \mathcal{C}^1(I, \mathbb{R}^d)$  pour bien faire apparaître la dépendance en les données de cette solution.

Pour manipuler plus convenablement ces quantités par rapport aux différentes variables, on introduit la définition suivante :

#### Définition I.34 (Flot d'un champ de vecteurs)

*Sous les hypothèses précédentes, on définit pour tout  $t, t_0, x_0$  la quantité*

$$\varphi(t, t_0, x_0) = x_{t_0, x_0}(t).$$

*L'application  $\varphi$  est appelée le flot associé au champ de vecteurs  $f$ .*

*On introduit également la famille d'applications*

$$\Phi(t, t_0) : x_0 \in \mathbb{R}^d \mapsto \varphi(t, t_0, x_0) \in \mathbb{R}^d$$

*qu'on appelle flot entre les instants  $t_0$  et  $t$ .*

Les instants initial  $t_0$  et final  $t$  étant fixés, il faut voir  $\Phi(t, t_0)$  comme l'application qui associe à chaque donnée initiale possible, la valeur correspondante de la solution à l'instant final.

## III.3.b Principales propriétés

**Théorème I.35 (Propriétés du flot)**

Avec les notations et hypothèses précédentes, on a les propriétés suivantes :

— **Propriété de groupe du flot :**

Pour tous  $t_0, t_1, t_2 \in I$ , on a

$$\Phi(t_0, t_0) = \text{Id},$$

$$\Phi(t_2, t_0) = \Phi(t_2, t_1) \circ \Phi(t_1, t_0).$$

En particulier, pour tous  $t_1, t_0 \in I$ , le flot  $\Phi(t_1, t_0)$  est un homéomorphisme bi-lipschitzien de  $\mathbb{R}^d$  dont la réciproque est donnée par

$$\Phi(t_1, t_0)^{-1} = \Phi(t_0, t_1).$$

— **Continuité par rapport aux données :**

L'application  $\varphi : (t, t_0, x_0) \in I \times I \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  est localement lipschitzienne.

**Remarque I.36 (A propos du cas autonome)**

Si le champ de vecteurs  $f$  est autonome (i.e. ne dépend pas du temps) alors on peut montrer (exercice) que l'application  $\Phi(t, t_0)$  ne dépend que de la quantité  $t - t_0$ . En d'autres termes, seules comptent les durées et pas les valeurs absolues des instants initiaux et finaux.

Dans ces conditions, on note plutôt

$$\varphi(t, x_0) := \varphi(t, 0, x_0),$$

$$\Phi(t) := \Phi(t, 0),$$

et les propriétés de groupe deviennent

$$\Phi(0) = \text{Id},$$

$$\Phi(t + s) = \Phi(t) \circ \Phi(s), \quad \forall t, s \in \mathbb{R},$$

$$(\Phi(t))^{-1} = \Phi(-t), \quad \forall t \in \mathbb{R}.$$

**Preuve :**

— On remarque d'abord que par définition on a  $\Phi(t, t) = \text{Id}$  pour tout  $t \in I$ .

Fixons maintenant  $x_0 \in \mathbb{R}^d$  et  $t_0$  et  $t_1$  dans  $I$ . On définit les applications

$$x(t) := \varphi(t, t_0, x_0), \quad \forall t \in I,$$

$$y(t) := \varphi(t, t_1, \varphi(t_1, t_0, x_0)), \quad \forall t \in I.$$

Par définition de  $\varphi$ , ces deux fonctions satisfont la même équation différentielle

$$x'(t) = f(t, x(t)), \quad \text{et} \quad y'(t) = f(t, y(t)).$$

Par ailleurs, nous avons

$$x(t_1) = \varphi(t_1, t_0, x_0), \quad y(t_1) = \varphi(t_1, t_1, \varphi(t_1, t_0, x_0)) = \varphi(t_1, t_0, x_0).$$

Ainsi,  $x$  et  $y$  sont solutions du même problème de Cauchy pour le champ de vecteurs  $f$ , ainsi la propriété d'unicité d'une telle solution nous donne que  $x$  et  $y$  sont identiques. En particulier, on a  $x(t_2) = y(t_2)$  qui est la propriété annoncée.

— Montrons maintenant que  $\Phi(t_1, t_0)$  est lipschitzien pour tout  $t_0, t_1 \in I$ . On fixe  $t_0 \in I$  et on prend  $x_0, \tilde{x}_0 \in \mathbb{R}^d$  puis on écrit la forme intégrale du problème de Cauchy sous la forme

$$\varphi(t, t_0, x_0) = x_0 + \int_{t_0}^t f(s, \varphi(s, t_0, x_0)) ds, \quad \forall t \in I,$$

$$\varphi(t, t_0, \tilde{x}_0) = \tilde{x}_0 + \int_{t_0}^t f(s, \varphi(s, t_0, \tilde{x}_0)) ds, \quad \forall t \in I$$

ce qui donne, par soustraction et inégalité triangulaire (pour  $t \geq t_0$ )

$$\|\varphi(t, t_0, x_0) - \varphi(t, t_0, \tilde{x}_0)\| \leq \|x_0 - \tilde{x}_0\| + \int_{t_0}^t L(s) \|\varphi(s, t_0, x_0) - \varphi(s, t_0, \tilde{x}_0)\| ds, \quad \forall t \in I, t \geq t_0.$$

En utilisant le lemme de Gronwall il vient

$$\|\varphi(t, t_0, x_0) - \varphi(t, t_0, \tilde{x}_0)\| \leq \|x_0 - \tilde{x}_0\| \exp\left(\int_{t_0}^t L(s) ds\right), \quad \forall t \in I, t \geq t_0.$$

Ceci montre bien que, pour  $t_0$  et  $t_1 \geq t_0$  fixés,  $\Phi(t_1, t_0)$  est lipschitzienne. Le cas  $t_1 \leq t_0$  se démontre de façon similaire.

- Fixons maintenant  $t, t_0 \in I$  et  $x_0 \in \mathbb{R}^d$ . On se donne un intervalle ouvert  $J$  contenant  $t_0$  et  $t$  et tel que  $\bar{J} \subset I$ . L'inégalité établie au point précédent montre que

$$\|\varphi(t, t_0, x_0) - \varphi(t, t_0, \tilde{x}_0)\| \leq C(J)\|x_0 - \tilde{x}_0\|, \quad \forall t_0, t \in J, \forall x_0 \in \mathbb{R}^n, \quad (\text{I.12})$$

où  $C(J) = \exp\left(\int_J L\right)$  ne dépend que de  $J$  et de la constante de lipschitz  $L$ .

L'estimation *a priori* que nous avons vue dans la preuve du théorème de Cauchy-Lipschitz global nous donne, pour tous  $t_1, t_2 \in J$ ,

$$\|\varphi(t_1, t_2, x_0)\| \leq \left(\|x_0\| + \int_J \|f(s, x_0)\| ds\right) \exp\left(\int_J L(s) ds\right) := R_1(x_0, J).$$

En combinant avec l'estimée du point précédent, on obtient que, pour tout  $\tilde{x}_0 \in B(x_0, 1)$  et pour tous  $t_1, t_2 \in J$ , on a

$$\|\varphi(t_1, t_2, \tilde{x}_0)\| \leq R_1(x_0, J) + \exp\left(\int_J L(s) ds\right) := R_2(x_0, J).$$

On note maintenant par  $M(x_0, J)$  une borne de  $f$  sur le compact  $\bar{J} \times \bar{B}(0, R_2(x_0, J))$ , et on utilise la forme intégrée de l'équation pour trouver, pour tous  $t_1, t_2, s \in J$  et tout  $\tilde{x}_0 \in B(x_0, 1)$ ,

$$\varphi(t_1, s, \tilde{x}_0) = \varphi(t_2, s, \tilde{x}_0) + \int_{t_2}^{t_1} f(\tau, \varphi(\tau, s, \tilde{x}_0)) d\tau,$$

et donc

$$\|\varphi(t_1, s, \tilde{x}_0) - \varphi(t_2, s, \tilde{x}_0)\| \leq M(x_0, J)|t_2 - t_1|. \quad (\text{I.13})$$

On peut désormais écrire, en utilisant l'inégalité triangulaire, la propriété de groupe du flot, puis l'inégalité (I.12) et enfin (I.13)

$$\begin{aligned} \|\varphi(\tilde{t}, \tilde{t}_0, \tilde{x}_0) - \varphi(t, t_0, \tilde{x}_0)\| &\leq \|\varphi(\tilde{t}, \tilde{t}_0, \tilde{x}_0) - \varphi(\tilde{t}, t_0, \tilde{x}_0)\| + \|\varphi(\tilde{t}, t_0, \tilde{x}_0) - \varphi(t, t_0, \tilde{x}_0)\| \\ &= \|\varphi(\tilde{t}, \tilde{t}_0, \tilde{x}_0) - \varphi(\tilde{t}, \tilde{t}_0, \varphi(\tilde{t}_0, t_0, \tilde{x}_0))\| + \|\varphi(\tilde{t}, t_0, \tilde{x}_0) - \varphi(t, t_0, \tilde{x}_0)\| \\ &= \|\varphi(\tilde{t}, \tilde{t}_0, \varphi(t_0, t_0, \tilde{x}_0)) - \varphi(\tilde{t}, \tilde{t}_0, \varphi(\tilde{t}_0, t_0, \tilde{x}_0))\| + \|\varphi(\tilde{t}, t_0, \tilde{x}_0) - \varphi(t, t_0, \tilde{x}_0)\| \\ &\leq C(J)\|\varphi(t_0, t_0, \tilde{x}_0) - \varphi(\tilde{t}_0, t_0, \tilde{x}_0)\| + \|\varphi(\tilde{t}, t_0, \tilde{x}_0) - \varphi(t, t_0, \tilde{x}_0)\| \\ &\leq C(J)\|\varphi(t_0, t_0, \tilde{x}_0) - \varphi(\tilde{t}_0, t_0, \tilde{x}_0)\| + \|\varphi(\tilde{t}, t_0, \tilde{x}_0) - \varphi(t, t_0, \tilde{x}_0)\| \\ &\leq M(x_0, J)(C(J)|t_0 - \tilde{t}_0| + |t - \tilde{t}|). \end{aligned}$$

Utilisant une dernière fois (I.12), on obtient le résultat suivant : pour tout  $\tilde{t}, \tilde{t}_0 \in J$  et pour tout  $\tilde{x}_0 \in B(x_0, 1)$ , on a l'estimation

$$\|\varphi(\tilde{t}, \tilde{t}_0, \tilde{x}_0) - \varphi(t, t_0, x)\| \leq C(J)\|\tilde{x}_0 - x_0\| + M(x_0, J)(C(J)|t_0 - \tilde{t}_0| + |t - \tilde{t}|),$$

ce qui montre bien le caractère localement lipschitzien de  $\varphi$ . ■

### III.3.c Dépendance par rapport à un paramètre

On suppose dans cette section que le champ de vecteur  $f$  dépend d'un paramètre  $\alpha \in \mathbb{R}^p$ . Plus précisément on suppose que

$$f : (t, x, \alpha) \in I \times \mathbb{R}^d \times \mathbb{R}^p \mapsto f(t, x, \alpha) \in \mathbb{R}^d,$$

est une application continue, globalement lipschitzienne par rapport à la variable d'état  $x$  et au paramètre  $\alpha$ .

On s'intéresse alors au système différentiel

$$x' = f(t, x, \alpha),$$

et à la dépendance des solutions par rapport au paramètre  $\alpha$ .



Une astuce pratique pour ramener cette étude aux cas précédents est de considérer une nouvelle variable d'état  $X = (x, \alpha) \in \mathbb{R}^{d+p}$  et le champ de vecteurs associé

$$F : (t, X = (x, \alpha)) \in I \times (\mathbb{R}^d \times \mathbb{R}^p) \mapsto (f(t, x, \alpha), 0) \in \mathbb{R}^d \times \mathbb{R}^p.$$

On vérifie aisément que  $F$  est un champ de vecteurs continu et globalement lipschitzien sur  $\mathbb{R}^{d+p}$  et qu'on peut donc lui appliquer le théorème de Cauchy-Lipschitz usuel et considérer le flot  $\psi : I \times I \times \mathbb{R}^{d+p} \rightarrow \mathbb{R}^{d+p}$ .

On décompose ce flot en deux parties

$$\psi(t, s, (x_0, \alpha_0)) = (\varphi(t, s, x_0, \alpha_0), \zeta(t, s, x_0, \alpha_0)) \in \mathbb{R}^d \times \mathbb{R}^p,$$

et on écrit l'équation vérifiée par  $\psi$

$$\partial_t \psi(t, s, (x_0, \alpha_0)) = F(t, \psi(t, s, (x_0, \alpha_0))),$$

ce qui donne, grâce à la définition de  $F$ , les deux équations

$$\partial_t \varphi(t, s, (x_0, \alpha_0)) = f(t, \varphi(t, s, (x_0, \alpha_0)), \zeta(t, s, (x_0, \alpha_0))),$$

$$\partial_t \zeta(t, s, (x_0, \alpha_0)) = 0,$$

avec les conditions initiales

$$\psi(s, s, (x_0, \alpha_0)) = x_0, \quad \zeta(s, s, (x_0, \alpha_0)) = \alpha_0.$$

L'équation sur  $\zeta$  montre que cette quantité ne dépend pas du temps et donc en particulier que

$$\forall t \in I, \quad \zeta(t, s, (x_0, \alpha_0)) = \alpha_0.$$

Si on reporte dans l'équation sur  $\varphi$ , on trouve

$$\partial_t \varphi(t, s, (x_0, \alpha_0)) = f(t, \varphi(t, s, (x_0, \alpha_0)), \alpha_0),$$

ce qui montre que  $\varphi(\cdot, \cdot, \cdot, \alpha_0)$  est bien le flot associé à l'équation  $x' = f(t, x, \alpha)$  pour la valeur  $\alpha = \alpha_0$ .

Comme  $\Phi$  est localement lipschitzien par rapport à toutes ses variables, il en est de même de  $\varphi$ . Autrement dit, on a obtenu le résultat suivant :

### Proposition I.37

*Le flot du système  $x' = f(t, x, \alpha)$  paramétré par  $\alpha$  et noté  $\varphi : I \times I \times \mathbb{R}^d \times \mathbb{R}^p \rightarrow \mathbb{R}^d$  est une application localement lipschitzienne par rapport à toutes ses variables.*

En pratique, on utilise souvent le résultat sous la forme suivante :

### Corollaire I.38

*On fixe  $t_0 \in I$ ,  $x_0 \in \mathbb{R}^d$  et un intervalle compact  $J \subset I$ . Si on note  $t \mapsto x_\alpha(t)$  la trajectoire associée au système  $x' = f(t, x, \alpha)$  pour la donnée de Cauchy  $(t_0, x_0)$ , alors l'application*

$$\alpha \in \mathbb{R}^p \mapsto x_\alpha \in \mathcal{C}^0(J, \mathbb{R}^d),$$

*est localement lipschitzienne, ce que l'on résume par la phrase « la solution du problème de Cauchy dépend de façon régulière du paramètre ».*

## III.3.d Le cas linéaire

Considérons un champ de vecteurs linéaire défini par

$$f(t, x) = A(t)x, \quad \forall t \in I, \forall x \in \mathbb{R}^d.$$

**Définition et Proposition I.39 (Résolvante)**

Dans le cas linéaire, pour tous  $t, s \in I$ , l'application  $\Phi(t, s)$  est linéaire, on l'identifie à une matrice notée  $R(t, s)$  et appelée la résolvante de l'équation.

On a donc

$$\varphi(t, s, x) = R(t, s).x, \quad \forall t, s \in I, \forall x \in \mathbb{R}^d.$$

Les propriétés du flot se traduisent sur la résolvante de la façon suivante :

- L'application  $(t, s) \in I \times I \mapsto R(t, s) \in M_d(\mathbb{R})$  est lipschitzienne.
- $R(t, t) = \text{Id}$  pour tout  $t \in I$ .
- $R(t_1, t_2).R(t_2, t_3) = R(t_1, t_3)$ , pour tous  $t_1, t_2, t_3 \in I$ .
- Pour tous  $t, s \in I$ ,  $R(t, s)$  est inversible et on a  $R(t, s) = R(s, t)^{-1}$ .

Enfin, dans le cas autonome ( $A$  ne dépend pas du temps), on a l'expression explicite

$$R(t, s) = e^{(t-s)A}, \quad \forall t, s \in I.$$

**Proposition I.40**

- Pour tout  $s \in I$ , l'application  $t \in I \mapsto R(t, s) \in M_d(\mathbb{R})$  est l'unique solution globale du problème de Cauchy (linéaire mais à valeurs dans  $M_d(\mathbb{R})$ ) suivant

$$\begin{cases} M'(t) = A(t)M(t), \\ M(s) = \text{Id}. \end{cases}$$

- Pour tout  $t \in I$ , l'application  $s \in I \mapsto R(t, s) \in M_d(\mathbb{R})$  est l'unique solution globale du problème de Cauchy suivant

$$\begin{cases} M'(s) = -M(s)A(s), \\ M(t) = \text{Id}. \end{cases}$$

**Proposition I.41 (Formule de Duhamel)**

L'unique solution du problème de Cauchy linéaire non-homogène suivant

$$\begin{cases} x'(t) = A(t)x(t) + b(t), \\ x(t_0) = x_0, \end{cases}$$

est donnée par la formule

$$x(t) = R(t, t_0)x_0 + \int_{t_0}^t R(t, s)b(s) ds.$$

**III.3.e Différentiabilité du flot**

Par définition, le flot  $\varphi$  d'un champ de vecteurs est différentiable par rapport à la variable de temps  $t$ , on a même le résultat suivant

$$t \mapsto (\varphi(t, \cdot, \cdot) \in \mathcal{C}^0(I \times \mathbb{R}^d, \mathbb{R}^d)),$$

est une application de classe  $\mathcal{C}^1$  dont la dérivée en un point  $t \in I$  est donnée par

$$f(t, \varphi(t, \cdot, \cdot)) \in \mathcal{C}^0(I \times \mathbb{R}^d, \mathbb{R}^d).$$

On s'intéresse maintenant à la différentiabilité du flot par rapport à la donnée initiale  $x_0$ .

### Théorème I.42

On suppose que  $f : I \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  est un champ de vecteurs globalement lipschitzien par rapport à la variable d'état et de classe  $C^1$  par rapport à cette même variable.

Alors, pour tout  $t, t_0 \in I$ , l'application  $\Phi(t, t_0)$  est de classe  $C^1$ . On note  $D_{x_0}(\Phi(t, t_0))(x_0)$  sa jacobienne en tout point  $x_0 \in \mathbb{R}^d$ .

Pour tout  $t_0$  et  $x_0$ , l'application  $t \mapsto D_{x_0}(\Phi(t, t_0))(x_0) \in M_d(\mathbb{R})$  est l'unique solution du problème de Cauchy linéaire matriciel suivant

$$\begin{cases} M'(t) = (D_x f)(t, \varphi(t, t_0, x_0)) \cdot M(t), \\ M(t_0) = \text{Id}. \end{cases} \quad (\text{I.14})$$

### Remarque I.43

Même si le champ de vecteurs  $f$  est autonome, le problème de Cauchy (I.14) satisfait par la jacobienne du flot est un problème non autonome.

### Preuve :

On remarque tout d'abord que les équations (I.14) s'obtiennent en différenciant formellement le problème de Cauchy initial et en supposant que l'on peut intervertir toutes les opérations de dérivation.

On va définir  $M$  comme l'unique solution de (I.14) et on va montrer que  $M$  est bien la jacobienne du flot par rapport à  $x_0$ .

Comme  $f$  est de classe  $C^1$ , on peut définir, pour tout  $s \in I, y \in \mathbb{R}^d$  et  $k \in \mathbb{R}^d$

$$R(s, y, k) := \frac{1}{\|k\|} (f(s, y + k) - f(s, y) - D_x f(s, y) \cdot k),$$

et, par définition, constater que

$$\forall s \in I, \forall y \in \mathbb{R}^d, \lim_{k \rightarrow 0} \|R(s, y, k)\| = 0. \quad (\text{I.15})$$

Du reste, par l'inégalité des accroissements finis, et la propriété de Lipschitz on a

$$\|R(s, y, k)\| \leq 2L(s), \quad \forall s \in I, \forall y \in \mathbb{R}^d, \forall k \in \mathbb{R}^d. \quad (\text{I.16})$$

Ecrivons les formulations intégrales suivantes

$$\begin{cases} \varphi(t, t_0, x_0 + h) = x_0 + h + \int_{t_0}^t f(s, \varphi(s, t_0, x_0 + h)) ds, \\ \varphi(t, t_0, x_0) = x_0 + \int_{t_0}^t f(s, \varphi(s, t_0, x_0)) ds, \\ M(t) = \text{Id} + \int_{t_0}^t D_x f(s, \varphi(s, t_0, x_0)) M(s) ds. \end{cases}$$

En les combinant on trouve

$$\begin{aligned} & \varphi(t, t_0, x_0 + h) - \varphi(t, t_0, x_0) - M(t) \cdot h \\ &= \int_{t_0}^t (f(s, \varphi(s, t_0, x_0 + h)) - f(s, \varphi(s, t_0, x_0)) - D_x f(s, \varphi(s, t_0, x_0)) M(s) \cdot h) ds, \\ &= \int_{t_0}^t (f(s, \varphi(s, t_0, x_0 + h)) - f(s, \varphi(s, t_0, x_0) + M(s) \cdot h)) ds \\ & \quad + \int_{t_0}^t (f(s, \varphi(s, t_0, x_0) + M(s) \cdot h) - f(s, \varphi(s, t_0, x_0)) - D_x f(s, \varphi(s, t_0, x_0)) M(s) \cdot h) ds. \end{aligned}$$

En utilisant le caractère lipschitzien du champ de vecteurs et la définition de  $R$ , on obtient

$$\begin{aligned} & \|\varphi(t, t_0, x_0 + h) - \varphi(t, t_0, x_0) - M(t) \cdot h\| \\ & \leq \int_{t_0}^t L(s) \|\varphi(s, t_0, x_0 + h) - \varphi(s, t_0, x_0) - M(s) \cdot h\| ds + \|h\| \int_{t_0}^t \|R(s, \varphi(s, t_0, x_0), M(s) \cdot h)\| \|M(s)\| ds. \end{aligned}$$

Le lemme de Gronwall nous donne alors

$$\|\varphi(t, t_0, x_0 + h) - \varphi(t, t_0, x_0) - M(t).h\| \leq \|h\| \left( \int_{t_0}^t \|R(s, \varphi(s, t_0, x_0), M(s).h)\| \|M(s)\| ds \right) \exp \left( \int_0^t L(s) ds \right),$$

et donc

$$\frac{1}{\|h\|} \|\varphi(t, t_0, x_0 + h) - \varphi(t, t_0, x_0) - M(t).h\| \leq C_t \left( \int_{t_0}^t \|R(s, \varphi(s, t_0, x_0), M(s).h)\| ds \right).$$

En utilisant (I.15) et (I.16), on peut utiliser le théorème de convergence dominée pour passer à la limite dans le terme intégral à droite de cette inégalité et ainsi conclure que

$$\frac{1}{\|h\|} \|\varphi(t, t_0, x_0 + h) - \varphi(t, t_0, x_0) - M(t).h\| \xrightarrow{h \rightarrow 0} 0.$$

En reprenant la mise en forme proposée dans la section III.3.c, on peut transcrire le résultat précédent en terme de dépendance  $C^1$  du flot par rapport à un paramètre  $\alpha \in \mathbb{R}^p$ .

## IV Théorème de Cauchy-Lipschitz local

Traisons maintenant le cas plus général où le champ de vecteurs est défini seulement sur un ensemble d'états admissibles  $\Omega$  et qu'il est seulement localement lipschitzien.

### IV.1 Enoncé et preuve du théorème principal

#### Théorème I.44 (Cauchy-Lipschitz local, version 1)

On suppose que le champ de vecteurs  $f : I \times \Omega \rightarrow \mathbb{R}^d$  est continu et **localement lipschitzien** par rapport à la variable d'état.

Alors, pour toute donnée de Cauchy  $(t_0, x_0) \in I \times \Omega$ , il **existe** un  $\delta > 0$  (dépendant des données!) tel qu'il existe une **unique** solution  $x$  du problème de Cauchy (I.1) définie sur  $J_\delta := ]t_0 - \delta, t_0 + \delta[ \cap I$ .

#### Preuve :

L'idée de la preuve est de se ramener au théorème de Cauchy-Lipschitz global par un argument de prolongement/restriction.

On commence par se donner un  $\varepsilon > 0$  tel que l'ensemble  $K_\varepsilon = ([t_0 - \varepsilon, t_0 + \varepsilon] \cap I) \times \bar{B}(x_0, \varepsilon)$  est contenu dans  $I \times \Omega$ .

De plus, on peut choisir  $\varepsilon$  assez petit pour que  $K_\varepsilon$  soit compact. Si  $t_0 \in \overset{\circ}{I}$ , il suffit de le choisir pour que  $[t_0 - \varepsilon, t_0 + \varepsilon] \subset I$ . Si  $t_0 \in \partial I$  alors on le choisit de sorte que  $[t_0 - \varepsilon, t_0 + \varepsilon] \cap I = [t_0, t_0 + \varepsilon]$  ou  $[t_0 - \varepsilon, t_0 + \varepsilon] \cap I = [t_0 - \varepsilon, t_0]$ .

On introduit l'ouvert de  $I \times \Omega$  défini par  $U_\varepsilon = ([t_0 - \varepsilon, t_0 + \varepsilon] \cap I) \times B(x_0, \varepsilon)$ .

D'après le corollaire I.19, il existe un champ de vecteurs continu  $f_\varepsilon : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ , globalement lipschitzien par rapport à la variable d'état et qui coïncide avec  $f$  sur  $K_\varepsilon$ . On peut donc lui appliquer le théorème de Cauchy-Lipschitz global et en déduire l'existence et l'unicité d'une solution globale  $(\mathbb{R}, x_\varepsilon)$  au problème de Cauchy suivant

$$\begin{cases} x'_\varepsilon(t) &= f_\varepsilon(t, x_\varepsilon(t)), \\ x_\varepsilon(t_0) &= x_0. \end{cases}$$

Comme  $x_\varepsilon(t_0) = x_0$  et que  $t \mapsto (t, x_\varepsilon(t))$  est une fonction continue, il existe  $0 < \delta < \varepsilon$  tel que

$$(t, x_\varepsilon(t)) \in U_\varepsilon, \quad \forall t \in I \cap ]t_0 - \delta, t_0 + \delta[.$$

Comme  $f_\varepsilon$  et  $f$  coïncident sur  $U_\varepsilon$ , on voit que  $(I \cap ]t_0 - \delta, t_0 + \delta[, x_\varepsilon)$  est bien une solution du problème de Cauchy initial. Par ailleurs, toute autre solution au problème de Cauchy initial va coïncider avec la solution  $x_\varepsilon$  sur un intervalle ouvert centré en 0. ■

#### Corollaire I.45 (Principe d'unicité)

Soient  $(J_1, x_1)$  et  $(J_2, x_2)$  deux solutions de l'équation  $x' = f(t, x)$  telles que  $J_1 \cap J_2 \neq \emptyset$ .

Les propositions suivantes sont équivalentes :

1. Il existe  $t^* \in J_1 \cap J_2$  tel que  $x_1(t^*) = x_2(t^*)$ .
2. Les deux fonctions  $x_1$  et  $x_2$  coïncident sur tout  $J_1 \cap J_2$ .

**Preuve :**

2.  $\Rightarrow$  1. Immédiat.

1.  $\Rightarrow$  2. On introduit l'ensemble

$$J := \{t \in J_1 \cap J_2, x_1(t) = x_2(t)\}.$$

—  $J$  est non vide car par hypothèse  $t^* \in J$ .

—  $J$  est fermé dans  $J_1 \cap J_2$  : car  $x_1$  et  $x_2$  sont continues.

—  $J$  est ouvert dans  $J_1 \cap J_2$  :

En effet, si on prend  $t_0 \in J$  alors  $(J_1, x_1)$  et  $(J_2, x_2)$  sont deux solutions du **même** problème de Cauchy posé sur  $J_1 \cap J_2$  :

$$\begin{cases} y'(t) &= f(t, y), \\ y(t_0) &= y_0. \end{cases} \quad (\text{I.17})$$

D'après le théorème de Cauchy-Lipschitz local, il existe  $\delta > 0$  t.q. (I.17) a une unique solution sur l'intervalle  $(J_1 \cap J_2) \cap ]t_0 - \delta, t_0 + \delta$ . En particulier,  $x_1 = x_2$  sur cet intervalle et donc

$$(J_1 \cap J_2) \cap ]t_0 - \delta, t_0 + \delta[ \subset J.$$

Par connexité de l'intervalle  $J_1 \cap J_2$ , nous en déduisons que

$$J = J_1 \cap J_2,$$

et donc que  $x_1$  et  $x_2$  coïncident sur  $J_1 \cap J_2$ . ■

**Théorème I.46 (Cauchy-Lipschitz local, version 2)**

*Sous les hypothèses du Théorème I.44, pour toute donnée de Cauchy il existe une unique solution maximale  $(J, x)$  au problème de Cauchy (I.1). De plus,  $J$  est un ouvert de  $I$ . Enfin, toute autre solution  $(\tilde{J}, \tilde{x})$  est la restriction de  $x$  sur  $\tilde{J}$ .*

**Preuve :**

On définit l'ensemble  $\mathcal{J}$  par

$$\mathcal{J} := \bigcup_{\substack{(J,x) \\ \text{solution de (I.1)}}} J.$$

D'après la première version du théorème de Cauchy-Lipschitz local (Théorème I.44), nous savons que  $\mathcal{J}$  est non vide.

Soit  $t \in \mathcal{J}$ . Considérons deux solutions  $(J_1, x_1)$  et  $(J_2, x_2)$  du problème de Cauchy (I.1) telles que  $t \in J_1$  et  $t \in J_2$ . Par définition, nous avons  $[t_0, t] \subset J_1 \cap J_2$  (ou  $[t, t_0]$  si  $t < t_0$  ...) et donc d'après le principe d'unicité donné par le Corollaire I.45, nous avons que  $x_1$  et  $x_2$  coïncident sur  $[t_0, t]$  (ou  $[t, t_0]$  ...) et donc en particulier  $x_1(t) = x_2(t)$ .

Par conséquent, nous pouvons définir à bon droit la quantité

$$y(t) = \text{la valeur en } t \text{ de n'importe quelle solution } (J, x) \text{ t.q. } t \in J.$$

Il est alors clair que  $(\mathcal{J}, y)$  est solution du problème de Cauchy (I.1) et même que c'est l'unique solution maximale du problème d'après la définition de  $\mathcal{J}$  !

Montrons maintenant que  $\mathcal{J}$  est ouvert dans  $I$ . Pour cela, on prend n'importe quel  $t^* \in \mathcal{J}$  et on définit  $y^* := y(t^*)$ . On considère alors le problème de Cauchy auxiliaire

$$\begin{cases} z'(t) &= f(t, z) \\ z(t^*) &= y^*, \end{cases} \quad (\text{I.18})$$

associé à l'instant initial  $t^*$ .

D'après la première version du théorème de Cauchy-Lipschitz local (Théorème I.44), il existe  $\delta > 0$  tel que le problème (I.18) admet une unique solution sur l'intervalle  $]t^* - \delta, t^* + \delta[ \cap I$ .

On peut alors poser  $J = \mathcal{J} \cup (]t^* - \delta, t^* + \delta[ \cap I)$  et définir

$$\tilde{y}(t) = \begin{cases} y(t), & \text{si } t \in \mathcal{J}, \\ z(t), & \text{si } t \in ]t^* - \delta, t^* + \delta[ \cap I. \end{cases}$$

Cette définition est cohérente par le principe d'unicité et on a bien construit une nouvelle solution  $(J, \tilde{y})$  du problème de Cauchy initial.

Par définition de  $\mathcal{J}$  nous avons  $J \subset \mathcal{J}$  ce qui prouve que

$$]t^* - \delta, t^* + \delta[ \cap I \subset \mathcal{J}.$$

On conclut bien que  $\mathcal{J}$  est ouvert dans  $I$ . ■

## IV.2 Critères de globalité

Le théorème de Cauchy-Lipschitz local (Théorème I.46) nous donne l'existence et l'unicité d'une solution maximale à un problème de Cauchy mais ne nous dit rien *a priori* sur l'intervalle  $J$  sur lequel celle-ci est définie. On peut par exemple se demander s'il est possible que  $J$  soit égal à  $I$  tout entier, autrement dit que la solution maximale soit globale. Bien entendu, ceci est vrai si le champ de vecteurs est globalement lipschitzien par rapport à la variable d'état (d'après le th. de Cauchy-Lipschitz global) mais en général, on a besoin de critères pour déterminer si oui ou non une solution donnée est, ou pas, globale.

On va commencer par un résultat dans le cas où l'espace d'états est l'espace entier  $\mathbb{R}^d$  puis on donnera le résultat plus général.

### Théorème I.47 (Théorème d'explosion en temps fini)

Soit  $f : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  un champ de vecteurs continu et localement lipschitzien par rapport à la variable d'état et  $(t_0, x_0)$  une donnée de Cauchy associée.

Soit  $(J, x)$  l'unique solution maximale du problème de Cauchy. On rappelle que  $J$  est nécessairement ouvert; on le note  $J = ]\alpha, \beta[$  avec  $\beta \in ]t_0, +\infty]$  et  $\alpha \in [-\infty, t_0[$ .

- Si  $\beta < +\infty$ , alors  $\sup_{t \in [t_0, \beta[} \|x(t)\| = +\infty$ .
- Si  $\alpha > -\infty$ , alors  $\sup_{t \in ]\alpha, t_0]} \|x(t)\| = +\infty$ .

### Remarque I.48

- Comme  $x$  est une application continue, elle est bornée sur tout intervalle compact et donc les conclusions ci-dessus montrent que  $x$  n'est pas bornée au voisinage de  $\beta$  (resp.  $\alpha$ ).
  - On peut en fait montrer une propriété un peu plus forte en remplaçant les conclusions par  $\lim_{t \rightarrow \beta} \|x(t)\| = +\infty$  (resp.  $\lim_{t \rightarrow \alpha} \|x(t)\| = +\infty$ ), c'est un bon exercice.
- La version ci-dessus est en général suffisante dans les applications.

### Preuve :

Supposons que  $\beta < +\infty$  et que  $R := \sup_{t \in [t_0, \beta[} \|x(t)\| < +\infty$ .

Par restriction et prolongement, on construit un champ de vecteurs  $f_R : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  continu et globalement lipschitzien par rapport à la variable d'état qui coïncide avec  $f$  sur  $[t_0, \beta + 1] \times \bar{B}(0, R + 1)$ . Le problème de Cauchy associé à  $f_R$  et à la donnée  $(t_0, x_0)$  admet une unique solution globale dont la restriction à  $J \cup \{\beta\}$  est solution du problème initial, ce qui contredit la maximalité de la solution  $(J, x)$ . ■

### Exercice I.1

Etudier (sans les calculer) l'ensemble des solutions de l'équation  $x' = x(1 - x)$ .

Une version plus générale du résultat précédent, dans le cas d'un champ de vecteurs seulement défini sur un ouvert de

$\mathbb{R}^d$ , est la suivante.

### **Théorème I.49 (de sortie de tout compact)**

Soit  $I$  un intervalle ouvert non vide,  $f : I \times \Omega \rightarrow \mathbb{R}^d$  un champ de vecteurs continu et localement lipschitzien par rapport à la variable d'état et  $(t_0, x_0)$  une donnée de Cauchy associée.

Soit  $(J, x)$  l'unique solution maximale du problème de Cauchy; on note  $J = ]\alpha, \beta[$  avec  $\beta > t_0$  et  $\alpha < t_0$ .

— Si  $\beta \in I$ , alors pour tout compact  $K \subset \Omega$ , on a

$$x([t_0, \beta]) \not\subset K.$$

— Si  $\alpha \in I$ , alors pour tout compact  $K \subset \Omega$ , on a

$$x([\alpha, t_0]) \not\subset K.$$

### **Exemple I.50**

L'équation scalaire suivante est une bonne illustration du phénomène de sortie de tout compact

$$x' = \frac{1}{x(1-x)}.$$

On peut faire les mêmes remarques que pour le résultat précédent.

### **Remarque I.51**

Plaçons nous dans le premier cas où  $\beta \in I$ , des remarques similaires pouvant être faites dans l'autre cas.

— Comme  $x$  est une application continue, on peut établir que pour tout  $t^* \in J$ , on a  $x([t^*, \beta]) \not\subset K$  donc  $x$  sort de  $K$  aussi proche que l'on soit de l'extrémité  $\beta$ .

— On peut en fait montrer qu'il existe  $t^* \in J$  tel que  $x([t^*, \beta]) \subset \Omega \setminus K$ , ce qui est bien sûr plus fort.

La version ci-dessus est en général suffisante dans les applications.

### **Preuve :**

Concentrons-nous sur le premier cas. On suppose que  $\beta \in I$  et que, pour un certain compact  $K \subset \Omega$ , on a  $x([t_0, \beta]) \subset K$ .

On construit un ouvert borné  $U$  tel que  $K \subset U$  et  $\overline{U} \subset \Omega$ <sup>5</sup>. De plus, comme  $\beta \in I$ , on peut trouver  $\delta > 0$  tel que  $[t_0, \beta + \delta] \subset I$ . Appliquons l'argument de restriction prolongement (Corollaire I.19) sur  $\mathcal{K} = [t_0, \beta + \delta] \times \overline{U}$ , ce qui nous donne un champ de vecteurs  $\tilde{f}$  continu et globalement lipschitzien sur  $\mathbb{R} \times \mathbb{R}^d$  qui coïncide avec  $f$  sur  $\mathcal{K}$ .

Il existe donc une unique solution globale  $\tilde{x} : \mathbb{R} \rightarrow \mathbb{R}^d$  au problème de Cauchy

$$\begin{cases} \tilde{x}'(t) = \tilde{f}(t, \tilde{x}(t)), & \forall t \in \mathbb{R}, \\ \tilde{x}(t_0) = x_0, \end{cases} \quad (\text{I.19})$$

— Comme  $\tilde{f}$  et  $f$  coïncident sur  $[t_0, \beta] \times K$  et que  $x([t_0, \beta]) \subset K$ , on voit que  $([t_0, \beta], x)$  est aussi solution de (I.19). Par unicité, on en déduit que  $x = \tilde{x}$  sur  $[t_0, \beta]$ .

— La fonction  $\tilde{x}$  étant continue, il existe  $\varepsilon > 0$  (que l'on choisit de sorte que  $\varepsilon \leq \delta$ ) tel que  $\tilde{x}([t_0, \beta + \varepsilon]) \subset U$ . Comme  $f$  et  $\tilde{f}$  coïncident sur  $[t_0, \beta + \varepsilon] \times U$ , on constate que  $\tilde{x}$  vérifie

$$\tilde{x}'(t) = \tilde{f}(t, \tilde{x}(t)) = f(t, \tilde{x}(t)), \quad \forall t \in [t_0, \beta + \varepsilon].$$

et donc  $([t_0, \beta + \varepsilon], \tilde{x})$  est une solution du problème de Cauchy initial (qui coïncide sur  $[t_0, \beta]$  avec la solution  $x$  de départ.

— La fonction  $\bar{x}$  définie par

$$\bar{x}(t) = \begin{cases} x(t), & \forall t \in ]\alpha, t_0], \\ \tilde{x}(t), & \forall t \in [t_0, \beta + \varepsilon], \end{cases}$$

est alors une solution du problème de Cauchy initial définie sur un intervalle strictement plus grand que  $J = ]\alpha, \beta[$ , ce qui contredit la maximalité de  $(J, x)$ .

5. Savez-vous construire un tel ouvert ?

## V Équilibres. Stabilité. Stabilité asymptotique

On considère dans cette partie un champ de vecteurs **autonome**  $f : \Omega \rightarrow \mathbb{R}^d$  de classe  $C^1$  (donc localement lipschitzien)<sup>6</sup>. L'instant initial pour la donnée de Cauchy n'a donc aucune importance et sera arbitrairement choisie comme étant  $t_0 = 0$ .

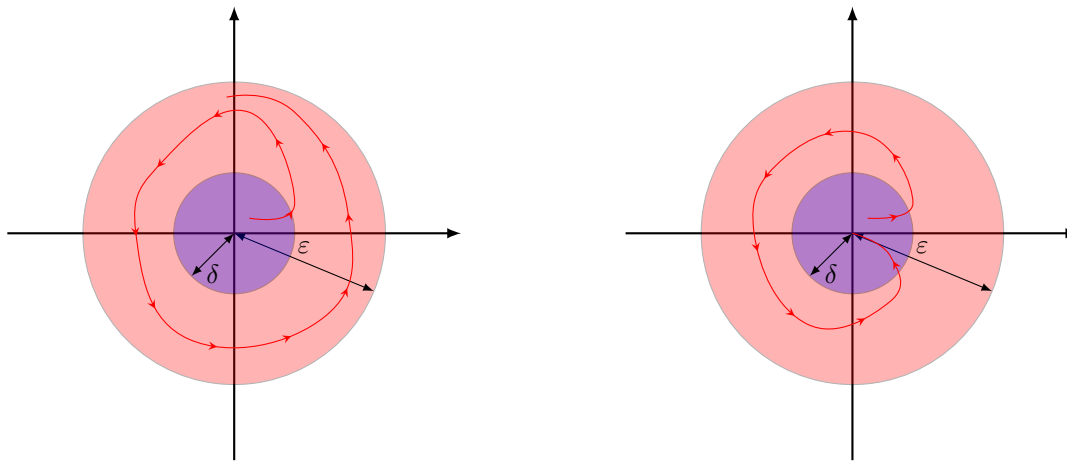
### Définition I.52 (Équilibres)

On appelle **équilibre du système** (on parle aussi d'un point critique du champ de vecteurs  $f$ ) tout point  $x^* \in \Omega$  tel que

$$f(x^*) = 0.$$

Ceci implique que la fonction constante  $t \in \mathbb{R} \mapsto x^*$  est une solution particulière de l'équation différentielle  $x' = f(x)$ .

On s'intéresse dans la suite aux propriétés de stabilité (sous-entendu quand  $t \rightarrow +\infty$ ) de ces solutions, ce qui se formalise de la façon suivante.



(a) L'origine est stable

(b) L'origine est asymptotiquement stable

FIGURE I.6 – Notions de stabilité d'un point d'équilibre

### Définition I.53

Soit  $x^*$  un équilibre du système différentiel  $x' = f(x)$ .

- On dit que  $x^*$  est un équilibre **stable** si : pour tout  $\varepsilon > 0$ , il existe  $\delta > 0$  tel que, pour toute donnée initiale  $x_0$  proche de  $x^*$  à  $\delta$  près, la solution du système partant de  $x_0$  est d'une part définie sur tout  $[0, +\infty[$  et d'autre part reste proche de  $x^*$  à  $\varepsilon$  près au cours du temps. Voir la figure I.6. En utilisant la notion de flot, cette propriété s'écrit

$$\forall \varepsilon > 0, \exists \delta > 0, \Phi(t)(B(x^*, \delta)) \subset B(x^*, \varepsilon), \quad \forall t \geq 0.$$

- On dit que  $x^*$  est un équilibre **asymptotiquement stable** s'il est stable et si de plus, pour toute donnée initiale  $x_0$  suffisamment proche de  $x^*$ , la solution associée converge vers  $x^*$  quand  $t$  tend vers l'infini.

En utilisant la notion de flot, ceci s'exprime de la façon suivante : il existe  $\delta > 0$  tel que

$$\forall x_0 \in B(x^*, \delta), \quad \varphi(t, x_0) \xrightarrow[t \rightarrow \infty]{} x^*.$$

6. On peut raffiner les hypothèses de régularité nécessaire mais ce n'est pas ce qui nous intéresse ici



Il faut prendre garde que dans la définition de la stabilité le rayon  $\delta$  peut être plus petit que le rayon  $\varepsilon$  : en particulier on ne dit pas que les boules sont invariantes par le flot ! Ceci est illustré sur la figure I.7.

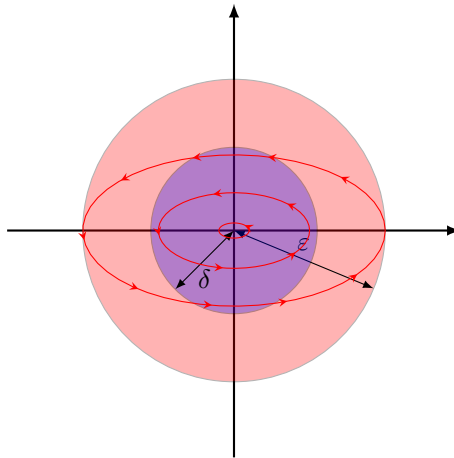


FIGURE I.7 – Un exemple de système stable mais non asymptotiquement stable

## V.1 Cas linéaire

Commençons par comprendre le cas linéaire  $x' = Ax$  avec  $A \in M_d(\mathbb{R})$ .

### V.1.a Le cas plan

Pour l'étude qui suivra il est important de bien comprendre la géométrie des solutions dans le cas  $d = 2$ . Trois cas peuvent se produire (voir par exemple [5, page 290] pour une description plus détaillée) :

— Premier cas :  $A$  est diagonalisable dans  $\mathbb{R}$ .

Si  $\lambda_1$  et  $\lambda_2$  sont les deux valeurs propres et  $e_1, e_2$  les vecteurs propres associés, les solutions de l'équation sont données par

$$y(t) = \alpha_1 e^{\lambda_1 t} e_1 + \alpha_2 e^{\lambda_2 t} e_2.$$

Selon le signe des valeurs propres on peut donc tracer les trajectoires, voir Figure I.8.

On voit que l'origine est stable (resp. asymptotiquement stable) si et seulement si les deux valeurs propres sont négatives (resp. strictement négatives).

— Deuxième cas :  $A$  a ses valeurs propres réelles mais n'est pas diagonalisable.

Dans ce cas, la valeur propre est nécessairement unique et non semi-simple. D'après le théorème de Jordan (ou la décomposition de Dunford, voir ci-dessous la proposition I.55), la matrice  $A$  s'écrit

$$P^{-1}AP = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

et dans ce cas, on peut voir que l'exponentielle s'écrit

$$P^{-1}e^{tA}P = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}.$$

Si on note  $e_1$  et  $e_2$  les deux colonnes de  $P$  (i.e. un vecteur propre de  $A$  et un vecteur propre généralisé), on trouve

$$y(t) = (\alpha_1 + \alpha_2 t)e^{\lambda t} e_1 + \alpha_2 e^{\lambda t} e_2.$$

La stabilité (et la stabilité asymptotique) du point d'équilibre est alors équivalente à la stricte négativité de la valeur propre.

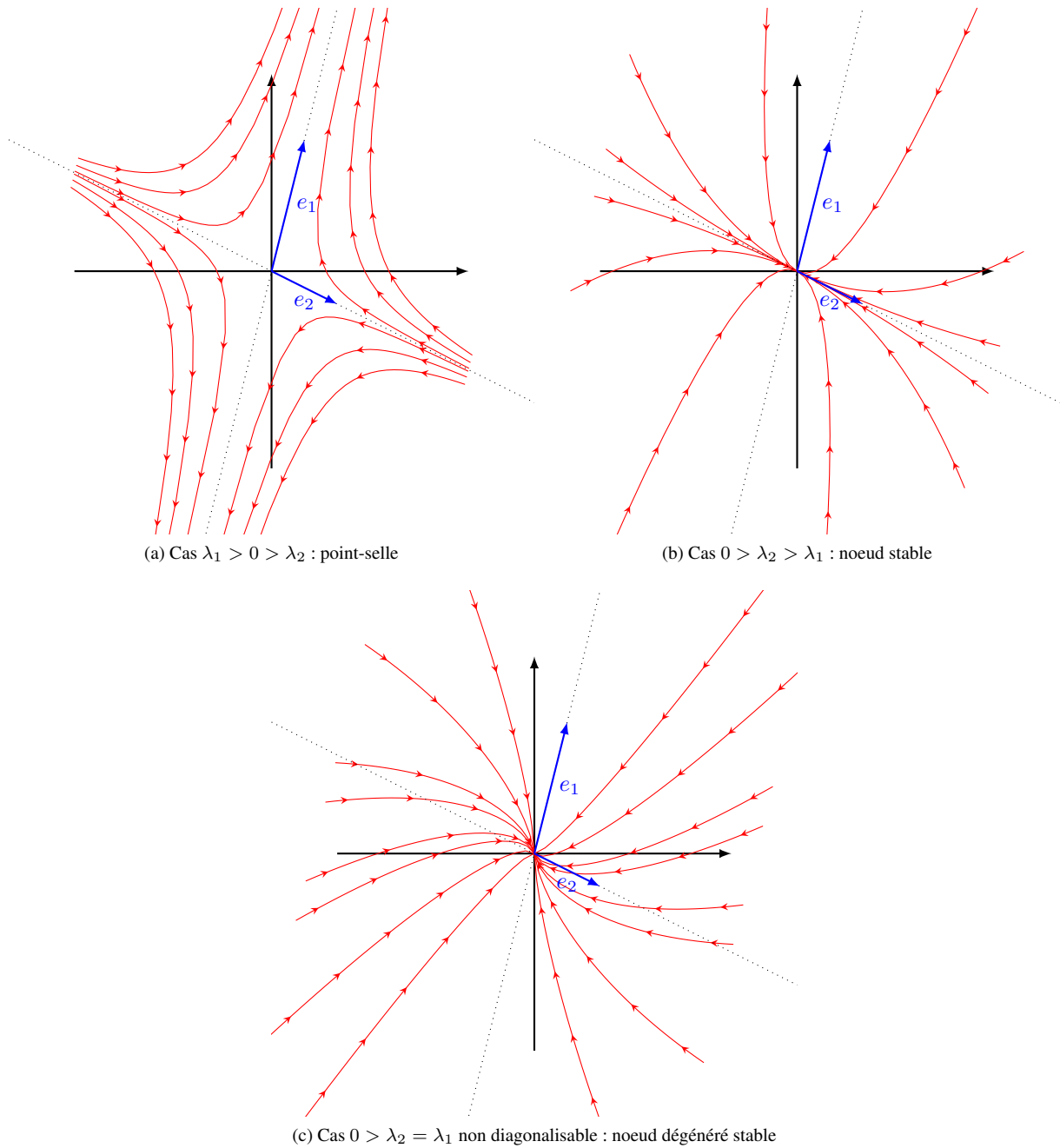


FIGURE I.8 – Trois portraits de phase typiques dans le cas de deux valeurs propres réelles

— Troisième cas :  $A$  a deux valeurs propres complexes conjuguées distinctes  $\lambda_{\pm} = a \pm bi$ , avec  $b \neq 0$ .

Dans ce cas, on peut vérifier que l'exponentielle de  $tA$  est donnée par une formule du genre

$$P^{-1}e^{tA}P = e^{ta} \begin{pmatrix} \cos(bt) & \sin(bt) \\ -\sin(bt) & \cos(bt) \end{pmatrix}.$$

La stabilité du point d'équilibre est alors donnée par la valeur de  $a$  : si  $a = 0$  les trajectoires sont périodiques, si  $a > 0$  le point d'équilibre est instable et si  $a < 0$  le point est stable. Qualitativement, les trajectoires s'enroulent autour de l'origine. Le sens de rotation est donné par l'analyse des vecteurs propres de  $A$ , voir Figure I.9.

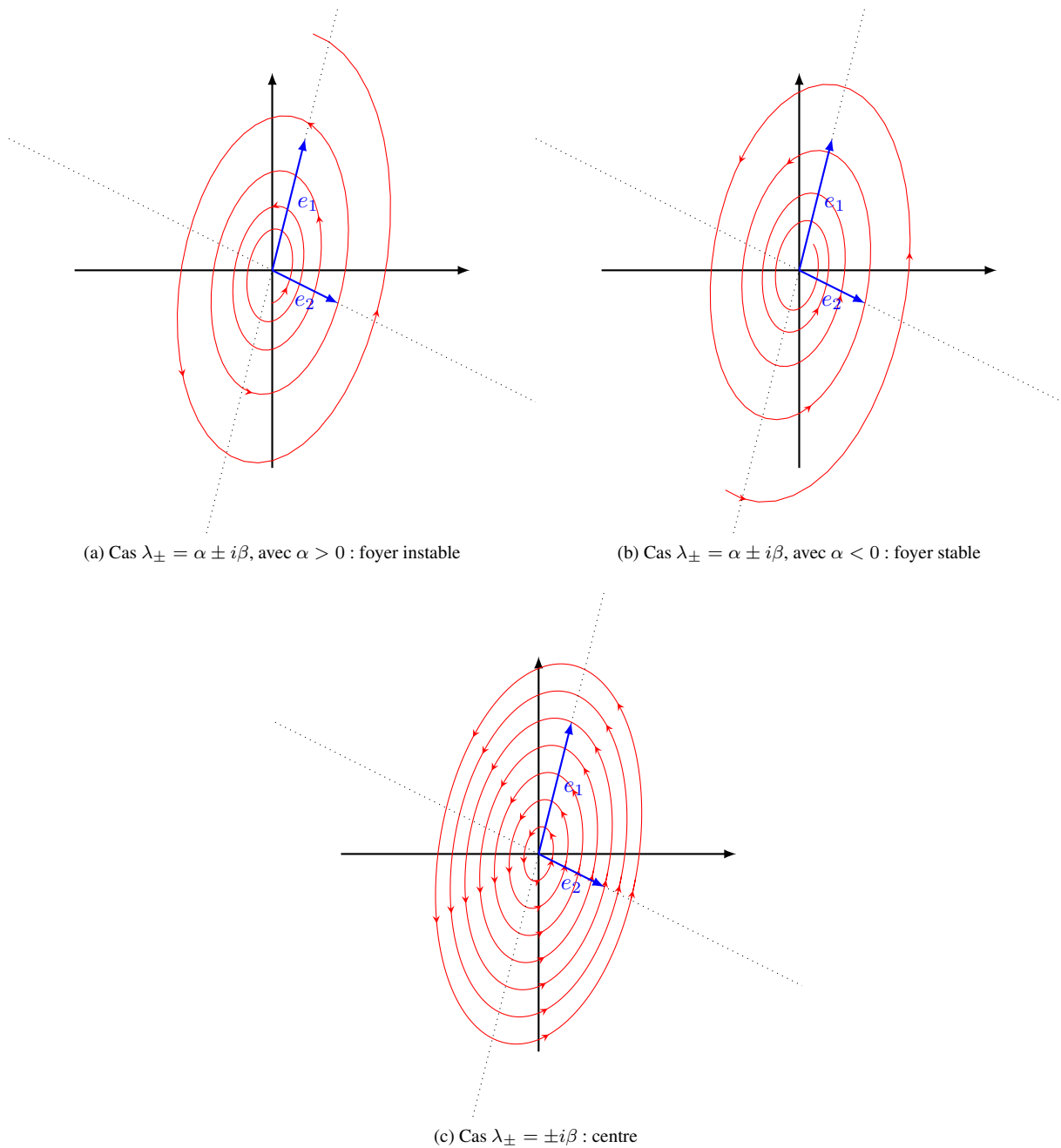


FIGURE I.9 – Portraits de phase typiques dans le cas de deux valeurs propres complexes conjuguées (le sens de rotation dépend du signe de  $\beta$  et de la position relative des vecteurs propres associés à  $\lambda^+$  et  $\lambda^-$ )

### V.1.b Stabilité de l'origine pour les systèmes d'EDO linéaires autonomes

On peut maintenant énoncer le théorème général. Avant cela on rappelle une définition d'algèbre linéaire et une propriété utile.

#### Définition I.54

Soit  $A \in M_n(\mathbb{C})$  une matrice quelconque et  $\lambda \in \text{Sp}(A)$ . On dit que  $\lambda$  est semi-simple si c'est une racine simple du polynôme minimal de  $A$ , ou encore (de manière équivalente) si  $\text{Ker}(A - \lambda I) = \text{Ker}(A - \lambda I)^2$ .

#### Proposition I.55 (Décomposition de Dunford)

Toute matrice  $A \in M_n(\mathbb{C})$  s'écrit de manière unique sous la forme

$$A = D + N,$$

avec  $D$  diagonalisable dans  $\mathbb{C}$ ,  $N$  nilpotente et  $DN = ND$ .

De plus,  $A$  et  $D$  ont les mêmes valeurs propres et on a la caractérisation suivante :

$$\lambda \text{ est une v.p. semi-simple de } A \iff \text{Ker}(D - \lambda) \subset \text{Ker } N,$$

c'est-à-dire que  $N$  est nulle sur les sous-espaces propres de  $D$  pour la valeur propre en question.

#### Preuve :

Pour l'existence et l'unicité de la décomposition, on renvoie à un cours standard d'algèbre linéaire <sup>7</sup>.

On montre juste la caractérisation des valeurs propres semi-simples.

- Soit  $\lambda$  une valeur propre non semi-simple. Il existe alors un  $v \in \mathbb{C}^n$ , tel que  $(A - \lambda)^2 v = 0$  et  $(A - \lambda)v \neq 0$ . Par la formule de Newton (valable car  $D$  et  $N$  commutent) nous avons

$$(D - \lambda)^{n+1} = \sum_{k=0}^{n+1} C_{n+1}^k (-N)^k (A - \lambda)^{n+1-k} = \sum_{k=0}^{n-1} C_{n+1}^k (-N)^k (A - \lambda)^{n+1-k},$$

où on a utilisé le caractère nilpotent de  $N$  pour éliminer les termes d'indices  $k \geq n$  dans la somme. On voit que la puissance sur  $(A - \lambda)$  dans formule est au moins égale à 2 et donc en appliquant cette égalité au vecteur  $v$  on trouve

$$(D - \lambda)^{n+1} v = 0.$$

Comme  $D$  est diagonalisable, cela implique que  $v$  est un vecteur propre de  $D$  pour la valeur propre  $\lambda$ . En écrivant que  $(A - \lambda) = (D - \lambda) + N$  et en appliquant  $v$  nous obtenons que

$$Nv = (A - \lambda)v \neq 0, \quad \text{par hypothèse.}$$

On a donc trouvé un élément de  $\text{Ker}(D - \lambda)$  qui n'est pas dans  $\text{Ker } N$ .

- Si maintenant  $\lambda$  est une valeur propre semi-simple, on commence par constater que l'hypothèse implique par une récurrence immédiate

$$\text{Ker}(A - \lambda)^k = \text{Ker}(A - \lambda)^{k+1}, \quad \forall k \geq 0,$$

et donc

$$\text{Ker}(A - \lambda)^k = \text{Ker}(A - \lambda), \quad \forall k \geq 0. \quad (\text{I.20})$$

On utilise à nouveau la formule de Leibniz pour écrire

$$(A - \lambda)^n = \sum_{k=0}^n C_n^k N^k (D - \lambda)^{n-k} = \sum_{k=0}^{n-1} C_n^k N^k (D - \lambda)^{n-k},$$

où on a à nouveau utilisé le caractère nilpotent de  $N$ . On observe que dans cette formule les puissances de  $(D - \lambda)$  sont toutes strictement positives et donc si on prend  $v \in \text{Ker}(D - \lambda)$ , la formule précédente montre que

$$(A - \lambda)^{n+1} v = 0,$$

et donc par (I.20), on conclut que  $(A - \lambda)v = 0$ .

L'égalité  $A - \lambda = (D - \lambda) + N$  permet de conclure que  $Nv = 0$  et le résultat est prouvé.

7. voir aussi les feuilles de TD!

**Théorème I.56**

On considère le système différentiel linéaire autonome  $x' = Ax$ .

- L'état d'équilibre  $x^* = 0$  est asymptotiquement stable si et seulement si toutes les valeurs propres de  $A$  sont de partie réelle strictement négative.

Dans ce cas, l'équilibre est exponentiellement stable : il existe  $C, \gamma > 0$  tel que

$$\|e^{tA}\| \leq Ce^{-\gamma t}, \quad \forall t > 0.$$

- L'état d'équilibre  $x^* = 0$  est stable si et seulement si toutes les valeurs propres de  $A$  sont de partie réelle négative ou nulle et que celles de partie réelle nulle sont semi-simples.

Dans ce cas, le flot est uniformément borné (en temps positif) : il existe  $C > 0$  tel que

$$\|e^{tA}\| \leq C, \quad \forall t > 0.$$

Notons que le même résultat est valable pour tout autre équilibre  $x^*$  (c'est-à-dire pour tout  $x^*$  dans le noyau de  $A$ ).

**Preuve :**

- Supposons tout d'abord que  $A$  admet une valeur propre  $\lambda = \alpha + i\beta$  de partie réelle strictement positive. Montrons alors que le point d'équilibre ne peut pas être stable.

Soit  $v = v_1 + iv_2$  un vecteur propre de  $A$  pour cette valeur propre. On suppose que  $v_1 \neq 0$  (ce qu'on peut toujours faire quitte à multiplier  $v$  par  $i$ ).

On choisit un nombre  $L$  tel que  $\beta L \in 2\pi\mathbb{Z}$  et on constate que, si on pose  $t_n = nL$ , alors on a

$$e^{t_n A} v = e^{t_n(\alpha + i\beta)} v = e^{t_n \alpha} e^{inL\beta} v = e^{t_n \alpha} v.$$

Comme la matrice  $A$  est réelle, on peut prendre la partie réelle de cette égalité et obtenir

$$e^{t_n A} v_1 = e^{t_n \alpha} v_1,$$

ce qui prouve, vu que  $\alpha > 0$ , que  $(e^{t_n A} v_1)_n$  n'est pas bornée et donc l'instabilité de l'origine pour ce système.

- Le même calcul montre que si  $A$  admet une valeur propre imaginaire pure  $i\beta$ , alors il existe  $v_1 \neq 0$  et une suite  $(t_n)_n$  tendant vers l'infini tels que

$$e^{t_n A} v_1 = v_1,$$

et donc l'équilibre ne saurait être asymptotiquement stable.

- Supposons maintenant que  $A$  admette une valeur propre  $\lambda = i\beta$  imaginaire pure et non semi-simple. On va également montrer que le point d'équilibre n'est pas stable (mais on remarquera que l'instabilité est moins violente que dans le cas précédent).

D'après la Proposition I.55, il existe un vecteur propre  $v$  de  $D$  pour la valeur propre  $\lambda$  tel que  $Nv \neq 0$ .

Calculons la solution de l'équation pour la donnée initiale  $v$  :

$$e^{tA} v = e^{tN} e^{tD} v = e^{t\lambda} e^{tN} v = e^{it\beta} \sum_{k=0}^n C_n^k t^k N^k v.$$

Comme le facteur  $e^{it\beta}$  est de module 1, on voit que  $e^{tA} v$  est borné si et seulement si le polynôme en  $t$

$$\sum_{k=0}^n C_n^k t^k N^k v,$$

est lui-même borné, ce qui n'est pas possible car son terme d'ordre 1 n'est pas nul vu que  $Nv \neq 0$ .

- Etudions maintenant les cas de stabilité. On commence par diagonaliser  $D$  (ce qui est possible par définition de la décomposition de Dunford). Il existe donc  $P \in GL_d(\mathbb{C})$  telle que

$$P^{-1}DP = \Lambda, \text{ avec } \Lambda = \text{diag}(\lambda_1, \dots, \lambda_d),$$

et comme  $\tilde{N} = P^{-1}NP$  est aussi nilpotente (et commute avec  $\Lambda$ ) on voit que dans ces conditions, on peut simplement calculer

$$e^{tA} = P e^{t\Lambda} e^{t\tilde{N}} P^{-1}.$$

La matrice de passage  $P$  ne joue pas de rôle dans la stabilité (asymptotique) du système et donc on est ramenés à étudier la stabilité du flot suivant

$$t \mapsto e^{t\Lambda} e^{t\tilde{N}}.$$

Sachant que

$$e^{t\Lambda} = \text{diag}(e^{t\lambda_1}, \dots, e^{t\lambda_d}),$$

et que

$$|e^{t\lambda_i}| = e^{t\mathcal{R}e\lambda_i},$$

on peut mener la discussion suivante :

- Si toutes les valeurs propres sont de parties réelles strictement négatives, on a

$$\|e^{t\Lambda}\| \leq e^{-\gamma^* t},$$

avec

$$\gamma^* := \inf\{(-\mathcal{R}e\lambda_i), 1 \leq i \leq d\},$$

qui est une quantité strictement positive par hypothèse.

Comme  $\tilde{N}$  est nilpotente,  $e^{t\tilde{N}}$  est une quantité polynomiale en temps et donc nous avons

$$\|e^{tA}\| \leq \|P\| \|P^{-1}\| \|e^{t\tilde{N}}\| e^{-\gamma^* t},$$

et donc, comme l'exponentielle l'emporte sur les termes polynomiaux, nous obtenons que, pour tout  $0 < \gamma < \gamma^*$ , il existe  $C_\gamma > 0$  telle que

$$\|e^{tA}\| \leq C_\gamma e^{-\gamma t}, \quad \forall t \geq 0,$$

ce qui implique la stabilité asymptotique de l'équilibre.

- S'il existe des valeurs propres de partie réelle nulle, on ne peut plus raisonner comme précédemment car  $\|e^{t\Lambda}\| = 1$  et les termes polynomiaux en temps sont potentiellement désagréables.

On sépare alors l'espace  $\mathbb{C}^n$  en deux parties  $\mathbb{C}^n = E_- \oplus E_0$  où  $E_-$  est la somme des sous-espaces propres de  $\Lambda$  associés aux valeurs propres de partie réelle strictement négative et  $E_0$  la somme des sous-espaces associés aux valeurs propres imaginaires pures. Comme  $\tilde{N}$  commute avec  $\Lambda$ , les espaces  $E_-$  et  $E_0$  sont stables par  $\tilde{N}$ . Par ailleurs, l'hypothèse de semi-simplicité assure que  $\tilde{N}(E_0) = 0$ .

On peut alors conclure en étudiant séparément la situation sur  $E_-$  et  $E_0$ .

■

## V.2 Cas nonlinéaire

### V.2.a Le théorème à connaître

#### Théorème I.57 (Critère spectral de stabilité)

Soit  $x^* \in \mathbb{R}^d$  tel que  $f(x^*) = 0$ . On note  $A = D_x f(x^*)$  la matrice jacobienne de  $f$  en  $x^*$ .

1. Si toutes les valeurs propres de  $A$  sont de partie réelle **strictement négative**, alors  $x^*$  est un équilibre asymptotiquement stable de l'équation  $x' = f(x)$ .
2. Si toutes les valeurs propres de  $A$  sont de partie réelle **strictement positive**, alors  $x^*$  est un équilibre instable de l'équation  $x' = f(x)$ .
3. Si  $A$  a toutes ses valeurs propres de partie réelle négative ou nulle et qu'au moins l'une d'entre elles est de partie réelle nulle, alors on ne peut pas conclure.

#### Preuve :

On commence par supposer que  $x^* = 0$  sans perte de généralité (on s'y ramène en changeant  $f$  en  $f - f(x^*)$ ).

1. D'après l'hypothèse sur les valeurs propres de  $A$  et les calculs qu'on a effectués dans le cas linéaire, nous savons qu'il existe  $\gamma > 0$  et  $C_1 > 0$  tels que

$$\|e^{tA}\| \leq C_1 e^{-\gamma t}, \quad \forall t \geq 0.$$

On pose maintenant  $g(x) := f(x) - Ax$ . Par définition de la jacobienne en 0, nous savons que  $\lim_{x \rightarrow 0} g(x)/\|x\| = 0$ . En particulier, il existe  $\delta > 0$  tel que

$$\|g(x)\| \leq \frac{\gamma}{2C_1} \|x\|, \quad \forall x \in \bar{B}(0, \delta).$$

On pose alors

$$g_\delta(x) := g(P_\delta x),$$

où  $P_\delta$  est la projection orthogonale sur  $\bar{B}(0, \delta)$ . Il est clair que  $g_\delta$  est globalement lipschitzienne et vérifie

$$\|g_\delta(x)\| \leq \frac{\gamma}{2C_1} \|x\|, \quad \forall x \in \mathbb{R}^d.$$

On pose alors  $f_\delta(x) := Ax + g_\delta(x)$  qui est une fonction globalement lipschitzienne sur  $\mathbb{R}^d$  et qui coïncide avec  $f$  sur  $\bar{B}(0, \delta)$ .

Soit maintenant  $x_\delta$  la solution du système  $x' = f_\delta(x)$  pour une donnée initiale  $x_0$  fixée. On écrit l'équation sous la forme

$$x'_\delta = Ax_\delta + g_\delta(x_\delta),$$

et on utilise alors la formule de Duhamel pour écrire

$$x_\delta(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A}g_\delta(x_\delta(s)) ds,$$

et en prenant la norme, pour  $t \geq 0$ ,

$$\|x_\delta(t)\| \leq C_1 e^{-\gamma t} \|x_0\| + C_1 \frac{\gamma}{2C_1} \int_0^t e^{-\gamma(t-s)} \|x_\delta(s)\| ds.$$

Réécrivons cette inégalité comme suit

$$e^{\gamma t} \|x_\delta(t)\| \leq C_1 \|x_0\| + \frac{\gamma}{2} \int_0^t e^{\gamma s} \|x_\delta(s)\| ds, \quad \forall t \geq 0,$$

ce qui permet d'appliquer le lemme de Gronwall à la fonction  $t \mapsto e^{\gamma t} \|x_\delta(t)\|$  et ainsi d'obtenir l'inégalité

$$e^{\gamma t} \|x_\delta(t)\| \leq C_1 \|x_0\| e^{\gamma t/2}, \quad \forall t \geq 0,$$

dont on déduit immédiatement

$$\|x_\delta(t)\| \leq C_1 \|x_0\| e^{-\gamma t/2}.$$

— La première conséquence de cette inégalité est que

$$\|x_\delta(t)\| \leq C_1 \|x_0\|, \quad \forall t \geq 0,$$

et donc que si on choisit des données initiales  $x_0$  vérifiant

$$\|x_0\| \leq \frac{\delta}{C_1},$$

alors la solution  $x_\delta$  du système perturbé vérifie

$$\|x_\delta(t)\| \leq \delta, \quad \forall t \geq 0,$$

et donc, comme  $f_\delta$  et  $f$  coïncident sur la boule centrée en zéro et de rayon  $\delta$ , nous avons

$$x'_\delta(t) = f_\delta(x_\delta(t)) = f(x_\delta(t)).$$

Ainsi  $t \in [0, \infty[ \mapsto x_\delta(t)$  n'est autre que l'unique solution maximale en temps positif du problème de Cauchy initial  $x' = f(x)$ . Ce qui prouve l'existence globale de ces solutions.

— Toujours pour des données initiales vérifiant  $\|x_0\| \leq \frac{\delta}{C_1}$ , nous avons finalement obtenu

$$\|x(t)\| \leq C_1 \|x_0\| e^{-\gamma t/2},$$

et donc le fait que  $\lim_{t \rightarrow \infty} x(t) = 0$ , ce qui prouve bien la stabilité asymptotique de l'équilibre 0 pour le système  $x' = f(x)$ .

On a montré en fait une propriété plus forte (la stabilité exponentielle) qui montre que le flot  $\varphi$  de l'équation vérifie, pour tout  $r > 0$  assez petit,

$$\varphi(t, B(0, r)) \subset e^{-\gamma t/2} B(0, rC_1).$$

2. On note  $(t, x) \mapsto \varphi(t, x) = \varphi(t, 0, x)$  le flot du système qui ne dépend que d'une seule variable de temps car le système est autonome (voir la remarque I.36). On observe immédiatement que  $\psi(t, x) = \varphi(-t, x)$  est le flot du système (en temps rétrograde)

$$y' = -f(y),$$

qui possède le même équilibre  $x^*$ . Par ailleurs on sait que  $\psi(t, \cdot) = \varphi(t, \cdot)^{-1}$ .

Comme la jacobienne de  $-f$  en  $x^*$  est égale à  $-A$  et que celle-ci a donc des valeurs propres de partie réelle strictement négative, nous savons que  $x^*$  est exponentiellement stable pour ce nouveau système et donc que, pour tout  $r > 0$  assez petit on a

$$\psi(t, B(0, r)) \subset e^{-\gamma t/2} B(0, rC_1) = B(0, rC_1 e^{-\gamma t/2}).$$

En appliquant  $\varphi(t, \cdot)$  à cette formule on trouve

$$B(0, r) \subset \varphi\left(t, B\left(0, rC_1 e^{-\gamma t/2}\right)\right), \quad \forall t \geq 0.$$

On fixe maintenant un  $r > 0$  pour lequel cette propriété est vraie et on se donne un  $\delta > 0$  quelconque. En posant

$$t_\delta = \frac{2}{\gamma} \log\left(\frac{rC_1}{\delta}\right),$$

on voit que l'inclusion précédente donne

$$B(0, r) \subset \varphi(t_\delta, B(0, \delta)).$$

En particulier, pour tout  $x$  tel que  $\|x\| = r/2$ , il existe  $x_0$  tel que  $\|x_0\| < \delta$  et tel que

$$\varphi(t_\delta, x_0) = x,$$

et donc en particulier

$$\|\varphi(t_\delta, x_0)\| = r/2,$$

Ceci prouve que, aussi petit que soit  $\delta$ , on finit toujours par sortir de la boule de rayon  $r/2$  et donc le point d'équilibre ne saurait être stable.

On remarque bien sûr que  $t_\delta \rightarrow +\infty$  quand  $\delta \rightarrow 0$ , autrement dit : plus  $\delta > 0$  est petit, plus on mettra du temps à sortir de la boule de rayon  $r/2$ , ce qui est naturel.

3. Un exemple très simple est le problème scalaire suivant :

$$x' = \alpha x^3,$$

avec  $\alpha \in \{-1, 0, 1\}$ . Noter que cette équation peut se résoudre explicitement sans difficulté<sup>8</sup> mais que nous allons utiliser dans ce qui suit des arguments un peu plus génériques.

La "jacobienne" de  $f$  est la dérivée  $f'(0) = 0$  qui est bien de partie réelle nulle.

- Pour  $\alpha = 0$ , le système se réduit à  $x' = 0$  et la stabilité de l'équilibre est claire.
- Pour  $\alpha = 1$ , on a  $x' = x^3$ . Si  $x_0 > 0$ ,  $x$  reste positive et donc  $x$  est strictement croissante. Tant que la solution existe, on a donc

$$x'(t) \geq x_0^3, \quad \forall t \geq 0, t \in J,$$

et ainsi

$$x(t) \geq x_0 + x_0^3 t, \quad \forall t \geq 0, t \in J.$$

Pour tout choix de  $x_0 > 0$ , même très petit, nous avons  $\lim_{t \rightarrow \sup J} x(t) = +\infty$  et donc l'équilibre est instable.

Pour  $x_0 < 0$  on montre de manière similaire que  $\lim_{t \rightarrow \inf J} x(t) = -\infty$ .

- Pour  $\alpha = -1$ , on a  $x' = -x^3$  et cette fois si  $x_0 > 0$ , la solution reste positive et décroissante. Par ailleurs, elle est minorée par 0. Le théorème d'explosion en temps fini montre que cette solution est bien définie sur  $[0, +\infty[$  et par les propriétés précédentes nous savons qu'il existe  $\bar{x} \in [0, x_0[$  tel que  $\lim_{t \rightarrow +\infty} x(t) = \bar{x}$ .

D'après l'équation différentielle nous avons donc

$$\lim_{t \rightarrow +\infty} x'(t) = -\bar{x}^3.$$

8. Le faire en exercice !



Si on utilise maintenant le lemme **L.58** ci-dessous, nous déduisons

$$-\bar{x}^3 = 0,$$

et donc

$$\bar{x} = 0.$$

On a donc montré que, pour toute donnée initiale positive la solution converge vers le point d'équilibre 0. On peut raisonner de façon similaire pour les données négatives et ainsi conclure à la stabilité asymptotique de cet équilibre. ■

### Lemme I.58

Soit  $x : [0, +\infty[ \rightarrow \mathbb{R}$  une fonction dérivable. Si  $\lim_{t \rightarrow +\infty} x(t)$  et  $\lim_{t \rightarrow +\infty} x'(t)$  existent, alors on a

$$\lim_{t \rightarrow +\infty} x'(t) = 0.$$

### Remarque I.59

Si on ne suppose pas que la limite de  $x'$  existe le résultat est faux comme on le voit par exemple sur la fonction  $x(t) = \frac{\sin(t^2)}{t}$ .

### Preuve :

On note  $\alpha = \lim_{t \rightarrow +\infty} x'(t)$  et on suppose que  $\alpha \neq 0$ . Quitte à changer  $x$  en  $-x$  on peut toujours supposer que  $\alpha > 0$ . Par définition de la limite, il existe un  $A \geq 0$  tel que

$$x'(t) \geq \frac{\alpha}{2}, \quad \forall t \geq A.$$

Par intégration entre  $A$  et  $t \geq A$ , nous obtenons

$$x(t) - x(A) \geq \frac{\alpha}{2}(t - A), \quad \forall t \geq A,$$

et donc

$$x(t) \geq x(A) + \frac{\alpha}{2}(t - A) \xrightarrow[t \rightarrow +\infty]{} +\infty,$$

ce qui implique que  $\lim_{t \rightarrow +\infty} x(t) = +\infty$ . C'est une contradiction avec les hypothèses du lemme. ■

## V.2.b Un théorème plus précis mais hors programme

Le théorème de la section précédente ne permet de conclure que si les parties réelles des valeurs propres de la jacobienne sont toutes strictement positives ou bien toutes strictement négatives. Evidemment cela ne permet pas de comprendre le comportement des solutions dans beaucoup de cas. Le théorème suivant permet de comprendre ce comportement dans un cas plus général (mais pas encore dans tous les cas).

### Théorème I.60 (Hartmann-Grobman)

Soit  $x' = f(x)$  un système différentiel autonome dont 0 est un équilibre. On introduit la jacobienne  $A = Df(0)$ . On suppose que 0 est un équilibre **hyperbolique** (ce qui signifie que toutes les valeurs propres de  $A$  sont de partie réelle **non nulle**).

Alors, au voisinage de 0, le flot  $\varphi$  de l'équation est conjugué à celui (noté  $\psi$ ) du système linéaire  $y' = Ay$ .

Cela signifie qu'il existe  $\delta > 0$  et deux ouverts  $U, V \subset \mathbb{R}^d$  contenant 0, et un homéomorphisme  $h : U \rightarrow V$  tel que

$$\forall t \in ]-\delta, \delta[, \quad \varphi(t, \cdot) = h^{-1} \circ \psi(t, \cdot) \circ h, \quad \text{dans } U.$$

Autrement dit,

$$\forall t \in ]-\delta, \delta[, \forall x \in U, \quad \varphi(t, x) = h^{-1}(e^{tA}h(x)).$$

En particulier, pour un tel équilibre hyperbolique, il suffit qu'au moins une des valeurs propres ait une partie réelle strictement positive pour que l'équilibre soit instable, ce qui est plus précis que le second point énoncé dans le théorème I.57.

A noter que ce résultat ne dit toujours rien sur le cas où certaines valeurs propres ont une partie réelle nulle et, de fait, le théorème est faux dans ce cas comme on l'a vu précédemment sur le système  $x' = \alpha x^3$ . Dans ce cas, le système linéarisé est  $y' = 0$  dont les solutions sont toutes constantes. Si le flot (trivial) associé  $\psi$  était conjugué au flot  $\varphi$  de l'équation non linéaire, on en déduirait que, au voisinage de 0, toutes les solutions de l'équation non linéaire sont constantes, ce qui n'est manifestement pas le cas.

Pour aller plus loin, de nouveaux outils sont nécessaires comme la notion de variété centrale (resp. stable, instable), ou la notion de forme normale d'un système au voisinage d'un équilibre. Cela sort totalement du programme de ce cours mais il est bon de savoir, en cas de besoin, que cette théorie existe et a été très développée.

### V.2.c Intégrales premières. Ensembles invariants

#### Définition I.61

Une fonction dérivable  $E : \Omega \rightarrow \mathbb{R}$  est appelée **intégrale première** du champ de vecteurs  $f$  si  $E$  est constante le long de toutes les trajectoires du système :

$$\frac{d}{dt} E(\varphi(t, x)) = 0, \quad \forall t \in I, \forall x \in \Omega.$$

#### Lemme I.62

La fonction  $E$  est une intégrale première de  $f$  si et seulement si on a

$$(f \cdot \nabla E)(x) = 0, \quad \forall x \in \Omega,$$

autrement dit si, en tout point, le vecteur  $f(x)$  et le gradient de  $E$  sont orthogonaux.

#### Définition I.63

Pour toute fonction  $E : \Omega \rightarrow \mathbb{R}$  et toute valeur  $\alpha$ , on appelle **ensemble de niveau de  $E$  pour le niveau  $\alpha$**  l'ensemble  $E_\alpha$  défini par

$$E_\alpha := \{x \in \Omega, E(x) = \alpha\} = E^{-1}(\{\alpha\}).$$

On appelle **sous-ensemble de niveau** (resp. **sous-ensemble de niveau strict**) de  $E$  pour le niveau  $\alpha$ , l'ensemble  $E_{\leq \alpha}$  (resp.  $E_{< \alpha}$ ) défini par

$$E_{\leq \alpha} := \{x \in \Omega, E(x) \leq \alpha\} = E^{-1}(]-\infty, \alpha]),$$

$$E_{< \alpha} := \{x \in \Omega, E(x) < \alpha\} = E^{-1}(]-\infty, \alpha[).$$

Si une telle fonction existe, on en déduit que toute trajectoire est complètement contenue dans l'ensemble de niveau  $E_\alpha$ .

Si la fonction  $E$  est constante, ceci ne nous apprend rien mais si elle a de bonnes propriétés, l'ensemble  $E_\alpha$  est une hypersurface (i.e. une surface en dimension  $d = 3$ , une courbe en dimension  $d = 2$ ) cela peut donner de précieuses informations. Considérons deux exemples :

— *L'équation du pendule :*

On considère l'équation du pendule (avec des constantes physiques égales à 1) :

$$\theta'' + \sin(\theta) = 0,$$

et on constate que la quantité (positive)

$$E(\theta, \theta') = \frac{1}{2}(\theta')^2 + (1 - \cos(\theta)),$$

est conservée au cours du temps, c'est donc bien une intégrale première de ce système.

Par ailleurs, pour  $\alpha \in [0, 2[$  les courbes de niveau  $E_\alpha$  sont des unions de courbes fermées. Les trajectoires du système correspondant à une donnée de Cauchy vérifiant  $E(\theta_0, \theta'_0) < 2$  sont donc contenues dans l'une de ces courbes.

— Lotka-Volterra : il s'agit d'un modèle de dynamique des populations qui prend la forme suivante

$$\begin{cases} x' = ax - bxy, \\ y' = -cy + dxy, \end{cases} \quad (\text{I.21})$$

où les constantes  $a, b, c$  et  $d$  sont strictement positives.

Ce système a été introduit pour modéliser l'évolution des effectifs de population de proies (disons des sardines, dont la population à l'instant  $t$  est  $x(t)$ ) et de prédateurs (disons des requins, dont la population à l'instant  $t$  est  $y(t)$ ). Le nombre  $a > 0$  représente le taux de croissance des sardines (elles disposent de ressources infinies, le plancton), le nombre  $-c < 0$  est le taux de disparition des requins en l'absence de tout apport de nourriture. Les termes  $-bxy$  et  $dxy$  représentent l'impact des rencontres requins/sardines sur les populations : les requins mangent les sardines donc le nombre de sardine diminuent et les requins se refont une santé, se reproduisent, etc ...

On peut montrer que ce système est globalement bien posé pour toutes données initiales positives  $x_0, y_0$  (voir la section VI pour une étude détaillée d'un exemple similaire).

On constate par ailleurs qu'il y a un unique point d'équilibre dans la zone d'intérêt qui est donné par  $x^* = c/d$  et  $y^* = a/b$ . Du reste, on peut vérifier que le système possède une intégrale première donnée par

$$E(x, y) = by + dx - a \log y - c \log x.$$

Une étude élémentaire montre que les lignes de niveau de  $E$  sont des courbes fermées que l'on peut tracer aisément et ainsi obtenir un portrait de phases du système

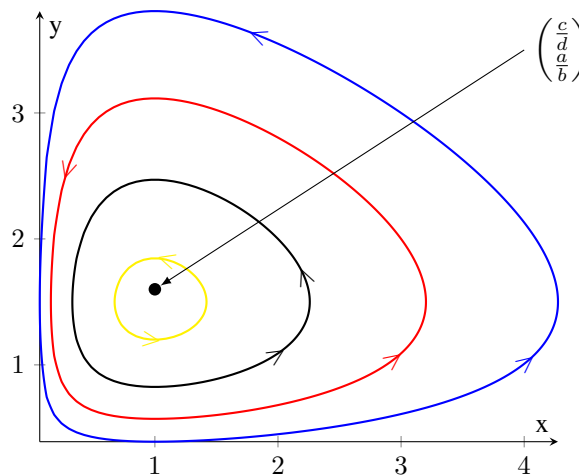


FIGURE I.10 – Portrait de phase pour le modèle de Lotka-Volterra

Il peut arriver que, même si la fonction  $E$  n'est pas constante sur les trajectoires du système, on puisse tirer des informations intéressantes sur ces trajectoires par l'argument suivant.

**Proposition I.64 (Sous-ensembles de niveau invariants)**

Soit  $E$  une fonction de classe  $\mathcal{C}^1$  et  $\alpha \in \mathbb{R}$  donné.

On suppose que

$$\forall x \in E_\alpha, \quad (f \cdot \nabla E)(x) < 0,$$

alors le sous-ensemble de niveau strict  $E_{<\alpha}$  est invariant par le flot (en temps positif), c'est-à-dire que

$$\Phi(t)(E_{<\alpha}) \subset E_{<\alpha}, \quad \forall t \geq 0.$$

**Exemple I.65**

On considère le système suivant en dimension 2

$$\begin{cases} x' = x(1 - ax^2 - by^2), \\ y' = y(1 - cx^2 - dy^2), \end{cases}$$

avec  $a, b, c, d > 0$ . On considère  $E(x, y) = x^2 + y^2$ .

Un calcul élémentaire nous donne

$$(f \cdot \nabla E)(x, y) = 2x^2(1 - ax^2 - by^2) + 2y^2(1 - cx^2 - dy^2).$$

Par ailleurs, il existe  $\alpha_0 > 0$  assez grand tel que

$$x^2 + y^2 = \alpha_0 \Rightarrow ax^2 + by^2 > 1,$$

$$x^2 + y^2 = \alpha_0 \Rightarrow cx^2 + dy^2 > 1.$$

Pour toute valeur de  $\alpha \geq \alpha_0$ , nous avons bien  $(f \cdot \nabla E) < 0$  sur  $E_\alpha$ . Ainsi les sous-ensembles de niveau  $E_{<\alpha}$  (qui ne sont autres que des boules ouvertes suffisamment grandes) sont invariants par le flot. On en déduit en particulier que toutes les solutions maximales du système restent bornées et sont donc bien définies sur  $[0, +\infty[$ , par le théorème d'explosion en temps fini.

Remarquons malgré tout que, d'après le critère spectral (Théorème I.57), l'origine est un point d'équilibre instable pour ce système.

**V.2.d Quelques mots sur la théorie de Liapounov**

On considère un champ de vecteurs  $f : \Omega \rightarrow \mathbb{R}^d$  localement lipschitzien.

Soit  $V : \Omega \rightarrow \mathbb{R}$  une fonction de classe  $C^1$  qui vérifie les propriétés suivantes<sup>9</sup> :

— **Monotonie/dissipation relativement à  $f$  :**

$$f(x) \cdot \nabla V(x) \leq 0, \quad \forall x \in \Omega,$$

— **Structure des points critiques de  $V$  relativement à  $f$  :**

On note

$$C := \{x \in \Omega \text{ t.q. } f(x) \cdot \nabla V(x) = 0\},$$

l'ensemble des points critiques de  $V$  relativement à  $f$ . On suppose que pour tout  $\alpha \in \mathbb{R}$ , l'intersection de  $C$  avec la ligne de niveau  $\alpha$  de  $V$  définie par

$$C_\alpha := C \cap V_\alpha,$$

est localement finie.

— **Compacité :**

Les sous-ensembles de niveau  $V_{\leq \alpha}$  ont des composantes connexes compactes.

**Théorème I.66 (de Liapounov)**

Sous les hypothèses précédentes, pour toute donnée initiale  $x_0$ , la solution associée  $t \mapsto \varphi(t, x_0)$  converge quand  $t \rightarrow +\infty$  vers un point d'équilibre du système.

En particulier, toutes les composantes connexes des  $V_{\leq \alpha}$  contiennent un point d'équilibre.

En pratique, on utilise souvent ce théorème à travers le corollaire suivant qui se déduit immédiatement.

**Corollaire I.67**

Soit  $x^*$  un point d'équilibre du système. S'il existe  $\alpha > V(x^*)$  tel que  $x^*$  est l'unique point d'équilibre du système dans sa composante connexe de  $V_{\leq \alpha}$  alors  $x^*$  est asymptotiquement stable.

En particulier, tout minimum local strict de  $V$  est un point d'équilibre asymptotiquement stable du système  $x' = f(x)$ .

9. d'autres jeux d'hypothèses sont envisageables mais on ne prétend pas ici à un résultat optimal

**Preuve (du Théorème I.66):**

On note  $(t, x) \mapsto \varphi(t, x)$  le flot du système (qui ne dépend que d'une variable de temps car il est autonome).

— Etape 1 : *Compacité des trajectoires.*

Par hypothèse nous avons, pour tout  $x \in \Omega$ ,

$$\frac{d}{dt}V(\varphi(t, x)) = \partial_t \varphi(t, x) \cdot \nabla V(\varphi(t, x)) = (f \cdot \nabla V)(\varphi(t, x)) \leq 0.$$

Ainsi la fonction  $V$  décroît le long des trajectoires, en particulier  $\varphi(t, x_0)$  appartient au sous-ensemble de niveau  $V_{\leq \alpha_0}$  avec  $\alpha_0 := V(x_0)$  pour tout  $t \geq 0$  pour lequel la solution est définie. Comme la trajectoire dépend continûment du temps, elle est même contenue dans une composante connexe de  $V_{\leq \alpha_0}$ .

Comme cette composante connexe est compacte par hypothèse, le théorème de sortie de tout compact prouve que la solution en question est définie sur tout  $[0, +\infty[$ , et ce pour toute donnée initiale.

Par ailleurs la demi-trajectoire  $(\varphi(t, x_0))_{t \geq 0}$  étant contenue dans un compact, elle admet des valeurs d'adhérence. On va montrer par la suite qu'elle n'en possède qu'une et qu'elle converge vers i-celle.

— Etape 2 : *Caractérisation des valeurs d'adhérence.*

Soit  $a = \lim_{n \rightarrow \infty} \varphi(t_n, x_0)$  une telle valeur d'adhérence, avec  $t_n \rightarrow \infty$ . Par monotonie et continuité, on obtient que

$$V(\varphi(t_n, x_0)) \xrightarrow{n \rightarrow \infty} V(a), \text{ en décroissant.}$$

Comme  $t \mapsto V(\varphi(t, x_0))$  est une fonction numérique décroissante qui admet une valeur d'adhérence, elle converge vers celle-ci. Autrement dit nous avons obtenu que

$$\lim_{t \rightarrow \infty} V(\varphi(t, x_0)) = V(a). \quad (\text{I.22})$$

Pour tout  $s \geq 0$ , nous avons par la propriété de groupe du flot et la continuité de celui-ci et de  $V$

$$V(\varphi(t_n + s, x_0)) = V(\varphi(s, \varphi(t_n, x_0))) \xrightarrow{n \rightarrow \infty} V(\varphi(s, a)),$$

et donc, en comparant à (I.22), on déduit

$$V(\varphi(s, a)) = V(a), \text{ pour tout } s.$$

La trajectoire  $s \mapsto \varphi(s, a)$  est donc contenue dans l'ensemble de niveau  $V_\alpha$ , pour  $\alpha := V(a)$ . Par ailleurs, en dérivant par rapport à  $s$  cette égalité, on obtient que  $\varphi(s, a) \in C$ , pour tout  $s$ . On a donc

$$\varphi(s, a) \in C_\alpha, \quad \forall s \geq 0.$$

Comme on a supposé que  $C_\alpha$  est localement fini, on obtient que  $\varphi(s, a) = a$  pour tout  $s$ , autrement dit que  $a$  est un équilibre du système, i.e.  $f(a) = 0$ .

— Etape 3 : *Convergence vers la valeur d'adhérence  $a$  et conclusion.*

On va raisonner par l'absurde en supposant que  $\varphi(t, x_0)$  ne converge pas vers  $a$  quand  $t \rightarrow \infty$ . Cela signifie qu'il existe  $\varepsilon_0 > 0$  et une suite strictement croissante  $(\tau_n)_n$  tendant vers l'infini telle que

$$\|\varphi(\tau_n, x_0) - a\| \geq \varepsilon_0.$$

Quitte à extraire des sous-suites convenables, on peut supposer que

$$\|\varphi(t_n, x_0) - a\| \leq \varepsilon_0/2,$$

et

$$t_n < \tau_n < t_{n+1}, \quad \forall n \geq 0.$$

On fixe  $\varepsilon_0/2 < \varepsilon < \varepsilon_0$ . Par construction on a

$$\|\varphi(t_n, x_0) - a\| < \varepsilon < \|\varphi(\tau_n, x_0) - a\|, \quad \forall n \geq 0.$$

Par le théorème des valeurs intermédiaires, on en déduit qu'il existe  $s_n \in ]t_n, \tau_n[$  vérifiant

$$\|\varphi(s_n, x_0) - a\| = \varepsilon.$$

La suite  $(\varphi(s_n, x_0))_n$  est dans un compact, elle admet donc une valeur d'adhérence, notée  $c_\varepsilon$ , qui est donc par le raisonnement ci-dessus un point de  $C_\alpha$  et qui, par ailleurs, vérifie  $\|a - c_\varepsilon\| = \varepsilon$ . Ceci étant vrai pour tout  $\varepsilon$ , cela contredit le fait que  $C_\alpha$  est localement fini et conclut le raisonnement par l'absurde. On a donc bien convergence de toute la trajectoire vers  $a$ . La preuve est complète.

Par la méthode Liapounov, on peut retrouver le critère linéaire de stabilité asymptotique donné dans la première partie du Théorème I.57 que l'on présente à nouveau ici sous la forme d'un corollaire, à noter que la méthode de Liapounov est un peu plus précise puisqu'elle peut permettre d'évaluer la taille du voisinage de l'équilibre à partir duquel les solutions vont converger vers celui-ci. Dans certains cas, elle permet même de prouver la stabilité asymptotique globale d'un équilibre (i.e. : pour toute donnée initiale la solution converge vers  $x^*$ ), ce que ne permettra jamais un critère de nature linéaire qui ne voit que ce qui se passe près du point d'équilibre.

### Corollaire I.68

Soit  $x^*$  un point d'équilibre du système  $x' = f(x)$  et  $A$  la matrice jacobienne de  $f$  au point  $x^*$ . Si  $A$  a toutes ses valeurs propres de partie réelle strictement négative alors il existe  $R > 0$  tel que que le système admette une fonction de Liapounov sur la boule  $\Omega = B(0, R)$  et en particulier cet équilibre est asymptotiquement stable.

#### Preuve :

On a déjà vu que, sous les hypothèses du théorème, il existe  $C, \gamma > 0$  tels que

$$\|e^{tA}\| \leq Ce^{-t\gamma}, \quad \forall t \geq 0.$$

On va donc définir à bon droit la fonction suivante

$$V(x) = \int_0^{+\infty} \|e^{sA}x\|^2 ds,$$

qui vérifie, pour des  $\alpha_{\min}, \alpha_{\max} > 0$  bien choisis

$$\alpha_{\min}\|x\|^2 \leq V(x) \leq \alpha_{\max}\|x\|^2, \quad \forall x \in \mathbb{R}^d. \quad (\text{I.23})$$

Un calcul simple montre que

$$\begin{aligned} \frac{d}{dt}V(\varphi(t, x)) &= 2 \int_0^{+\infty} (e^{sA}f(\varphi(t, x)), e^{sA}\varphi(t, x)) ds \\ &= 2 \int_0^{+\infty} (e^{sA}A\varphi(t, x), e^{sA}\varphi(t, x)) ds + 2 \int_0^{+\infty} (e^{sA}g(\varphi(t, x)), e^{sA}\varphi(t, x)) ds \\ &= \int_0^{+\infty} \frac{d}{ds} \|e^{sA}\varphi(t, x)\|^2 ds + 2\sqrt{V(g(\varphi(t, x)))}\sqrt{V(\varphi(t, x))} \\ &= -\|\varphi(t, x)\|^2 + 2C_2\|g(\varphi(t, x))\|\|\varphi(t, x)\|. \end{aligned}$$

En évaluant en  $t = 0$ , on trouve

$$(f \cdot \nabla V)(x) \leq -\|x\|^2 + 2C_2\|g(x)\|\|x\|.$$

Soit  $\delta > 0$  tel que

$$\|g(y)\| \leq \frac{1}{4C_2}\|y\|, \quad \forall y \in \bar{B}(0, \delta),$$

puis on pose  $\Omega = B(0, \delta)$ . D'après le calcul précédent, pour tout  $x \in \Omega$ , on a

$$(f \cdot \nabla V)(x) \leq -\|x\|^2 + \frac{1}{2}\|x\|^2 = -\frac{1}{2}\|x\|^2.$$

Cela montre la monotonie relativement à  $f$  de  $V$ , que l'ensemble des points critiques de  $V$  relativement à  $f$  est simplement constitué du point 0 et, par construction de  $V$ , les sous-ensembles de niveau de  $V$  sont contenues dans des boules et donc compacts.

Le théorème de Liapounov s'applique alors dans le domaine  $\Omega$  et montre que 0 est un équilibre asymptotiquement stable.

A noter que l'inégalité ci-dessus, avec les propriétés de  $V$  (I.23), peut se préciser en

$$(f \cdot \nabla V)(x) \leq -\frac{1}{2}\|x\|^2 \leq -\frac{1}{2\alpha_{\min}}V(x).$$

Ceci va impliquer que  $V$ , le long des trajectoires, va vérifier

$$\frac{d}{dt}V(\varphi(t, x)) \leq -\frac{1}{2\alpha_{\min}}V(\varphi(t, x)),$$

et donc par le lemme de comparaison sur les inégalités différentielles I.22, on en déduit

$$V(\varphi(t, x)) \leq V(x)e^{-\frac{1}{2\alpha_{\min}}t}, \quad \forall t \geq 0,$$

et donc on retrouve la convergence exponentielle des trajectoires grâce à (I.23). ■

Donnons deux exemples.

#### 1. Flot de gradient :

Le premier cas de système pour lequel l'existence d'une fonction de Liapounov est claire est le cas des **flots de gradients**. On se donne une fonction  $V : \Omega \rightarrow \mathbb{R}$  de classe  $C^2$  et on considère le système différentiel

$$x' = -\nabla V(x).$$

Autrement dit, le champ de vecteurs qui définit le problème est exactement l'opposé du gradient d'un *potentiel*  $V$ . Ce type de systèmes intervient souvent en physique.

On vérifie aisément que si  $V$  est coercif avec un nombre fini de points critiques, alors c'est une fonction de Liapounov et les équilibres stables du système sont ceux qui sont des minima locaux de  $V$ .

#### 2. Système du pendule amorti :

On considère l'équation

$$\theta'' + \mu\theta' + \sin \theta = 0,$$

qui modélise l'évolution d'un pendule amorti ( $\theta$  est l'angle que fait le pendule avec la verticale), dans lequel le coefficient  $\mu > 0$  mesure l'amortissement lié, par exemple, aux frottements qui s'exercent sur le système au niveau du point d'attache du pendule.

On peut montrer que tous les équilibres de la forme  $(\theta, \theta') = (2k\pi, 0)$  sont des équilibres asymptotiquement stables par analyse spectrale de la jacobienne.

On peut être plus précis en considérant une fonction de Liapounov qui n'est autre que la fonctionnelle d'énergie du système

$$V(\theta, \theta') = \frac{1}{2}|\theta'|^2 + (1 - \cos \theta).$$

Son évolution au cours du temps, le long des solutions vérifie

$$\frac{d}{dt}V(\theta, \theta') = -\mu|\theta'|^2,$$

ce qui prouve la monotonie de  $V$  relativement au champ de vecteurs du pendule amorti.

Par ailleurs, les points critiques relatifs de  $V$  sont exactement les points du plan de phase vérifiant  $\theta' = 0$ , c'est-à-dire les points de la forme  $(\theta, 0)$ . Pour tout  $\alpha \in \mathbb{R}$  l'intersection de cet ensemble avec la ligne de niveau  $\{V = \alpha\}$  est un ensemble discret (c'est l'ensemble des  $(\theta, 0)$  avec  $\cos(\theta) = 1 - \alpha$ ) donc la propriété de structure des points critiques de  $V$  relativement au champ énoncée plus haut est bien érigée. Enfin, la propriété de compacité des composantes connexes des sous-ensembles de niveau  $\{V \leq \alpha\}$  est bien vérifiée (au moins pour  $\alpha$  strictement plus petit que 2).

On déduit de tout cela que les points d'équilibres de la forme  $X_k = (2k\pi, 0)$ , qui vérifient  $V(X_k) = 0$  sont asymptotiquement stables et que pour toute donnée initiale vérifiant  $V(\theta_0, \theta'_0) < 2$ , la solution du système va converger vers l'unique point d'équilibre appartenant à la même composante connexe des sous-ensembles de niveau de  $V$ . Les autres équilibres  $Y_k = ((2k+1)\pi, 0)$  vérifient  $V(Y_k) = 2$  et rentrent donc pas dans l'étude précédente et de fait, on observe que la jacobienne du champ en ces points est

$$A = \begin{pmatrix} 0 & 1 \\ 1 & -\mu \end{pmatrix},$$

qui a deux valeurs propres réelles de signe opposé, et en particulier, l'une d'entre elles est strictement positive ce qui montre la non stabilité de cet équilibre.

### V.2.e Barrières

Concluons cette discussion par un outil qui peut être utile pour décrire un peu plus précisément le comportement des solutions d'un système différentiel **scalaire**. Il s'agit d'un résultat permettant de comparer des fonctions vérifiant certaines

inégalités différentielles. On peut le voir comme une généralisation de la proposition I.22 qui traitait le cas linéaire.

### Théorème I.69

Soit  $f : I \times \mathbb{R} \rightarrow \mathbb{R}$  un champ de vecteurs continu, localement lipschitzien. On se donne deux fonctions dérivables  $\alpha, \beta : J \rightarrow \mathbb{R}$ , avec  $J \subset I$  et qui vérifient

$$\beta'(t) \geq f(t, \beta(t)), \quad \forall t \in J,$$

$$\alpha'(t) \leq f(t, \alpha(t)), \quad \forall t \in J.$$

Soit  $t_0 \in J$ , on a les propriétés suivantes

$$\alpha(t_0) \leq \beta(t_0) \implies \alpha(t) \leq \beta(t), \quad \forall t \in J, t \geq t_0,$$

$$\alpha(t_0) \geq \beta(t_0) \implies \alpha(t) \geq \beta(t), \quad \forall t \in J, t \leq t_0.$$

### Preuve :

- Si l'une des deux inégalités de l'hypothèse est stricte alors la preuve du résultat est plus simple. En effet, dans ce cas nous avons en tout  $t \in J$

$$\beta'(t) - \alpha'(t) > f(t, \beta(t)) - f(t, \alpha(t)),$$

et donc en tout point  $t$  où  $\alpha$  et  $\beta$  coïncident, nous savons que  $(\beta - \alpha)'(t) > 0$ .

Commençons par prouver qu'il existe  $\varepsilon > 0$  tel que  $\alpha < \beta$  sur  $]t_0, t_0 + \varepsilon[$  :

- Si  $\beta(t_0) > \alpha(t_0)$  le résultat est clair par continuité de ces deux fonctions.
- Si  $\beta(t_0) = \alpha(t_0)$ , on peut appliquer la propriété déduite ci-dessus de l'hypothèse pour obtenir que  $\beta - \alpha$  est strictement croissante au voisinage de  $t_0$  et on conclut également à la propriété attendue.

On peut maintenant introduire l'ensemble

$$T = \{t \in ]t_0, +\infty[ \cap J, \beta(t) = \alpha(t)\}.$$

On va montrer par l'absurde que cet ensemble est vide ce qui prouvera que  $\alpha - \beta$  ne peut pas changer de signe et donc reste positif sur  $[t_0, +\infty[$ . Supposons donc que  $T$  est non vide. Comme on sait par ailleurs qu'il est minoré par  $\varepsilon$ , on aurait

$$t^* := \inf T \in ]\varepsilon, +\infty[.$$

Par continuité, nous savons que  $t^* \in T$  et donc par la remarque préliminaire nous avons  $(\beta - \alpha)'(t^*) > 0$  et donc  $\beta - \alpha$  est strictement négative sur un intervalle de la forme  $]t^* - \delta, t^*[$ . Par le théorème des valeurs intermédiaires, on déduit l'existence d'un zéro  $\bar{t}$  de  $\beta - \alpha$  dans  $]\varepsilon, t^*[$ , qui est donc naturellement un élément de  $T$ , ce qui contredit la minimalité de  $t^*$ .

- Le cas général est un peu plus délicat et peut utiliser par exemple le résultat de la Proposition I.10. Pour cela on introduit la fonction suivante<sup>10</sup>

$$z(t) := (\alpha(t) - \beta(t)) + |\alpha(t) - \beta(t)|.$$

Cette fonction est positive et a pour propriété que  $z(t) = 0$  si et seulement si  $\alpha(t) \leq \beta(t)$ .

Ecrivons l'équation satisfaite par  $z$

$$z(t) = z(t_0) + \int_{t_0}^t (1 + \operatorname{sgn}(\alpha(s) - \beta(s)))(\alpha - \beta)'(s) ds.$$

Comme  $1 + \operatorname{sgn}$  est positif et que  $\alpha' \leq f(t, \alpha(t))$  et  $\beta' \geq f(t, \beta(t))$  on obtient

$$z(t) \leq z(t_0) + \int_{t_0}^t (1 + \operatorname{sgn}(\alpha(s) - \beta(s)))(f(s, \alpha(s)) - f(s, \beta(s))) ds.$$

Par le caractère lipschitzien de  $f$  on déduit

$$z(t) \leq z(t_0) + \int_{t_0}^t L(s)(1 + \operatorname{sgn}(\alpha(s) - \beta(s)))|\alpha(s) - \beta(s)| ds.$$

10. A un facteur près, il s'agit de la **partie positive** de la différence  $\alpha - \beta$



Comme par ailleurs on a

$$(1 + \operatorname{sgn}(\alpha(s) - \beta(s)))|\alpha(s) - \beta(s)| = z(s),$$

on arrive à l'inégalité suivante

$$z(t) \leq z(t_0) + \int_{t_0}^t L(s)z(s) ds.$$

Par le lemme de Gronwall, on obtient

$$z(t) \leq z(t_0)e^{\int_{t_0}^t L(s) ds}.$$

Ainsi, si  $\alpha(t_0) \leq \beta(t_0)$  alors  $z(t_0) = 0$  et l'inégalité ci-dessus montre que  $z(t) = 0$  pour tout  $t$ , autrement dit  $\alpha(t) \leq \beta(t)$ , ce qu'il fallait démontrer. ■

## VI Etude détaillée d'un exemple : un modèle de propagation d'épidémie

**Remarque préliminaire.** Les résultats numériques donnés par la suite ont été obtenus par un programme Python disponible sous la forme d'un notebook à l'adresse suivante

<https://gist.github.com/FranckBoyer/df2f2f839aef35f60e68aa7878e69a33>

**Présentation du modèle.** On considère ici le système d'équations différentielles ordinaires suivant

$$\begin{cases} S'(t) = -\beta I(t)S(t) + \gamma R(t), \\ R'(t) = \nu I(t) - \gamma R(t), \\ I'(t) = -\nu I(t) + \beta I(t)S(t), \end{cases} \quad (\text{I.24})$$

avec les conditions initiales

$$\begin{cases} S(0) = S_0, \\ R(0) = R_0, \\ I(0) = I_0. \end{cases} \quad (\text{I.25})$$

Dans ce système  $\beta$ ,  $\gamma$  et  $\nu$  sont trois paramètres strictement positifs. Ce système décrit l'évolution selon le temps (l'unité de temps étant *le jour*) d'une population soumise à une maladie contagieuse. Les hypothèses et les notations sont les suivantes :

- La population totale est constante au cours du temps : on ne tient pas compte dans ce modèle des effets de la natalité et de la mortalité, dont on suppose qu'ils se compensent exactement.
- $I(t) \in [0, 1]$  représente la proportion au temps  $t$  des individus qui sont *Infectés* par la maladie. Ces individus sont par ailleurs contagieux et sont le vecteur de la propagation de la maladie.
- $R(t) \in [0, 1]$  représente la proportion au temps  $t$  des individus qui sont *Résistants*, c'est-à-dire qui ont contracté la maladie, en ont guéri et en sont maintenant immunisés.
- $S(t) \in [0, 1]$  représente la proportion au temps  $t$  des individus qui sont *Susceptibles* de contracter la maladie. Il s'agit donc de tous les individus qui ne sont pas malades et qui ne sont pas immunisés.
- La signification des coefficients est la suivante :
  - $1/\nu$  est le temps moyen de guérison d'un patient infecté.
  - $\beta$  est le facteur de contamination, c'est-à-dire la probabilité que, si un individu *Infecté* est en contact avec un individu *Susceptible* pendant 1 unité de temps, alors ce dernier va contracter la maladie.
  - $1/\gamma$  est le temps moyen durant lequel un individu ayant contracté la maladie va en être immunisé.

### VI.1 Existence et unicité d'une solution globale en temps positif

#### Théorème I.70

Pour toutes données initiales  $S_0, R_0, I_0 \geq 0$ , il existe une unique solution au système (I.24)- (I.25) définie sur  $[0, +\infty[$ . De plus, celle-ci vérifie

$$\forall t \geq 0, \quad S(t) \geq 0, R(t) \geq 0, I(t) \geq 0,$$

et

$$\forall t \geq 0, \quad S(t) + R(t) + I(t) = S_0 + R_0 + I_0.$$

**Preuve :**

- Le champ de vecteurs (autonome)  $f$  qui définit l'équation est polynomiale en toutes ces variables, donc de classe  $\mathcal{C}^1$  et localement lipschitzienne. Ainsi, le théorème de Cauchy-Lipschitz local s'applique et on obtient donc l'existence d'une unique solution maximale  $(J, (S, R, I))$  au problème de Cauchy proposé.
- Si on somme les trois équations, on voit de suite que

$$S'(t) + R'(t) + I'(t) = 0, \quad \forall t \in J,$$

et donc la somme des trois inconnues est constante au cours du temps.

- On va maintenant montrer que les solutions restent positives sur  $J$ .
  - On commence par étudier le cas de la variable  $I$  :

On voit qu'elle est solution de l'équation différentielle linéaire non autonome suivante

$$I'(t) = (-\nu + \beta S(t))I(t),$$

dont la solution est donnée par

$$I(t) = I_0 e^{-\int_0^t (-\nu + \beta S(s)) ds}.$$

On voit que la solution est bien positive et que si  $I_0 > 0$ , on a  $I(t) > 0$  pour tout  $t \in J$ .

Si  $I = 0$ , le système se réduit à

$$S' = \gamma R, \quad R' = -\gamma R,$$

d'où la résolution explicite

$$\begin{aligned} R(t) &= R_0 e^{-\gamma t}, \\ S(t) &= S_0 + R_0 (1 - e^{-\gamma t}), \end{aligned}$$

qui répond bien à la propriété énoncée.

A partir de maintenant, on supposera donc que  $I(t) > 0$  pour tout  $t \geq 0$ .

- Prenons maintenant l'équation sur  $R$  et utilisons que  $I$  est positive pour déduire que

$$R'(t) \geq -\gamma R(t), \quad \forall t \in J,$$

et donc par la proposition I.22, on déduit que pour  $t \in J, t \geq 0$  on a l'inégalité

$$R(t) \geq R_0 e^{-\gamma t},$$

qui prouve en particulier que  $R$  est positive sur  $[0, +\infty[ \cap J$ .

- On prend enfin l'équation sur  $S$  qui, en utilisant le signe de  $R$  nous donne

$$S'(t) \geq -\beta I(t)S(t),$$

et donc, toujours par la proposition I.22, on trouve

$$S(t) \geq S_0 e^{-\beta \int_0^t I(s) ds},$$

et donc  $S$  est positive sur  $[0, +\infty[ \cap J$ .

- On a donc montré que  $R, S$  et  $I$  sont positives sur tout  $[0, +\infty[ \cap J$  et que, par ailleurs, leur somme est constante. Il s'en suit que  $R, S$  et  $I$  sont bornées sur  $[0, +\infty[ \cap J$ . Le théorème d'explosion en temps fini permet de conclure à l'existence globale en temps positif des solutions, c'est-à-dire que  $[0, +\infty[ \subset J$ .

■

## VI.2 Etats d'équilibre

On va maintenant supposer que  $S_0 + R_0 + I_0 = 1$  car ce sont supposées être des proportions.<sup>11</sup>

Comme  $S + R + I = 1$  au cours du temps, on peut se ramener au système de taille 2 suivant

$$\begin{cases} S' = -\beta IS + \gamma(1 - S - I), \\ I' = -\nu I + \beta IS. \end{cases} \quad (\text{I.26})$$

11. Toutefois l'analyse qui suit fonctionnerait de façon similaire si on remplaçait la valeur 1 par n'importe quelle autre valeur non nulle.

— On trouve un premier point d'équilibre à ce système donnée par

$$I = 0, S = 1.$$

La jacobienne du champ de vecteurs réduit en ce point est donnée par

$$A = \begin{pmatrix} -\gamma & -\beta - \gamma \\ 0 & \beta - \nu \end{pmatrix}.$$

Elle admet donc une valeur propre strictement négative et une autre dont le signe dépend des paramètres :

- Si  $\beta > \nu$ ,  $A$  a une valeur propre réelle strictement positive et le système est donc instable (via Hartman Grobman), dans ce cas on peut développer une épidémie à partir d'un tout petit nombre d'individus infectés.
- Si  $\beta < \nu$ , les deux v.p. de  $A$  sont strictement négatives et donc le point d'équilibre en question est asymptotiquement stable.  
Cette situation correspond à l'impossibilité de développer une épidémie.
- Dans le cas  $\beta = \nu$ , il y a une valeur propre nulle et on ne peut pas conclure a priori. Il vaut étudier le système particulier ainsi obtenu, que l'on écrit dans les nouvelles variables  $\tilde{S} = 1 - S$  et  $\tilde{I} = I$ . On notera que ce choix de nouvelles variables implique que  $\tilde{I} \geq 0$  et  $\tilde{S} \geq 0$ ,

$$\begin{cases} \tilde{S}' = (\beta + \gamma)\tilde{I} - \beta\tilde{I}\tilde{S} - \gamma\tilde{S}, \\ \tilde{I}' = -\beta\tilde{I}\tilde{S}. \end{cases}$$

On considère la fonction suivante

$$V(\tilde{S}, \tilde{I}) := \beta\tilde{S}^2 + 2(\beta + \gamma)\tilde{I}.$$

Cette fonction vérifie

$$\frac{d}{dt}V(\tilde{S}, \tilde{I}) = -2\beta^2\tilde{I}\tilde{S}^2 - 2\beta\gamma\tilde{S}^2 \leq 0.$$

Cette fonction est donc décroissante le long des trajectoires et l'ensemble  $C$  des points critiques de  $V$  relatifs au champ de vecteur considéré est de la forme

$$C := \{(0, \tilde{I}), \tilde{I} \geq 0\}.$$

Sur cet ensemble, la fonction  $V$  prend comme valeurs  $V(0, \tilde{I}) = 2(\beta + \gamma)\tilde{I}$  et donc l'intersection de  $C$  avec les courbes de niveau de  $V$  sont réduites à un point.

Donc les hypothèses du théorème de Liapounov sont vérifiées et toutes les trajectoires du système réduit vont converger vers l'unique point d'équilibre  $(0, 0)$ , ce qui prouve la stabilité asymptotique globale de l'unique point d'équilibre du système initial.

#### Remarque I.71 (Une preuve sans Liapounov, ou presque ...)

On peut prouver la stabilité asymptotique **globale** de l'équilibre  $I^* = 0, S^* = 1$  sous l'hypothèse  $\beta \leq \nu$  par une analyse directe sans utiliser le théorème général de type Liapounov.

On passe à nouveau en variables  $\tilde{S}, \tilde{I}$  (mais sans supposer  $\beta = \nu$ )

$$\begin{cases} \tilde{S}' = (\beta + \gamma)\tilde{I} - \beta\tilde{I}\tilde{S} - \gamma\tilde{S}, \\ \tilde{I}' = -\beta\tilde{I}\tilde{S} - (\nu - \beta)\tilde{I}. \end{cases}$$

et on calcule à nouveau la dérivée de  $V(\tilde{S}, \tilde{I})$  le long de la trajectoire. On obtient

$$\frac{d}{dt}V(\tilde{S}, \tilde{I}) = -2\beta^2\tilde{I}\tilde{S}^2 - 2\beta\gamma\tilde{S}^2 - 2(\beta + \gamma)(\nu - \beta)\tilde{I}. \quad (I.27)$$

— Si  $\nu > \beta$ , on obtient pour un certain  $\delta > 0$ , l'inégalité

$$\frac{d}{dt}V(\tilde{S}, \tilde{I}) \leq -\delta V(\tilde{S}, \tilde{I}),$$

qui nous donne par la Proposition I.22

$$V(\tilde{S}(t), \tilde{I}(t)) \leq V_0 e^{-\delta t}, \quad \forall t \geq 0.$$

In fine, on obtient  $\lim_{t \rightarrow +\infty} V(\tilde{S}(t), \tilde{I}(t)) = 0$  ce qui montre que  $\tilde{I}$  et  $\tilde{S}$  tendent vers 0 en  $+\infty$ .

— Si  $\nu = \beta$ , on commence toujours par remarquer que  $t \mapsto V(\tilde{S}(t), \tilde{I}(t))$  est décroissante et positive, donc convergente vers une limite notée  $V_\infty \geq 0$ . Il s'en suit par intégration entre 0 et  $+\infty$  de (I.27), que

$$\int_0^{+\infty} (2\beta\gamma\tilde{S}^2 + 2\beta^2\tilde{I}\tilde{S}^2) dt = V(S_0, I_0) - V_\infty < +\infty.$$

Par ailleurs, l'équation sur  $\tilde{I}$  nous donne immédiatement que  $t \mapsto \tilde{I}$  est également décroissante et positive donc convergente, vers une limite notée  $I_\infty$ . Comme  $V(\tilde{S}, \tilde{I})$  a aussi une limite, on en déduit que  $\tilde{S}^2$  a également une limite quand  $t \rightarrow +\infty$  et comme  $\tilde{S}^2$  est intégrable en  $+\infty$ , cette limite ne peut être qu'égal à 0. Infine, on a prouvé que  $\tilde{S} \rightarrow 0$  quand  $t \rightarrow \infty$ .

D'après l'équation sur  $\tilde{S}$ , on a

$$\lim_{t \rightarrow +\infty} \tilde{S}'(t) = (\beta + \gamma)I_\infty.$$

D'après le lemme I.58, comme cette limite existe et que  $\tilde{S}$  a également une limite, on doit nécessairement avoir  $I_\infty = 0$ . On a donc bien prouvé que la solution du système réduit vérifie  $(\tilde{S}, \tilde{I}) \rightarrow (0, 0)$ , ce qu'il fallait démontrer.

A noter qu'on ne démontre pas ici la convergence exponentielle de la solution. On peut d'ailleurs se convaincre que cette propriété est fausse dans ce cas. Voir les illustrations ci-dessous

— Si  $\beta \neq \nu$ , il y a un autre équilibre non trivial

$$S^* = \frac{\nu}{\beta},$$

$$I^* = \frac{\gamma(\beta - \nu)}{\beta(\nu + \gamma)}.$$

Celui-ci est positif (seuls cas acceptables pour le modèle) si et seulement si  $\nu < \beta$  (si  $\nu = \beta$ , on retrouve le cas précédent), ce qui est exactement le cas où l'équilibre trivial est instable.

On suppose donc désormais que  $\nu < \beta$ . La jacobienne du système en ce point d'équilibre est alors

$$\begin{aligned} A &= \begin{pmatrix} -\beta I^* - \gamma & -\beta S^* - \gamma \\ \beta I^* & -\nu + \beta S^* \end{pmatrix} \\ &= \begin{pmatrix} \gamma \frac{\nu - \beta}{\nu + \gamma} - \gamma & -\nu - \gamma \\ -\gamma \frac{\nu - \beta}{\nu + \gamma} & 0 \end{pmatrix} \\ &= - \begin{pmatrix} \gamma \frac{\gamma + \beta}{\nu + \gamma} & \nu + \gamma \\ \gamma \frac{\nu - \beta}{\nu + \gamma} & 0 \end{pmatrix}. \end{aligned}$$

Le déterminant de  $A$  vaut

$$\det A = \gamma(\beta - \nu),$$

et il est strictement positif d'après l'hypothèse sur les paramètres.

Par ailleurs, la trace de  $A$  vaut  $\text{Tr } A = -\frac{\gamma(\gamma + \beta)}{\nu + \gamma}$  et elle est donc strictement négative.

Ainsi, nous déduisons que la matrice  $A$  a deux valeurs propres dont les parties réelles sont strictement négatives, on a donc affaire à un point d'équilibre asymptotiquement stable.

### VI.3 Quelques exemples de trajectoires

On montre ici quelques exemples de trajectoires du système. Les figures sont été obtenues grâce à un notebook jupyter/python disponible à l'adresse <https://git.io/fhH9L>. Dans ce notebook vous pourrez faire varier les paramètres de façon interactive pour visualiser les différents comportements possibles des solutions.

La partie gauche des figures montre l'évolution des trois quantités  $S$ ,  $R$  et  $I$  en fonction du temps, la partie droite montre la trajectoire dans le plan de phase  $(I, S)$ . Le point bleu est le point initial et les gros points rouges sont les équilibres du système.

### VI.3.a Extinction d'une épidémie de force modérée

On suppose que :

- La moitié de la population est infectée à l'instant initial :  $I_0 = S_0 = 1/2$ ,  $R_0 = 0$ .
- La durée moyenne d'infection est de 2 jours :  $\nu = 1/2$
- La probabilité de contagion est  $\beta = 1/30$
- L'immunité disparaît au bout d'une dizaine de jours :  $\gamma = 1/10$

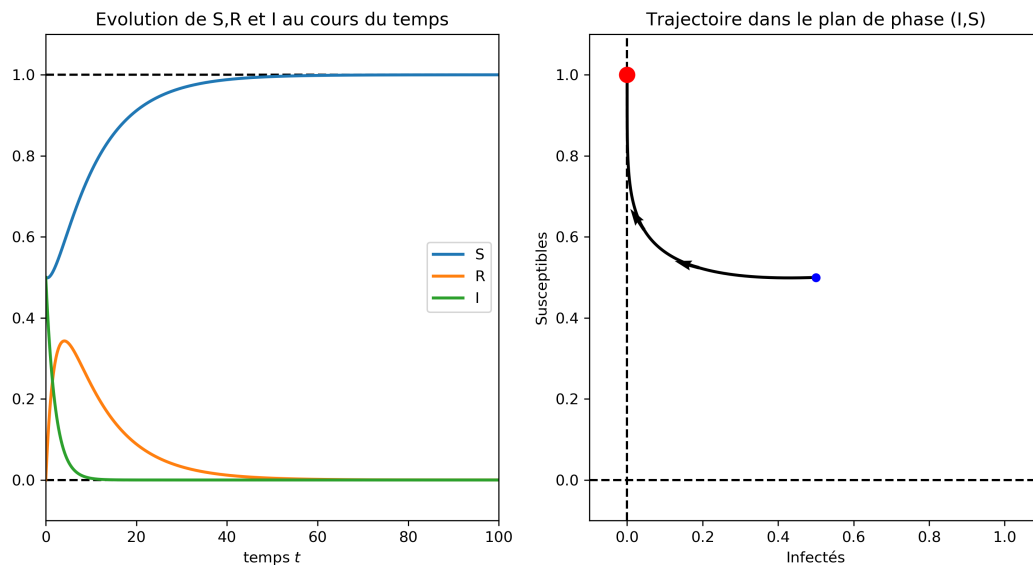


FIGURE I.11 – Epidémie de force modérée

Dans ces conditions on obtient la solution du système donnée dans la figure I.11. On observe que l'épidémie ne se propage pas (cela correspond à la stabilité asymptotique de l'équilibre  $(S = 1, I = 0)$  du système) et qu'elle finit par s'éteindre au bout de quelques dizaines de jours.

### VI.3.b Déclenchement d'une épidémie de grippe à New-York dans les années 60

Cette grippe très virulente a démarré avec 10 individus infectés sur une population totale de 8 millions d'habitants, ce qui donne  $I_0 = 1.25 \cdot 10^{-6}$ .

La durée moyenne d'infection constatée était de 3 jours et la probabilité de contagion était de  $1/2$ . On suppose que l'immunité acquise est permanente, i.e.  $\gamma = 0$  (ceci modifie un peu l'étude du système par rapport à ce qu'on a fait plus haut : en particulier tous les états de la forme  $(I = 0, S = S^*)$  sont des équilibres du système).

On observe que l'épidémie a mis presque 50 jours à prendre de l'ampleur et qu'elle a duré environ 150 jours au total. Comme l'immunité est permanente, la limite à l'infini de la quantité  $R$  dans cet exemple nous donne la proportion totale de la population qui a été infectée à un moment ou à un autre par le virus. On observe donc qu'au bout de 6 mois environ, près de 60% de la population New-Yorkaise a été affectée.

### VI.3.c Et si le virus mute ?

Reprenons l'exemple précédent en supposant que le virus est capable petit à petit de muter. On le prend en compte dans le modèle en supposant cette fois que l'immunité acquise par les patients infectés n'est pas permanente. On reprend donc les mêmes paramètres que précédemment en supposant cette fois que  $\gamma = 1/365$  (on estime qu'il faut un an au virus pour muter). Les résultats obtenus sont indiqués dans la figure I.13.

Si le début de la dynamique est essentiellement le même, on voit ensuite que la population est de moins en moins immunisée et permet ainsi à l'épidémie de connaître un nouveau pic (moins important toutefois que le précédent) et ainsi de suite pendant plusieurs années. In fine, la solution converge vers l'équilibre attendu (pour lequel la proportion d'individus infectés est petite mais pas nulle, ici environ 0.25%).

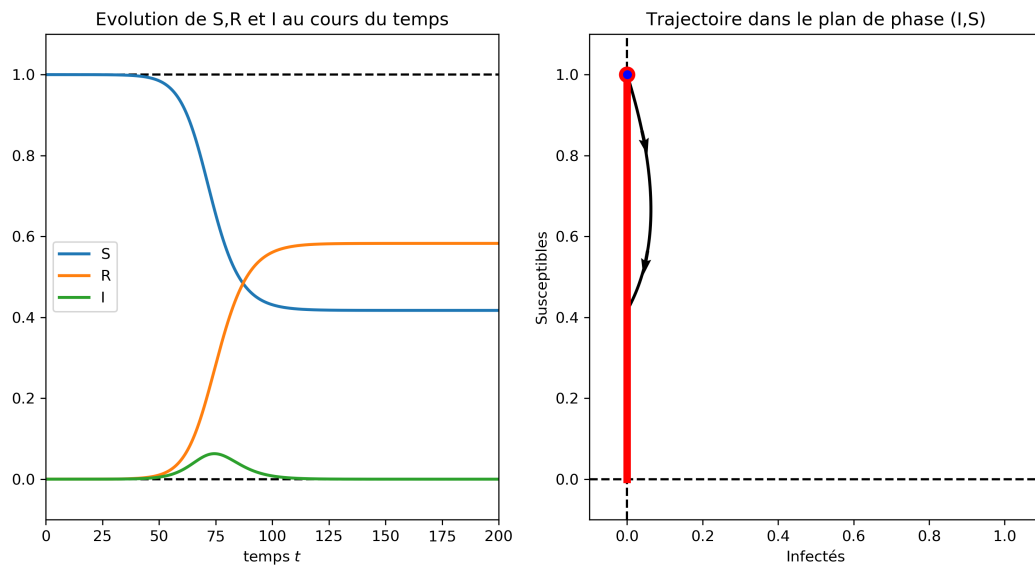


FIGURE I.12 – Epidémie virulente à New-York

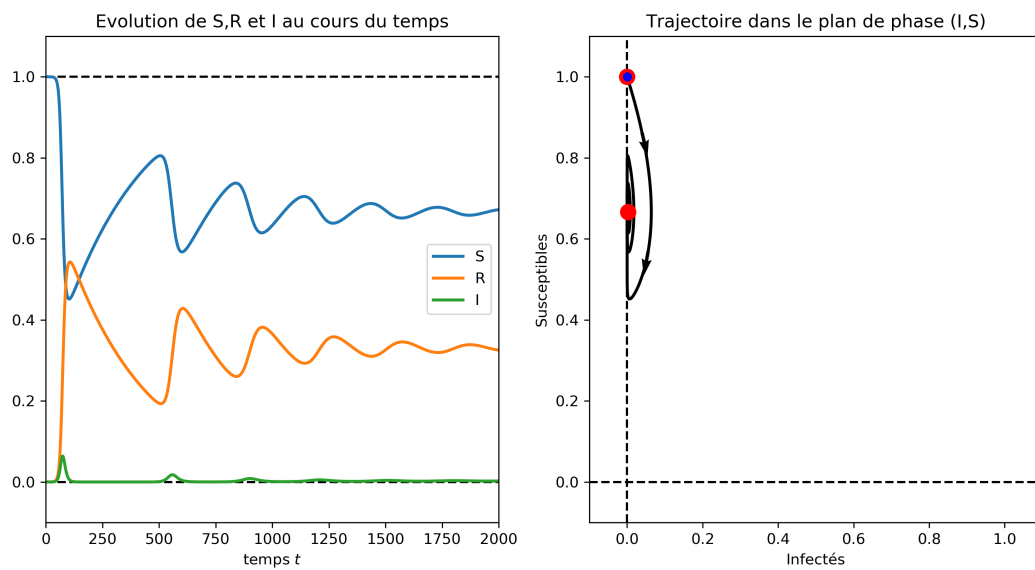
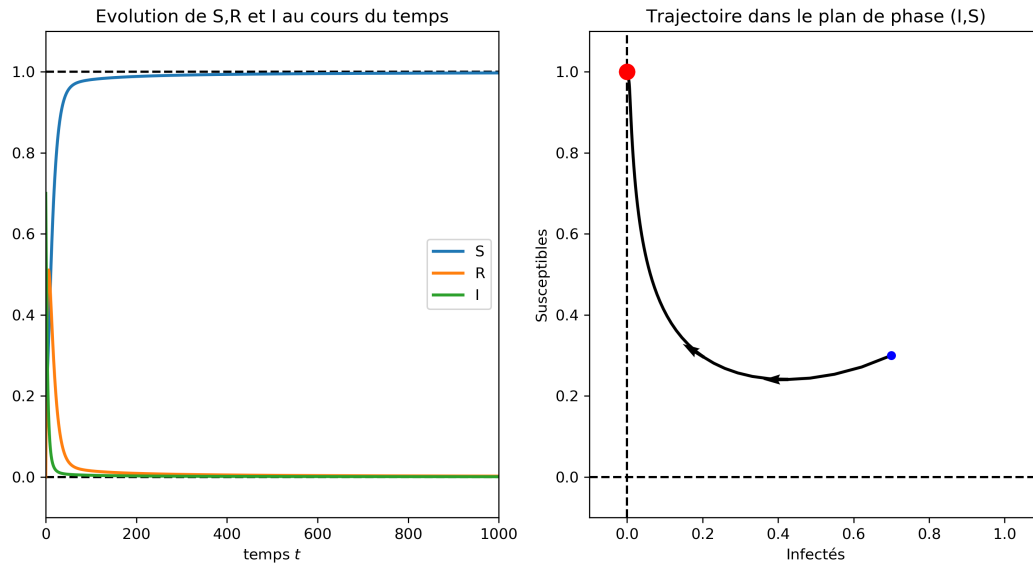
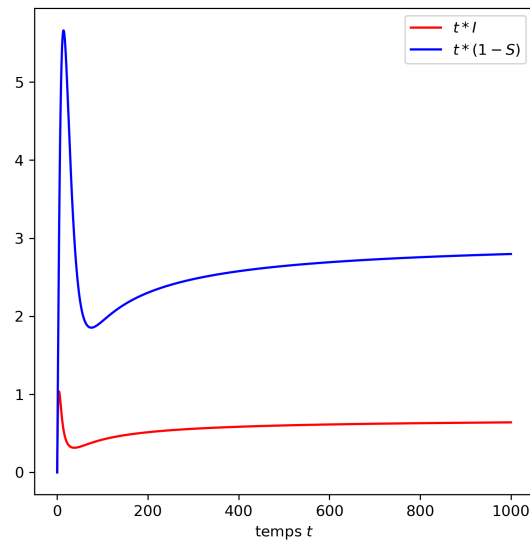


FIGURE I.13 – Epidémie virulente à New-York avec mutation

**VI.3.d Le cas  $\beta = \nu$** 

Dans ce cas, on a vu que la stabilité asymptotique globale de l'état d'équilibre  $(S^*, I^*) = (1, 0)$  est encore vraie mais qu'elle n'est plus exponentielle. C'est ce qu'on observe dans les figures I.14 et I.15 où la vitesse de retour à l'équilibre est de la forme

$$I(t) \sim_{t \rightarrow +\infty} \frac{C}{t}, \quad 1 - S(t) \sim_{t \rightarrow +\infty} \frac{C'}{t}.$$

FIGURE I.14 – Epidémie dans le cas  $\beta = \nu$ FIGURE I.15 – Epidémie dans le cas  $\beta = \nu$ . Vitesse de retour à l'équilibre

## A retenir du chapitre

### En priorité

- Les définitions principales (champs de vecteur et leurs régularité, problème de Cauchy, solutions locales, maximales, globales, flot, résolvante, stabilité, stabilité asymptotique ...)
- Enoncé et preuve du lemme de Gronwall.
- Résolution des équations différentielles linéaires autonomes. Formule de Duhamel.
- Etre capable de dessiner (au moins l'allure !) du portrait de phases pour un système linéaire  $2 \times 2$  autonome.
- Enoncés des théorèmes de Cauchy-Lipschitz global et local, des théorèmes d'explosion en temps fini et de sortie de tout compact.  
Savoir les utiliser convenablement pour prouver des propriétés qualitatives des solutions d'une EDO (positivité, etc ...)
- Propriétés du flot.
- Enoncé du théorème de stabilité/stabilité asymptotique dans le cas linéaire autonome.  
Connaître la preuve dans le cas de valeurs propres à parties réelles strictement négatives.
- Enoncé du critère spectral de stabilité asymptotique pour une EDO non-linéaire autonome.  
Savoir l'utiliser et connaître les cas où on ne peut pas conclure.

### Pour approfondir

- Travailler les preuves des principaux théorèmes. Observer que beaucoup d'entre elles utilisent le lemme de Gronwall de façon cruciale.
- Enoncé du théorème de Hartmann-Grobman.
- Théorie de Liapounov et théorème des barrières.



## Chapitre II

# Equations de transport

Dans ce chapitre et le suivant, on va étudier des équations aux dérivées partielles. On pourra utiliser de façon équivalente différentes notations pour les dérivées partielles

$$\partial_t, \frac{\partial}{\partial t}, \partial_x, \frac{\partial}{\partial x}, \partial_{x_i}, \partial_i, \dots$$

selon le contexte dans le but d'alléger le plus possible les formules.

Par ailleurs, on utilisera dans ce chapitre quelques opérateurs différentiels classiques (essentiellement la divergence et le gradient) dont on rappelle les définitions et les propriétés de base au début de l'annexe A. On utilisera également (très peu) le formalisme de la théorie des distributions qui est également rappelé dans cette même annexe que l'on pourra consulter en cas de besoin.

De façon très usuelle en EDP et sauf mention explicite du contraire les opérateurs  $\nabla$ ,  $\text{div}$  (et aussi  $\Delta$  que l'on rencontrera dans le prochain chapitre) n'agissent que sur les variables d'espace et pas de la variable de temps. Les dérivées par rapport au temps sont toujours explicitement donnée par les opérateurs  $\partial_t$ ,  $\partial_t^2$ , ... Ainsi par exemple si  $(t, x) \in \mathbb{R} \times \mathbb{R}^d \mapsto f(t, x) \in \mathbb{R}$  est une fonction régulière, on notera

$$\nabla f(t, x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(t, x) \\ \vdots \\ \frac{\partial f}{\partial x_d}(t, x) \end{pmatrix}.$$

## I Modèles de transport en 1D

### I.1 Trafic routier

On s'intéresse à la modélisation mathématique du trafic routier. C'est un sujet vaste que nous effleurons à peine afin de motiver l'étude des équations de type « transport ». Un des enjeux de ces modèles (dans leur version plus sophistiquée) est, par exemple, la compréhension de la formation des bouchons comme on peut le voir dans les vidéos suivantes qui montrent une expérience et la simulation numérique correspondante

Expérience : [https://www.youtube.com/watch?v=7wm-pZp\\_mi0](https://www.youtube.com/watch?v=7wm-pZp_mi0)

Simulation : <https://www.youtube.com/watch?v=Q78Kb4uLAdA>

- **Hypothèses de modélisation :** On modélise une route nationale rectiligne, avec une seule voie de circulation. On suppose que la route est infinie (autrement dit, on ne regarde pas les problèmes de bord) et qu'il n'y a pas de dépassements. De plus, on ne regarde pas les problèmes d'entrées et sorties.

On note  $\rho(t, x)$  la densité de véhicules (nombre moyen de véhicules par unité de longueur au voisinage de la position  $x$  et à l'instant  $t$ ) et  $v(t, x)$  la vitesse moyenne des véhicules à l'instant  $t$  et au point  $x$ .

- **Trajectoire d'un véhicule particulier :** Supposons connues  $\rho$  et  $v$  en tout point et à tout instant. On considère un véhicule qui se situerait à la position  $x_0$  et à l'instant  $t_0$ . On note  $X(t, t_0, x_0)$  sa position à un autre instant  $t$ . Par définition du champ de vitesse on a la relation

$$\partial_t X(t, t_0, x_0) = v(t, X(t, t_0, x_0)), \quad \forall t \geq t_0, \quad (\text{II.1})$$

et bien sûr

$$X(t_0, t_0, x_0) = x_0.$$

On voit donc que, si on connaît  $v$ , la trajectoire d'un véhicule est donnée par la solution de l'équation différentielle (II.1) associée à la donnée initiale  $(t_0, x_0)$ . Autrement dit,  $X$  est le flot associé au champ de vecteurs  $v$ .

On supposera dans toute la suite de ce chapitre que  $v$  est suffisamment régulière pour qu'on puisse appliquer le théorème de Cauchy-Lipschitz global et ainsi assurer que les trajectoires  $X$  (=le flot) sont définies de façon globale sans qu'on ait à s'inquiéter du domaine de définition de celles-ci.

— La loi de conservation des véhicules :

**Proposition II.1**

Si elles sont régulières, les deux fonctions  $\rho$  et  $v$  sont liées par l'équation aux dérivées partielles

$$\partial_t \rho + \partial_x(\rho v) = 0. \quad (\text{II.2})$$

**Preuve :**

On va donner plusieurs façons de démontrer ce résultat (elles sont bien sûr totalement équivalentes mais c'est plutôt l'esprit du calcul qui est différent).

— La preuve des physiciens.

On écrit un bilan de l'évolution du nombre de véhicules entre  $a$  et  $b$  entre les instants  $t$  et  $t + \delta t$ , où  $\delta t$  est supposé petit. Ce bilan est donc écrit *au premier ordre* ce qui signifie qu'on néglige tous les termes petits devant  $\delta t$ .

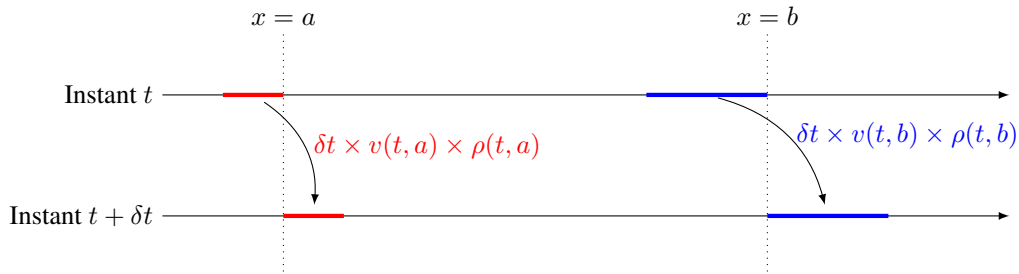


FIGURE II.1 – Preuve des physiciens

$$\int_a^b \rho(t + \delta t, x) dx \approx \int_a^b \rho(t, x) dx + \delta t v(t, a) \rho(t, a) - \delta t v(t, b) \rho(t, b).$$

On écrit par ailleurs que  $\rho(t + \delta t, x) \approx \rho(t, x) + \delta t \partial_t \rho(t, x)$  et on constate que la contribution des deux derniers termes s'écrit aussi  $-\delta t \int_a^b \partial_x(\rho v)(t, x) dx$ . On a donc obtenu

$$\delta t \int_a^b \partial_t \rho(t, x) dx \approx -\delta t \int_a^b \partial_x(\rho v)(t, x) dx,$$

et donc, par simplification par  $\delta t$

$$\int_a^b [\partial_t \rho + \partial_x(\rho v)] dx \approx 0.$$

Comme cette « égalité » doit être vraie pour tout  $t$ , tout  $a$ , tout  $b$ , on peut dériver par rapport à  $b$  et conclure que

$$\partial_t \rho + \partial_x(\rho v) = 0,$$

ce qui est le résultat attendu.

— La preuve des mathéux.

On écrit la relation suivante

$$\int_a^b \rho(t, x) dx = \int_{X(t+s, t, a)}^{X(t+s, t, b)} \rho(t + s, x) dx, \quad \forall s \in \mathbb{R},$$

qui décrit le fait que le nombre de véhicules compris entre deux véhicules donnés (ceux qui étaient en  $a$  et en  $b$  à l'instant  $t$ ) reste constant au cours du temps. On utilise alors le lemme II.2 ci-dessous pour dériver cette quantité par rapport à  $s$  et obtenir (en évaluant le résultat en  $s = 0$ )

$$0 = \int_a^b \partial_t \rho(t, x) dx + \partial_t X(t, t, b) \rho(t, X(t, t, b)) - \partial_t X(t, t, a) \rho(t, X(t, t, a)).$$

Utilisant la définition des trajectoires (II.1), on obtient

$$0 = \int_a^b \partial_t \rho(t, x) dx + v(t, b) \rho(t, b) - v(t, a) \rho(t, a) = \int_a^b \partial_t \rho + \partial_x(\rho v) dx.$$

La fin de la preuve est identique à la précédente. ■

Dans la preuve ci-dessus on a utilisé le lemme suivant.

### Lemme II.2

Soient  $f : (t, x) \in \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^1$  et  $\alpha, \beta : \mathbb{R} \rightarrow \mathbb{R}$  deux autres fonctions de classe  $\mathcal{C}^1$ . On a la formule de dérivation suivante

$$\frac{d}{dt} \int_{\alpha(t)}^{\beta(t)} f(t, x) dx = \int_{\alpha(t)}^{\beta(t)} \partial_t f(t, x) dx + \beta'(t) f(t, \beta(t)) - \alpha'(t) f(t, \alpha(t)).$$

#### Preuve :

On pose

$$\Phi : (a, b, t) \in \mathbb{R}^3 \mapsto \int_a^b f(t, x) dx.$$

Les théorèmes usuels d'analyse montrent<sup>1</sup> que  $\Phi$  admet des dérivées partielles données par

$$\partial_a \Phi(a, b, t) = -f(t, a),$$

$$\partial_b \Phi(a, b, t) = f(t, b),$$

$$\partial_t \Phi(a, b, t) = \int_a^b \partial_t f(t, x) dx.$$

Ces dérivées partielles sont des fonctions continues des trois variables  $(a, b, t)$  et donc la fonction  $\Phi$  est de classe  $\mathcal{C}^1$ . Le résultat du lemme provient donc juste de la dérivation de la fonction composée  $t \mapsto \Phi(\alpha(t), \beta(t), t)$ . ■

Notons que la loi de conservation (II.2) est valable en toute généralité sous les hypothèses faites au départ.

— **Modélisation du comportement des véhicules :** Pour fermer<sup>2</sup> le système, il faut trouver un lien entre la densité  $\rho$  et la vitesse  $v$ .

1. Premier modèle : tous les conducteurs vont à la même vitesse (la vitesse maximale autorisée) quelles que soient les conditions de trafic :

$$v(t, x) = V_{max} = cte.$$

Dans ces conditions,  $v$  n'est bien sûr plus une inconnue et il reste l'équation

$$0 = \partial_t \rho + \partial_x(\rho V_{max}) = \partial_t \rho + V_{max} \partial_x \rho, \quad (\text{II.3})$$

qui est appelée l'équation de transport à vitesse constante  $V_{max}$ .

2. Second modèle : La vitesse maximale autorisée varie le long de la route selon une loi donnée  $V_{max}(x)$ , l'équation vérifiée est alors

$$0 = \underbrace{\partial_t \rho + \partial_x(\rho V_{max}(x))}_{\text{forme conservative}} = \underbrace{\partial_t \rho + V_{max}(x) \partial_x \rho + V'_{max}(x) \rho}_{\text{forme non conservative}}. \quad (\text{II.4})$$

3. Troisième modèle : les conducteurs adaptent leur vitesse au trafic. Plus le trafic est dense, plus la vitesse est réduite, ce qui donne, par exemple, une loi du type

$$v = V_{max}(1 - \rho),$$

et donc l'équation aux dérivées partielles **non-linéaire**

$$0 = \partial_t \rho + V_{max} \partial_x(\rho(1 - \rho)). \quad (\text{II.5})$$

On peut aussi l'écrire en fonction de la vitesse

$$0 = \partial_t v + \partial_x(v(v - V_{max})).$$

En fonction de la modélisation retenue, la densité de véhicules vérifiera donc l'une des trois équations (II.3), (II.4) ou (II.5).

1. Vérifiez que vous savez le faire !

2. C'est-à-dire obtenir autant d'équations que d'inconnues

## I.2 Dynamique des gaz simplifiée

Etudions un contexte simple de mécanique des fluides : on étudie l'évolution d'un gaz parfait en une dimension d'espace (dans un tuyau rectiligne et de section petite par exemple). Les variables qui décrivent le fluide dans ces conditions sont : la densité du gaz  $\rho(t, x)$  (i.e. la masse volumique moyenne du gaz à l'instant  $t$  et au voisinage de la position  $x$ ), le champ de vitesse  $v(t, x)$  (i.e. la vitesse **moyenne** des particules de gaz à l'instant  $t$  et à la position  $x$ ) et la pression  $p(t, x)$ . On néglige les effets de gravité et on supposera ces quantités suffisamment régulières pour justifier les calculs qui suivent<sup>3</sup>

La première équation qui régit l'évolution du système est l'équation de conservation de la masse qui s'obtient exactement comme pour l'équation de conservation des véhicules dans le modèle de trafic routier :

$$\partial_t \rho + \partial_x(\rho v) = 0.$$

Pour obtenir la deuxième équation, il faut modéliser le comportement des particules de gaz. Ce comportement est donné par la loi de Newton qui nous dit que la variation de quantité de mouvement de tout système mécanique est égale à la somme des forces extérieures appliquées au système. La quantité de mouvement à l'instant  $t_0$  de la portion de gaz entre les abscisses  $a$  et  $b$  est donnée par

$$\int_a^b \rho(t_0, x) v(t_0, x) dx.$$

Au cours de l'évolution, la portion de gaz en question est comprise entre les abscisses  $X(t, t_0, a)$  et  $X(t, t_0, b)$  et la quantité de mouvement à l'instant  $t$  est donc donnée par

$$\left( \int_{X(t, t_0, a)}^{X(t, t_0, b)} \rho(t, x) v(t, x) dx \right).$$

La seule force à laquelle est soumise cette portion de gaz est la force de pression (exercée par les particules de gaz à l'extérieur de la portion en question) donc la loi de Newton s'écrit dans ce contexte sous la forme

$$\frac{d}{dt} \left( \int_{X(t, t_0, a)}^{X(t, t_0, b)} \rho(t, x) v(t, x) dx \right) = \text{Somme des forces} = -p(X(t, t_0, b)) + p(X(t, t_0, a)).$$

On utilise à nouveau le lemme II.2 et la définition des caractéristiques pour calculer la dérivée du membre de gauche

$$\begin{aligned} \frac{d}{dt} \left( \int_{X(t, t_0, a)}^{X(t, t_0, b)} \rho(t, x) v(t, x) dx \right) &= (\rho v)(t, X(t, t_0, b)) \frac{d}{dt} X(t, t_0, b) - (\rho v)(t, X(t, t_0, a)) \frac{d}{dt} X(t, t_0, a) \\ &\quad + \int_{X(t, t_0, a)}^{X(t, t_0, b)} \partial_t(\rho v)(t, x) dx \\ &= (\rho v)(t, X(t, t_0, b)) v(t, X(t, t_0, b)) - (\rho v)(t, X(t, t_0, a)) v(t, X(t, t_0, a)) \\ &\quad + \int_{X(t, t_0, a)}^{X(t, t_0, b)} \partial_t(\rho v)(t, x) dx \\ &= \int_{X(t, t_0, a)}^{X(t, t_0, b)} [\partial_t(\rho v) + \partial_x(\rho v^2)](t, x) dx. \end{aligned}$$

On écrit le terme de droite comme l'intégrale de la dérivée de la pression pour arriver à

$$\int_{X(t, t_0, a)}^{X(t, t_0, b)} [\partial_t(\rho v) + \partial_x(\rho v^2) + \partial_x p](t, x) dx = 0.$$

On spécialise cette égalité à  $t_0 = t$  pour obtenir

$$\int_a^b [\partial_t(\rho v) + \partial_x(\rho v^2) + \partial_x p](t, x) dx = 0.$$

Ceci étant vrai pour tous réels  $a < b$ , et la fonction sous l'intégrale étant supposée continue par rapport à  $x$ , on peut dériver cette égalité, pour  $t$  fixé quelconque, par rapport à  $b$  et obtenir l'équation aux dérivées partielles

$$\partial_t(\rho v) + \partial_x(\rho v^2) + \partial_x p = 0, \forall t > 0, \forall x \in \mathbb{R}.$$

Ce dernier argument est spécifique à la dimension 1 d'espace. De façon plus générale, on peut arriver à la même conclusion en utilisant le lemme de Du Bois-Reymond (ou plus exactement sa variante - Lemme A.5).

3. Cette hypothèse n'est en fait pas anodine mais cette discussion sort du cadre de cette introduction au sujet...

L'équation aux dérivées partielles obtenue s'écrit souvent sous la forme un peu plus compacte

$$\partial_t(\rho v) + \partial_x(\rho v^2 + p) = 0.$$

Le système de deux équations aux dérivées partielles obtenu

$$\begin{cases} \partial_t \rho + \partial_x(\rho v) = 0, \\ \partial_t(\rho v) + \partial_x(\rho v^2 + p) = 0, \end{cases}$$

est appelé : **équations d'Euler**. Il reste à modéliser le comportement thermodynamique du gaz car nous avons seulement deux équations et trois inconnues (densité, vitesse, pression).

1. Cas numéro 1 : gaz sans pression, ou plus exactement à pression constante. Il s'agit d'une hypothèse très simplificatrice, on trouve alors le système

$$\begin{cases} \partial_t \rho + \partial_x(\rho v) = 0, \\ \partial_t(\rho v) + \partial_x(\rho v^2) = 0 \end{cases}$$

qui est formellement équivalent (pour des solutions régulières telles que  $\rho \neq 0$ ) à l'équation scalaire

$$\partial_t v + v \partial_x v = 0,$$

couplé à l'équation de conservation de la masse.

2. Cas numéro 2 : On va se placer dans un cadre simple dit *isentropique*. Dans ces conditions, le cours de thermodynamique<sup>4</sup> nous dit que la pression est liée à la densité par la relation

$$p = p_0 \rho^\gamma,$$

où  $p_0$  et  $\gamma$  sont des constantes positives données (par exemple  $\gamma = 1.4$  pour un gaz diatomique  $O_2$ ,  $H_2$ , ...).

Le système d'équations devient alors le système des équations d'Euler isentropique :

$$\begin{cases} \partial_t \rho + \partial_x(\rho v) = 0, \\ \partial_t(\rho v) + \partial_x(\rho v^2 + p_0 \rho^\gamma) = 0. \end{cases}$$

Ce système est fortement non-linéaire et hautement non trivial à résoudre et à comprendre.

3. Cas numéro 3 : On suppose désormais que l'écoulement est isotherme<sup>5</sup>. Dans le cas des gaz parfaits, le cours de thermodynamique<sup>6</sup> nous apprend que la loi liant la pression et la densité est linéaire de la forme  $p = c^2 \rho$ . On trouve donc un système complet de la forme

$$\begin{cases} \partial_t \rho + \partial_x(\rho v) = 0, \\ \partial_t(\rho v) + \partial_x(\rho v^2 + c^2 \rho) = 0. \end{cases}$$

On peut récrire ce système en posant  $\eta = \log \rho$  comme nouvelle variable, ce qui donne

$$\begin{cases} \partial_t \eta + v \partial_x \eta + \partial_x v = 0, \\ \partial_t v + v \partial_x v + c^2 \partial_x \eta = 0. \end{cases}$$

Si on regarde des régimes de petites vitesses (on dit qu'on linéarise autour de  $v = 0$ ) on peut considérer le système approché

$$\begin{cases} \partial_t \eta + \partial_x v = 0, \\ \partial_t v + c^2 \partial_x \eta = 0. \end{cases}$$

Et donc  $\eta$  et  $v$  vérifient la même équation des ondes

$$\begin{aligned} \partial_t^2 \eta - c^2 \partial_x^2 \eta &= 0, \\ \partial_t^2 v - c^2 \partial_x^2 v &= 0. \end{aligned}$$

Le réel  $c$  est la vitesse des ondes, dans le cas présent il s'agit de la vitesse du son.

Tous ces systèmes sont dits *hyperboliques* ce qui signifie, *grosso modo* que localement, on peut toujours trouver des changements de variable grâce auxquels le système ressemble à des équations de type transport.

4. que nous admettrons ici bien sûr...

5. la température reste constante durant l'évolution

6. Toujours lui !

## II Modèles de transport en dimension quelconque

On se place cette fois dans  $\mathbb{R}^d$  et on suppose donné un champ de vitesse  $v : (t, x) \in \mathbb{R} \times \mathbb{R}^d \mapsto v(t, x) \in \mathbb{R}^d$  que l'on suppose régulier et borné.

Dans le contexte de l'étude des équations de transport, il est d'usage d'utiliser un autre nom pour le flot associé au champ de vitesse, que nous avons rencontré dans la définition I.34.

### Définition II.3

Dans tout le chapitre les trajectoires associées au champ  $v$  sont appelées **courbes caractéristiques** (ou plus simplement **caractéristiques**) et sont notées  $X$ . Elles sont, par définition, solution de

$$\begin{cases} \frac{\partial}{\partial t} X(t, t_0, x_0) = v(t, X(t, t_0, x_0)), \\ X(t_0, t_0, x_0) = x_0. \end{cases}$$

### II.1 Théorème de Liouville

On commence par démontrer un théorème important, issu de la mécanique<sup>7</sup>. Celui-ci concerne l'évolution au cours du temps du déterminant jacobien de  $X$ . Cette quantité mesure les changements de volume sous l'action du flot.

Pour tout  $t \in \mathbb{R}, y \in \mathbb{R}^d$ , on va noter  $J(t, y)$  le déterminant Jacobien de  $y \mapsto X(t, 0, y)$  entre les instants 0 et  $t$ , au point  $y$ .

### Théorème II.4 (de Liouville)

Le jacobien  $J$  vérifie l'équation différentielle linéaire scalaire

$$\partial_t J(t, y) = (\operatorname{div} v)(t, X(t, 0, y)) J(t, y). \quad (\text{II.6})$$

Noter que  $y$  joue ici le rôle d'un paramètre.

#### Preuve :

On utilise le théorème de différentiabilité du flot qui nous donne l'équation (I.14) satisfaite par la matrice jacobienne

$$\partial_t (D_y X) = (D_x v)(t, X(t, 0, y)) \cdot (D_y X).$$

On rappelle maintenant que l'application déterminant  $\det : GL_d(\mathbb{R}) \rightarrow \mathbb{R}$  est différentiable et que sa différentielle est donnée par

$$(D \det)(M) \cdot H = (\det M) \operatorname{Tr}(M^{-1} \cdot H), \quad \forall M \in GL_d(\mathbb{R}), \forall H \in M_d(\mathbb{R}).$$

On peut donc calculer la dérivée de  $J$  comme composée de  $\det$  et  $D_y X$

$$\begin{aligned} \partial_t J(t, y) &= \partial_t (\det(D_y X)(t, 0, y)) \\ &= (\det(D_y X)(t, 0, y)) \operatorname{Tr}((D_y X(t, 0, y))^{-1} \cdot \partial_t (D_y X)(t, 0, y)) \\ &= J(t, y) \operatorname{Tr}((D_y X(t, 0, y))^{-1} \cdot (D_x v)(t, X(t, 0, y)) \cdot (D_y X(t, 0, y))) \\ &= J(t, y) \operatorname{Tr}((D_x v)(t, X(t, 0, y))) \\ &= J(t, y) (\operatorname{div} v)(t, X(t, 0, y)). \end{aligned}$$

■

### Corollaire II.5 (Champs à divergence nulle)

Si  $v$  est un champ de vecteurs à divergence nulle i.e.  $\operatorname{div} v = 0$ , alors le flot associé préserve les volumes.

#### Preuve :

Le théorème de Liouville montre que  $J(t, y)$  ne dépend pas de  $t$  et comme à l'instant initial nous avons  $X(0, \cdot) = \operatorname{Id}$ , dont le déterminant vaut 1, nous obtenons que

$$J(t, y) = 1, \quad \forall t \in \mathbb{R}, \forall y \in \mathbb{R}^d.$$

Le théorème de changement de variable montre alors que pour tout borélien  $A \subset \mathbb{R}^d$ , et tout temps  $t \in \mathbb{R}$ , le borélien  $X(t, 0, A)$  a le même volume (i.e. la même mesure) que  $A$ . Par la propriété de groupe du flot, on a immédiatement que  $X(t, s, A)$  a le même volume que  $A$  pour tous les instants  $s$  et  $t$ . ■

7. et qu'on retrouve sous différentes formes, par exemple en mécanique hamiltonienne

## II.2 Théorème de Reynolds

Pour tout ouvert borné régulier  $\Omega \subset \mathbb{R}^d$ , on note

$$\Omega_t := X(t, 0, \Omega), \quad \forall t \in \mathbb{R},$$

l'image de  $\Omega$  par le flot  $X$  entre les instants 0 et  $t$ . On a donc affaire à des domaines de  $\mathbb{R}^d$  qui évoluent avec le temps en suivant le flot comme illustré dans la figure II.2.

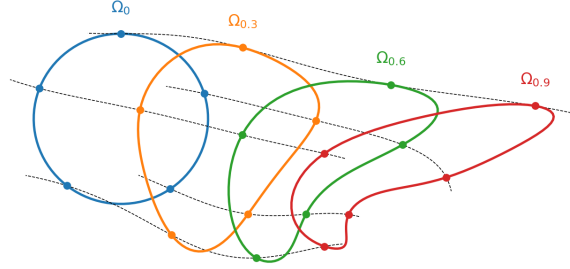


FIGURE II.2 – Tracé aux instants  $t = 0, 0.3, 0.6, 0.9$  d'un domaine  $\Omega_t$  évoluant avec le flot. Quelques trajectoires sont tracées en pointillés.

On se donne maintenant une fonction scalaire  $f : (t, x) \in \mathbb{R} \times \mathbb{R}^d \mapsto f(t, x) \in \mathbb{R}$ , de classe  $\mathcal{C}^1$ , et on souhaite pouvoir exprimer l'évolution en temps d'une quantité du type

$$t \mapsto \int_{\Omega_t} f(t, x) dx.$$

Pour ce faire on va pouvoir calculer sa dérivée grâce au théorème suivant

### Théorème II.6 (de Reynolds)

*Sous les hypothèses précédentes, on a pour tout  $t$*

$$\frac{d}{dt} \left( \int_{\Omega_t} f(t, x) dx \right) = \left( \int_{\Omega_t} [\partial_t f + \operatorname{div}(fv)](t, x) dx \right).$$

#### **Preuve :**

Donnons un nom à la quantité d'intérêt

$$F(t) := \int_{\Omega_t} f(t, x) dx.$$

La preuve du théorème consiste à effectuer le changement de variables  $x = X(t, 0, y)$  dans l'intégrale que l'on souhaite dériver, ce qui donne

$$F(t) = \int_{\Omega} f(t, X(t, 0, y)) J(t, y) dy,$$

où  $J$  est le déterminant Jacobien défini plus haut. Notons qu'on a utilisé ici que  $J > 0$  !

On dérive ensuite sous l'intégrale (toutes les quantités étant suffisamment régulières, ce calcul est bien justifié). Il vient

$$F'(t) = \int_{\Omega} \left[ \partial_t f(t, X(t, 0, y)) + (D_x f)(t, X(t, 0, y)) \cdot \partial_t X(t, 0, y) \right] J(t, y) + f(t, X(t, 0, y)) \partial_t J(t, y) dy$$

puis par le théorème de Liouville et (II.1)

$$= \int_{\Omega} \left[ \partial_t f(t, X(t, 0, y)) + v(t, X(t, 0, y)) \cdot (\nabla f)(t, X(t, 0, y)) + f(t, X(t, 0, y)) (\operatorname{div} v)(t, X(t, 0, y)) \right] J(t, y) dy.$$

Par le changement de variables inverse<sup>8</sup>, on revient à la variable  $x$  et on obtient

$$F'(t) = \int_{\Omega_t} \left[ \partial_t f(t, x) + v(t, x) \cdot (\nabla f)(t, x) + f(t, x)(\operatorname{div} v)(t, x) \right] dx,$$

ce qui donne le résultat voulu d'après la formule  $\operatorname{div}(fv) = f(\operatorname{div} v) + v \cdot \nabla f$  (voir (A.2) dans l'annexe A). ■

Il est intéressant de constater, à l'aide de la formule de Stokes (Théorème B.1 de l'annexe B), que l'on peut aussi écrire

$$\frac{d}{dt} \left( \int_{\Omega_t} f(t, x) dx \right) = \int_{\Omega_t} \partial_t f dx + \int_{\partial\Omega_t} f(v \cdot n) dx, \quad (\text{II.7})$$

où  $n$  désigne la normale unitaire sortante au bord de  $\Omega_t$ .

Cela montre que la variation de  $t \mapsto F(t)$  est due à deux phénomènes : la variation de la fonction  $f$  elle-même en fonction du temps et le fait que le bord du domaine d'intégration  $\Omega_t$  va se déplacer au cours du temps selon le champ  $v$ . En particulier, on voit que cette variation ne dépend que des valeurs de  $v \cdot n$  au bord de  $\Omega_t$ , ce qui est assez naturel<sup>9</sup>.

Sous la forme (II.7), la formule est très similaire à celle du Lemme II.2.

### Remarque II.7

Si applique le Théorème II.6 à la fonction constante égale à 1, on obtient une équation pour la variation du volume (=de la mesure de Lebesgue) du domaine  $\Omega_t$  au cours du temps

$$\frac{d}{dt} |\Omega_t| = \int_{\Omega_t} (\operatorname{div} v) dx.$$

Ceci redonne bien entendu le corollaire II.5 mais de façon peut-être plus pratique à manipuler. On voit également que si  $\operatorname{div} v \geq 0$  alors le volume de  $\Omega_t$  augmente au cours du temps (on a affaire à un phénomène de dilatation) et si  $\operatorname{div} v \leq 0$  alors le volume de  $\Omega_t$  diminue au cours du temps (on a alors affaire à un phénomène de compression).

## II.3 Exemple d'application à l'établissement d'une loi de conservation

On considère un système constitué d'un grand nombre de particules dans une description macroscopique<sup>10</sup> (comme nous l'avons fait dans la section I). Cela pourrait être des particules de fluide, des piétons dans la modélisation d'un mouvement de foules, etc ...

Soit  $\rho(t, x)$  la densité massique de particules (=masse des particules par unité de volume au temps  $t$  et au voisinage du point  $x$ ) et  $v(t, x)$  le champ de vitesse moyen des particules. On suppose que  $\rho$  et  $v$  sont des fonctions inconnues suffisamment régulières.

### Théorème II.8 (Loi de conservation de la masse)

On suppose qu'il n'y a ni création ni disparition de masse au cours de l'évolution, alors  $\rho$  et  $v$  sont reliées par l'équation aux dérivées partielles suivante

$$\partial_t \rho + \operatorname{div}(\rho v) = 0.$$

Cette équation s'appelle **équation de continuité** ou **équation de conservation de la masse**.

Si on suppose de plus que le système comporte deux types de particules (des rouges et des bleues par exemple), on peut alors définir  $\alpha(t, x) \in [0, 1]$  la fraction massique de la première espèce à l'instant  $t$  et au point  $x$ .

### Théorème II.9 (Loi de conservation des espèces)

Si la densité totale  $\rho$  est strictement positive, alors  $\alpha$  et  $v$  sont liées par l'équation aux dérivées partielles suivante

$$\partial_t \alpha + v \cdot \nabla \alpha = 0.$$

Cette équation s'appelle **équation de transport** ou **équation de convection**.

8. Observez que le jacobien  $J$  est désormais en facteur de tout ce qui est sous l'intégrale

9. Imaginez par exemple ce que signifierait que  $v \cdot n = 0$  et pourquoi il est clair alors que seul le premier terme doit demeurer dans l'égalité

10. Cela signifie qu'on se place à une échelle où les particules ne peuvent pas être étudiées individuellement et sont plutôt considérées comme un milieu continu



**Preuve (du théorème II.8):**

Pour tout ouvert  $\Omega$  initial, la quantité totale de matière dans le domaine  $\Omega_t$  reste constante au cours du temps puisque le domaine d'intégration évolue selon le champ  $v$  et donc contient toujours exactement les mêmes particules. On a donc

$$0 = \frac{d}{dt} \int_{\Omega_t} \rho(t, x) dx = \int_{\Omega_t} \partial_t \rho + \operatorname{div}(\rho v) dx.$$

Ceci étant vrai pour tout  $t$  et tout  $\Omega$ , on en déduit que

$$\int_U (\partial_t \rho + \operatorname{div}(\rho v)) dx = 0, \quad \forall t \in \mathbb{R}, \quad \forall U \text{ ouvert.}$$

D'après le lemme de Du Bois-Reymond (ou plutôt sa variante - Lemme A.5), on en déduit que l'équation suivante est vérifiée

$$\partial_t \rho + \operatorname{div}(\rho v) = 0.$$

**Preuve (du théorème II.9):**

Si on applique le même raisonnement à la conservation de la masse de la première espèce on trouve

$$\partial_t(\alpha \rho) + \operatorname{div}(\alpha \rho v) = 0.$$

En utilisant les formules de dérivation des produits on obtient

$$0 = \alpha(\partial_t \rho + \operatorname{div}(\rho v)) + \rho(\partial_t \alpha + v \cdot \nabla \alpha)$$

et on reconnaît l'équation de conservation de la masse dans le premier terme qui est non nul. Comme on a supposé que  $\rho$  était partout non nulle, on a finalement bien obtenu l'équation attendue pour  $\alpha$ . ■

Ce type de raisonnement est à la base de l'obtention de très nombreux modèles dans lesquels intervient un phénomène de *transport* de certaines quantités par un champ de vitesse.

### III Solutions classiques des équations de transport

On va maintenant résoudre les équations obtenues dans les sections précédentes.

#### III.1 Cas général de l'équation de convection

On se donne un champ de vitesse  $v : \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}^d$  que l'on suppose  $C^1$  et borné.

Pour  $u_0 \in C^1(\mathbb{R}^d)$ , on cherche à résoudre le problème de Cauchy suivant

$$\begin{cases} \partial_t u + v(t, x) \cdot \nabla u = 0, & \forall t > 0, \forall x \in \mathbb{R}^d, \\ u(t = 0, x) = u_0(x), & \forall x \in \mathbb{R}^d, \end{cases} \quad (\text{II.8})$$

c'est-à-dire que l'on cherche une fonction  $u \in C^1([0, +\infty[ \times \mathbb{R}^d)$  solution de ces équations.

**Remarque II.10 (Attention!)**

Bien que le problème (II.8) porte le nom de **problème de Cauchy** de part le fait qu'il s'agit d'une équation d'évolution en temps à laquelle on ajoute une donnée initiale, il faut bien se garder d'invoquer ici un quelconque théorème de Cauchy-Lipschitz. L'équation n'est pas une équation différentielle ordinaire car si on l'écrit de façon abstraite sous la forme

$$\partial_t u = \mathcal{F}(t, u),$$

alors  $\mathcal{F}$  n'est pas une fonction au sens usuel mais plutôt un opérateur différentiel qui fait intervenir des dérivées (par rapport à  $x$ ) de la fonction inconnue.

Les résultats obtenus dans le chapitre I ne peuvent donc pas être appliqués dans ce contexte pour résoudre des équations aux dérivées partielles d'évolution.

**Théorème II.11**

Pour tout champ  $v$  régulier et borné, pour toute donnée initiale  $u_0 \in C^1(\mathbb{R}^d)$ , il existe une unique solution  $u$  de (II.8) et celle-ci est donnée par la formule

$$u(t, x) = u_0(X(0, t, x)), \quad \forall t > 0, \forall x \in \mathbb{R}^d, \quad (\text{II.9})$$

où  $X$  désigne les caractéristiques associées au champ  $v$ .

Dans le cadre de l'étude des équations de transport, les courbes intégrales du champ  $t \mapsto X(t, 0, x_0)$  sont appelées les **courbes caractéristiques** du problème. Le théorème ci-dessus consiste à montrer que la solution est constante le long des caractéristiques, voir Figure II.3 (de façon traditionnelle dans ce domaine des mathématiques, le temps est représenté sur l'axe des ordonnées et la position sur l'axe des abscisses).

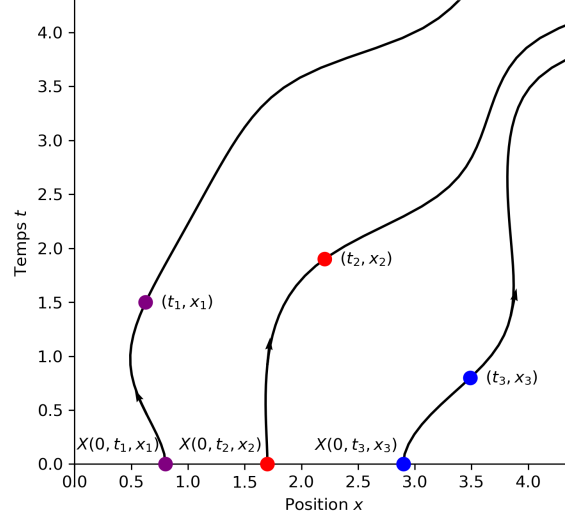


FIGURE II.3 – Illustration de la méthode des caractéristiques en dimension  $d = 1$

**Preuve :**

- Commençons par montrer que, si une solution existe, elle vérifie nécessairement la formule explicite (II.9). Cela montrera l'unicité de la solution par la même occasion.

Soit donc  $u$  une solution de (II.8). On fixe un point  $x_0 \in \mathbb{R}^d$  et on étudie les valeurs de  $u$  le long de la courbe caractéristique  $t \mapsto (t, X(t, 0, x_0))$  dans le plan  $(t, x)$ . Plus précisément, si on pose

$$\varphi(t) := u(t, X(t, 0, x_0)), \quad \forall t \geq 0,$$

on observe que

$$\varphi(0) = u(0, x_0) = u_0(x_0),$$

et que (vu la définition de  $X$ )

$$\begin{aligned} \varphi'(t) &= \partial_t u(t, X(t, 0, x_0)) + (\partial_t X(t, 0, x_0)) \cdot \nabla u(t, X(t, 0, x_0)) \\ &= \partial_t u(t, X(t, 0, x_0)) + v(t, X(t, 0, x_0)) \cdot \nabla u(t, X(t, 0, x_0)) \\ &= (\partial_t u + v \cdot \nabla u)(t, X(t, 0, x_0)) \\ &= 0. \end{aligned}$$

Ainsi, la fonction  $\varphi$  est constante, ce qui donne

$$u(t, X(t, 0, x_0)) = u_0(x_0).$$

Pour tout  $x \in \mathbb{R}^d$ , on peut inverser la caractéristique grâce à la propriété de groupe du flot et poser  $x_0 = X(0, t, x)$ , pour obtenir

$$u(t, x) = u_0(X(0, t, x)),$$

qui est bien la formule attendue.

- Réciproquement, si on définit  $u$  par la formule (II.9), alors nous savons que  $u$  est de classe  $\mathcal{C}^1$  grâce aux propriétés de régularité du flot et de  $u_0$ . Par ailleurs, nous avons pour tout  $t$

$$u(t, X(t, 0, x_0)) = u_0(X(0, t, X(t, 0, x_0))) = u_0(x_0),$$

et donc la fonction  $u(t, X(t, 0, x_0))$  ne dépend pas du temps. Un calcul de dérivée identique au précédent, montre donc que, pour tout  $t$  et tout  $x_0$ , nous avons

$$(\partial_t u + v \cdot \nabla u)(t, X(t, 0, x_0)) = 0.$$

Pour tout  $x \in \mathbb{R}^d$ , on peut appliquer cette formule à  $x_0 = X(0, t, x)$  et ainsi obtenir

$$(\partial_t u + v \cdot \nabla u)(t, x) = 0, \quad \forall t \geq 0, \forall x \in \mathbb{R}^d,$$

ce qui est bien l'équation de transport attendue. ■

### Remarque II.12 (Ingrédient principal)

*L'essentiel de la preuve ci-dessus réside dans le calcul de la dérivée de la fonction composée*

$$\frac{d}{dt}(u(t, X(t, 0, x_0))) = (\partial_t u + v \cdot \nabla u)(t, X(t, 0, x_0)), \quad (\text{II.10})$$

*valable pour toute fonction  $u$  dérivable et pour tout champ de vecteurs régulier  $v$  et dont on note  $X$  le flot. On retrouvera ce calcul à plusieurs reprises dans la suite.*

## III.2 Cas particuliers importants

Il y a quelques cas particuliers où les caractéristiques du champ de vitesse peuvent se calculer explicitement et donc pour lesquels on a des formules explicites des solutions de l'équation de transport.

— **Vitesse constante :** Si  $v(t, x) = v$  est indépendant de  $t$  et  $x$ , alors les caractéristiques se calculent aisément

$$X(t, s, x_0) = x_0 + (t - s)v,$$

et donc la solution du problème de transport s'écrit

$$u(t, x) = u_0(x - tv), \quad \forall t \in \mathbb{R}, \forall x \in \mathbb{R}^d.$$

La figure II.4 montre un exemple d'évolution en temps de la solution d'un problème de transport 1D à vitesse constante  $v = 2$  pour une donnée initiale gaussienne.

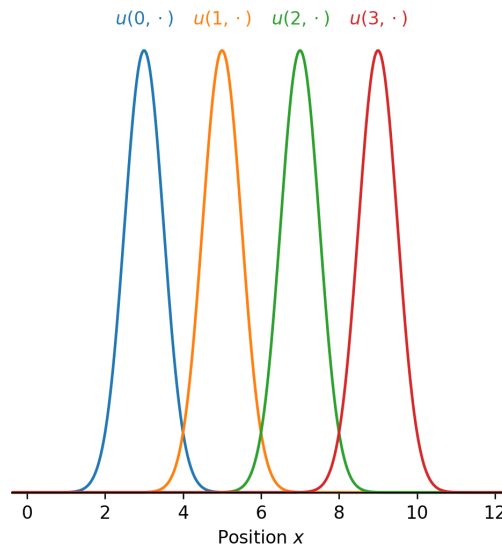


FIGURE II.4 – Solution d'un problème de transport à vitesse constante aux temps  $t = 0, 1, 2, 3$

— **Vitesse uniforme en espace :** On suppose que  $v(t, x) = v(t)$  ne dépend que du temps. Les caractéristiques sont alors données par

$$X(t, s, x_0) = x_0 + \int_s^t v(\tau) d\tau.$$

On trouve donc

$$u(t, x) = u_0\left(x + \int_t^0 v(\tau) d\tau\right) = u_0\left(x - \int_0^t v(\tau) d\tau\right).$$

La figure II.5 montre un exemple d'évolution en temps de la solution d'un problème de transport 1D à vitesse uniforme mais dépendant du temps pour la même donnée initiale que ci-dessus. On observe que la dynamique de la solution est différente mais que celle-ci ne se déforme pas au cours du temps.

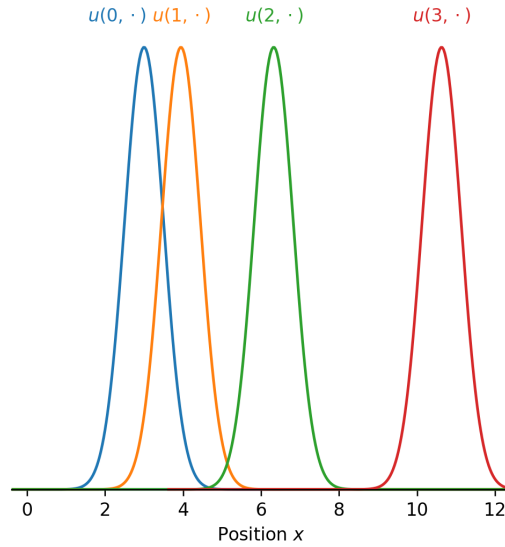


FIGURE II.5 – Solution d'un problème de transport à vitesse uniforme mais non constante  $v(t) = 0.7 + 0.5t$  toujours aux temps  $t = 0, 1, 2, 3$

A titre de comparaison, on donne dans la figure II.6, l'allure de la solution lorsque le champ de vitesse n'est plus uniforme par rapport à  $x$ . On observe que cette fois la solution se déforme au cours du temps sous l'influence du transport.

- **Champ tournant en 2D :** On se place en dimension 2. On suppose qu'il existe  $\omega \in \mathbb{R}$  tel que

$$v(x) = \omega \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}, \quad \forall x \in \mathbb{R}^2.$$

Les caractéristiques associées sont alors des cercles concentriques donnés par

$$X(t, s, x_0) = \begin{pmatrix} \cos(\omega(t-s)) & -\sin(\omega(t-s)) \\ \sin(\omega(t-s)) & \cos(\omega(t-s)) \end{pmatrix} x_0.$$

La solution de l'équation est alors donnée par

$$u(t, x) = u_0(R_{-\omega t}x),$$

où  $R_\theta$  est la matrice de la rotation d'angle  $\theta$  dans le plan orienté dans le sens trigonométrique.

- **Champ tournant en 3D :** On se place en dimension 3 et on suppose qu'il existe un vecteur non nul  $\omega \in \mathbb{R}^3$  tel que

$$v(t, x) = \omega \wedge x, \quad \forall t \in \mathbb{R}, \forall x \in \mathbb{R}^3.$$

On note  $P$  le plan vectoriel orthogonal à  $\omega$  ( $P = \omega^\perp$ ). On constate que la composante selon  $\omega$  des caractéristiques est constante

$$X(t, s, x_0) \cdot \omega = x_0 \cdot \omega, \quad \forall t, s \in \mathbb{R}, \forall x_0 \in \mathbb{R}^3.$$

Par ailleurs, la projection sur  $P$  de la caractéristique  $X(t, s, x_0)$  vérifie la même équation que celle du cas précédent modulo le choix judicieux de l'orientation. Les trajectoires sont donc des cercles tournant autour de l'axe porté par  $\omega$  et de vitesse angulaire  $|\omega|$ .

### III.3 Autres applications de la méthode des caractéristiques

La méthode présentée ci-dessus et essentiellement basée sur la formule (II.10) permet de résoudre beaucoup d'autres équations dont la partie principale<sup>11</sup> est de type *transport*. On va donner ci-dessous quelques exemples.

On se donne toujours un champ  $v$  de classe  $\mathcal{C}^1$  et borné et  $X$  désigne à nouveau les caractéristiques associées à ce champ.

11. c'est-à-dire l'ensemble des termes de l'équation qui comportent des dérivées de la fonction inconnue

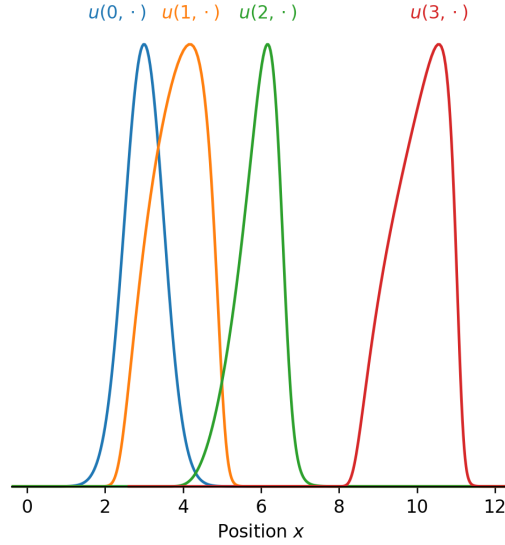


FIGURE II.6 – Solution d'un problème de transport à vitesse variable  $v(t, x) = 0.7 + 0.5t + 0.5 \sin(2x)$  toujours aux temps  $t = 0, 1, 2, 3$

### III.3.a Termes sources

#### Théorème II.13

Pour tout  $u_0 \in C^1(\mathbb{R}^d)$ , et tout  $f \in C^0(\mathbb{R} \times \mathbb{R}^d)$ , il existe une unique solution  $u \in C^1([0, +\infty[ \times \mathbb{R}^d)$  au problème de Cauchy suivant

$$\begin{cases} \partial_t u + v(t, x) \cdot \nabla u = f(t, x), & \forall t > 0, \forall x \in \mathbb{R}, \\ u(t = 0, x) = u_0(x), & \forall x \in \mathbb{R}. \end{cases} \quad (\text{II.11})$$

#### Preuve :

On procède à nouveau par analyse/synthèse en supposant que la solution  $u$  existe. On fixe alors  $x_0 \in \mathbb{R}^d$  et on s'intéresse à l'évolution de la solution  $u$ , le long de la caractéristique issue de la donnée de Cauchy  $(0, x_0)$ , c'est-à-dire de la quantité

$$\varphi(t) := u(t, X(t, 0, x_0)), \quad \forall t \in \mathbb{R}.$$

On a toujours  $\varphi(0) = u(0, X(0, 0, x_0)) = u(0, x_0) = u_0(x_0)$  et on peut calculer la dérivée de  $\varphi$  comme dans (II.10)

$$\varphi'(t) = (\partial_t u + v \cdot \nabla u)(t, X(t, 0, x_0)).$$

Cette fois cette quantité n'est pas nulle mais on peut utiliser l'équation vérifiée par  $u$  pour obtenir la relation

$$\varphi'(t) = f(t, X(t, 0, x_0)),$$

qui ne dépend que des données du problème et permet donc de déterminer  $\varphi$  par intégration

$$\varphi(t) = u_0(x_0) + \int_0^t f(s, X(s, 0, x_0)) ds.$$

Revenant à la définition de  $\varphi$ , on voit qu'on a obtenu

$$u(t, X(t, 0, x_0)) = u_0(x_0) + \int_0^t f(s, X(s, 0, x_0)) ds, \quad \forall t \in \mathbb{R}, \forall x_0 \in \mathbb{R}^d.$$

Il reste à inverser la caractéristique : on fixe  $t$  et  $x$  et on pose  $x_0 = X(0, t, x)$  pour obtenir grâce aux propriétés de groupe du flot

$$u(t, x) = u_0(X(0, t, x)) + \int_0^t f(s, X(s, t, x)) ds.$$

Cela montre l'unicité de la solution  $u$  du problème et fournit une formule (relativement) explicite de la solution. On peut maintenant reprendre tous les calculs à l'envers pour établir que la formule ci-dessus définit bien une fonction de classe  $C^1$  qui est solution de (II.11). ■

### III.3.b Termes de réaction. Equation de continuité

On suppose toujours le champ  $v$  donné régulier et borné. On s'intéresse d'abord à l'équation de continuité

$$\partial_t \rho + \operatorname{div}(\rho v) = 0,$$

que l'on a discuté plus haut. Cette équation n'est pas une équation de transport usuelle car on a un terme  $\operatorname{div}(\rho v)$  à la place de  $v \cdot \nabla \rho$ . On peut néanmoins utiliser les formules de dérivation de produits (formule (A.2) de l'annexe A) pour écrire l'équation sous la forme

$$\partial_t \rho + v \cdot \nabla \rho + (\operatorname{div} v) \rho = 0.$$

Comme  $v$  est une donnée du problème,  $\operatorname{div} v$  est un champ scalaire connu et donc la partie principale de cette équation est bien du type *transport*. Remarquons que si  $\operatorname{div} v = 0$  on a en fait exactement une équation de transport. En général toutefois, cette équation est un cas particulier d'un problème de la forme

$$\partial_t u + v \cdot \nabla u + au = 0,$$

où  $a : (t, x) \in \mathbb{R} \times \mathbb{R}^d \mapsto a(t, x) \in \mathbb{R}$  est une fonction continue donnée. Le terme  $au$  dans cette équation est parfois appelé terme de **réaction**.

#### Théorème II.14

Pour tout  $u_0 \in C^1(\mathbb{R}^d)$ , et tout  $a \in C^0(\mathbb{R} \times \mathbb{R}^d)$ , il existe une unique solution  $u \in C^1([0, +\infty[ \times \mathbb{R}^d)$  au problème de Cauchy suivant

$$\begin{cases} \partial_t u + v(t, x) \cdot \nabla u + a(t, x)u = 0, & \forall t > 0, \forall x \in \mathbb{R}^d, \\ u(t = 0, x) = u_0(x), & \forall x \in \mathbb{R}^d. \end{cases} \quad (\text{II.12})$$

De plus, si  $u_0(x) \geq 0$  pour tout  $x$ , alors  $u(t, x) \geq 0$  pour tout  $t, x$ .

#### Preuve :

On reprend la démarche maintenant usuelle en essayant de déterminer l'évolution de  $u$  le long des caractéristiques du champ  $v$  en posant

$$\varphi(t) := u(t, X(t, 0, x_0)),$$

où  $u$  est une solution potentielle du système. On a toujours la donnée initiale  $\varphi(0) = u_0(x_0)$ , et on trouve

$$\varphi'(t) = (\partial_t u + v \cdot \nabla u)(t, X(t, 0, x_0)) = -a(t, X(t, 0, x_0))u(t, X(t, 0, x_0)).$$

On reconnaît dans le membre de droite la définition de  $\varphi$  ce qui fournit

$$\varphi'(t) = -a(t, X(t, 0, x_0))\varphi(t).$$

Cette fois, on n'obtient pas une formule explicite pour  $\varphi$ , mais une équation différentielle linéaire du premier ordre que l'on peut résoudre pour obtenir

$$\varphi(t) = u_0(x_0) \exp \left( - \int_0^t a(s, X(s, 0, x_0)) ds \right).$$

On a donc obtenu

$$u(t, X(t, 0, x_0)) = u_0(x_0) \exp \left( - \int_0^t a(s, X(s, 0, x_0)) ds \right),$$

ce qui permet, comme à chaque fois, de retrouver  $u$ , par inversion de la caractéristique

$$u(t, x) = u_0(X(0, t, x)) \exp \left( - \int_0^t a(s, X(s, t, x)) ds \right).$$

On obtient encore une fois une formule qui prouve l'unicité d'une éventuelle solution. Il reste à nouveau à remonter les calculs pour prouver que la formule obtenue est bien une solution du problème.

Il est clair sur la formule obtenue que si  $u_0$  est positive, alors  $u$  l'est aussi puisque c'est le produit d'une valeur de  $u$  avec une exponentielle. ■

## IV Solutions faibles de l'équation de transport

On va chercher à définir une notion un peu plus générale de solution pour les équations de transport. Il y a plusieurs motivations pour cela (dont certaines, et pas des moindres, dépassent le cadre de ce cours) mais on peut par exemple citer le fait que les données initiales et/ou au bord pour le problème peuvent ne pas être dérivables (et même pas continues) : on peut penser par exemple à la modélisation du démarrage à un feu rouge dans le cadre du trafic routier. Au moment où le feu passe au vert, la densité est nulle en aval et non nulle en amont du feu, il y a donc une discontinuité dans la donnée initiale.

Cet objectif est d'autant plus raisonnable que l'on constate que, par exemple, la formule

$$u(t, x) = u_0(X(0, t, x)),$$

obtenue dans le théorème II.11 pour la solution générale de l'équation de transport est parfaitement bien définie même si  $u_0$  est seulement une fonction continue et même seulement une fonction mesurable ...

On dispose d'autre part de la théorie des distributions qui permet de donner un sens très général à la notion de solution d'une EDP. On va tenter ici de faire un lien entre tous ces concepts.

### IV.1 Définition des solutions faibles

#### Définition II.15

Soit  $v$  un champ de vecteur de classe  $C^1$  et borné. Pour toute donnée initiale  $u_0 \in L^1_{loc}(\mathbb{R}^d)$  on dit qu'une fonction  $u \in L^1_{loc}([0, +\infty[ \times \mathbb{R}^d)$  est une solution faible du problème de Cauchy (II.8) si elle vérifie, pour toute fonction test  $\varphi \in C^\infty(\mathbb{R} \times \mathbb{R}^d)$  l'égalité

$$\int_0^{+\infty} \int_{\mathbb{R}^d} u(t, x) [\partial_t \varphi + \operatorname{div}(v\varphi)](t, x) dx dt + \int_{\mathbb{R}^d} u_0(x) \varphi(0, x) dx = 0. \quad (\text{II.13})$$

On se convainc que l'intégrabilité locale de  $u$  et de  $u_0$  suffit à donner un sens à tous les termes de cette égalité car  $\varphi$  est à support compact.

#### Remarque II.16

Il est assez clair dans la définition que les valeurs de  $\varphi(t, x)$  pour  $t < 0$  ne jouent aucun rôle dans la définition. Toute solution faible  $u$  vérifie donc la formule (II.13) dès que  $\varphi \in C^\infty(\mathbb{R} \times \mathbb{R}^d)$  vérifie

$$\operatorname{Supp}(\varphi) \cap ([0, +\infty[ \times \mathbb{R}^d) \text{ est compact.}$$

Commençons par faire le lien avec les autres notions de solution que nous connaissons et ainsi justifier la pertinence de cette définition.

#### Proposition II.17

- Si  $u \in C^1([0, +\infty[ \times \mathbb{R}^d)$  est une solution classique du problème de transport (II.8), alors  $u$  est aussi une solution faible du problème.
- Si  $u \in L^1_{loc}([0, +\infty[ \times \mathbb{R}^d)$  est une solution faible du problème de transport alors elle vérifie, au sens des distributions, l'équation

$$\partial_t u + \operatorname{div}(uv) - (\operatorname{div} v)u = 0, \text{ dans } \mathcal{D}'([0, +\infty[ \times \mathbb{R}^d). \quad (\text{II.14})$$

Si on sait de plus que  $u \in C^1([0, +\infty[ \times \mathbb{R}^d)$  alors elle vérifie l'équation de transport au sens classique

$$\partial_t u + v \cdot \nabla u = 0.$$

Remarquez bien la forme particulière de l'équation (II.14) qu'il faut bien se garder de mettre sous la forme  $\partial_t u + v \cdot \nabla u = 0$  car le produit  $v \cdot \nabla u$  n'est a priori pas défini si  $u$  est une distribution et  $v$  un champ de vecteur qui n'est pas de classe  $C^\infty$ .

#### Preuve :

- Il s'agit d'une simple intégration par parties. On prend une fonction test  $\varphi$  comme dans la définition, on multiplie l'équation vérifiée par  $u$  par  $\varphi$ , on intègre sur  $]0, +\infty[ \times \mathbb{R}^d$  puis on intègre par parties, en utilisant les formules données en annexe.

$$\begin{aligned}
0 &= \int_0^{+\infty} \int_{\mathbb{R}^d} \varphi(\partial_t u + v \cdot \nabla u) dx dt \\
&= \int_{\mathbb{R}^d} \left( \int_0^{+\infty} \varphi \partial_t u dt \right) dx + \int_0^{+\infty} \left( \int_{\mathbb{R}^d} \varphi v \cdot \nabla u dx \right) dt \\
&= \int_{\mathbb{R}^d} \left( - \int_0^{+\infty} \partial_t \varphi u dt - \varphi(0, x) u(0, x) \right) dx - \int_0^{+\infty} \left( \int_{\mathbb{R}^d} \operatorname{div}(\varphi v) u dx \right) dt.
\end{aligned}$$

En utilisant la donnée initiale  $u(0, x) = u_0(x)$  et en rassemblant les autres termes, on trouve l'équation attendue.

- On prend une fonction test  $\varphi \in \mathcal{D}(]0, +\infty[ \times \mathbb{R}^d)$ . Son prolongement par 0 à tout  $\mathbb{R} \times \mathbb{R}^d$  (noté encore  $\varphi$ ) est bien une fonction de  $\mathcal{D}(\mathbb{R} \times \mathbb{R}^d)$  et on peut donc appliquer la définition, tout en constatant que  $\varphi(0, x) = 0$  pour tout  $x$ . On trouve

$$0 = \int_0^{+\infty} \int_{\mathbb{R}^d} u(t, x) [\partial_t \varphi + \operatorname{div}(\varphi v)](t, x) dx dt,$$

et si on développe le terme  $\operatorname{div}(\varphi v)$  on obtient

$$0 = \int_0^{+\infty} \int_{\mathbb{R}^d} u \partial_t \varphi + uv \nabla \varphi + u(\operatorname{div} v) \varphi dx dt,$$

et donc

$$0 = \langle u, \partial_t \varphi \rangle_{\mathcal{D}', \mathcal{D}} + \langle uv, \nabla \varphi \rangle_{\mathcal{D}', \mathcal{D}} + \langle u(\operatorname{div} v), \varphi \rangle_{\mathcal{D}', \mathcal{D}}.$$

Par définition des dérivées au sens des distributions (plus exactement de la divergence au sens des distributions) on obtient

$$0 = \langle -\partial_t u, \varphi \rangle_{\mathcal{D}', \mathcal{D}} - \langle \operatorname{div}(uv), \varphi \rangle_{\mathcal{D}', \mathcal{D}} + \langle u(\operatorname{div} v), \varphi \rangle_{\mathcal{D}', \mathcal{D}},$$

soit encore

$$0 = \langle -\partial_t u - \operatorname{div}(uv) + u(\operatorname{div} v), \varphi \rangle_{\mathcal{D}', \mathcal{D}},$$

ce qui donne bien l'équation attendue.

Si  $u$  est en fait une fonction régulière, alors les dérivées au sens de distributions, coïncident avec les dérivées usuelles (Proposition A.18) et on a donc

$$\partial_t u + \operatorname{div}(uv) - u(\operatorname{div} v) = 0,$$

et on reconnaît l'équation de transport grâce à la formule (A.2). ■

En dimension  $d = 1$ , on peut déterminer des solutions faibles régulières par morceaux comme suit.

### Proposition II.18

On se donne une abscisse  $a \in \mathbb{R}$ , et on définit les deux sous-domaines

$$\Omega_g = \{(t, x), x < X(t, 0, a)\}, \text{ et } \Omega_d = \{(t, x), x > X(t, 0, a)\},$$

délimités par une interface  $\Gamma$  qui n'est autre que la caractéristique issue de  $a$

$$\Gamma = \{(t, X(t, 0, a)), t > 0\}.$$

Soit  $u$  une fonction mesurable définie sur  $]0, +\infty[ \times \mathbb{R}$  et  $\mathcal{C}^1$  par morceaux sur les domaines  $\Omega_g$  et  $\Omega_d$ . On rappelle que cela signifie qu'il existe deux fonctions  $u_g, u_d \in \mathcal{C}^1(\mathbb{R} \times \mathbb{R})$  telles que

$$u(t, x) = \begin{cases} u_g(t, x), & \text{si } (t, x) \in \Omega_g, \\ u_d(t, x), & \text{si } (t, x) \in \Omega_d. \end{cases}$$

On impose aucune relation a priori entre  $u_g$  et  $u_d$  et donc  $u$  peut être (la classe d') une fonction discontinue.

Si  $u_g$  (resp.  $u_d$ ) est solution classique de l'équation de transport dans  $\Omega_g$  (resp. dans  $\Omega_d$ ), alors  $u$  est une solution faible de l'équation de transport dans le domaine entier  $\mathbb{R} \times \mathbb{R}$ .



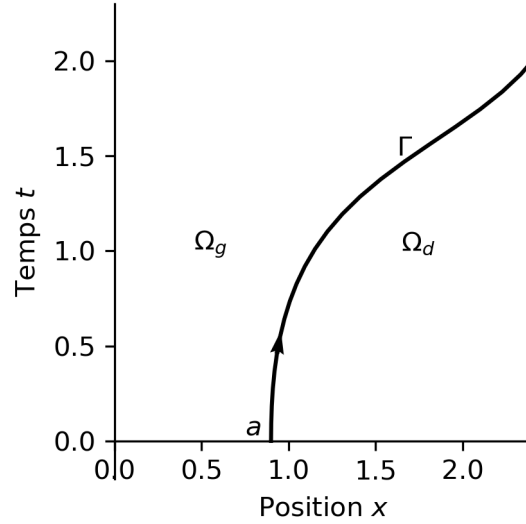


FIGURE II.7 – Solutions faibles régulières par morceaux

La situation décrite ici est illustrée dans la figure II.7.

Par exemple la fonction constante par morceaux définie par

$$u(t, x) = \begin{cases} u_g, & \text{si } (t, x) \in \Omega_g, \\ u_d, & \text{si } (t, x) \in \Omega_d, \end{cases}$$

avec  $u_g, u_d \in \mathbb{R}$  quelconque est solution faible du problème de Cauchy (II.8) pour la donnée initiale

$$u_0(x) = \begin{cases} u_g, & \text{si } x < a, \\ u_d, & \text{si } x > a. \end{cases}$$

**Preuve :**

Comme  $u = u_g$  dans  $\Omega_g$  et qu'elle vérifie l'équation dans  $\Omega_g$ , on peut écrire

$$\begin{aligned} 0 &= \int_{\Omega_g} \varphi(\partial_t u + v \partial_x u) dx dt \\ 0 &= \int_{\Omega_g} \varphi(\partial_t u_g + v \partial_x u_g) dx dt \\ &= \int_0^{+\infty} \int_{-\infty}^{X(t,0,a)} \varphi(\partial_t u_g + v \partial_x u_g) dx dt \\ &= \int_0^{+\infty} \left( \int_{-\infty}^{X(t,0,a)} \varphi \partial_t u_g dx \right) dt + \int_0^{+\infty} \left( \int_{-\infty}^{X(t,0,a)} \varphi v \partial_x u_g dx \right) dt. \end{aligned} \tag{II.15}$$

L'idée est maintenant de faire une intégration par parties en temps dans le premier terme et en espace dans le second. Traitons les deux cas séparément :

— *Intégration par parties en temps.*

On utilise le Lemme II.2 pour effectuer le calcul suivant

$$\begin{aligned} \frac{d}{dt} \left( \int_{-\infty}^{X(t,0,a)} u_g \varphi dx \right) &= \int_{-\infty}^{X(t,0,a)} (\partial_t u_g) \varphi dx + \int_{-\infty}^{X(t,0,a)} u_g (\partial_t \varphi) dx \\ &\quad + \left( \frac{d}{dt} X(t, 0, a) \right) u_g(t, X(t, 0, a)) \varphi(t, X(t, 0, a)) \\ &= \int_{-\infty}^{X(t,0,a)} (\partial_t u_g) \varphi dx + \int_{-\infty}^{X(t,0,a)} u_g (\partial_t \varphi) dx \\ &\quad + (\varphi u_g v)(t, X(t, 0, a)). \end{aligned}$$

On intègre maintenant cette égalité par rapport au temps entre 0 et  $+\infty$  pour obtenir

$$-\int_{-\infty}^a u_g(0, x) \varphi(0, x) dx = \int_0^{+\infty} \int_{-\infty}^{X(t, 0, a)} (\partial_t u_g) \varphi dx dt + \int_0^{+\infty} \int_{-\infty}^{X(t, 0, a)} u_g (\partial_t \varphi) dx dt + \int_0^{+\infty} (\varphi u_g v)(t, X(t, 0, a)) dt.$$

Remettant les termes dans l'ordre, on peut exprimer le premier terme du membre de droite de (II.15) sous la forme

$$\int_0^{+\infty} \int_{-\infty}^{X(t, 0, a)} \varphi (\partial_t u_g) dx dt = - \int_0^{+\infty} \int_{-\infty}^{X(t, 0, a)} u_g (\partial_t \varphi) dx dt - \int_{-\infty}^a u_g(0, x) \varphi(0, x) dx - \int_0^{+\infty} (\varphi u_g v)(t, X(t, 0, a)) dt. \quad (\text{II.16})$$

— *Intégration par parties en espace.*

Le calcul est plus simple et utilise simplement que  $\varphi$  est à support compact en espace. On trouve

$$\int_0^{+\infty} \left( \int_{-\infty}^{X(t, 0, a)} \varphi v \partial_x u_g dx \right) dt = - \int_0^{+\infty} \left( \int_{-\infty}^{X(t, 0, a)} \partial_x (\varphi v) u_g dx \right) dt + \int_0^{+\infty} (\varphi u_g v)(t, X(t, 0, a)) dt. \quad (\text{II.17})$$

Si on somme (II.16) et (II.17), on observe que les deux derniers termes (ceux portés par l'interface  $\Gamma$ ) se compensent. En reportant dans (II.15), on obtient finalement

$$0 = - \int_{\Omega_g} u_g (\partial_t \varphi + \partial_x (\varphi v)) dx dt - \int_{-\infty}^a u_g(0, x) \varphi(0, x) dx.$$

Un calcul similaire dans le domaine  $\Omega_d$  donne

$$0 = - \int_{\Omega_d} u_d (\partial_t \varphi + \partial_x (\varphi v)) dx dt - \int_a^{+\infty} u_d(0, x) \varphi(0, x) dx.$$

Comme  $u = u_g$  dans  $\Omega_g$  et  $u = u_d$  dans  $\Omega_d$ , on peut sommer ces équations et obtenir finalement (en enlevant tous les signes  $-$ )

$$0 = \int_0^{+\infty} \int_{\mathbb{R}} u (\partial_t \varphi + \partial_x (\varphi v)) dx dt + \int_{-\infty}^a u_g(0, x) \varphi(0, x) dx + \int_a^{+\infty} u_d(0, x) \varphi(0, x) dx.$$

On trouve donc bien la définition d'une solution faible (II.13) avec pour donnée initiale la fonction

$$u_0(x) = \begin{cases} u_g(0, x), & \text{si } x < a, \\ u_d(0, x), & \text{si } x > a. \end{cases}$$

■

## IV.2 Validité de la formule de représentation par les caractéristiques

Le but de cette section est de prouver que la formule (II.9) définit bien une solution faible de l'équation de transport pour toute donnée initiale  $u_0$  peu régulière, disons  $L^1_{loc}(\mathbb{R}^d)$ .

**Cas de la vitesse constante** On commence par le cas de la vitesse constante qui est plus simple.

Soit donc  $u_0 \in L^1_{loc}(\mathbb{R}^d)$  quelconque et  $v \in \mathbb{R}^d$  un vecteur constant. On définit

$$u(t, x) := u_0(x - tv), \quad \forall t > 0, \forall x \in \mathbb{R}^d.$$

Remarquons que cette définition est bien consistante avec les classes de fonction presque partout (i.e. la classe de  $u$  ne dépend pas du choix de  $u_0$  parmi tous les éléments de sa classe d'équivalence).

Prenons une fonction test  $\varphi \in \mathcal{D}(\mathbb{R} \times \mathbb{R}^d)$  et effectuons le calcul suivant (en utilisant le changement de variable  $(t, x) \mapsto (t, y = x - tv)$ )

$$\begin{aligned} \int_0^{+\infty} \int_{\mathbb{R}^d} u(t, x) (\partial_t \varphi + v \cdot \nabla \varphi)(t, x) dx dt &= \int_0^{+\infty} \int_{\mathbb{R}^d} u_0(x - tv) (\partial_t \varphi + v \cdot \nabla \varphi)(t, x) dx dt \\ &= \int_0^{+\infty} \int_{\mathbb{R}^d} u_0(y) (\partial_t \varphi + v \cdot \nabla \varphi)(t, y + tv) dy dt \end{aligned}$$

puis par le calcul (II.10) de la dérivée de la fonctions composée  $t \mapsto \varphi(y + tv)$

$$= \int_0^{+\infty} \int_{\mathbb{R}^d} u_0(y) \frac{\partial}{\partial t} [\varphi(t, y + tv)] dy dt$$

enfin on utilise le théorème de Fubini

$$= \int_{\mathbb{R}^d} u_0(y) \left( \int_0^{+\infty} \frac{\partial}{\partial t} [\varphi(t, y + tv)] dt \right) dy$$

puis le fait que  $\varphi(t, \cdot)$  est nulle pour  $t$  assez grand

$$= - \int_{\mathbb{R}^d} u_0(y) \varphi(0, y) dy,$$

ce qui donne bien la formule annoncée.

**Cas général** Dans le cas d'un champ de vitesse quelconque le calcul est similaire mais un tout petit peu plus compliqué car il faut tenir compte du jacobien du changement de variable  $(t, x) \mapsto (t, y = X(0, t, x))$ . Plus précisément on va utiliser le Théorème de Liouville (Théorème II.4) pour décrire l'évolution de ce jacobien  $J(t, y) = (\det D_y X)(t, 0, y)$ . On pose donc

$$u(t, x) := u_0(X(0, t, x))$$

et on effectue le calcul suivant

$$\begin{aligned} & \int_0^{+\infty} \int_{\mathbb{R}^d} u(t, x) (\partial_t \varphi + \operatorname{div}(\varphi v))(t, x) dx dt \\ &= \int_0^{+\infty} \int_{\mathbb{R}^d} u_0(X(0, t, x)) (\partial_t \varphi + \operatorname{div}(\varphi v))(t, x) dx dt \\ &= \int_0^{+\infty} \int_{\mathbb{R}^d} u_0(y) (\partial_t \varphi + v \cdot \nabla \varphi + (\operatorname{div} v) \varphi)(t, X(t, 0, y)) J(t, y) dy dt \end{aligned}$$

on utilise (II.10)

$$= \int_0^{+\infty} \int_{\mathbb{R}^d} u_0(y) \left( \frac{\partial}{\partial t} [\varphi(t, X(t, 0, y))] J(t, y) + (\operatorname{div} v)(t, X(t, 0, y)) J(t, y) \varphi(t, X(t, 0, y)) \right) dy dt$$

puis grâce au théorème de Liouville (Théorème II.4)

$$\begin{aligned} &= \int_0^{+\infty} \int_{\mathbb{R}^d} u_0(y) \left( \frac{\partial}{\partial t} [\varphi(t, X(t, 0, y))] J(t, y) + [\partial_t J(t, y)] \varphi(t, X(t, 0, y)) \right) dy dt \\ &= \int_0^{+\infty} \int_{\mathbb{R}^d} u_0(y) \left( \frac{\partial}{\partial t} [J(t, y) \varphi(t, X(t, 0, y))] \right) dy dt \end{aligned}$$

enfin on applique le théorème de Fubini

$$\begin{aligned} &= \int_{\mathbb{R}^d} u_0(y) \left( \int_0^{+\infty} \frac{\partial}{\partial t} [J(t, y) \varphi(t, X(t, 0, y))] dt \right) dy \\ &= - \int_{\mathbb{R}^d} u_0(y) J(0, y) \varphi(0, y) dy \end{aligned}$$

et comme  $J(0, y) = 1$  vu que  $X(0, 0, \cdot) = \operatorname{Id}$ ,

$$= - \int_{\mathbb{R}^d} u_0(y) \varphi(0, y) dy.$$

Le résultat est bien établi.

### IV.3 Unicité

Dès que l'on affaiblit la notion de solution d'une équation quelconque on agrandit l'ensemble des solutions possibles et il faut bien souvent s'assurer que l'on n'a pas créé de nouvelles solutions non désirées afin de conserver une théorie utilisable.

#### Théorème II.19

*On suppose que le champ de vitesse est de classe  $C^1$ . Pour toute donnée initiale  $u_0 \in L^1_{loc}(\mathbb{R}^d)$ , la formule (II.9) fournit l'unique solution faible du problème de Cauchy (II.8).*

On a déjà établi que cette formule donne une solution faible. Il reste à montrer que c'est la seule (au sens de l'égalité presque partout). Comme le problème est linéaire, il suffit de montrer<sup>12</sup> que, si  $u_0(x) = 0$  pour presque tout  $x$ , alors la seule solution faible est  $u = 0$ .

**Preuve (dans le cas d'une vitesse constante):**

Soit donc  $u \in L^1_{loc}([0, +\infty[ \times \mathbb{R}^d)$  vérifiant la formulation faible, avec  $u_0 = 0$ , c'est-à-dire

$$\int_0^{+\infty} \int_{\mathbb{R}^d} u(t, x) (\partial_t \varphi + v \cdot \nabla \varphi) dx dt = 0, \quad (\text{II.18})$$

pour toute fonction test  $\varphi \in \mathcal{D}(\mathbb{R} \times \mathbb{R}^d)$ . L'idée est de montrer que  $u = 0$  à l'aide du lemme de Du Bois-Reymond et pour cela, étant donnée une fonction  $\psi \in \mathcal{D}([0, +\infty[ \times \mathbb{R}^d)$ , on va chercher une fonction  $\varphi \in C^\infty(\mathbb{R} \times \mathbb{R}^d)$  à support compact dans  $[0, +\infty[ \times \mathbb{R}^d$  telle que

$$\partial_t \varphi + v \cdot \nabla \varphi = \psi, \quad \text{dans } ]0, +\infty[ \times \mathbb{R}^d.$$

On a déjà résolu ce problème dans le théorème II.13 et obtenu

$$\varphi(t, x) = \varphi_0(x - tv) + \int_0^t \psi(s, x + (s - t)v) ds,$$

mais pour l'instant  $\varphi_0$  n'est pas déterminé.

En particulier on a

$$\varphi(t, x + tv) = \varphi_0(x) + \int_0^t \psi(s, x + sv) ds,$$

et donc, comme on souhaite que  $\varphi$  soit à support compact dans  $[0, +\infty[ \times \mathbb{R}^d$ , la limite de la quantité ci-dessus quand  $t \rightarrow +\infty$  doit être nulle, ce qui fournit nécessairement

$$\varphi_0(x) = - \int_0^{+\infty} \psi(s, x + sv) ds.$$

Au final, on a prouvé que le bon choix de  $\varphi$  est donné par

$$\varphi(t, x) = - \int_t^{+\infty} \psi(s, x + (s - t)v) ds.$$

Cette fonction vérifie bien les bonnes propriétés et on peut la choisir comme fonction test dans (II.18) (voir la remarque II.16), ce qui fournit

$$\int_0^{+\infty} \int_{\mathbb{R}^d} u(t, x) \psi(t, x) dx dt = 0,$$

pour toute fonction  $\psi$  et on peut donc appliquer le lemme de Du Bois Reymond (Lemme A.4) et conclure. ■

Dans le cas général d'une vitesse non constante, la preuve fonctionne de manière similaire mais avec des calculs un peu plus compliqués du fait du caractère non explicite des caractéristiques du champ. Nous laissons cette preuve en exercice.

### IV.4 Le problème des conditions aux limites / conditions au bord

On se donne à nouveau un champ de vecteurs  $v$  régulier et borné défini sur  $\mathbb{R} \times \mathbb{R}^d$  tout entier (ses caractéristiques seront toujours notées  $X$ ).

Soit maintenant un domaine  $\Omega$  de  $\mathbb{R}^d$  (ouvert régulier au sens de la définition B.8), et une donnée initiale  $u_0 \in C^1(\Omega)$  et on s'intéresse au problème de Cauchy posé cette fois dans seulement dans le domaine  $\Omega$  :

$$\begin{cases} \partial_t u + v(t, x) \cdot \nabla u = 0, & \forall t > 0, \forall x \in \Omega, \\ u(t = 0, x) = u_0(x), & \forall x \in \Omega, \end{cases} \quad (\text{II.19})$$

12. s'en convaincre !

#### IV.4.a Champ tangent

On commence par supposer que le champ  $v$  est tangent au bord du domaine. Cela signifie que

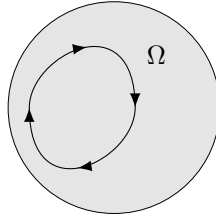
$$v(t, x) \cdot n(x) = 0, \quad \forall t \in \mathbb{R}, \forall x \in \partial\Omega.$$

On va alors montrer que les caractéristiques issues de  $\Omega$  sont complètement contenues dans  $\Omega$ .

##### Proposition II.20

*Sous les hypothèses précédentes, on a*

$$X(t, s, x_0) \in \Omega, \quad \forall t, s \in \mathbb{R}, \quad \forall x_0 \in \Omega.$$



Ceci nous permet immédiatement de résoudre le problème de Cauchy ci-dessus.

##### Corollaire II.21

*Pour un champ de vitesse  $v$  tangent au bord du domaine, le problème (II.19) admet une unique solution régulière donnée à nouveau par la formule*

$$u(t, x) = u_0(X(0, t, x)), \quad \forall t > 0, \quad \forall x \in \Omega.$$

##### Preuve (de la Proposition II.20):

On suppose (pour simplifier un peu, ça n'est absolument pas utile ...) que l'ouvert  $\Omega$  est borné. Dans ces conditions, son bord  $\partial\Omega$  est compact et en utilisant les propriétés de régularité de  $\Omega$ , on peut montrer qu'il existe  $\varepsilon > 0$  tel que l'application

$$\Phi : (s, y) \in ]-\varepsilon, \varepsilon[ \times \partial\Omega \times \mathbb{R} \mapsto y - \varepsilon n(y) \in \mathbb{R}^d,$$

est un  $\mathcal{C}^1$  difféomorphisme sur son image  $U_\varepsilon$ , celle-ci étant donnée par

$$U_\varepsilon := \{x \in \mathbb{R}^d, d(x, \partial\Omega) < \varepsilon\}.$$

Cet ensemble est appelé un  $\varepsilon$ -voisinage tubulaire de  $\partial\Omega$ , les notations sont décrites dans la Figure II.8. Par ailleurs, on peut montrer que l'application réciproque de  $\Phi$  est de la forme

$$\Phi^{-1}(x) = (\delta(x, \partial\Omega), P(x)) \in ]-\varepsilon, \varepsilon[ \times \partial\Omega,$$

où  $\delta(x)$  est la **distance signée** d'un point  $x$  au bord de  $\Omega$  définie comme suit

$$\delta(x) := \begin{cases} d(x, \partial\Omega), & \text{si } x \in \overline{\Omega}, \\ -d(x, \partial\Omega), & \text{si } x \in \overline{\Omega}^c, \end{cases}$$

et  $P(x)$  est la projection de  $x$  sur  $\partial\Omega$ , c'est-à-dire l'unique <sup>13</sup> point du compact  $\partial\Omega$  vérifiant

$$\|x - P(x)\| = d(x, \partial\Omega).$$

Enfin, on peut prouver que

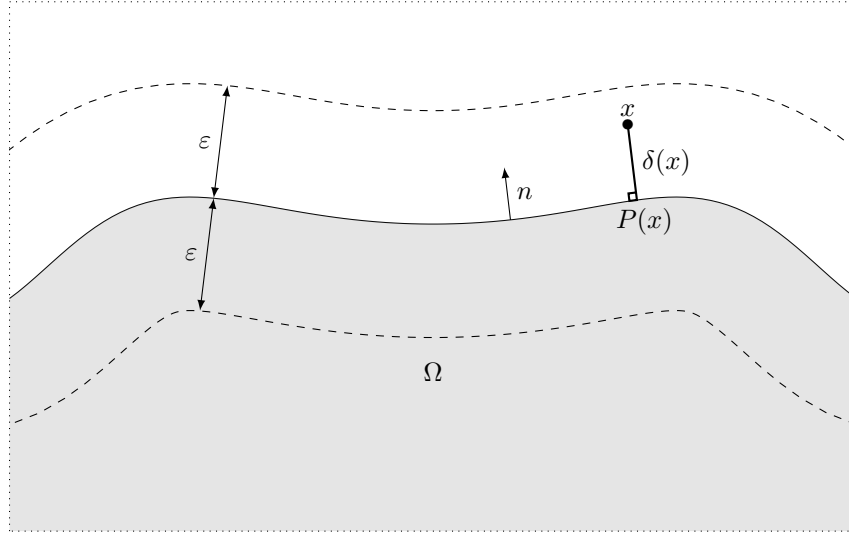
$$\nabla\delta(x) = -n(P(x)).$$

Ces préliminaires un peu techniques mais assez intuitifs étant acquis, on peut passer à la preuve de la propriété annoncée.

Pour tout point  $x \in U_\varepsilon$ , comme  $v$  et  $\delta$  sont régulières, nous avons par le théorème des accroissements finis

$$|(v \cdot \nabla\delta)(t, x) - (v \cdot \nabla\delta)(t, P(x))| \leq C\|x - P(x)\| = C|\delta(x)|, \quad \forall t \in [-T, T], \forall x \in U_\varepsilon,$$

13. L'unicité vient ici du fait que  $\Omega$  est régulier et  $\varepsilon$  est petit, ... en général il n'y a pas unicité d'un projection sur un compact quelconque !

FIGURE II.8 – Notations dans le voisinage tubulaire du bord de  $\Omega$ 

la constante  $C$  dépendant de  $T > 0$  fixé. Par ailleurs, par hypothèse, nous savons que sur le bord du domaine, on a

$$(v \cdot \nabla \delta)(t, y) = (v \cdot n)(t, y) = 0, \quad \forall t, \forall y \in \partial\Omega.$$

En combinant les deux résultats ci-dessus, nous obtenons

$$|(v \cdot \nabla \delta)(t, x)| \leq C\delta(x), \quad \forall t \in [-T, T], \forall x \in U_\varepsilon. \quad (\text{II.20})$$

On se donne maintenant un  $x_0 \in \partial\Omega$  et on s'intéresse à l'évolution de la quantité  $\delta(X(t, 0, x_0))$  c'est-à-dire de la distance de la caractéristique issue de  $x_0$  au bord du domaine. On calcule simplement

$$\begin{aligned} \frac{d}{dt} \delta(X(t, 0, x_0)) &= \nabla \delta(X(t, 0, x_0)) \cdot \frac{d}{dt} X(t, 0, x_0) \\ &= \nabla \delta(X(t, 0, x_0)) \cdot v(t, X(t, 0, x_0)). \end{aligned}$$

Par intégration, et avec (II.20), on trouve que, tant que la caractéristique reste dans  $U_\varepsilon$ , on a

$$|\delta(X(t, 0, x_0))| \leq |\delta(x_0)| + C \left| \int_0^t |\delta(X(s, 0, x_0))| ds \right|, \quad \forall t \in [-T, T].$$

Le lemme de Gronwall nous donne alors

$$|\delta(X(t, 0, x_0))| \leq |\delta(x_0)| e^{C|t|}, \quad \forall t \in [-T, T].$$

Comme on est parti d'un point sur le bord de  $\Omega$ , on a  $\delta(x_0) = 0$  et donc on déduit finalement

$$\delta(X(t, 0, x_0)) = 0, \quad \forall t \in [-T, T].$$

Comme ceci est vrai pour tout  $T$ , on a bien montré que la totalité de la caractéristique  $t \mapsto X(t, 0, x_0)$  est contenue dans le bord de  $\Omega$ .

Si maintenant  $x_0 \in \Omega$ , alors  $t \mapsto X(t, 0, x_0)$  est entièrement contenue dans  $\Omega$ . En effet, si ça n'était pas le cas, cette trajectoire devrait intersecter le bord de  $\Omega$  et donc, d'après le raisonnement précédent, elle serait contenue dans le bord, ce qui est exclu vu que  $X(0, 0, x_0) = x_0$  est dans l'ouvert  $\Omega$ . ■

#### IV.4.b Champ non tangent

Dans le cas où le champ  $v$  n'est plus tangent au bord du domaine, la proposition II.20 n'est plus valable et la méthode des caractéristiques montre qu'il manque une information pour résoudre le problème et que celle-ci doit être apportée sur le bord du domaine, et plus précisément sur la partie du bord où le champ est entrant.

Si pour simplifier on suppose que  $v$  est indépendant du temps, alors on peut introduire le sous-ensemble du bord

$$\Gamma_{\text{ent}} = \{x \in \partial\Omega, v(x) \cdot n(x) < 0\}.$$

Le problème de Cauchy associé à l'équation de transport dans  $\Omega$  doit maintenant être complété par une donnée au bord de la façon suivante : étant données  $u_0 \in \mathcal{C}^1(\Omega)$  et  $u_{\text{ent}} \in \mathcal{C}^1(\mathbb{R} \times \Gamma_{\text{ent}})$ , on cherche  $u$  solution de

$$\begin{cases} \partial_t u + v(t, x) \cdot \nabla u = 0, & \forall t > 0, \forall x \in \Omega, \\ u(t = 0, x) = u_0(x), & \forall x \in \Omega, \\ u(t, x) = u_{\text{ent}}(t, x), & \forall t > 0, \forall x \in \Gamma_{\text{ent}}. \end{cases} \quad (\text{II.21})$$

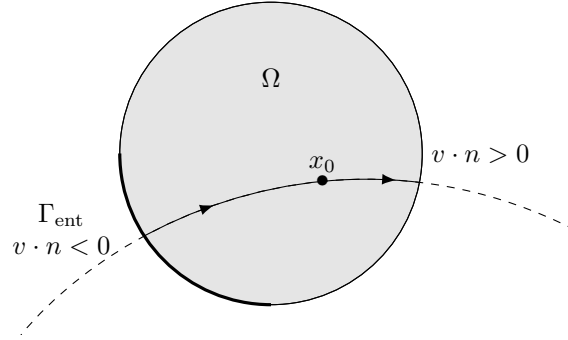


FIGURE II.9 – Les caractéristiques associées à un champ non tangent au bord de  $\Omega$

On donne alors sans démonstration le résultat suivant :

#### **Théorème II.22**

*Pour toutes données  $u_0, u_{\text{ent}}$  comme ci-dessus, il existe une unique solution faible  $u$  au problème (II.21).*

Même si la démonstration rigoureuse n'est pas simple, on peut aisément comprendre la structure de la solution en remontant les caractéristiques comme on l'a fait précédemment, voir la Figure II.9. Pour  $t > 0$  et  $x \in \Omega$  fixés, on s'intéresse à la caractéristique  $s \mapsto X(s, t, x)$  en remontant le temps en partant de  $s = t$  jusqu'à  $s = 0$ , deux cas se présentent alors :

- Si  $X(s, t, x) \in \Omega$  pour tout  $s \in [0, t]$ , alors la solution au point  $(t, x)$  sera à nouveau donnée par

$$u(t, x) = u_0(X(0, t, x)),$$

et ne dépend donc pas de la donnée au bord.

- Dans le cas contraire, si on note  $\tau(t, x)$  le plus grand temps dans  $[0, t]$  pour lequel  $X(\tau(t, x), t, x)$  appartient au bord de  $\Omega$ , alors la solution au point  $(t, x)$  sera cette fois donnée par

$$u(t, x) = u_{\text{ent}}(\tau(t, x), X(\tau(t, x), t, x)),$$

et ne dépend pas cette fois de la donnée initiale.

### **IV.5 Un exemple de modèle en dynamique des populations**

**Mise en équation** On considère un modèle de division cellulaire structuré en âge (voir [7]). La variable « d'espace »  $x$  est ici l'âge de chaque cellule, et en particulier c'est une quantité positive. L'inconnue du problème est une densité de cellule  $n(t, x)$  ayant l'âge  $x$  à l'instant  $t$ .

En l'absence de tout autre phénomène, le vieillissement naturel des cellules (dont l'âge augmente à vitesse 1 en fonction du temps<sup>14</sup>) est régi par l'équation de transport libre

$$\partial_t n + \partial_x n = 0, \quad \forall t > 0, \forall x > 0.$$

On souhaite maintenant modéliser la division cellulaire. On suppose que celle-ci a lieu avec un taux  $k(x)$  dépendant de l'âge des cellules<sup>15</sup>.

14. Comme pour tout le monde ...

15. typiquement les cellules les plus jeunes ne se divisent pas ou très peu

Fixons deux âges  $0 < a < b$  et regardons la quantité de cellules dont l'âge à l'instant  $t$  est compris entre  $a + t$  et  $b + t$

$$F(t) := \int_{a+t}^{b+t} n(t, x) dx.$$

La variation de cette quantité n'est due qu'au phénomène de division cellulaire. Plus précisément, parmi les cellules concernées dans  $F(t)$  le nombre de celles qui vont se diviser par unité de temps est donné par hypothèse par

$$\int_{a+t}^{b+t} k(x)n(t, x) dx.$$

Comme ces cellules se divisent, elles disparaissent et donnent naissance à deux cellules d'âge 0. On déduit donc que

$$F'(t) = - \int_{a+t}^{b+t} k(x)n(t, x) dx.$$

Après le calcul de la dérivée de  $F$  par le Lemme II.2, et l'argument usuel de type du Bois-Reymond, on arrive à l'équation aux dérivées partielles

$$\partial_t n + \partial_x n + k(x)n = 0.$$

Mais on doit aussi prendre en compte le fait que, à chaque instant, un certain nombre de cellules d'âge nul sont créées. Il y en a deux fois plus que de cellules qui se sont divisées, ce qui donne la relation

$$n(t, 0) = 2 \int_0^{+\infty} k(x)n(t, x) dx.$$

Si on rassemble toutes ces équations, on aboutit au modèle final suivant

$$\begin{cases} \partial_t n + \partial_x n + k(x)n = 0, \forall t > 0, \forall x > 0, \\ n(t, 0) = 2 \int_0^{+\infty} k(y)n(t, y) dy, \forall t > 0, \\ n(0, x) = n_0(x), \forall x > 0. \end{cases} \quad (\text{II.22})$$

## Résolution

### Théorème II.23

*On suppose que  $n_0$  est à support compact et que  $k$  est positive et bornée. Il existe alors une unique solution faible de (II.22).*

### Preuve :

On définit la primitive  $K$  de  $k$  par

$$K(x) = \int_0^x k(y) dy, \quad \forall x \geq 0.$$

Puis on fait le changement de variable

$$N(t, x) = e^{K(x)} n(t, x),$$

de sorte que  $N$  vérifie le nouveau système

$$\begin{cases} \partial_t N + \partial_x N = 0, \forall t > 0, \forall x > 0, \\ N(t, 0) = 2 \int_0^{+\infty} k(y)e^{-K(y)} N(t, y) dy, \forall t > 0, \\ N(0, x) = N_0(x), \forall x > 0, \end{cases} \quad (\text{II.23})$$

avec  $N_0$  défini par

$$N_0(x) = e^{K(x)} n_0(x).$$

Ainsi  $N$  vérifie une équation de transport à vitesse constante (égale à 1) avec une donnée initiale connue et une donnée au bord  $t \mapsto N(t, 0)$  inconnue. On la note  $\rho(t) := N(t, 0)$  de sorte que tout le problème se ramène à déterminer  $\rho$ . En effet, une fois  $\rho$  connue, on peut complètement déterminer  $N$  par la méthode des caractéristiques en prenant en compte le bord comme dans la section IV.4.b, en distinguant deux zones :

— Si  $x > t$ , on obtient immédiatement que

$$N(t, x) = N(0, x - t) = N_0(x - t).$$



— Si  $x < t$ , on obtient que

$$N(t, x) = N(t - x, 0) = \rho(t - x).$$

Traduisons maintenant la condition au bord pour tout  $t > 0$

$$\rho(t) = 2 \int_0^{+\infty} k(y) e^{-K(y)} N(t, y) dy.$$

Découpons l'intégrale en deux morceaux

$$\rho(t) = 2 \int_0^t k(y) e^{-K(y)} N(t, y) dy + 2 \int_t^{+\infty} k(y) e^{-K(y)} N(t, y) dy,$$

et utilisons les résultats obtenus ci-dessus pour exprimer  $N$  dans les deux intégrales

$$\rho(t) = 2 \int_0^t k(y) e^{-K(y)} \rho(t - y) dy + 2 \int_t^{+\infty} k(y) e^{-K(y)} N_0(y - t) dy.$$

On observe que le second terme est une fonction de  $t$  complètement déterminée par les données du problème, on va la noter  $\rho_0$ . Le problème revient donc à trouver une fonction  $\rho$  qui vérifie le problème intégral suivant,

$$\rho(t) = 2 \int_0^t k(y) e^{-K(y)} \rho(t - y) dy + \rho_0(t), \quad \forall t > 0. \quad (\text{II.24})$$

On va utiliser une méthode de point fixe, de façon similaire à la preuve du théorème de Cauchy-Lipschitz global donnée dans le chapitre I. On pose  $\gamma = 4\|k\|_\infty$  et on définit l'espace

$$E = \left\{ z \in \mathcal{C}^0([0, +\infty[, \mathbb{R}), \quad \|z\|_E := \sup_{t \geq 0} e^{-\gamma t} |z(t)| < +\infty \right\},$$

qui est un espace de Banach comme on l'a déjà vu.

On définit ensuite une application  $\Theta : E \rightarrow E$  par

$$\Theta : z \in E \mapsto \Theta(z) : \left( t \in [0, +\infty[ \mapsto 2 \int_0^t k(y) e^{-K(y)} z(t - y) dy + \rho_0(t) \right) \in E.$$

On vérifie que  $\Theta$  envoie bien  $E$  dans lui-même en utilisant que  $\rho_0$  est bornée et le lemme de Gronwall.

On va maintenant montrer que  $\Theta$  est contractante. Pour cela, on prend  $z_1, z_2 \in E$  et, pour tout  $t \geq 0$ , on majore

$$\begin{aligned} |\Theta(z_1)(t) - \Theta(z_2)(t)| &\leq 2 \int_0^t k(y) e^{-K(y)} |z_1(t - y) - z_2(t - y)| dy \\ &\leq 2\|k\|_\infty \int_0^t |z_1(s) - z_2(s)| ds \\ &\leq 2\|k\|_\infty \int_0^t e^{\gamma s} e^{-\gamma s} |z_1(s) - z_2(s)| ds \\ &\leq 2\|k\|_\infty \left( \int_0^t e^{\gamma s} ds \right) \|z_1 - z_2\|_E \\ &\leq \frac{2\|k\|_\infty}{\gamma} e^{\gamma t} \|z_1 - z_2\|_E. \end{aligned}$$

On fait passer le terme  $e^{\gamma t}$  à gauche de l'inégalité et, d'après le choix de  $\gamma$ , on obtient

$$\|\Theta(z_1) - \Theta(z_2)\|_E \leq \frac{1}{2} \|z_1 - z_2\|_E.$$

Ainsi,  $\Theta$  admet un unique point fixe dans  $E$  noté  $\rho$  qui est bien l'unique solution de l'équation (II.24). Observons que comme  $\rho_0$  est positive, il apparaît clairement que  $\rho$  est également une fonction positive.

Maintenant que  $\rho$  est connue, l'unique solution faible  $N$  de notre problème est complètement déterminée. Elle est de classe  $\mathcal{C}^1$  dans les deux domaines  $\{t < x\}$  et  $\{t > x\}$  mais possiblement discontinue à l'interface (voir la proposition II.18). ■

On montre dans la figure II.10 un exemple de la dynamique de ce système. La donnée initiale  $n_0$  est une gaussienne centrée en  $x = 1.5$  et la fonction  $k$  est une fonction en escalier qui vaut 0 pour  $x \leq 3$  et 1 pour  $x \geq 3$ .

On observe simultanément le phénomène de vieillissement et celui de division cellulaire qui a pour conséquence la disparition progressive des cellules d'âge supérieur à 3 et l'apparition concomitante de nouvelles cellules filles d'âge nul. Globalement la taille de la population augmente (ce qu'on peut, du reste, vérifier directement sur l'équation). On observe également que la solution est continue mais présente une discontinuité de la dérivée au point  $x = 3$ .

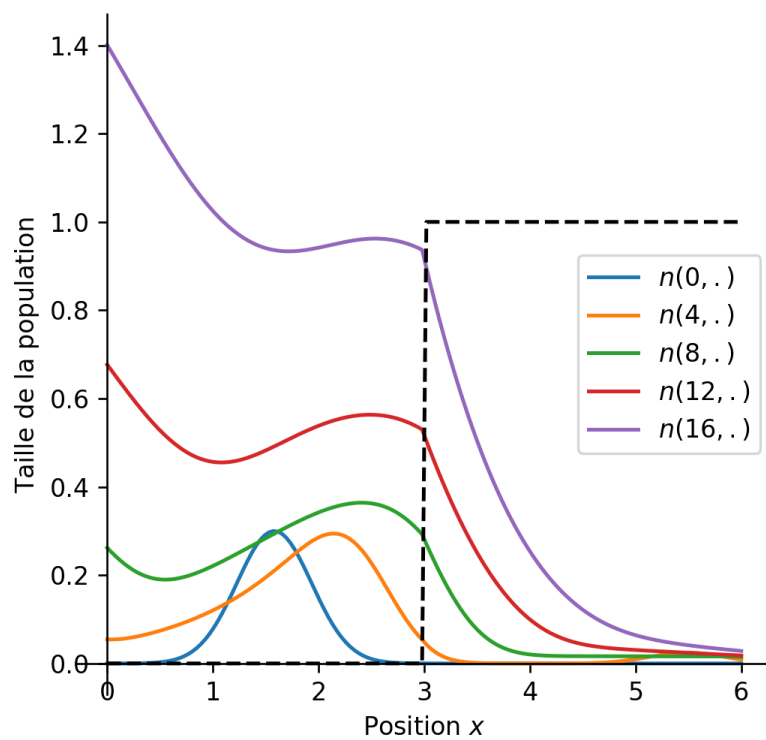


FIGURE II.10 – Un exemple d'évolution en temps du modèle de division cellulaire

## A retenir du chapitre

### En priorité

- Dérivation des intégrales à paramètre mono-dimensionnelles comme dans le lemme II.2.
- Reconnaître des situations de modélisation (prioritairement en dimension 1 d'espace) qui aboutissent à une équation de type transport (ou de continuité).
- Connaître la méthode des caractéristiques pour résoudre le problème de Cauchy pour des équations de type transport (et variantes) à données régulières : en particulier le cas 1D doit être parfaitement maîtrisé (voir TD).
- La solution explicite de ces problèmes doit être connue dans les cas simples où les caractéristiques sont faciles à calculer.

### Pour approfondir

- Théorème de Reynolds.
- Preuves de l'existence et unicité des solutions faibles. Solutions faibles régulières par morceaux.
- La question des conditions au bord pour le problème de transport.



## Chapitre III

# Formulations variationnelles de problèmes aux limites elliptiques

Le but de ce chapitre est de présenter, à partir d'un exemple simple (et relativement concret), quelques notions de calcul des variations. Cela nous mènera à la notion de formulation *variationnelle* d'un problème aux limites (c'est-à-dire une EDP elliptique + des conditions aux limites). Nous verrons également pourquoi il est naturel, en commençant par la dimension 1, d'introduire des espaces fonctionnels adaptés à ce type d'approche.

Par la suite nous généraliserons ces concepts pour attaquer des problèmes plus complexes, en particulier en dimension quelconque.

## I Le problème de la corde/membrane élastique à l'équilibre

### I.1 Présentation

On considère une membrane élastique qui au repos est représentée par une région compacte  $\overline{\Omega}$  du plan horizontal  $\mathbb{R}^2$  (on fixe l'origine des altitudes sur ce plan), où  $\Omega$  est un ouvert **connexe** de  $\mathbb{R}^2$  (voir Figure III.1)

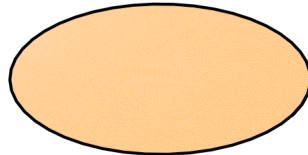


FIGURE III.1 – Une membrane plane circulaire au repos

On applique une force verticale sur la membrane ce qui a pour effet de la déformer. Cette force est exercée en tout point de la membrane, et plus précisément on note  $f(x)$  la densité de forces qui va être exercée en chaque point  $x$  de la membrane. Autrement dit, une surface élémentaire  $dx$  autour du point  $x$  va subir une force d'intensité  $f(x) dx$  ce qui va impliquer que la force totale  $F_A$  exercée sur une partie  $A$  de la membrane sera donnée par l'intégrale

$$F_A = \int_A f(x) dx.$$

Sous l'effet de la force appliquée, pour tout  $x \in \overline{\Omega}$ , le point initial de l'espace physique  $(x, 0) \in \mathbb{R}^3$  est déplacé en un nouveau point de  $\mathbb{R}^3$  noté  $\tilde{u}(x) = (x, u(x))$ , où  $u$  est une fonction de  $\overline{\Omega}$  dans  $\mathbb{R}$  que l'on va chercher à déterminer et qui représente le déplacement vertical. La question est donc : parmi l'infinité de positions d'équilibres envisageables de la membrane (voir trois exemples dans la Figure III.2) laquelle va-t-on réellement observer pour une force  $f$  donnée ? Pour répondre à cette question il va falloir écrire en termes mathématiques un principe fondamental qui nous est donné par la physique : sous l'effet d'une force extérieure, le système va *choisir* la position  $u$  qui va minimiser l'énergie totale de la membrane. Cette énergie est la somme de deux énergies potentielles : l'une liée au travail des forces exercées et l'autre liée à l'élasticité du matériau qui compose la membrane (typiquement quand on étire un élastique on lui communique une certaine énergie - que l'on peut libérer en lâchant l'élastique !).

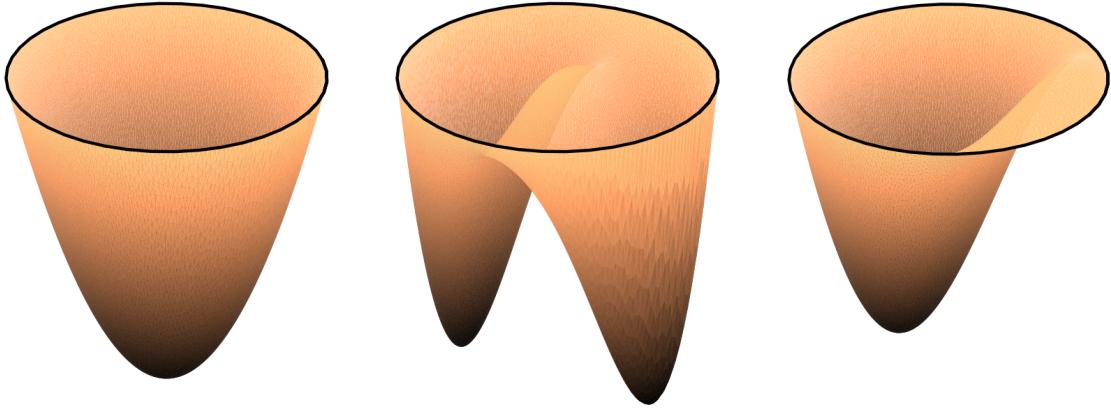


FIGURE III.2 – Différentes positions envisageables de la membrane

On se placera désormais dans l'hypothèse de petits déplacements (i.e.  $f$  et  $u$  seront supposés *petits*, nous verrons plus loin ce que cela implique). De plus, on va considérer la situation dans laquelle la membrane est attachée par son bord à un référentiel fixe (penser à la peau d'un tambour). Ceci implique que l'on cherchera une fonction  $u$  qui est nulle sur le bord  $\partial\Omega$ .

Faisons le bilan d'énergie potentielle du système :

- Le système acquiert une énergie potentielle virtuelle qui vaut l'opposé du travail des forces extérieures  $f$  par rapport au déplacement  $u$ , c'est-à-dire :

$$E_1(u) := - \int_{\Omega} u(x)f(x) dx.$$

La présence du signe  $-$  est naturelle : si la force est orientée vers le bas ( $f \leq 0$ ) alors les zones de forte énergie potentielle sont les zones les plus hautes (donc pour les  $u \geq 0$  grand) (comme pour le champ de gravité par exemple).

- Par ailleurs le système contient de l'énergie potentielle élastique due à la déformation de la membrane. On admet que cette énergie est proportionnelle au changement d'aire de la membrane<sup>1</sup>, c'est-à-dire de la forme

$$k(\text{Aire déformée} - \text{Aire au repos}) = k(|\tilde{u}(\Omega)| - |\Omega|),$$

où  $k > 0$  est une constante de *raideur* caractéristique du matériau considéré.

Par changement de variable, on trouve que cette énergie s'écrit

$$k \int_{\Omega} \left( \sqrt{1 + |\nabla u|^2} - 1 \right) dx.$$

Dans l'hypothèse des petits déplacements,  $u$  est petite ainsi que ses dérivées. Par un développement limité usuel, on approche alors cette expression par la quantité *quadratique* :

$$E_2(u) := \frac{k}{2} \int_{\Omega} |\nabla u|^2 dx.$$

L'énergie potentielle totale du système en fonction du déplacement  $u$  est donnée par

$$E(u) := E_1(u) + E_2(u) = \frac{k}{2} \int_{\Omega} |\nabla u|^2 dx - \int_{\Omega} fu dx.$$

Dans cette formule,  $u : \Omega \rightarrow \mathbb{R}$  est une fonction et  $E(u)$  est un nombre.

Le principe fondamental de la mécanique Lagrangienne nous dit alors que, sous l'effet du champ de forces  $f$ , la membrane va se déformer selon un déplacement  $u$  qui va minimiser l'énergie potentielle totale (qu'on devrait plutôt appeler *l'action* selon le vocabulaire *ad hoc* de la mécanique). On s'intéresse donc au problème suivant : trouver  $u \in X$  tel que

$$E(u) = \inf_{v \in X} E(v) = \inf_{v \in X} \left( \frac{k}{2} \int_{\Omega} |\nabla v|^2 dx - \int_{\Omega} fv dx \right), \quad (\text{III.1})$$

1. Ceci peut se "démontrer", par exemple, en assimilant la membrane à un réseau de petits ressorts

où, *a priori*  $X$  est l'espace fonctionnel suivant (on a représenté trois éléments de  $X$  dans la figure III.2 ; le bord noir de la membrane est la partie fixée) :

$$X = \{v : \bar{\Omega} \mapsto \mathbb{R}, \text{ dérivable, } v = 0, \text{ sur } \partial\Omega\},$$

constitué de l'ensemble des positions **admissibles** de la membrane.

### Remarque III.1

*Noter que le modèle a été établi sous une condition de petitesse des déplacements qui ne se retrouve pas dans la formulation présentée ici ; dans les applications il peut donc être pertinent de se poser a posteriori la question de la validité de la solution obtenue après résolution du problème. Si celle-ci s'avère trop grande (par rapport à quoi ?) alors il est possible que les simplifications du modèle qu'on a utilisées ne soient pas, en fait, pertinentes. Nous avons déjà rencontré cette problématique pour discuter la validité de l'équation du pendule linéarisé.*

Nous verrons par la suite que l'espace  $X$  ci-dessus n'est pas nécessairement le bon choix pour faire fonctionner les méthodes mathématiques que nous allons présenter. Une partie de ce chapitre va d'ailleurs consister à comprendre quel est le bon espace  $X$  à considérer.

## I.2 Les questions mathématiques que l'on veut résoudre

Dans la suite, on va essayer de répondre aux questions suivantes :

1. Le problème (III.1) admet-il une solution ? En particulier, est-ce que l'infimum de  $E$  sur  $X$  est fini ?
2. Si une telle solution existe, est-elle unique ?
3. Si une telle solution existe, est-ce qu'on peut la caractériser au moyen d'une équation "simple" que l'on pourra éventuellement résoudre ?

### I.2.a Réécriture dans le cas de la dimension 1

A partir de maintenant, et ce jusqu'à la section IV, on va se placer dans le cas plus simple de la dimension 1. Le modèle correspond alors à une corde élastique (et non plus une membrane). Pour simplifier encore, on considère  $\Omega = ]0, 1[$  de sorte que le problème s'écrit de la façon suivante : trouver  $u : [0, 1] \mapsto \mathbb{R}$  dérivable, nulle en  $x = 0$  et  $x = 1$  et telle que

$$E(u) = \inf_{v \in X} E(v), \quad (\text{III.2})$$

avec

$$X = \{v : [0, 1] \mapsto \mathbb{R}, \text{ dérivable et tq } v(0) = v(1) = 0\},$$

et l'énergie s'écrit maintenant

$$E(v) = \frac{k}{2} \int_0^1 |v'(x)|^2 dx - \int_0^1 f(x)v(x) dx.$$

On supposera également que  $f$  est, au minimum, une fonction intégrable.

### I.2.b Unicité

On va commencer par traiter le problème de l'unicité, qui est finalement le problème le plus simple. Cette propriété découle naturellement de la **stricte convexité** de la fonctionnelle  $E$  (et de la convexité de l'ensemble  $X$  qui, ici, est un espace vectoriel).

Supposons données deux fonctions  $u_1, u_2 \in X$  solutions du problème (III.2). On suppose donc, en particulier, que l'infimum de  $E$  est fini. On notera sa valeur

$$I_E = \inf_{v \in X} E(v),$$

et on a donc, par hypothèse  $E(u_1) = E(u_2) = I_E$ .

On pose alors  $u = \frac{u_1 + u_2}{2}$ , qui est bien dans  $X$  car cet ensemble est convexe, et on calcule l'énergie de  $u$ , en utilisant l'identité du parallélogramme<sup>2</sup>

$$\begin{aligned} E(u) &= \frac{k}{2} \int_0^1 \left| \frac{u'_1(x) + u'_2(x)}{2} \right|^2 dx - \int_0^1 f(x) \frac{u_1(x) + u_2(x)}{2} dx \\ &= \frac{k}{4} \int_0^1 |u'_1(x)|^2 dx + \frac{k}{4} \int_0^1 |u'_2(x)|^2 dx - \frac{k}{2} \int_0^1 \left| \frac{u'_1(x) - u'_2(x)}{2} \right|^2 dx \\ &\quad - \frac{1}{2} \int_0^1 f(x) u_1(x) dx - \frac{1}{2} \int_0^1 f(x) u_2(x) dx \\ &= \frac{1}{2} E(u_1) + \frac{1}{2} E(u_2) - \frac{k}{2} \int_0^1 \left| \frac{u'_1(x) - u'_2(x)}{2} \right|^2 dx. \end{aligned}$$

Or, par hypothèse,  $u_1$  et  $u_2$  sont solutions du problème (III.2) et donc, on a finalement

$$E(u) = I_E - \frac{k}{2} \int_0^1 \left| \frac{u'_1(x) - u'_2(x)}{2} \right|^2 dx.$$

Par ailleurs, par définition de l'infimum, on a  $E(u) \geq I_E$  (c'est ici qu'on utilise que  $u \in X$ ). Comme la constante de raideur  $k$  est strictement positive, on déduit que l'on a nécessairement

$$\frac{k}{2} \int_0^1 \left| \frac{u'_1(x) - u'_2(x)}{2} \right|^2 dx = 0.$$

La fonction sous l'intégrale étant positive, on obtient immédiatement que

$$\forall x \in [0, 1], \quad u'_1(x) = u'_2(x).$$

Ceci implique que  $u_1 - u_2$  est une fonction constante, mais comme  $u_1(0) = u_2(0) = 0$  (par définition des conditions au bord dans  $X$ ), on a finalement montré

$$\forall x \in [0, 1], \quad u_1(x) = u_2(x),$$

ce qui montre bien l'unicité d'une éventuelle solution de (III.2).

### I.2.c Caractérisation de la solution

On continue à supposer dans ce paragraphe que la solution  $u \in X$  du problème (III.2) existe. On va montrer que, sous de bonnes hypothèses, on peut la caractériser par une équation aux dérivées partielles (ici en 1D donc avec une seule variable).

La méthode ci-dessous est standard en optimisation et calcul des variations. Il s'agit d'établir les équations **d'Euler-Lagrange** associées au problème de minimisation (III.2).

La preuve de ce résultat n'est finalement rien d'autre que la traduction en dimension infinie ( $\dim X = \infty$ ) du résultat élémentaire mais fondamental suivant :

#### Lemme III.2

Soit  $\varphi : \mathbb{R} \mapsto \mathbb{R}$  une fonction dérivable. On suppose qu'il existe  $t^* \in \mathbb{R}$  tel que

$$\varphi(t^*) = \inf_{t \in \mathbb{R}} \varphi(t), \tag{III.3}$$

alors on a

$$\varphi'(t^*) = 0.$$

Il ne semble pas inutile de rappeler la démonstration de ce lemme pour comprendre comment intervient l'hypothèse.

**Preuve :**

D'après (III.3), pour tout  $h > 0$  (le signe de  $h$  joue ici un rôle crucial !), on a

$$\varphi(t^* + h) \geq \varphi(t^*).$$

Comme  $h > 0$ , on en déduit

$$\frac{\varphi(t^* + h) - \varphi(t^*)}{h} \geq 0.$$

2. Rappel : dans un Hilbert  $H$ , pour tous  $a, b \in H$  on a  $\left\| \frac{a+b}{2} \right\|^2 + \left\| \frac{a-b}{2} \right\|^2 = \frac{\|a\|^2}{2} + \frac{\|b\|^2}{2}$ .



On passe maintenant à la limite quand  $h \rightarrow 0^+$  dans cette inégalité, ce qui donne par définition du nombre dérivé

$$\varphi'(t^*) \geq 0.$$

Si maintenant on reprend ce calcul avec  $h < 0$ , on a toujours

$$\varphi(t^* + h) \geq \varphi(t^*),$$

mais par contre (comme  $h < 0$  !) on a cette fois

$$\frac{\varphi(t^* + h) - \varphi(t^*)}{h} \leq 0.$$

En passant à la limite quand  $h \rightarrow 0^-$ , on trouve

$$\varphi'(t^*) \leq 0,$$

ce qui donne le résultat. ■

Revenons à notre problème de corde élastique. Supposons donc qu'une solution  $u$  de (III.2) existe. On se donne un  $v$  quelconque dans  $X$ , de sorte que, pour tout  $t \in \mathbb{R}$ , on a  $u + tv \in X$  (car  $X$  est un espace vectoriel !). Ainsi, par définition de l'infimum, on a

$$\forall t \in \mathbb{R}, \quad E(u + tv) \geq E(u),$$

ce qui montre que la fonction  $\varphi_v : \mathbb{R} \mapsto \mathbb{R}$  définie par  $\varphi_v(t) = E(u + tv)$  admet un minimum en  $t^* = 0$ . Par ailleurs, cette fonction est dérivable (on va même voir ci-dessous que c'est un polynôme de degré 2 dans la variable  $t$ ). D'après le lemme précédent, on en déduit que nécessairement  $\varphi'_v(0) = 0$ .

Il reste à calculer  $\varphi'_v(0)$ . Pour cela, on écrit  $\varphi_v$  sous la forme suivante

$$\varphi_v(t) = E(u) + t \left[ k \int_0^1 u'(x)v'(x) dx - \int_0^1 f(x)v(x) dx \right] + \frac{t^2}{2} k \int_0^1 |v'(x)|^2 dx, \quad (\text{III.4})$$

et donc, la relation  $\varphi'_v(0) = 0$  revient à écrire que le coefficient de  $t$  dans ce polynôme est nul, c'est-à-dire

$$k \int_0^1 u'(x)v'(x) dx = \int_0^1 f(x)v(x) dx.$$

Comme ceci est vrai pour toute fonction  $v \in X$ , on a donc démontré que la solution, si elle existe, du problème (III.2) vérifie les équations d'Euler-Lagrange suivantes :

$$\forall v \in X, \quad k \int_0^1 u'(x)v'(x) dx = \int_0^1 f(x)v(x) dx. \quad (\text{III.5})$$

En revenant à (III.4), on observe que **dans cet exemple particulier** la réciproque est vraie : si  $u \in X$  vérifie (III.5), alors  $u$  est solution du problème (III.2), car le terme en  $t^2$  dans (III.4) est toujours positif.

**ATTENTION** : ceci est faux en général car on sait bien que la réciproque du lemme III.2 n'est pas vraie : si  $\varphi'(t^*) = 0$ ,  $t^*$  n'est pas nécessairement un extremum local de  $\varphi$ .

Au bilan, on a donc montré le résultat suivant

### Proposition III.3

*Une fonction  $u \in X$  vérifie (III.5) si et seulement si  $u$  est solution du problème (III.2).*

Jusqu'à présent nous n'avons eu besoin d'aucune hypothèse particulière sur la solution  $u$  de notre problème. Si on admet que celle-ci est un peu plus régulière que simplement dérivable, alors on peut aller plus loin dans l'analyse.

### Théorème III.4

*On suppose que le problème (III.2) admet une solution  $u \in X$ .*

*Si on suppose, de plus, que cette solution vérifie  $u \in C^2([0, 1])$  et que  $f \in C^0([0, 1])$ , alors  $u$  vérifie le problème aux limites suivant :*

$$\begin{cases} -k\partial_x^2 u = f, & \text{dans } ]0, 1[, \\ u(0) = u(1) = 0. \end{cases} \quad (\text{III.6})$$

*Notons que la réciproque est également vraie : toute solution de (III.6) vérifie (III.5) et donc (III.2)*

Le problème (III.6) est appelé *problème de Poisson avec condition de Dirichlet homogène*.

**Preuve :**

Pour montrer la réciproque, il suffit de multiplier l'équation dans (III.6) par la fonction test  $v$  et d'intégrer par parties. Les termes de bord sont bien nuls car  $v$  est nulle au bord.

Pour le sens direct, il s'agit, dans un premier temps, également d'une simple intégration par parties. Pour tout  $v \in X$ , comme  $u$  est de classe  $\mathcal{C}^2$ , on peut en effet intégrer par parties l'équation (III.5) et obtenir

$$\int_0^1 (k\partial_x^2 u(x) + f(x))v(x) dx - [k(\partial_x u)v]_0^1 = 0.$$

On ne sait rien de la valeur de  $\partial_x u$  en  $x = 0$  et  $x = 1$ , par contre on a  $v(0) = v(1) = 0$ , ce qui montre que le dernier terme de cette formule est nul.

Si on pose  $G(x) = k\partial_x^2 u(x) + f(x)$ , on a donc obtenu

$$\forall v \in X, \int_0^1 G(x)v(x) dx = 0.$$

Comme  $X$  contient en particulier toutes les fonctions  $\mathcal{C}^\infty$  à support compact, le lemme de Du Bois-Reymond (Lemme A.4) permet de conclure que  $G = 0$ . Il est important de noter que l'on ne peut pas se contenter de prendre  $v = G$  dans l'intégrale ci-dessus car  $G$  n'est pas dans  $X$  (en particulier, il n'y a aucune raison qu'elle soit nulle au bord). On a bien montré (III.6) ■

#### Vocabulaire

- (III.6) est un problème aux limites (une EDP + des conditions aux limites). Même s'il n'y a qu'une seule variable ici, ce problème **n'est pas** un problème de Cauchy !
- (III.5) est une formulation variationnelle de ce problème aux limites dont  $u$  est la solution.
- La fonction  $v$  est appelée fonction-test. Elle vit, en général, dans le même espace que la solution  $u$  recherchée.

Tous les calculs précédents peuvent être effectués en dimension  $d$  quelconque (i.e. dans le cas de la membrane). Le problème aux limites que l'on obtient formellement à la place de (III.6) est

$$\begin{cases} -k\Delta u = f, & \text{dans } \Omega, \\ u = 0, & \text{sur } \partial\Omega, \end{cases} \quad (\text{III.7})$$

et la formulation variationnelle associée est

$$\forall v \in X, \quad k \int_{\Omega} \nabla u(x) \cdot \nabla v(x) dx = \int_{\Omega} f(x)v(x) dx. \quad (\text{III.8})$$

On étudiera plus précisément ce cas dans la section IV.

#### Remarque III.5

Si on reprend toute l'analyse précédente en supposant que la tension  $k$  de la corde/membrane dépend du point  $x$  où l'on se place, alors on aboutit aux équations suivantes

$$-\partial_x(k(x)\partial_x u) = f(x), \quad \text{dans } ]0, 1[,$$

en dimension 1, et

$$-\operatorname{div}(k(x)\nabla u) = f, \quad \text{dans } \Omega,$$

en dimension  $d$  toujours assorties des conditions aux limites que l'on ne récrit pas ici.

Pour l'instant nous avons démontré que si le problème de minimisation de  $E$  sur  $X$  admet une solution, alors elle est unique et que de plus si celle-ci est régulière, alors elle est solution de l'équation de Poisson.

Inversement toute solution régulière de l'équation de Poisson est solution de la formulation variationnelle et minimise l'énergie sur  $X$ .

Il nous reste à montrer l'existence d'un minimiseur, et c'est là que le choix de l'espace fonctionnel  $X$  va être crucial.

### I.3 Comment montrer l'existence d'un minimiseur ?

Le principe général de la preuve de l'existence d'un minimiseur pour une fonctionnelle  $E$  sur un ensemble  $X$  est le suivant :

1. On démontre que  $\inf_X E > -\infty$ . Si ceci est faux, le problème de chercher un minimiseur n'a évidemment pas de sens, c'est donc une étape indispensable.
2. On considère une *suite minimisante*, c'est-à-dire une suite  $(u_n)_n \subset X$  qui vérifie  $\lim_{n \rightarrow \infty} E(u_n) = \inf_X E$ . Une telle suite existe **toujours**, c'est juste la définition de l'infimum qui nous la fournit.
3. On essaie de démontrer que cette suite (ou l'une de ses sous-suites) converge (en un sens à préciser : idéalement pour la norme naturelle dans  $X$ ). On note  $u$  la limite obtenue.
4. On essaie de vérifier que  $u$  réalise bien le minimum recherché. En général, pour faire cela, on essaie de montrer que  $E(u_n) \rightarrow E(u)$  ce qui fournira le résultat grâce au point 2. On a donc besoin d'une certaine forme de continuité de la fonctionnelle  $E$  dans  $X$ .

Il est bon de remarquer que les deux premiers points ne dépendent que de l'ensemble  $X$  (et de  $E$  bien sûr) alors que les deux derniers points requièrent une topologie sur  $X$ , puisqu'il y est question de suites convergentes, de fonctions continues, etc ...

Ainsi, pour faire fonctionner ce programme de travail, le choix du bon espace  $X$  et de la bonne topologie sur cet espace sont cruciales. En effet, l'espace  $X$  étant, en général, de dimension infinie, le choix d'une topologie sur  $X$  (même une topologie d'e.v.n.) n'est pas du tout trivial.

De façon générale, pour réaliser la troisième étape du programme ci-dessus, on va essayer de montrer que la suite minimisante est de Cauchy, donc convergente, **à condition que l'espace  $X$  soit complet**. C'est pourquoi la complétude de l'espace est une notion centrale. Une première idée serait donc de modifier un tout petit peu l'espace  $X$  considéré plus haut en choisissant cette fois

$$X = \{v : [0, 1] \rightarrow \mathbb{R}, \text{ de classe } \mathcal{C}^1 \text{ et tq } v(0) = v(1) = 0\},$$

muni de la norme  $\|v\|_X = \|v\|_{L^\infty} + \|v'\|_{L^\infty}$ .

On sait en effet que cet espace est un Banach, il est donc susceptible de convenir.

**Finitude de l'infimum** Commençons par vérifier que l'infimum est fini. Pour cela, on utilise le résultat suivant

#### Lemme III.6

Si  $v \in X$ , on a

$$\|v\|_{L^\infty} \leq \|v'\|_{L^2}.$$

#### Preuve :

On écrit, pour tout  $x \in [0, 1]$ ,

$$v(x) = \underbrace{v(0)}_{=0, \text{ car } v \in X} + \int_0^x v'(t) dt,$$

puis on majore l'intégrale par l'inégalité de Cauchy-Schwarz. ■

Ainsi, pour tout  $v \in X$ ,

$$E(v) = \frac{k}{2} \|v'\|_{L^2}^2 - \int_0^1 f v dx \geq \frac{k}{2} \|v'\|_{L^2}^2 - \|f\|_{L^1} \|v\|_{L^\infty} \geq \frac{k}{2} \|v'\|_{L^2}^2 - \|f\|_{L^1} \|v'\|_{L^2}.$$

Or la fonction (polynômiale)  $y \mapsto \frac{k}{2} y^2 - \|f\|_{L^1} y$  est minorée sur  $\mathbb{R}$  par  $-\frac{\|f\|_{L^1}^2}{2k}$ .

On a donc montré le résultat attendu

$$\inf_X E \geq -\frac{\|f\|_{L^1}^2}{2k}.$$

**Etude d'une suite minimisante** Comment peut-on alors montrer la convergence d'une suite minimisante dans cet espace  $X$  ?

Une des seules choses que l'on sait sur la suite minimisante, c'est que la suite de nombre réels  $(E(u_n))_n$  est bornée et converge vers l'infimum. Est-ce qu'on peut en déduire que la suite  $(u_n)_n$  est bornée dans  $X$  ?

La réponse à cette question est, en général, non ! On peut par exemple facilement construire une suite  $(u_n)_n$  de fonctions de  $X$  qui ne soit pas bornée et telle que  $(E(u_n))_n$  soit bornée (voir le TD).

On voit donc qu'il n'est pas du tout évident de tirer une information utile sur la suite  $(u_n)_n$  uniquement à partir de la connaissance de la suite des énergies  $(E(u_n))_n$ .

En théorie de l'optimisation, on dit que la fonctionnelle  $E$  n'est pas **coercive** sur  $(X, \|\cdot\|_X)$ .

On pourrait envisager de changer la norme sur  $X$  pour mieux correspondre au problème étudié. Ainsi, si on munit  $X$  de la norme suivante

$$\|u\|_{H^1} = \sqrt{\|u\|_{L^2}^2 + \|u'\|_{L^2}^2},$$

alors il est immédiat de voir que l'on récupère une forme de coercivité de  $E$  sur  $(X, \|\cdot\|_{H^1})$ .

### Proposition III.7

*Si  $(u_n)_n$  est une suite d'éléments de  $X$  telle que  $(E(u_n))_n$  est bornée, alors  $(u_n)_n$  est bornée dans  $(X, \|\cdot\|_{H^1})$ .*

#### Preuve :

On note  $C = \sup_n E(u_n) < +\infty$ . On utilise ensuite la définition de l'énergie

$$\frac{k}{2} \|u'_n\|_{L^2}^2 = E(u_n) + \int_0^1 f(x)u_n(x) dx \leq C + \|f\|_{L^1} \|u_n\|_{L^\infty} \leq C + \|f\|_{L^1} \|u'_n\|_{L^2}.$$

On utilise ensuite l'inégalité de Young suivante<sup>3</sup>

$$\forall \varepsilon > 0, \forall a, b \in \mathbb{R}, \quad ab \leq \varepsilon a^2 + \frac{1}{4\varepsilon} b^2.$$

pour majorer le dernier terme de la façon suivante

$$\frac{k}{2} \|u'_n\|_{L^2}^2 \leq C + \underbrace{\frac{k}{4}}_{=\varepsilon} \|u'_n\|_{L^2}^2 + \frac{\|f\|_{L^1}^2}{k}.$$

On en déduit

$$\frac{k}{4} \|u'_n\|_{L^2}^2 \leq C + \frac{\|f\|_{L^1}^2}{k},$$

ce qui donne une borne sur  $\|u'_n\|_{L^2}$ . Il nous faut maintenant une borne sur  $\|u_n\|_{L^2}$  mais celle-ci s'obtient immédiatement avec le lemme III.6

$$\|u_n\|_{L^2}^2 = \int_0^1 |u_n|^2 dx \leq \|u_n\|_{L^\infty}^2 \leq \|u'_n\|_{L^2}^2.$$

■

Ainsi, avec cette nouvelle norme, on a la coercivité de  $E$  et l'on peut en déduire que toute suite minimisante est bornée pour la norme  $\|\cdot\|_{H^1}$ . C'est un premier pas vers la convergence de la suite. On peut même montrer que cette suite est de Cauchy toujours pour la norme  $\|\cdot\|_{H^1}$  (voir plus loin la démonstration).

Malheureusement, en changeant la norme sur  $X$ , on a perdu une propriété essentielle : la complétude. On ne peut donc rien déduire du fait que la suite minimisante est de Cauchy.

**Bilan :** Un choix pour lequel la démarche va fonctionner sera de remplacer  $X$  par son **complété** pour la norme  $\|\cdot\|_{H^1}$ . Cet espace est, *a priori*, un espace abstrait. On va voir dans la suite du chapitre, qu'on peut en fait le construire de façon relativement explicite et travailler dans cet espace de façon finalement naturelle.

## II Espaces de Sobolev en dimension 1

### II.1 L'espace $H^1([a, b])$

Bien qu'on se consacre pour l'instant au cas monodimensionnel à fin de comprendre les idées mises en place, il faut retenir que toutes les méthodes se généralisent au cas multi-D, comme nous le verrons dans la suite du cours. C'est d'ailleurs en dimension quelconque que ces outils d'analyse fonctionnelle sont réellement utiles et efficaces.

**ATTENTION quand même :** les résultats théoriques sur les espaces de Sobolev peuvent varier selon la dimension de l'espace.

3. Savez-vous la démontrer ? c'est un très bon exercice dont la preuve prend quelques lignes ...

On se place donc désormais en dimension 1 et on travaille sur un intervalle ouvert borné  $I = ]a, b[$ .

### Définition III.8

On appelle **espace de Sobolev**  $H^1(I)$ , l'ensemble des fonctions  $u$  de  $L^2(I)$  dont la dérivée au sens des distributions est une fonction de  $L^2(I)$ .

On rappelle que cela signifie qu'il existe une fonction  $g \in L^2(I)$  (nécessairement unique presque partout, d'après le lemme de du Bois-Reymond) telle que

$$\text{Pour toute fonction test } \varphi \in \mathcal{C}_c^\infty(I), \text{ on a } \int_a^b u(x)\varphi'(x) dx = - \int_a^b g(x)\varphi(x) dx. \quad (\text{III.9})$$

Cette fonction  $g$  est désormais notée  $u'$  (ou parfois  $\partial_x u$  ou  $\nabla u$  par analogie avec la dimension supérieure).  
On munit l'espace vectoriel  $H^1(I)$  de la norme

$$\|u\|_{H^1} = \sqrt{\|u\|_{L^2}^2 + \|\nabla u\|_{L^2}^2}.$$

Ce qu'il faut retenir, c'est la formule "d'intégration par parties" suivante, qui est en fait la **définition** de la dérivée faible

$$\forall \varphi \in \mathcal{C}_c^\infty(I), \text{ on a } \int_a^b u(x)\varphi'(x) dx = - \int_a^b u'(x)\varphi(x) dx.$$

On va maintenant établir que cette définition généralise la notion usuelle de dérivée.

### Proposition III.9

Si  $u$  est une fonction de classe  $\mathcal{C}^1$  sur  $\bar{I} = [a, b]$ , alors  $u \in H^1(I)$  et sa dérivée faible  $u'$  coïncide presque partout avec sa dérivée usuelle (ce qui justifie la notation).

On a donc l'inclusion algébrique

$$\mathcal{C}^1(\bar{I}) \subset H^1(I),$$

mais on a aussi continuité de cette injection, c'est-à-dire plus précisément que

$$\forall u \in \mathcal{C}^1(\bar{I}), \quad \|u\|_{H^1} \leq \sqrt{2(b-a)} \|u\|_{\mathcal{C}^1}.$$

#### Preuve :

Si  $u$  est de classe  $\mathcal{C}^1$ , la formule (III.9) est vraie avec  $g = u'$  (la dérivée usuelle) par une simple intégration par parties. Celle-ci ne comporte aucun terme de bord car  $\varphi$  est choisie à support compact.

Pour comparer les normes, on utilise simplement que pour toute fonction  $f$  continue sur  $[a, b]$ , on a  $\|f\|_{L^2} \leq \sqrt{b-a} \|f\|_{L^\infty}$ . ■

### Théorème III.10 (Principales propriétés de $H^1(I)$ )

1. L'espace  $(H^1(I), \|\cdot\|_{H^1})$  introduit dans la Définition III.8 est un espace de Hilbert.

2. Toute fonction  $u \in H^1(I)$  admet un représentant (au sens presque partout) continu sur  $\bar{I}$ , qu'on note toujours  $u$ , et on a

$$u(\beta) - u(\alpha) = \int_\alpha^\beta \partial_x u dx, \quad \forall \alpha, \beta \in \bar{I}.$$

3. L'ensemble  $\mathcal{C}^\infty(\bar{I})$  est dense dans  $H^1(I)$ .

Notons d'ores et déjà que la propriété 2 (la continuité des fonctions de  $H^1(I)$ ) n'est pas vraie en dimension supérieure.

#### Preuve :

1. Il est clair que la norme  $\|\cdot\|_{H^1}$  est une norme Hilbertienne associée au produit scalaire  $H^1$  défini par

$$(u, v)_{H^1} = (u, v)_{L^2} + (\partial_x u, \partial_x v)_{L^2} = \int_I uv dx + \int_I \partial_x u \partial_x v dx.$$

Il reste juste à vérifier la complétude de  $H^1(I)$ . Soit donc une suite de Cauchy  $(u_n)_n$  dans  $H^1(I)$ . Par définition de la norme  $H^1$ , on déduit que  $(u_n)_n$  et  $(\partial_x u_n)_n$  sont de Cauchy dans  $L^2(I)$ . Comme  $L^2(I)$  est complet, il existe  $u, g \in L^2(I)$  telles que

$$u_n \rightarrow u, \quad \text{et} \quad \partial_x u_n \rightarrow g.$$

Pour toute fonction test  $\varphi$ , et pour tout  $n$ , on a

$$\int_I u_n(x) \varphi'(x) dx = - \int_I \partial_x u_n(x) \varphi(x) dx,$$

et il est clair que les convergences établies plus haut permettent de passer à la limite dans cette formule (la fonction test  $\varphi$  étant fixée !). Ceci prouve que  $u \in H^1(I)$  et que  $\partial_x u = g$ .

La convergence de  $(u_n)_n$  vers  $u$  pour la norme  $H^1$  est alors immédiate.

2. Soit  $v$  la fonction définie sur  $I$  par

$$v(x) = \int_a^x \partial_x u(t) dt, \quad \forall x \in I.$$

D'après le théorème de convergence dominée de Lebesgue, cette fonction est continue sur  $\bar{I}$ . Vérifions que  $v \in H^1(I)$  et que  $\partial_x v = \partial_x u$ . Pour cela, on choisit une fonction-test  $\varphi$  et on calcule, en utilisant le théorème de Fubini

$$\int_a^b \left( \int_a^x \partial_x u(t) dt \right) \varphi'(x) dx = \int_a^b \left( \int_t^b \varphi'(x) dx \right) \partial_x u(t) dt = - \int_a^b \varphi(t) \partial_x u(t) dt.$$

Donc,  $u - v$  est une fonction de  $H^1(I)$  à dérivée nulle au sens des distributions. D'après la proposition A.20, il existe donc une constante  $C$  telle que  $u = C + v$  presque partout.

Ceci montre que  $C + v$  est un représentant continu de  $u$ . En identifiant  $u$  à ce représentant continu et en utilisant la formule qui définit  $v$  on obtient bien la formule annoncée.

3. Ce résultat, un peu technique, est admis. Il se démontre de la même façon que la densité des fonctions régulières dans  $L^2(I)$  par extension puis convolution.

Bien entendu la construction de cet espace se justifie par le fait (entre autres) qu'elle permet de prendre en compte plus de fonctions que les fonctions de classe  $C^1$  tout en contenant les fonctions dérivables au sens usuel.

### Proposition III.11 (Lien entre dérivées faibles et dérivées classiques)

Soit  $u \in H^1(I)$  et  $J \subset I$  un sous-intervalle ouvert.

1. Si la restriction  $u|_J$  est une fonction de classe  $C^1$  alors la dérivée faible de  $u$  sur  $J$  coïncide presque partout avec la dérivée usuelle de  $u|_J$ .
2. Si la restriction de la dérivée faible de  $u$  sur  $J$  coïncide presque partout avec une fonction de classe  $C^k$  ( $k \geq 0$ ), alors la restriction de  $u$  à  $J$  est une fonction de classe  $C^{k+1}$  dont la dérivée usuelle coïncide avec la dérivée faible de  $u$ .

#### Preuve :

1. On prend  $\varphi$  à support compact dans le sous-intervalle  $J$ , et comme  $u|_J$  est de classe  $C^1$ , on peut intégrer par parties sur  $J$  (sans termes de bord)

$$\int_I u \varphi' = \int_J u \varphi' = - \int_J (u|_J)' \varphi = - \int_I (u|_J)' \varphi.$$

Mais par définition de la dérivée faible de  $u$  on a aussi

$$\int_I u \varphi' = - \int_I (\partial_x u) \varphi.$$

Par comparaison on a donc établi que

$$\int_I (u|_J' - \partial_x u) \varphi = 0,$$

pour toute fonction test à support dans  $J$ . D'après le lemme A.4, on en déduit que  $\partial_x u = u|_J'$  presque partout dans  $J$ .

2. On fixe un point quelconque  $y_0 \in J$  et on utilise la propriété 3 du Théorème III.10 pour écrire

$$u(y) = u(y_0) + \int_{y_0}^y (\partial_x u) dy.$$

Par hypothèse,  $\partial_x u$  est  $C^k$  sur  $J$  et  $u$  est donc une de ses primitives sur  $J$ . Le résultat est démontré.

**Exemples :** On prend  $I = ]-1, 1[$ .

— La fonction  $x \mapsto u(x) = |x|$  est dans  $H^1(I)$  et  $\partial_x u$  est la fonction de Heaviside  $\mathcal{H}$  définie par

$$\mathcal{H}(x) = \begin{cases} -1, & \text{sur } ]-1, 0[, \\ 1, & \text{sur } ]0, 1[. \end{cases}$$

— La fonction de Heaviside  $\mathcal{H}$  elle-même n'appartient pas à  $H^1(I)$  car elle n'admet pas de représentant continu (voir Théorème III.10). En fait, la dérivée de  $\mathcal{H}$  au sens des distributions est une masse de Dirac

$$\partial_x \mathcal{H} = 2\delta_0.$$

En effet, un simple calcul montre que l'on a

$$\int_{-1}^1 H(x) \varphi'(x) dx = -2\varphi(0) = -2\langle \delta_0, \varphi \rangle_{\mathcal{D}', \mathcal{D}}.$$

Il est important de se convaincre que la distribution  $\delta_0$  ne peut pas s'écrire comme un élément de  $L^2(I)$  : autrement dit il n'existe pas de fonction  $g \in L^2(I)$  telle que

$$\varphi(0) = \int_{-1}^1 g \varphi dx, \quad \forall \varphi \in \mathcal{C}_c^\infty(]-1, 1[).$$

— La fonction  $x \mapsto u(x) = |x|^\alpha$  est dans  $H^1(I)$  si et seulement si  $\alpha > 1/2$  et on a alors  $\partial_x u(x) = \alpha|x|^{\alpha-2}x$ , pour  $x \neq 0$ .

### Corollaire III.12

*L'injection canonique (i.e. l'application identité!) de  $H^1(I)$  dans  $C^0(\bar{I})$  est continue et plus précisément on a*

$$\|u\|_{L^\infty} \leq \left( \frac{1}{|b-a|} + |b-a| \right) \|u\|_{H^1}, \quad \forall u \in H^1(I).$$

*Conséquence : si  $(u_n)_n \subset H^1(I)$  vérifie  $u_n \xrightarrow{n \rightarrow \infty} u$  dans  $H^1(I)$ , alors  $(u_n)_n$  converge vers  $u$  uniformément sur  $\bar{I}$ .*

### Preuve :

D'après l'inégalité de Cauchy-Schwarz on a pour tout  $x, y$

$$|u(y)| \leq |u(x)| + |b-a|^{\frac{1}{2}} \|\partial_x u\|_{L^2}.$$

Si on intègre ceci par rapport à  $x$ , on trouve

$$|b-a| |u(y)| \leq \int_I |u(x)| dx + |b-a|^{\frac{3}{2}} \|\partial_x u\|_{L^2}.$$

Utilisant à nouveau l'inégalité de Cauchy-Schwarz on obtient

$$|b-a| |u(y)| \leq |b-a|^{\frac{1}{2}} \|u\|_{L^2} + |b-a|^{\frac{3}{2}} \|\partial_x u\|_{L^2}.$$

En prenant finalement le sup par rapport à  $y$  on trouve

$$\|u\|_{C^0(\bar{I})} \leq |b-a|^{-\frac{1}{2}} \|u\|_{L^2} + |b-a|^{\frac{1}{2}} \|\partial_x u\|_{L^2},$$

ce qui permet de conclure. ■

On peut montrer par exemple les propriétés suivantes (voir le TD).

### Proposition III.13

— Le produit de deux fonctions  $u, v$  dans  $H^1(I)$  est encore dans  $H^1(I)$  et on a l'égalité presque partout

$$\partial_x(uv) = u(\partial_x v) + (\partial_x u)v.$$

— Plus généralement, si  $u \in H^1(I)$  et  $F \in \mathcal{C}^1(\mathbb{R})$ , alors  $F(u) \in H^1(I)$  et on a l'égalité presque partout

$$\partial_x(F(u)) = F'(u)\partial_x u.$$

## II.2 L'espace $H_0^1(I)$

### Définition III.14

On appelle  $H_0^1(I)$ , le sous-espace fermé de  $H^1(I)$  constitué des fonctions de  $H^1(I)$  dont les valeurs sont nulles au bord.

Cette définition a bien un sens car on a vu que les fonctions de  $H^1(I)$  ont un unique représentant continu sur  $\bar{I}$ , ce qui légitime la notion de "valeur au bord" pour des fonctions *a priori* définies seulement presque partout.

Cette espace est un fermé car c'est le noyau de l'application **continue** (d'après le corollaire III.12) définie par

$$u \in H^1(I) \mapsto \begin{pmatrix} u(a) \\ u(b) \end{pmatrix} \in \mathbb{R}^2.$$

Le résultat suivant est fondamental dans la théorie. Il est à rapprocher du Lemme III.6.

### Proposition III.15 (Inégalité de Poincaré)

Pour tout  $u \in H_0^1(I)$ , on a l'inégalité suivante

$$\|u\|_{L^2} \leq |b-a| \|\partial_x u\|_{L^2}.$$

### Corollaire III.16

L'application  $u \mapsto \|\partial_x u\|_{L^2}$  est donc une norme sur  $H_0^1(I)$  équivalente à la norme de  $H^1(I)$ . La structure de Hilbert correspondante est équivalente à la structure héritée de  $H^1$ .

#### Notation

Pour  $u \in H^1(I)$ , on note

$$|u|_{H^1} \stackrel{\text{def}}{=} \|\partial_x u\|_{L^2},$$

que l'on appelle la **semi-norme**  $H^1$ . Le corollaire précédent dit que, sur  $H_0^1(I)$ , cette semi-norme est en fait une norme équivalente à la norme  $H^1$  standard.

La démonstration du corollaire est immédiate. Montrons l'inégalité de Poincaré :

#### Preuve :

Comme  $u(a) = u(b) = 0$ , on peut appliquer la troisième propriété du Théorème III.10, et obtenir que pour tout  $x \in I$ , on a

$$|u(x)| \leq \int_a^x |\partial_x u| dt \leq |b-a|^{\frac{1}{2}} \|\partial_x u\|_{L^2}.$$

En élevant cette inégalité au carré, puis en l'intégrant sur  $I$ , on trouve

$$\int_a^b |u(x)|^2 dx \leq |b-a|^2 \|\partial_x u\|_{L^2}^2,$$

ce qui donne le résultat. On peut remarquer qu'on a seulement utilisé le fait que  $u(a) = 0$  ici. ■

### Remarque III.17

La constante  $|b-a|$  qui apparaît dans l'inégalité de Poincaré ci-dessus n'est pas optimale. On peut montrer que la valeur optimale (i.e. la plus petite) de cette constante vaut  $\frac{|b-a|}{\pi}$ .

On a enfin le résultat suivant qui est admis.

### Théorème III.18

L'ensemble des fonctions  $\mathcal{C}_c^\infty(I)$  est dense dans  $H_0^1(I)$ .



Ce résultat démontre que  $H_0^1(I)$  est bien le complété de l'espace  $X = \{v \in \mathcal{C}^1([0, 1], \mathbb{R}), v(0) = v(1)\}$  introduit plus haut pour la norme  $H^1$ . En effet, on a les inclusions

$$\mathcal{C}_c^\infty(I) \subset X \subset H_0^1(I),$$

et donc la proposition ci-dessus établit la densité de  $X$  dans  $H_0^1(I)$  pour la norme  $H^1$ . Cet espace étant complet, on a bien affaire à l'unique complété (à isomorphisme près) de  $X$  pour cette norme.

L'espace  $H_0^1(I)$  semble donc un bon candidat pour notre analyse.

### II.3 Résolution du problème variationnel pour la corde élastique

Revenons au problème qui a motivé toute cette théorie. On a maintenant compris qu'il faut travailler dans l'espace  $H_0^1(I)$  muni de sa norme naturelle et non pas dans l'espace (plus petit)  $X$  introduit plus haut. Le problème est donc reformulé de la façon suivante :

Soit  $f \in L^2(I)$  (ou  $f \in \mathcal{C}^0(\bar{I})$ ), trouver  $u \in H_0^1(I)$  vérifiant

$$E(u) = \inf_{v \in H_0^1(I)} E(v).$$

L'espace  $H_0^1(I)$  est un espace vectoriel et dans ces conditions la preuve de l'unicité d'un éventuel minimum, qu'on a déjà effectuée est encore valable.

Démontrons maintenant l'existence d'un minimum. Pour cela, on va commencer par démontrer que la fonctionnelle  $E$  est minorée sur  $H_0^1(I)$ . En effet pour tout  $v \in H_0^1(I)$ , nous avons

$$\begin{aligned} \int_I f v \, dx &\leq \|f\|_{L^2} \|v\|_{L^2}, \quad \text{par l'inégalité de Cauchy-Schwarz} \\ &\leq |b-a| \|f\|_{L^2} \|\partial_x v\|_{L^2}, \quad \text{par l'inégalité de Poincaré, car } v \in H_0^1(I) \\ &\leq \frac{1}{2k} |b-a|^2 \|f\|_{L^2}^2 + \frac{k}{2} \|\partial_x v\|_{L^2}^2, \quad \text{par l'inégalité de Young.} \end{aligned}$$

Tout ceci montre que

$$E(v) \geq -\frac{|b-a|^2}{2k} \|f\|_{L^2}^2.$$

Comme  $E$  est minorée, son infimum est fini et par définition de celui-ci, il existe une suite minimisante, c'est-à-dire une suite de fonctions  $(u_n)_n$  dans  $H_0^1(I)$  telle que la suite de nombre réels  $(E(u_n))_n$  converge vers  $\inf_{H_0^1(I)} E$ .

Il faut maintenant démontrer que la suite  $(u_n)_n$  converge. Pour cela, on va exploiter la complétude de  $H_0^1(I)$  (on a tout fait pour cela !) et directement montrer que  $(u_n)_n$  est de Cauchy dans  $H_0^1(I)$ .

On note  $I_E = \inf_{v \in H_0^1(I)} E(v)$ . On reprend la démonstration de la propriété d'unicité (avec l'identité du parallélogramme), pour montrer

$$I_E \leq E\left(\frac{u_n + u_{n+p}}{2}\right) = \frac{1}{2}E(u_n) + \frac{1}{2}E(u_{n+p}) - \frac{k}{8} \|\partial_x u_n - \partial_x u_{n+p}\|_{L^2}^2,$$

d'où l'on déduit

$$\frac{k}{8} \|\partial_x u_n - \partial_x u_{n+p}\|_{L^2}^2 \leq \frac{1}{2}E(u_n) + \frac{1}{2}E(u_{n+p}) - I_E. \quad (\text{III.10})$$

Soit  $\varepsilon > 0$ , comme  $E(u_n)$  tend vers  $I_E$  sur  $H_0^1(I)$ , il existe  $n_0 \geq 0$  tel que

$$\forall n \geq n_0, \quad |E(u_n) - I_E| \leq \varepsilon.$$

Ainsi, pour  $n \geq n_0$  et  $p \geq 0$ , l'inégalité (III.10) fournit

$$\frac{k}{8} \|\partial_x u_n - \partial_x u_{n+p}\|_{L^2}^2 \leq \varepsilon.$$

Cela s'écrit encore

$$\|u_n - u_{n+p}\|_{H^1}^2 \leq \frac{8\varepsilon}{k}, \quad \forall n \geq n_0, \forall p \geq 0.$$

Comme la semi-norme  $H^1$  est une norme sur  $H_0^1(I)$  équivalente à la norme  $H^1$ . Ceci montre bien que  $(u_n)_n$  est de Cauchy dans  $H_0^1(I)$ , donc par complétude de cet espace, nous savons qu'elle converge vers une certaine fonction  $u \in H_0^1(I)$ .

On montre ensuite que  $E(u) = \lim_{n \rightarrow \infty} E(u_n)$  en passant à la limite dans tous les termes de  $E(u_n)$  (dans le premier c'est une conséquence de la convergence  $H^1$  et dans le second c'est la convergence  $L^2$  qui donne le résultat). Autrement dit, on démontre ici la continuité de la fonction  $E$ .

Finalement, comme  $(u_n)_n$  est une suite minimisante, on a bien  $E(u) = I_E$ .

On peut maintenant reformuler les équations d'Euler-Lagrange associées à ce problème de minimisation exactement de la façon dont on l'a fait dans la section I.2.c et obtenir

$$u \in H_0^1(I) \text{ et vérifie } \forall v \in H_0^1(I), \int_I k \partial_x u \partial_x v \, dx = \int_I f v \, dx. \quad (\text{III.11})$$

Notons que cette formulation variationnelle admet, elle-aussi, une unique solution. En effet, si  $u_1$  et  $u_2$  sont deux telles solutions, alors leur différence  $\bar{u} \in H_0^1(I)$  vérifie la même formulation avec terme source nul. En prenant  $v = \bar{u}$  dans la formulation, on obtient

$$k \int_I |\partial_x \bar{u}|^2 \, dx = 0,$$

ce qui prouve que  $\bar{u}$  est une constante qui ne peut être que 0 car  $\bar{u}$  est nulle au bord.

**Régularité de la solution :** Dans le cadre de la dimension 1, il est facile maintenant de démontrer que la solution  $u \in H_0^1(I)$  obtenue est suffisamment régulière et vérifie le problème de Poisson au sens usuel.

En effet, comme  $\mathcal{C}_c^\infty(I) \subset H_0^1(I)$ , on peut prendre tous les éléments de  $\mathcal{C}_c^\infty(I)$  comme fonctions test dans les équations d'Euler-Lagrange (III.11) pour obtenir

$$\forall \varphi \in \mathcal{C}_c^\infty(I), \int_0^1 (\partial_x u) \partial_x \varphi \, dx = \int_0^1 \frac{f}{k} \varphi \, dx.$$

Par définition, ceci montre que  $\partial_x u$  est elle-même un élément de l'espace de Sobolev  $H^1(I)$  qui admet  $-\frac{1}{k}f$  comme dérivée faible dans  $L^2$ . On a donc, avec des notations évidentes,

$$\partial_x^2 u = \partial_x(\partial_x u) = -\frac{f}{k},$$

ce qui montre bien que  $u$  vérifie l'équation aux dérivées partielles  $-k\partial_x^2 u = f$  en un sens faible et bien sûr aussi les conditions aux limites (car  $u$  est dans l'espace  $H_0^1(I)$  des fonctions nulles au bord).

On peut en dire davantage sur la régularité de  $u$ . En effet, comme  $\partial_x u \in H^1(I)$ , le théorème III.10 nous dit que  $\partial_x u \in \mathcal{C}^0(\bar{I})$  (ou en tout cas, admet un représentant continu). Ainsi, la Proposition III.11 montre que  $u$  est de classe  $\mathcal{C}^1$  sur  $\bar{I}$ .

#### Remarque III.19

*La solution  $u$  ainsi obtenue appartient à l'espace  $X$  défini au début de l'analyse, et comme  $X \subset H_0^1(I)$ , la fonction  $u$  réalise aussi le minimum de  $E$  sur  $X$  :*

$$E(u) = \inf_{v \in X} E(v).$$

*Néanmoins, bien que  $u$  réalise aussi le minimum de  $E$  sur  $X$ , il faut bien comprendre que l'on ne pouvait pas établir directement cette propriété en travaillant dans  $X$  et que l'introduction des espaces de Sobolev pour résoudre ce problème est absolument cruciale !*

Jusqu'à présent, on a seulement utilisé le fait que  $f \in L^2(I)$ . Si maintenant, on fait l'hypothèse supplémentaire que le terme source  $f$  est une fonction continue sur  $\bar{I}$ , on peut aller plus loin car on a alors

$$\partial_x^2 u = -\frac{f}{k} \in \mathcal{C}^0(\bar{I}).$$

La Proposition III.11 montre à nouveau que  $\partial_x u$  est de classe  $\mathcal{C}^1$  et donc  $u$  est de classe  $\mathcal{C}^2$ . Ainsi, la fonction  $u$  résout bien le problème aux dérivées partielles suivant, au sens classique :

$$\begin{cases} -k\partial_x^2 u = f, & \text{dans } I \\ u(0) = u(1) = 0. \end{cases}$$

### III Formulation variationnelle d'un problème aux limites linéaires. Théorème de Lax-Milgram.

On a vu sur l'exemple de la corde élastique que, en partant d'un problème d'optimisation à résoudre, on peut écrire les équations d'Euler-Lagrange qui caractérisent les points critiques de la fonctionnelle étudiée. Si ces équations ont une solution  $u$ , alors on peut essayer d'en déduire l'équation aux dérivées partielles vérifiée par la solution  $u$  ainsi que les conditions aux limites. Cela permet donc de construire une solution d'un problème aux limites.

De façon plus générale, on cherche maintenant à inverser le processus et à établir l'existence et l'unicité de solutions de problèmes aux limites *via* une "formulation variationnelle" même si on verra qu'elle n'est pas nécessairement issue d'un problème de calcul des variations.

### III.1 Principe général

On se donne un problème aux limites **linéaire** de la forme abstraite suivante

$$\begin{cases} Au = f, \text{ posée dans un domaine } \Omega \text{ de } \mathbb{R}^d, \\ + \text{ conditions aux limites sur } u \text{ ou ces dérivées } \dots, \end{cases} \quad (\text{III.12})$$

où  $A$  est un opérateur différentiel linéaire (i.e.  $Au$  est une fonction de  $\Omega$  dans  $\mathbb{R}$  construite à partir de différentes dérivées de  $u$ ),  $f$  est un terme source donné.

Soit  $X$  un espace fonctionnel dans lequel on va chercher la solution  $u$ . On choisit également un espace  $V$  de fonctions tests et on multiplie l'équation  $Au = f$  par un élément de  $v$  qu'on intègre ensuite sur le domaine  $\Omega$ . Ainsi toute solution éventuelle  $u \in X$  va vérifier

$$\int_{\Omega} Au.v \, dx = \int_{\Omega} f v \, dx, \quad \forall v \in V. \quad (\text{III.13})$$

On a vu que si  $V$  est suffisamment gros (s'il contient les fonctions de  $C_c^\infty(I)$  par exemple) alors le lemme de du Bois-Reymond (Lemme A.4) permettra de remonter de la formulation (III.13) à (III.12).

Maintenant, on intègre par parties un certain nombre de fois le membre de gauche (ou pas, plusieurs choix sont possibles ici) pour mettre ce problème sous la forme suivante

$$\int_{\Omega} (B_1 u).(B_2 v) \, dx = \int_{\Omega} f v \, dx + \dots \text{termes de bord éventuels}, \quad \forall v \in V. \quad (\text{III.14})$$

On a fait porter certaines dérivées sur la fonction test  $v$  et on a conservé les autres dérivées sur la fonction inconnue  $u$ . A ce stade, il s'agit de calculs formels : on ne se pose pas encore la question de l'existence de  $u$ , ni de sa régularité ... en bref, tous les coups sont permis ici.

Bien entendu, il ne faut pas faire n'importe quoi et le but du jeu est maintenant de faire en sorte que les équations obtenues après ces manipulations formelles aient une solution unique et que l'on puisse ensuite montrer que l'on a bien résolu le problème initial. Il s'agit donc de bien choisir :

- L'espace des solutions  $X$ .
- L'espace des fonctions test  $V$ .
- La forme exacte des intégrations par parties, c'est-à-dire des opérateurs  $B_1$  et  $B_2$  obtenus à partir de l'opérateur  $A$  de départ.

Notons qu'un même problème aux limites peut avoir plusieurs formulations variationnelles de la forme (III.14). Dans les cas usuels qu'on verra par la suite il y a souvent un choix *naturel*.

Pour choisir ces objets, il faut se raccrocher à un résultat abstrait permettant ensuite d'assurer l'existence de la solution au problème variationnel (III.14). On observe que les équations de (III.14) comportent des termes de deux types : des termes bilinéaires en  $(u, v)$ , qui sont les termes principaux et qu'on met usuellement dans le membre de gauche de l'égalité, et des termes linéaires en  $v$  et qui ne contiennent pas  $u$ , qui correspondent à la contribution des termes sources et éventuellement des termes de bord ; on les met usuellement dans le membre de droite de l'égalité.

Voici le résultat de base de cette théorie. Vous l'avez sûrement rencontré dans le cours d'Analyse Fonctionnelle.

### Théorème III.20 (Lax-Milgram)

Soit  $H$  un espace de Hilbert réel,  $a$  une forme bilinéaire sur  $H$ ,  $L$  une forme linéaire sur  $H$ . On suppose que

1.  $a$  est continue :

$$\forall u, v \in H, |a(u, v)| \leq \|a\| \|u\|_H \|v\|_H.$$

2.  $a$  est coercive : il existe  $\alpha > 0$  tel que

$$\forall u \in H, a(u, u) \geq \alpha \|u\|_H^2.$$

3.  $L$  est continue :

$$\forall u \in H, |L(u)| \leq \|L\| \|u\|_H.$$

Alors, il existe un unique  $u$  dans  $H$  qui vérifie

$$\forall v \in H, a(u, v) = L(v), \quad (\text{III.15})$$

et celui-ci vérifie

$$\|u\|_H \leq \frac{\|L\|}{\alpha}.$$

Si de plus  $a$  est symétrique, alors  $u$  est aussi l'unique élément de  $H$  qui minimise la fonctionnelle

$$J(v) = \frac{1}{2} a(v, v) - L(v).$$

Ce théorème donne des conditions suffisantes de résolubilité d'un jeu d'équations de la forme (III.15). Insistons sur le fait que ce ne sont pas des conditions suffisantes. Il existe un résultat du même type donnant des conditions nécessaires et suffisantes d'existence et unicité (y compris dans des espaces de Banach) mais il sort du cadre de ce cours.

Si on veut pouvoir appliquer le théorème de Lax-Milgram dans le cadre abstrait introduit précédemment on voit donc qu'il va falloir :

- Choisir les mêmes espaces  $X = V$  et prendre un espace de Hilbert, noté dorénavant  $H$ .
- Poser (en oubliant pour l'instant les termes de bord qui selon les cas interviennent ou bien dans  $a$  ou bien dans  $L$ )

$$a(u, v) = \int_{\Omega} (B_1 u) \cdot (B_2 v) dx, \quad \forall u, v \in H,$$

$$L(v) = \int_{\Omega} f v dx, \quad \forall v \in H.$$

- Pour assurer que ces objets soient bien définis et continus sur  $H$  il faut typiquement que  $B_1$  et  $B_2$  soient bien définis et continus sur  $H$  à valeurs dans  $L^2(\Omega)$ . Si on imagine que  $B_1$  et  $B_2$  sont des opérateurs différentiels du premier ordre, il faudra donc que  $H$  contienne au moins l'espace de Sobolev  $H^1$ .
- Pour assurer la coercivité de  $a$ , il faut que

$$a(u, u) = \int_{\Omega} (B_1 u) \cdot (B_2 u) dx \geq \alpha \|u\|_H^2,$$

ceci montre que la norme de  $H$  doit être contrôlée par la quantité  $a(u, u)$  donc il ne faut pas que cette norme contienne plus de dérivées que ce que  $B_1$  et  $B_2$  n'en contiennent.

En résumé, si on se restreint aux problèmes aux limites du second ordre, en dimension 1 pour l'instant. On voit qu'on va pouvoir systématiquement choisir pour  $H$  un sous-espace fermé de l'espace  $H^1(I)$ . Celui-ci dépendra notamment des conditions aux limites choisies. Le fait qu'il soit fermé est indispensable pour qu'il soit lui-même un espace de Hilbert.

## III.2 Exemples en dimension 1

### III.2.a Problème de Poisson avec conditions de Dirichlet

Soit  $I = [0, 1]$ . On se donne une fonction  $k : I \rightarrow \mathbb{R}$  mesurable positive, bornée et telle que  $\alpha = \inf_I k > 0$  et un terme source  $f \in L^2(I)$ .

On considère le problème aux limites suivant

$$\begin{cases} -\partial_x(k(x)\partial_x u) = f(x), & x \in I, \\ u(0) = u(1) = 0. \end{cases} \quad (\text{III.16})$$

Ce n'est autre que le problème de la corde élastique déjà étudié plus haut que l'on va reprendre avec l'éclairage nouveau ci-dessus.

Le cadre fonctionnel adéquat est  $H = H_0^1(I)$  et on pose

$$a(u, v) = \int_I k(x)(\partial_x u)(\partial_x v) dx,$$

et

$$L(v) = \int_I f(x)v dx.$$

L'espace  $H$  est bien complet, la forme bilinéaire  $a$  est continue sur  $H$  car  $k$  est bornée et l'inégalité de Cauchy-Schwarz donne

$$|a(u, v)| \leq \|k\|_\infty \|\partial_x u\|_{L^2} \|\partial_x v\|_{L^2} \leq \|k\|_\infty \|u\|_H \|v\|_H.$$

De même, on trouve

$$|L(v)| \leq \|f\|_{L^2} \|v\|_{L^2} \leq \|f\|_{L^2} \|v\|_H.$$

Pour la coercivité de  $a$ , on utilise les hypothèses sur  $k$  pour obtenir

$$a(u, u) = \int_I k |\partial_x u|^2 dx \geq \alpha \|\partial_x u\|_{L^2}^2,$$

et d'après l'inégalité de Poincaré (Proposition III.15 et son Corollaire III.16), on obtient bien

$$a(u, u) \geq C \|u\|_H^2.$$

Les hypothèses du théorème de Lax-Milgram sont donc vérifiées et on peut en déduire l'existence et l'unicité d'une solution  $u \in H_0^1(I)$  de la formulation variationnelle de (III.16) donnée par

$$\int_I k(x)(\partial_x u)(\partial_x v) dx = \int_I f(x)v dx, \quad \forall v \in H_0^1(I).$$

**Conditions aux limites non homogènes** Essayons maintenant de résoudre le même problème mais avec des données au bord non nulles. Les conditions aux limites sont donc de même nature (Dirichlet) mais sont non-homogènes. On se donne deux valeurs  $u_0, u_1 \in \mathbb{R}$  et on s'intéresse à

$$\begin{cases} -\partial_x(k(x)\partial_x u) = f(x), & x \in I, \\ u(0) = u_0, \\ u(1) = u_1. \end{cases} \quad (\text{III.17})$$

Il est clair que cette fois, on ne peut pas chercher  $u \in H_0^1(I)$  car sinon elle serait nulle au bord. On va donc *relever* les conditions aux limites de la façon suivante :

On prend une fonction  $R : I \rightarrow \mathbb{R}$  (comme *relèvement*) suffisamment régulière mais quelconque, qui vérifie

$$R(0) = u_0, \quad R(1) = u_1.$$

L'idée est alors de chercher  $u$  sous la forme  $u = R + w$  avec cette fois  $w$  qui est nulle au bord. Ecrivons l'équation satisfaite par  $w$

$$\begin{cases} -\partial_x(k(x)\partial_x w) = f + \partial_x(k(x)\partial_x R), & x \in I, \\ w(0) = w(1) = 0. \end{cases}$$

C'est donc un problème de Poisson avec conditions homogène et le nouveau terme source

$$\tilde{f} = f + \partial_x(k(x)\partial_x R).$$

On a choisi  $R$  régulière, donc  $\partial_x R$  est également régulière.

- Si  $k$  est régulière, la fonction  $\tilde{f}$  est bien définie et appartient à  $L^2$  donc l'existence et l'unicité de  $w \in H_0^1$  solution du problème ci-dessus est conséquence de l'analyse du cas homogène. On en déduit donc l'existence et unicité de la solution  $u \in H^1$  du problème initial.

- Si par contre on ne suppose pas de régularité particulière de  $k$  (si  $k$  est discontinue par exemple), la fonction  $\tilde{f}$  n'est peut-être pas bien définie ... l'idée est alors d'intégrer par parties le terme gênant dans la définition de la forme linéaire  $L$ .

Plus précisément, au lieu de définir  $L(v) = \int_I \tilde{f}v \, dx$ , on va poser

$$L(v) = \int_I f v \, dx - \int_I k(x) \partial_x R \partial_x v \, dx.$$

On vérifie aisément que  $L$  est linéaire continue sur  $H$  et on peut à nouveau appliquer le théorème de Lax-Milgram à ce nouveau cadre et ainsi obtenir l'existence et unicité de la solution  $w \in H_0^1$  du problème

$$\int_I k(x) \partial_x w \partial_x v \, dx = \int_I f v \, dx - \int_I k(x) \partial_x R \partial_x v \, dx, \quad \forall v \in H_0^1,$$

ce qui s'écrit encore

$$\int_I k(x) \partial_x (w + R) \partial_x v \, dx = \int_I f v \, dx, \quad \forall v \in H_0^1,$$

et donne donc la solution  $u = w + R$ .

### III.2.b Ajout d'un terme de réaction linéaire

Sous les mêmes hypothèses que précédemment, on se donne maintenant un nombre réel  $\gamma \in \mathbb{R}$  quelconque et on s'intéresse au problème suivant

$$\begin{cases} -\partial_x(k(x)\partial_x u) + \gamma u = f(x), & x \in I, \\ u(0) = u(1) = 0. \end{cases} \quad (\text{III.18})$$

On peut reproduire le formalisme précédent en posant  $H = H_0^1(I)$ ,

$$a(u, v) = \int_I k(x) (\partial_x u) (\partial_x v) \, dx + \gamma \int_I u v \, dx,$$

$$L(v) = \int_I f v \, dx.$$

Toutes les hypothèses du théorème de Lax-Milgram sont clairement vérifiées sauf peut-être la coercivité de  $a$ . En effet, on calcule aisément que

$$a(u, u) = \int_I k |\partial_x u|^2 \, dx + \gamma \int_I |u|^2 \, dx.$$

- Si  $\gamma \geq 0$ , alors le second terme est positif ou nul et on peut donc à nouveau minorer  $a(u, u)$  par  $\|u\|_H^2$  à une constante multiplicative près.
- Si  $\gamma < 0$ , le second terme a le mauvais signe et il est donc nécessaire de le contrôler.

On peut déjà remarquer que si  $\gamma$  est "trop" négatif alors  $a$  n'est pas coercive. En effet, pour  $\bar{u} \in H_0^1(I)$  quelconque non nul, on peut toujours trouver un  $\gamma < 0$  tel que

$$a(\bar{u}, \bar{u}) = \int_I k |\partial_x \bar{u}|^2 \, dx + \gamma \int_I |\bar{u}|^2 \, dx < 0,$$

ce qui implique la non coercivité de  $a$ .

Il y a là un obstacle tout à fait réel car on peut démontrer<sup>4</sup>, par exemple, qu'il n'existe aucune solution du problème aux limites

$$\begin{cases} -\partial_x^2 u - \pi^2 u = \sin(\pi x), & x \in (0, 1), \\ u(0) = u(1) = 0. \end{cases}$$

### III.2.c Problème de convection-diffusion avec conditions de Dirichlet

Considérons maintenant le problème au bord suivant

$$\begin{cases} -\partial_x^2 u + \partial_x u = f(x), & x \in I, \\ u(0) = u(1) = 0. \end{cases} \quad (\text{III.19})$$

4. Ce n'est d'ailleurs pas très difficile ...

Par rapport au problème de Poisson vu plus haut, on a ajouté un terme d'ordre 1 dit *de convection* ou de *de transport*. Celui-ci étant d'ordre inférieur au terme principal, on peut conserver l'espace fonctionnel  $H = H_0^1(I)$  comme espace de travail et adopter la formulation faible suivante

$$a(u, v) = \int_0^1 (\partial_x u)(\partial_x v) dx + \int_0^1 (\partial_x u)v dx,$$

$$L(v) = \int_0^1 f v dx.$$

Là encore la continuité de  $a$  et  $L$  sur  $H$  ne fait aucun doute ; étudions la coercivité de  $a$ . Pour  $u \in H$ , nous avons

$$a(u, u) = \int_0^1 |\partial_x u|^2 dx + \int_0^1 (\partial_x u)u dx.$$

A priori le second terme n'a pas de signe clair et cela peut donc poser un problème. En réalité, on va voir que ce terme est nul. En effet, nous savons (voir TD) que  $u^2 \in H_0^1(I)$  et que  $\partial_x(u^2) = 2u\partial_x u$  de sorte que

$$\int_0^1 (\partial_x u)u dx = \frac{1}{2} \int_0^1 \partial_x(u^2) dx = \frac{1}{2} u^2(1) - \frac{1}{2} u^2(0) = 0.$$

Il s'en suit que  $a$  est coercive et donc que l'on peut appliquer le théorème de Lax-Milgram. La fin de l'histoire est la même que pour le problème de Dirichlet : on prend  $v = \varphi \in \mathcal{C}_c^\infty(I)$  dans la formulation faible,

$$\int_0^1 \partial_x u \partial_x \varphi dx = \int_0^1 (f - \partial_x u) \varphi dx.$$

Comme  $f \in L^2$  et  $\partial_x u \in L^2$ , on en déduit que  $\partial_x u$  est en fait dans  $H^1(I)$  et qu'elle vérifie, au sens faible,

$$\partial_x(\partial_x u) = -(f - \partial_x u).$$

Si maintenant on suppose que  $f \in \mathcal{C}^0(\bar{I})$ , sachant que  $\partial_x u \in \mathcal{C}^0(\bar{I})$  d'après le Corollaire III.12, nous obtenons que  $\partial_x(\partial_x u)$  est continue et donc  $\partial_x u \in \mathcal{C}^1(\bar{I})$  et donc  $u \in \mathcal{C}^2(\bar{I})$ . Ainsi la fonction  $u$  ainsi construite est bien une solution au sens usuel du problème (III.19).

### Remarque III.21

*Dans ce cas précis, la forme bilinéaire  $a$  n'est pas symétrique sur  $H$  (s'en convaincre !) et donc, bien que l'on parle encore de méthode variationnelle, il faut bien comprendre que ce problème n'est pas issu d'un problème de minimisation d'une fonctionnelle !*

Dans l'analyse précédente, on a bénéficié d'un *miracle* lié à l'annulation du terme nouveau dans l'estimation de coercivité. On va voir que, malheureusement, ceci n'est pas toujours vrai. Considérons par exemple le problème suivant

$$\begin{cases} -\partial_x^2 u + b(x)\partial_x u = f(x), & x \in I, \\ u(0) = u(1) = 0, \end{cases} \quad (\text{III.20})$$

où  $b : \bar{I} \rightarrow \mathbb{R}$  est une fonction régulière donnée. La forme bilinéaire associée à ce problème s'écrit (l'espace  $H$  et la forme linéaire  $L$  restant inchangés)

$$a(u, v) = \int_0^1 (\partial_x u)(\partial_x v) dx + \int_0^1 b(x)(\partial_x u)v dx,$$

et le calcul de coercivité devient

$$a(u, u) = \int_0^1 |\partial_x u|^2 dx + \int_0^1 b(x)(\partial_x u)u dx.$$

On peut, par exemple, intégrer par parties le second terme et obtenir

$$\int_0^1 b(x)(\partial_x u)u dx = \frac{1}{2} \int_0^1 b(x)\partial_x(u^2) dx = -\frac{1}{2} \int_0^1 b'(x)u^2 dx.$$

On observe que ce terme n'est plus nul et peut donc poser problème. Regardons quelques exemples :

- Si  $b$  est décroissante, alors le terme en question est positif ou nul et on a donc

$$a(u, u) \geq \int_0^1 |\partial_x u|^2 dx \geq C \|u\|_H^2,$$

et le théorème de Lax-Milgram s'applique.

- Si  $b$  est de la forme  $b(x) = -\gamma x$  alors on a

$$a(u, u) = \int_0^1 |\partial_x u|^2 dx + \frac{\gamma}{2} \int_0^1 u^2 dx.$$

Comme on l'a vu plus haut, pour  $\gamma$  suffisamment négatif, cette quantité peut devenir négative et ainsi on ne peut pas appliquer le théorème de Lax-Milgram.

- Supposons maintenant que  $b$  est “petite” en un sens qui sera précisé ci-dessous. On peut donc écrire avec Cauchy-Schwarz

$$\left| \int_0^1 b(x)(\partial_x u)u dx \right| \leq \|b\|_\infty \|\partial_x u\|_{L^2} \|u\|_{L^2}$$

puis en utilisant l'inégalité de Poincaré (Proposition III.15)

$$\left| \int_0^1 b(x)(\partial_x u)u dx \right| \leq \|b\|_\infty \|\partial_x u\|_{L^2}^2,$$

de sorte que

$$a(u, u) \geq (1 - \|b\|_\infty) \|\partial_x u\|_{L^2}^2.$$

Ainsi, on vient de montrer que si  $\|b\|_\infty < 1$  alors la forme bilinéaire  $a$  est coercive et toute la théorie peut se dérouler sans soucis.

- Remarquons que les deux cas dans lesquels nous avons pu appliquer le théorème de Lax-Milgram sont réellement différents :
  - Cas 1 :  $b$  est décroissante mais elle peut être aussi grande que l'on veut.
  - Cas 2 :  $b$  est petite mais sans aucune hypothèse de monotonie.

Concluons cette discussion en mentionnant que, bien que le théorème de Lax-Milgram ne s'applique pas toujours pour ce problème, il est possible de montrer l'existence d'une solution pour n'importe quelle fonction  $b$  mais cela utilise d'autres techniques qui dépassent le cadre de ce cours. Il faut seulement retenir que le théorème de Lax-Milgram, s'il est très utile, ne fonctionne pas à tous les coups.

### III.2.d Conditions aux limites de Neumann

Les conditions de Neumann consistent à imposer les valeurs des dérivées de la solution au bord du domaine, plutôt que de la solution elle-même. Ainsi, dans le modèle de la corde élastique par exemple, imposer  $\partial_x u(0) = 0$  revient à dire que les forces de tension à l'extrémité gauche de la corde sont nulles ou encore que la corde est fixée sur un axe vertical sur lequel elle peut glisser librement (sans frottement).

**Le cas mixte Dirichlet-Neumann** Essayons donc de résoudre le nouveau problème suivant

$$\begin{cases} -\partial_x^2 u = f(x), & x \in I, \\ \partial_x u(0) = 0 \\ u(1) = 0. \end{cases} \quad (\text{III.21})$$

On commence par supposer l'existence d'une solution  $u$  assez régulière puis par prendre une fonction test  $v$  (suffisamment régulière pour l'instant et sans préciser l'espace dans lequel elle vit). On multiplie l'équation par  $v$ , on intègre et on essaie de le mettre sous une forme du type “Lax-Milgram”.

$$\begin{aligned} \int_0^1 f v dx &= \int_0^1 (-\partial_x^2 u) v dx \\ &= [-(\partial_x u) v]_0^1 + \int_0^1 (\partial_x u)(\partial_x v) dx \\ &= (\partial_x u)(0)v(0) - (\partial_x u)(1)v(1) + \int_0^1 (\partial_x u)(\partial_x v) dx. \end{aligned}$$



Observons les termes de bord. Le premier doit être nul car on demande que  $\partial_x u(0) = 0$  (du coup la valeur de  $v(0)$  n'a aucune importance dans ce calcul). Le second n'a *a priori* aucune raison d'être nul. En revanche, on se souvient que, pour faire fonctionner le théorème de Lax-Milgram, il faut que la solution  $u$  et les fonctions test  $v$  habitent dans le même espace. Ainsi, comme on veut que  $u(1) = 0$  (et que cette condition va être directement prise en compte dans la définition de l'espace fonctionnel) on va aussi se restreindre à des fonctions test vérifiant  $v(1) = 0$ , ce qui va donc faire *disparaître* le second terme de bord.

Résumons : on va prendre

$$H = \{v \in H^1(I), v(1) = 0\},$$

$$a(u, v) = \int_0^1 (\partial_x u)(\partial_x v) dx, \quad \forall u, v \in H,$$

$$L(v) = \int_0^1 f v dx.$$

**Attention :** Il peut sembler que  $a$  et  $L$  sont les mêmes formes (bi-)linéaires que dans le cas de conditions de Dirichlet mais il n'en est rien car elles ne sont pas définies sur les mêmes espaces.

On observe que  $H$  est un sous-espace fermé de  $H^1$  (donc un Hilbert), et que  $a$  et  $L$  sont continues sur  $H$ . Encore une fois il reste à étudier la coercivité de  $a$ . Comme précédemment nous avons

$$a(u, u) = \int_0^1 |\partial_x u|^2 dx,$$

mais comme nous ne sommes plus sur l'espace  $H_0^1(I)$ , nous ne savons pas si on peut utiliser l'inégalité de Poincaré pour contrôler toute la norme  $H^1$  par ce terme. En réalité, on peut le faire à bon droit car, si on reprend la même preuve que dans la Proposition III.15, on peut montrer que

$$\|u\|_{L^2} \leq \|\partial_x u\|_{L^2}, \quad \forall u \in H.$$

Autrement dit, l'inégalité de Poincaré est vraie aussi pour les fonctions ne s'annulant qu'en un des deux points du bord.

Au final, on a donc montré que  $a$  est coercive, et le théorème de Lax-Milgram nous donne donc l'existence et unicité d'une solution  $u \in H$  du problème

$$a(u, v) = L(v), \quad \forall v \in H.$$

Qu'avons nous résolu ? Pour l'instant nous savons que  $u \in H^1(I)$  et que  $u(1) = 0$  car c'est la définition de  $H$ . Il reste à montrer que  $u$  vérifie l'EDP et la seconde condition aux limites. Nous avons encore  $\mathcal{C}_c^\infty(I) \subset H$  et donc nous pouvons prendre  $v = \varphi \in \mathcal{C}_c^\infty(I)$  comme fonction test dans la formulation variationnelle. C'est exactement le même calcul que précédemment. Il nous montre que  $\partial_x u$  est en fait une fonction de  $H^1(I)$  dont la dérivée faible vaut  $f$ . En particulier  $\partial_x u$  est continue.

Prenons maintenant  $v \in H$  quelconque et calculons en intégrant par parties

$$\begin{aligned} - \int_0^1 f v dx &= \int_0^1 \partial_x (\partial_x u) v dx \\ &= [(\partial_x u) v]_0^1 - \int_0^1 \partial_x u \partial_x v dx. \end{aligned}$$

Comme  $v(1) = 0$ , la formule ci-dessus s'écrit

$$a(u, v) - L(v) = (\partial_x u)(0)v(0),$$

et doit être vérifiée pour tous les  $v \in H$ . Mais on sait, par la définition de la formulation variationnelle vérifiée par  $u$ , que le membre de gauche de l'égalité est nul pour tout  $v \in H$ . On a donc finalement montré que

$$(\partial_x u)(0)v(0) = 0, \quad \forall v \in H.$$

Comme  $H$  contient des fonctions non nulles en 0 (par exemple  $v(x) = 1 - x$ ) on conclut que

$$\partial_x u(0) = 0,$$

et le problème initial est bien résolu.

**Le cas Neumann partout** Que se passe-t-il si on impose une condition de Neumann aux deux bords de l'intervalle.

$$\begin{cases} -\partial_x^2 u = f(x), & x \in I, \\ \partial_x u(0) = \partial_x u(1) = 0. \end{cases}$$

Deux remarques importantes sont nécessaires :

- La solution n'intervient que par l'intermédiaire de sa dérivée dans le problème, ce qui montre que si  $u$  est solution, alors  $u + c$  est solution pour tout  $c \in \mathbb{R}$ . Il ne peut donc en aucun cas y avoir unicité de la solution (si elle existe). Pour “fixer” la constante  $c$ , on est souvent amenés à imposer une condition supplémentaire, par exemple le fait que  $u$  doit être d'intégrale nulle (i.e. de moyenne nulle).
- S'il existe une solution, on peut intégrer l'équation sur l'intervalle  $I$  et on trouve

$$\int_0^1 f(x) dx = \int_0^1 (-\partial_x^2 u) dx = \partial_x u(0) - \partial_x u(1) = 0.$$

Donc, il ne peut y avoir de solution que pour les termes sources de moyenne nulle. On parle de condition de compatibilité entre le terme source et les conditions au bord.

On peut donc maintenant préciser les choses et considérer le problème complet suivant

$$\begin{cases} -\partial_x^2 u = f(x), & x \in I, \\ \partial_x u(0) = \partial_x u(1) = 0, \\ \int_0^1 u dx = 0, \end{cases} \quad (\text{III.22})$$

sous l'hypothèse

$$\int_0^1 f dx = 0. \quad (\text{III.23})$$

L'expérience des exemples précédents nous montre que les conditions aux limites de Neumann ne doivent pas apparaître dans la définition de l'espace fonctionnel<sup>5</sup>. En revanche, il semble naturel d'imposer “en dur” la contrainte de moyenne nulle.

On va donc introduire l'espace des fonctions à moyenne nulle,  $H = H_m^1(I)$  où

$$H_m^1(I) = \{v \in H^1(I), m(v) = 0\}, \text{ ou on a noté } m(v) = \int_0^1 v dx.$$

C'est un sous-espace fermé de  $H^1(I)$  et donc un Hilbert.

Posons

$$a(u, v) = \int_0^1 (\partial_x u)(\partial_x v) dx, \quad L(v) = \int_0^1 f v dx.$$

Encore une fois, c'est la coercivité de  $a$  qui peut poser soucis. On a à nouveau

$$a(u, u) = \int_0^1 |\partial_x u|^2 dx,$$

et on a donc besoin d'une inégalité de type Poincaré. Il se trouve que cette inégalité est vraie :

**Proposition III.22 (inégalité de Poincaré-moyenne ou de Poincaré-Wirtinger)**

Il existe  $C > 0$  telle que

$$\|u\|_{L^2} \leq C \|\partial_x u\|_{L^2}, \quad \forall u \in H_m^1.$$

**Preuve :**

Pour  $s, t \in I$ , on écrit

$$u(t) = u(s) + \int_s^t \partial_x u(x) dx,$$

puis on intègre cette égalité par rapport à  $s$  pour pouvoir utiliser la condition de moyenne nulle sur  $u$

$$u(t) = \underbrace{\int_0^1 u(s) ds}_{=m(u)=0} + \int_0^1 \left( \int_s^t \partial_x u dx \right) ds.$$

5. D'ailleurs pour une fonction quelconque de  $H^1$ , la valeur de  $\partial_x u(0)$  n'est pas bien définie...

On a donc, pour tout  $t$ ,

$$|u(t)| \leq \int_0^1 \left| \int_s^t |\partial_x u| dx \right| ds \leq \int_0^1 |\partial_x u| dx \leq \|\partial_x u\|_{L^2}.$$

On élève au carré et on intègre par rapport à  $t$  pour trouver le résultat. ■

D'après cette inégalité, la forme  $a$  est donc coercive sur  $H_m^1$  et le théorème de Lax-Milgram nous fournit l'existence et unicité d'une solution  $u \in H_m^1$  de

$$a(u, v) = L(v), \quad \forall v \in H_m^1.$$

Pour l'instant nous savons que  $u \in H^1$  et qu'elle est à moyenne nulle. Il nous faut montrer qu'elle vérifie l'EDP et les conditions aux limites.

Nous sommes confrontés ici à une nouvelle difficulté liée au fait que les éléments de  $C_c^\infty(I)$  ne sont pas tous dans  $H_m^1$  car ils ne sont pas tous à moyenne nulle ! On ne peut donc pas prendre  $v = \varphi$  dans la formulation variationnelle sans prendre garde. On peut par contre prendre  $v = \varphi - m(\varphi)$  qui est bien un élément de  $H_m^1$  et ainsi écrire

$$a(u, \varphi - m(\varphi)) = L(\varphi - m(\varphi)), \quad \forall \varphi \in C_c^\infty(I).$$

Comme la dérivée faible des constantes est nulle, cette formule s'écrit

$$\int_0^1 \partial_x u \partial_x \varphi dx = \int_0^1 f(\varphi - m(\varphi)) dx.$$

Le second membre peut aussi s'écrire

$$\int_0^1 f \varphi dx - m(\varphi) \int_0^1 f dx = \int_0^1 f \varphi dx,$$

grâce à l'hypothèse de compatibilité (III.23). Au final, on a bien établi que

$$\int_0^1 \partial_x u \partial_x \varphi dx = \int_0^1 f \varphi dx, \quad \forall \varphi \in C_c^\infty(I).$$

De façon maintenant habituelle ceci montre que  $u$  est solution de l'EDP au sens faible.

On prend alors n'importe quelle fonction test  $v \in H_m^1$  que l'on test contre l'équation. En utilisant la formulation variationnelle résolue plus haut on trouve in fine

$$(\partial_x u)(0)v(0) - (\partial_x u)(1)v(1) = 0, \quad \forall v \in H_m^1.$$

On choisit alors n'importe quelle fonction  $v$  telle que  $v(0) = 0$ ,  $v(1) = 1$  et  $\int_0^1 v dx = 0$  pour en déduire que  $\partial_x u(1) = 0$  puis n'importe quelle fonction telle que  $v(0) = 1$ ,  $v(1) = 0$  et  $\int_0^1 v dx = 0$  pour obtenir  $\partial_x u(0) = 0$ .

On a bien résolu le problème de Neumann initial.

### III.3 Preuve du théorème de Lax-Milgram

Il s'agit ici de redonner la preuve du théorème III.20 que vous avez déjà rencontré dans le cours d'analyse fonctionnelle.

**Le cas symétrique** Ce cas peut se traiter de deux façons (très proches en réalité) :

**Preuve 1 :** on introduit directement la fonctionnelle  $J$  de l'énoncé et on y applique exactement la même analyse que pour le cas de la corde élastique :

1. On montre que  $J$  est minorée (on utilise la continuité de  $L$  et la coercivité de  $a$ ).
2. On prend une suite minimisante  $(u_n)_n$ .
3. On démontre que  $(u_n)_n$  est de Cauchy en utilisant l'identité du parallélogramme pour la forme bilinéaire  $a$  et la coercivité de  $a$ .
4. On en déduit que  $(u_n)_n$  converge vers une limite  $u \in H$  (on utilise la complétude de  $H$ ).
5. On montre que  $J(u) = \inf_H J$ , on utilise la continuité de  $L$  et  $a$ .
6. Enfin, on écrit que  $J(u + tv) \geq J(u)$  pour tout  $v \in H$  et  $t \in \mathbb{R}$  et on en déduit les équations d'Euler-Lagrange du problème qui sont exactement les équations recherchées.

**Preuve 2 :** Comme  $a$  est symétrique et coercive (en particulier définie positive), c'est un produit scalaire sur  $H$ , qui possède donc maintenant deux produits scalaires. Par continuité et coercivité, la norme induite par le produit scalaire  $a$  est équivalente à la norme initiale. Ainsi  $L$  est continue aussi dans  $H$  muni de ce nouveau produit scalaire. On peut donc appliquer le théorème de représentation de Riesz dans ce nouvel espace et obtenir immédiatement le résultat.

**Le cas non symétrique** Ce cas est un peu plus délicat et ne peut en aucun cas se prouver par une méthode de type *variationnel*. En effet, si on suppose qu'une fonctionnelle quadratique  $J$  de la forme  $J(v) = \frac{1}{2}a(v, v) - L(v)$ , admet un minimiseur dans  $H$ , alors nous pouvons montrer que  $u$  vérifie la formulation suivante

$$\frac{1}{2}(a(u, v) + a(v, u)) = L(v), \quad \forall v \in H,$$

c'est-à-dire  $\tilde{a}(u, v) = L(v)$  pour une certaine forme bilinéaire **symétrique**  $\tilde{a}$ .

Pour prouver le théorème, il nous faut donc introduire une reformulation du problème sous la forme d'un opérateur dont on doit montrer l'inversibilité.

Constatons que pour tout  $u \in H$ , la forme linéaire  $v \in H \mapsto a(u, v)$  est continue et donc il existe un unique  $Au \in H$  tel que  $a(u, v) = (Au, v)_H$ , par application du théorème de représentation de Riesz. Il est clair que le  $A$  ainsi construit est un opérateur linéaire.

Par hypothèse sur  $a$ , on a  $\forall u \in H, \|Au\|_H \leq \|a\| \|u\|_H$  et donc  $A$  est un opérateur continu. Par ailleurs, on a

$$(Au, u)_H = a(u, u) \geq \alpha \|u\|_H^2, \quad \forall u \in H,$$

donc

$$\|Au\|_H \geq \alpha \|u\|_H, \quad \forall u \in H.$$

Ceci implique en particulier que  $A$  est injectif. D'après le théorème de Riesz, on peut représenter  $L$  par un élément  $l \in H$  et on est maintenant amenés à prouver qu'il existe un  $u \in H$  tel que  $Au = l$ .

Soit  $\rho > 0$ , on introduit l'application  $T : u \in H \mapsto Tu = u - \rho(Au - l)$  et on voit que notre problème se ramène à l'existence d'un point fixe pour  $T$ . Dans un espace de Hilbert, le résultat sera montré si on prouve que  $T$  est contractante. Pour cela, on mène le calcul suivant

$$\begin{aligned} \|Tu - Tv\|_H^2 &= \|u - v\|_H^2 + \rho^2 \|Au - Av\|_H^2 - 2\rho(u - v, A(u - v))_H \\ &\leq \|u - v\|_H^2 + \rho^2 \|a\|^2 \|u - v\|_H^2 - 2\rho\alpha \|u - v\|_H^2 \\ &= (1 - 2\rho\alpha + \rho^2 \|a\|^2) \|u - v\|_H^2. \end{aligned}$$

En choisissant  $\rho$  suffisamment petit, et comme  $\alpha > 0$ , on voit que  $T$  est bien contractante et le résultat est démontré.

## IV Espaces de Sobolev et problèmes elliptiques sur un domaine de $\mathbb{R}^d$

### IV.1 Espaces de Sobolev sur un domaine de $\mathbb{R}^d$

Grâce au formalisme introduit dans le chapitre précédent, on peut introduire les espaces de Sobolev en toute généralité.

#### Définition III.23 (Espace $H^1(\Omega)$ )

Soit  $\Omega$  un ouvert de  $\mathbb{R}^d$ . On note  $H^1(\Omega)$  le sous-espace de  $L^2(\Omega)$  constitué des fonctions dont les dérivées premières au sens des distributions sont des éléments de  $L^2(\Omega)$ , ce qui donne

$$H^1(\Omega) = \{u \in L^2(\Omega), \text{ tel que } \partial_{x_i} u \in L^2(\Omega), \forall i \in \{1, \dots, d\}\}.$$

On le munit de la norme

$$\|u\|_{H^1(\Omega)} = \sqrt{\|u\|_{L^2}^2 + \sum_{i=1}^d \|\partial_{x_i} u\|_{L^2}^2},$$

qui en fait un espace de Hilbert.

Il est aisé de se convaincre que, en dimension  $d = 1$ , cette définition coïncide avec celle vue au premier chapitre et que ce

qu'on avait appelé *dérivée faible* n'est rien d'autre que la dérivée au sens des distributions.

### Définition III.24 (Espace $H^k(\Omega)$ )

Soit  $\Omega$  un ouvert de  $\mathbb{R}^d$  et  $k \geq 1$  un entier. On définit l'espace  $H^k(\Omega)$  par la formule

$$H^k(\Omega) = \{u \in L^2(\Omega), \text{ tel que } \forall \alpha \in \mathbb{N}^d, \text{ avec } |\alpha| \leq k, \partial^\alpha u \in L^2(\Omega)\}.$$

On le munit de la norme

$$\|u\|_{H^k} = \sqrt{\sum_{\substack{\alpha \in \mathbb{N}^d \\ |\alpha| \leq k}} \|\partial^\alpha u\|_{L^2}^2},$$

qui en fait un espace de Hilbert.

### Théorème III.25 (densité des fonctions régulières)

1. Soit  $\Omega$  un ouvert borné quelconque de  $\mathbb{R}^d$ . Alors l'ensemble

$$C^\infty(\Omega) \cap H^k(\Omega)$$

est dense dans  $H^k(\Omega)$ .

2. Soit  $\Omega$  un domaine borné régulier (de classe  $C^1$ ). Alors l'ensemble

$$C^\infty(\overline{\Omega})$$

est dense dans  $H^k(\Omega)$ .

### Remarque III.26

Quand  $d \geq 2$ , contrairement à la dimension 1, nous avons

$$H^1(\Omega) \not\subset C^0(\overline{\Omega}).$$

En dimension  $d = 2$  par exemple, prenons  $\Omega = B(0, 1)$ , et posons

$$u(x) = \log |\log r|, \quad r = |x|.$$

Cette fonction n'est clairement pas continue (elle explose au voisinage de 0).

En revanche, par passage en polaire, nous avons

$$\int_{\Omega} |u(x)|^2 dx = \int_{\Omega} |\log |\log r||^2 dx = 2\pi \int_0^1 r |\log |\log r||^2 dr < +\infty,$$

et

$$\int_{\Omega} |\nabla u(x)|^2 dx = \int_{\Omega} \left| \frac{1}{r \log r} \right|^2 dx = 2\pi \int_0^1 \left| \frac{1}{r \log r} \right|^2 r dr = 2\pi \int_0^1 \frac{1}{r |\log r|^2} dr < +\infty,$$

d'après les résultats bien connus sur les intégrales de Bertrand.

Ainsi, cette fonction  $u$  est bien dans  $H^1(\Omega)$  sans être continue sur son domaine de définition.

On a néanmoins le résultat plus faible suivant.

### **Théorème III.27 (Injections de Sobolev)**

Soit  $\Omega$  un domaine borné régulier de  $\mathbb{R}^d$ ,  $d \geq 2$ . On pose

$$2^* = \begin{cases} \frac{2d}{d-2}, & \text{si } d \geq 3, \\ q, & \text{pour n'importe quel } q < +\infty, \text{ si } d = 2. \end{cases}$$

Alors nous avons

$$\forall u \in H^1(\Omega), \quad u \in L^{2^*}(\Omega), \quad \text{et } \|u\|_{L^{2^*}(\Omega)} \leq C \|u\|_{H^1(\Omega)},$$

pour un  $C > 0$  ne dépendant que de  $\Omega$ .

Ces injections sont optimales.

Malgré tout, on va essayer de définir la trace sur le bord du domaine  $\Omega$  d'une fonction dans un espace de Sobolev. Pour cela, on commence par définir l'espace des fonctions de carré intégrable sur le bord de  $\Omega$ . En supposant que  $(U, \gamma)$  est une paramétrisation de  $\partial\Omega$  (éventuellement locale ...) cet espace s'écrit

$$L^2(\partial\Omega) = \{f : \partial\Omega \rightarrow \mathbb{R}, \text{ telles que } f \circ \gamma \in L^2(U)\},$$

que l'on munit alors de la norme définie par

$$\|f\|_{L^2(\partial\Omega)}^2 = \int_{\partial\Omega} |f|^2 d\sigma.$$

On donne ainsi sans démonstration les résultats suivants.

### **Théorème III.28 (Traces)**

Soit  $\Omega$  un domaine borné régulier (de classe au moins  $\mathcal{C}^1$ ). L'application

$$\gamma_0 : u \in \mathcal{C}^1(\overline{\Omega}) \mapsto u|_{\partial\Omega} \in \mathcal{C}^0(\partial\Omega),$$

vérifie, pour un  $C > 0$  ne dépendant que de  $\Omega$ , l'inégalité

$$\|\gamma_0 u\|_{L^2(\partial\Omega)} \leq C \|u\|_{H^1(\Omega)}, \quad \forall u \in \mathcal{C}^1(\overline{\Omega}). \quad (\text{III.24})$$

Ainsi  $\gamma_0$  se prolonge de façon unique en un opérateur linéaire continu

$$\gamma_0 : H^1(\Omega) \rightarrow L^2(\partial\Omega),$$

appelé **opérateur de traces** associé à l'espace  $H^1(\Omega)$ .

### **Proposition III.29 (Formules de Stokes dans les espaces de Sobolev)**

Si  $u \in H^1(\Omega)$  et  $V \in (H^1(\Omega))^d$  alors nous avons

$$\int_{\Omega} u(\operatorname{div} V) dx = - \int_{\Omega} (\nabla u) \cdot V dx + \int_{\partial\Omega} (\gamma_0 u)(\gamma_0 V) \cdot n d\sigma.$$

**Théorème III.30 (Espace de traces et relèvement)**

L'opérateur de traces  $\gamma_0$  **n'est pas surjectif**. On notera conventionnellement  $H^{\frac{1}{2}}(\partial\Omega)$  son image

$$H^{\frac{1}{2}}(\partial\Omega) = \gamma_0(H^1(\Omega)).$$

On le munit de la norme quotient

$$\|v\|_{H^{\frac{1}{2}}} = \inf_{\substack{u \in H^1(\Omega) \\ \gamma_0(u)=v}} \|u\|_{H^1(\Omega)},$$

qui en fait un espace de Hilbert.

L'espace  $H^{\frac{1}{2}}(\partial\Omega)$  ainsi obtenu contient  $C^1(\partial\Omega)$  et il est dense dans  $L^2(\partial\Omega)$ .

Enfin, il existe un (non unique) opérateur linéaire continu  $R_0 : H^{\frac{1}{2}}(\partial\Omega) \rightarrow H^1(\Omega)$  (appelé **opérateur de relèvement**) qui vérifie

$$\gamma_0(R_0 v) = v, \quad \forall v \in H^{\frac{1}{2}}(\partial\Omega).$$

**Remarque III.31**

1. Dans ce théorème on a utilisé l'espace  $C^1(\partial\Omega)$  des fonctions de classe  $C^1$  sur le bord de  $\Omega$ . Si  $(U, \gamma)$  est une paramétrisation (éventuellement locale) de  $\Gamma = \partial\Omega$ , on dira qu'une fonction  $f : \Gamma \rightarrow \mathbb{R}$  est de classe  $C^1$  si  $f \circ \gamma \in C^1(U, \mathbb{R})$ .

Sous les hypothèses du théorème, on peut montrer que  $f : \partial\Omega \rightarrow \mathbb{R}$  est de classe  $C^1$  si et seulement s'il existe  $F : \mathbb{R}^d \rightarrow \mathbb{R}$  de classe  $C^1$  telle que

$$F|_{\Gamma} = f.$$

2. Il est difficile, et hors de propos dans ce cours, d'expliquer pourquoi cet espace de traces s'appelle  $H^{1/2}(\partial\Omega)$ . Moralement, cela signifie que les fonctions dans cet espace possèdent des demi-dérivées .... On peut aussi le comprendre en raffinant l'inégalité (III.24). En effet, on peut montrer qu'il existe  $C > 0$  telle que

$$\|\gamma_0 u\|_{L^2(\partial\Omega)} \leq C \|u\|_{L^2(\Omega)}^{\frac{1}{2}} \|u\|_{H^1(\Omega)}^{\frac{1}{2}}.$$

Ainsi la trace de  $u$  peut s'estimer en norme  $L^2$  en utilisant seulement la racine carrée de la norme  $H^1$  (celle qui contient des dérivées de  $u$  dans  $\Omega$ ).

**Définition III.32 (Espace  $H_0^1(\Omega)$ )**

On définit  $H_0^1(\Omega)$  comme le sous-espace **fermé** de  $H^1(\Omega)$  constitué des fonctions nulles au bord, au sens de la trace. Autrement dit

$$H_0^1(\Omega) = \text{Ker } \gamma_0.$$

**Proposition III.33**

Soit  $\Omega$  un domaine borné régulier de  $\mathbb{R}^d$ . L'espace  $\mathcal{D}(\Omega)$  est alors dense dans  $H_0^1(\Omega)$ .

**Théorème III.34 (Inégalité de Poincaré)**

Soit  $\Omega$  un domaine borné régulier de  $\mathbb{R}^d$ , il existe alors une constante  $C > 0$  ne dépendant que de  $\Omega$  vérifiant

$$\|u\|_{L^2(\Omega)} \leq C \|\nabla u\|_{L^2(\Omega)}, \quad \forall u \in H_0^1(\Omega).$$

**Théorème III.35 (Inégalité de Poincaré-Wirtinger)**

Soit  $\Omega$  un domaine borné régulier **connexe** de  $\mathbb{R}^d$ , il existe alors une constante  $C > 0$  ne dépendant que de  $\Omega$  vérifiant

$$\|u\|_{L^2(\Omega)} \leq C \|\nabla u\|_{L^2(\Omega)}, \quad \forall u \in H_m^1(\Omega),$$

où  $H_m^1(\Omega)$  est l'ensemble des fonction de  $H^1$  à moyenne nulle.

## IV.2 Problèmes aux limites elliptiques

On va reprendre les exemples vus dans le premier chapitre mais cette fois en dimension quelconque. Dans toute la suite,  $\Omega$  sera un domaine borné régulier de  $\mathbb{R}^d$ .

### IV.2.a Problème de Poisson avec conditions de Dirichlet

Soit  $k : \Omega \rightarrow \mathbb{R}$  une fonction mesurable bornée et telle que  $\alpha = \inf_{\Omega} k > 0$  et  $f \in L^2(\Omega)$ . On pose  $H = H_0^1(\Omega)$ , et on définit

$$a(u, v) = \int_{\Omega} k(x)(\nabla u) \cdot (\nabla v) dx, \quad L(v) = \int_{\Omega} f v dx.$$

La continuité de  $a$  et  $L$  ne fait aucun doute. La coercivité de  $a$  provient de l'inégalité de Poincaré et de l'hypothèse sur le coefficient  $k$ .

On peut donc appliquer le théorème de Lax-Milgram et ainsi obtenir une fonction  $u \in H_0^1(\Omega)$  telle que

$$\int_{\Omega} k(x) \nabla u \cdot \nabla v dx = \int_{\Omega} f v dx, \quad \forall v \in H_0^1(\Omega).$$

En prenant  $v = \varphi \in \mathcal{D}(\Omega)$ , cette formule devient

$$\sum_{i=1}^d \langle k \partial_{x_i} u, \partial_{x_i} \varphi \rangle_{\mathcal{D}', \mathcal{D}} = \langle f, \varphi \rangle_{\mathcal{D}', \mathcal{D}},$$

et on a donc résolu le problème

$$-\operatorname{div}(k \nabla u) = f, \quad \text{dans } \mathcal{D}'(\Omega),$$

avec la condition de Dirichlet homogène sur le bord.

#### Remarque III.36

Si  $k$  est constante, on a résolu le problème suivant

$$-k \Delta u = f,$$

avec conditions au bord de Dirichlet homogène.

En dimension 1, nous avons déduit que la dérivée seconde de  $u$  était dans  $L^2$  mais ici on va admettre le résultat (difficile) suivant

#### Théorème III.37 (Régularité elliptique)

Si  $\Omega$  est suffisamment régulier et que  $k$  est de classe  $\mathcal{C}^1$  sur  $\overline{\Omega}$ , alors la solution du problème précédent vérifie

$$u \in H^2(\Omega), \quad \text{avec } \|u\|_{H^2} \leq C \|f\|_{L^2},$$

pour un  $C > 0$  ne dépendant que de  $\Omega$ .

Pour le cas non-homogène, on cherche  $u \in H^1(\Omega)$  vérifiant

$$\begin{cases} -\operatorname{div}(k(x) \nabla u) = f, & \text{dans } \Omega, \\ u = u_b, & \text{sur } \partial\Omega. \end{cases}$$

D'après le théorème de trace, cela nécessite que  $u_b \in H^{\frac{1}{2}}(\partial\Omega)$ . De fait, si on fait cette hypothèse, le théorème de relèvement nous dit qu'il existe une fonction  $R_0 u_b \in H^1(\Omega)$  dont la trace est égale à  $u_b$ .

On modifie alors la forme linéaire  $L$  de la façon suivante

$$\tilde{L}(v) = \int_{\Omega} f v dx - \int_{\Omega} k(x) \nabla(R_0 u_b) \cdot \nabla v dx.$$

Le nouveau terme est toujours continue pour la norme  $H^1(\Omega)$ . On résout alors le problème

$$w \in H_0^1(\Omega), \quad a(w, v) = \tilde{L}(v), \quad \forall v \in H_0^1(\Omega),$$

de sorte que  $u = w + R_0 u_b$  vérifie bien l'équation de Poisson  $-\operatorname{div}(k \nabla u) = f$  et la condition au bord  $\gamma_0 u = u_b$  prescrite au départ.



### IV.2.b Problème de diffusion-advection

Soit  $b : \Omega \rightarrow \mathbb{R}^d$  un champ de vecteurs régulier. On s'intéresse au problème

$$\begin{cases} -\Delta u + b \cdot \nabla u = f, & \text{dans } \Omega, \\ u = 0, & \text{sur } \partial\Omega. \end{cases}$$

On propose la formulation variationnelle suivante

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx + \int_{\Omega} (b \cdot \nabla u) v \, dx = \int_{\Omega} f v \, dx, \quad \forall v \in H_0^1(\Omega).$$

Comme en dimension 1, la difficulté principale est d'établir la coercivité de la partie bilinéaire et donc d'estimer le terme d'advection

$$\int_{\Omega} (b \cdot \nabla u) u \, dx,$$

lequel peut être contrôlé dès que  $\|b\|_{\infty}$  est assez petite. On peut aussi l'intégrer par parties et l'écrire

$$-\frac{1}{2} \int_{\Omega} (\operatorname{div} b) u^2 \, dx,$$

ce qui, par exemple, permet de le minorer si  $\operatorname{div} b \leq 0$ .

Encore une fois c'est un exemple d'utilisation du théorème de Lax-Milgram dans un cadre non symétrique.

### IV.2.c Conditions aux limites de Neumann

On suppose que  $\Omega$  est connexe et on se donne un  $f \in L^2(\Omega)$  et un  $g \in L^2(\partial\Omega)$  tels que

$$\int_{\Omega} f \, dx + \int_{\partial\Omega} g \, d\sigma = 0. \quad (\text{III.25})$$

On définit l'espace  $H_m^1(\Omega)$  par

$$H_m^1(\Omega) = \left\{ u \in H^1(\Omega), \int_{\Omega} u \, dx = 0 \right\},$$

sur lequel on admet que l'inégalité de Poincaré-Wirtinger suivante est vraie

$$\|u\|_{L^2(\Omega)} \leq C \|\nabla u\|_{L^2(\Omega)}, \quad \forall u \in H_m^1(\Omega).$$

On définit ensuite les formes bilinéaires et linéaires suivantes

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx, \quad L(v) = \int_{\Omega} f v \, dx + \int_{\partial\Omega} g \gamma_0(v) \, d\sigma, \quad \forall v \in H_m^1(\Omega).$$

- La continuité de  $a$  est claire et la coercivité vient de l'inégalité de Poincaré-Wirtinger.
- La continuité de  $L$  provient de l'inégalité de Cauchy-Schwarz et du théorème de traces

$$|L(v)| \leq \|f\|_{L^2} \|v\|_{L^2} + \|g\|_{L^2(\partial\Omega)} \|\gamma_0(v)\|_{L^2(\partial\Omega)} \leq (\|f\|_{L^2(\mathcal{O})} + C\|g\|_{L^2(\partial\Omega)}) \|v\|_{H^1}.$$

Le théorème de Lax-Milgram s'applique et nous donne un unique  $u \in H_m^1(\Omega)$  tel que

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx + \int_{\partial\Omega} g \gamma_0(v) \, d\sigma, \quad \forall v \in H_m^1(\Omega).$$

Pour  $v \in H^1(\Omega)$ , on constate que  $v - \frac{1}{|\Omega|} \int_{\Omega} v \, dx$  est dans  $H_m^1(\Omega)$ . On peut donc appliquer la propriété ci-dessus et utiliser la condition de compatibilité (III.25) pour obtenir

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx + \int_{\partial\Omega} g \gamma_0(v) \, d\sigma, \quad \forall v \in H^1(\Omega). \quad (\text{III.26})$$

On peut maintenant prendre  $v = \varphi \in \mathcal{D}(\Omega)$  dans (III.26) et obtenir

$$\int_{\Omega} \nabla u \cdot \nabla \varphi \, dx = \int_{\Omega} f \varphi \, dx.$$

Ceci montre que

$$-\Delta u = -\operatorname{div}(\nabla u) = f, \text{ au sens des distributions.}$$

Si on suppose que  $u \in H^2(\Omega)$  (ce qu'il est possible de démontrer si les données sont suffisamment régulières) alors on peut utiliser la formule d'intégration par parties donnée dans la proposition III.29 pour obtenir pour tout  $v \in H_m^1(\Omega)$

$$\int_{\Omega} f v \, dx = \int_{\Omega} -\operatorname{div}(\nabla u) v \, dx = \int_{\Omega} \nabla u \cdot \nabla v \, dx - \int_{\partial\Omega} \gamma_0(v) \gamma_0(\nabla u) \cdot n \, d\sigma.$$

En comparant cette égalité avec (III.26) on obtient

$$\int_{\partial\Omega} (g - \gamma_0(\nabla u) \cdot n) \gamma_0(v) \, d\sigma = 0, \quad \forall v \in H^1(\Omega).$$

Comme l'image de l'opérateur de trace (l'espace  $H^{\frac{1}{2}}(\partial\Omega)$ ) est dense dans  $L^2(\partial\Omega)$ , nous déduisons que

$$\gamma_0(\nabla u) \cdot n = g.$$

Nous avons donc résolu le problème de Poisson avec la condition de Neumann non-homogène ci-dessus.

## A retenir du chapitre

### En priorité

- Définition et propriétés des espaces de Sobolev en dimension 1 :  $H^1(I)$  et  $H_0^1(I)$ .
- Théorème de Lax-Milgram. Preuve du cas symétrique par minimisation d'une fonctionnelle d'énergie.
- Mise sous forme variationnelle et résolution d'un problème aux limites linéaire en dimension 1 par utilisation du Th. de Lax-Milgram.
- Régularité de la solution faible. Interprétation *a posteriori* de l'équation résolue et des conditions aux limites.

### Pour approfondir

- Espaces de Sobolev et formulations variationnelles en dimensions supérieures.



## Annexe A

# Eléments de la théorie des distributions

Le but de cette annexe est de rappeler les principaux éléments de la théorie des distributions qui permet de travailler confortablement avec une notion de dérivée faible. Il n'est pas question de traiter la théorie complète mais seulement les principales définitions et propriétés qui sont indispensables à l'étude de toutes les équations aux dérivées partielles, y compris à leur analyse numérique.

### I Intégration par parties en dimension $d$ : le cas des fonctions à support compact

On souhaite désormais disposer d'une formule d'intégration par parties pour les fonctions de plusieurs variables qui ressemble à celle que l'on connaît pour les fonctions d'une variable.

Commençons par rappeler ce qui fait marcher le calcul en dimension 1 pour les fonctions définies sur  $\mathbb{R}$  tout entier mais à support compact.

1. On montre que pour toute fonction  $f \in \mathcal{C}_c^1(\mathbb{R})$  on a

$$\int_{\mathbb{R}} f'(x) dx = 0. \quad (\text{A.1})$$

2. On utilise que pour toutes fonctions  $u, v \in \mathcal{C}^1(\mathbb{R})$  on a

$$(uv)' = u'v + uv'.$$

3. On applique la première formule à  $f = uv$  en supposant que l'une des deux fonctions  $u$  ou  $v$  (éventuellement les deux) est à support compact. Ainsi  $f$  est à support compact et on a

$$0 = \int_{\mathbb{R}} f'(x) dx = \int_{\mathbb{R}} u'(x)v(x) + u(x)v'(x) dx,$$

ce qui fournit

$$\int_{\mathbb{R}} u'(x)v(x) dx = - \int_{\mathbb{R}} u(x)v'(x) dx,$$

ce qui est bien une intégration par parties pour des fonctions à support compact.

Essayons le même processus pour les fonctions de plusieurs variables réelles. Soit  $\Omega$  un ouvert de  $\mathbb{R}^d$ .

#### Définition A.1

Pour toute fonction scalaire  $u \in \mathcal{C}^1(\Omega, \mathbb{R})$  et tout champ de vecteurs  $F \in \mathcal{C}^1(\Omega, \mathbb{R}^d)$  on définit

$$\nabla u = \begin{pmatrix} \partial_{x_1} u \\ \vdots \\ \partial_{x_d} u \end{pmatrix},$$

$$\Delta u = \sum_{i=1}^d \partial_{x_i}^2 u,$$

$$\operatorname{div} F = \sum_{i=1}^d \partial_{x_i} F_i.$$

**Proposition A.2 (Quelques formules utiles)**

$$\begin{aligned}\operatorname{div}(uF) &= F \cdot (\nabla u) + (\operatorname{div} u)F, \\ \Delta u &= \operatorname{div}(\nabla u).\end{aligned}\tag{A.2}$$

**Proposition A.3 (Intégration par parties pour des fonctions à support compact)**

1. Soit  $F \in \mathcal{C}_c^1(\Omega, \mathbb{R}^d)$  un champ de vecteurs à support compact, alors on a

$$\int_{\Omega} \operatorname{div} F \, dx = 0.$$

2. Soient  $u \in \mathcal{C}^1(\Omega, \mathbb{R})$  et  $V \in \mathcal{C}^1(\Omega, \mathbb{R}^d)$  tels que  $u$  ou  $V$  soit à support compact, alors on a

$$\int_{\Omega} u(\operatorname{div} V) \, dx = - \int_{\Omega} (\nabla u) \cdot V \, dx.$$

3. Soient  $u, v \in \mathcal{C}^2(\Omega, \mathbb{R})$  tels que  $u$  ou  $v$  soit à support compact, alors on a

$$\int_{\Omega} u(-\Delta v) \, dx = \int_{\Omega} (\nabla u) \cdot (\nabla v) \, dx = \int_{\Omega} (-\Delta u)v \, dx.$$

**Preuve :**

On se contente de montrer le cas  $\Omega = \mathbb{R}^d$ . En effet dans le cas général, au moins l'une des fonctions en jeu est à support compact dans  $\Omega$  et on peut donc la prolonger par 0 à l'espace entier sans changer son caractère  $\mathcal{C}^1$  et ainsi se ramener au cas de l'espace entier.

1. Par définition nous avons  $\operatorname{div} F = \sum_{i=1}^d \partial_{x_i} F_i$ , on peut donc s'intéresser à l'un quelconque des termes et voir que par le théorème de Fubini on a

$$\int_{\mathbb{R}^d} \partial_{x_i} F_i \, dx = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \partial_{x_i} F_i \, dx_1 \cdots dx_d = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \left( \int_{\mathbb{R}} \partial_{x_i} F_i \, dx_i \right) dx_1 \cdots \widehat{dx_i} \cdots dx_d,$$

et d'après (A.1) l'intégrale par rapport à  $x_i$  est nulle pour tout  $x_1, \dots, \widehat{x_i}, \dots, x_d$ .

2. On applique le résultat ci-dessus au produit  $F = uV$  et on utilise la formule (A.2) pour en déduire

$$0 = \int_{\mathbb{R}^d} \operatorname{div}(uV) \, dx = \int_{\mathbb{R}^d} u(\operatorname{div} V) \, dx + \int_{\mathbb{R}^d} (\nabla u) \cdot V \, dx.$$

3. Il suffit d'appliquer la formule ci-dessus avec  $V = \nabla v$  par exemple et remarquer ensuite que  $u$  et  $v$  jouent des rôles symétriques. ■

On détaille dans l'annexe B ce qu'il convient de faire pour pouvoir "intégrer par parties" dans le cas le plus général où les fonctions en jeu ne sont plus à support compact. Le résultat à retenir est que

$$\int_{\Omega} \operatorname{div} F \, dx = \int_{\partial\Omega} F \cdot n \, d\sigma,$$

où  $n$  désigne la normale unitaire sortante à  $\Omega$  et  $\int_{\partial\Omega}$  est l'intégrale de surface. Ces objets sont définis dans cette annexe et nécessitent un peu de régularité du bord du domaine  $\Omega$ , qui ne peut donc pas être un ouvert quelconque de  $\mathbb{R}^d$ .

**II Un lemme important de la théorie de l'intégration**

Dans la suite, on va avoir besoin du lemme suivant

**Lemme A.4 (de Du Bois-Reymond)**

Soit  $\Omega \subset \mathbb{R}^d$  un ouvert non vide et  $f \in L_{loc}^1(\Omega)$  telle que

$$\int_{\Omega} f \varphi \, dx = 0, \quad \forall \varphi \in \mathcal{C}_c^\infty(\Omega),\tag{A.3}$$

alors  $f = 0$ .

**Preuve (du Lemme A.4):**

On va donner différentes preuves selon la régularité de  $f$ . Seule la dernière est la plus générale.

- Etudions d'abord le cas le plus simple où on suppose que  $f$  est continue.

On raisonne par l'absurde en supposant que  $f$  est non identiquement nulle. Comme  $f$  est continue, et quitte à changer  $f$  en  $-f$ , il existe une constante  $C > 0$  et une boule non triviale  $B(\alpha, r) \subset \Omega$  telle que

$$\forall x \in B(\alpha, r), f(x) \geq C.$$

On construit alors une fonction  $v$  de classe  $C^\infty$ , positive, identiquement nulle en dehors de  $B(\alpha, r)$  et telle que  $\int_{B(\alpha, r)} v(x) dx > 0$ . La construction d'une telle fonction est classique mais on n'a pas besoin de connaître la formule exacte pour faire la démonstration.

D'après (A.3) et les propriétés de  $f$  et  $v$ , on a

$$0 = \int_{\Omega} f(x)v(x) dx = \int_{B(\alpha, r)} f(x)v(x) dx \geq C \int_{B(\alpha, r)} v(x) dx > 0,$$

ce qui constitue une contradiction manifeste.

- Considérons le cas où  $f \in L^2(\Omega)$ .

On utilise la propriété suivante (voir le cours d'intégration, ou d'analyse fonctionnelle) :

L'ensemble des fonctions  $C^\infty$  à support compact dans  $\Omega$  est dense dans  $L^2(\Omega)$ .

Comme  $f \in L^2$ , cela signifie qu'il existe une suite  $(\varphi_n)_n$  d'éléments de  $C_c^\infty$  telle que  $\|\varphi_n - f\|_{L^2} \rightarrow 0$ .

On a alors

$$\int_{\Omega} |f(x)|^2 dx = \int_{\Omega} f(x)(f(x) - \varphi_n(x)) dx + \int_{\Omega} f(x)\varphi_n(x) dx.$$

Comme  $\varphi_n \in C_c^\infty$ , le second terme est nul d'après l'hypothèse (A.3), et on peut majorer le premier terme comme suit

$$\int_{\Omega} |f(x)|^2 dx \leq \|f\|_{L^2} \|f - \varphi_n\|_{L^2} \xrightarrow{n \rightarrow \infty} 0.$$

Ainsi,  $|f|^2$  est une fonction positive et d'intégrale nulle, elle est donc bien nulle presque partout.

- Faisons maintenant la preuve dans le cas général. On considère n'importe quel ouvert borné  $U \subset \Omega$  et la fonction  $g = \text{sgn}(f)1_U$  qui est donc une fonction mesurable bornée et intégrable.

Par densité de  $C_c^\infty(\Omega)$  dans  $L^1(\Omega)$ , il existe une suite  $(\varphi_n)_n$  d'éléments de  $C_c^\infty(\Omega)$  telle que  $\|g - \varphi_n\|_{L^1} \rightarrow 0$ .

On aimerait utiliser ces fonctions  $\varphi_n$  dans l'hypothèse mais on aura alors une difficulté pour passer à la limite dans l'intégrale car on ne dispose pas de domination uniforme pour les produits  $f\varphi_n$ . On va donc transformer légèrement les  $\varphi_n$  pour assurer une borne uniforme sans modifier leur convergence vers  $g$ .

On considère donc une fonction  $\beta : \mathbb{R} \rightarrow \mathbb{R}$  de classe  $C^\infty$  telle que

$$\begin{aligned} \beta(s) &= s, \text{ pour } s \in [-1, 1], \\ |\beta(s)| &\leq 2, \text{ pour tout } s \in \mathbb{R}. \end{aligned}$$

On pose maintenant  $\psi_n = \beta \circ \varphi_n$ . Par construction, on a  $\psi_n \in C_c^\infty(\Omega)$  et  $\|\psi_n\|_{L^\infty} \leq 2$ .

Quitte à extraire une sous-suite, on peut supposer également que  $\psi_n$  converge presque partout vers  $g$ .

On constate que

$$|f(x)\psi_n(x)| \leq |f(x)|\|\varphi_n\|_{L^\infty} = 2|f(x)|, \text{ pour presque tout } x,$$

ce qui permet d'utiliser le théorème de convergence dominée pour établir la convergence

$$0 = \int_{\Omega} f(x)\psi_n(x) dx \xrightarrow{n \rightarrow \infty} \int_{\Omega} f(x)g(x) dx = \int_U f(x)\text{sgn}(f)(x) dx = \int_U |f(x)| dx.$$

Ceci prouve donc que  $\int_U |f| = 0$  et donc  $f = 0$  sur  $U$ . Ceci étant valable pour tout ouvert borné  $U$ , on a bien montré que  $f = 0$ . ■

Le point crucial de la preuve est la densité de l'ensemble de fonctions tests utilisé dans l'espace  $L^1(\Omega)$ . Il existe donc d'autres versions de ce lemme avec d'autres ensembles de fonctions tests et qui peuvent être utiles dans d'autres contextes comme par exemple le résultat suivant.

**Lemme A.5**

Si  $f \in L^1_{loc}(\Omega)$  vérifie

$$\int_B f(x) dx = 0, \text{ pour toute boule } B \text{ contenue dans } \Omega,$$

alors  $f = 0$  presque partout.

Il faut bien voir que c'est formellement le même énoncé que le lemme précédent mais en prenant les fonctions test  $\varphi$  de la forme  $\varphi = 1_B$  dans (A.3), au lieu des fonctions régulières à support compact.

**Preuve :**

— Ici encore la preuve est sensiblement plus simple si  $f$  est continue. Nous allons commencer par étudier ce cas.

On se donne  $x_0 \in \Omega$  et on applique l'hypothèse à  $B = B(x_0, \varepsilon)$ , avec  $\varepsilon > 0$  assez petit pour que  $B(x_0, \varepsilon) \subset \Omega$ . Il vient

$$\begin{aligned} f(x_0) &= f(x_0) - \frac{1}{|B(x_0, \varepsilon)|} \underbrace{\left( \int_{B(x_0, \varepsilon)} f(x) dx \right)}_{=0}, \\ &= \frac{1}{|B(x_0, \varepsilon)|} \int_{B(x_0, \varepsilon)} (f(x_0) - f(x)) dx, \end{aligned}$$

d'où l'on déduit

$$|f(x_0)| \leq \frac{1}{|B(x_0, \varepsilon)|} \int_{B(x_0, \varepsilon)} |f(x_0) - f(x)| dx.$$

Il s'en suit que

$$|f(x_0)| \leq \sup_{x \in B(x_0, \varepsilon)} |f(x_0) - f(x)| \xrightarrow{\varepsilon \rightarrow 0} 0,$$

car  $f$  est continue en  $x_0$ . La preuve est complète.

— Etudions le cas général maintenant.

La preuve consiste à remarquer que, par linéarité, l'égalité

$$\int_U f(x) dx = 0$$

est valable pour tout  $U$  qui s'écrit comme une union finie de boules dans  $\Omega$ , puis, par passage à la limite (convergence dominée par exemple), pour tout ouvert  $U$  borné inclus dans  $\Omega$ <sup>1</sup>.

On prend maintenant un borélien borné quelconque  $A$  et on définit l'ouvert

$$A_\varepsilon = \bigcup_{x \in A} B(x, \varepsilon),$$

qui contient  $A$  par construction et vérifie donc

$$0 = \int_{A_\varepsilon} f(x) dx = \int_{\Omega} 1_{A_\varepsilon}(x) f(x) dx.$$

Par convergence dominée, la convergence  $\varepsilon \rightarrow 0$  dans l'intégrale est justifiée<sup>2</sup> et on obtient finalement

$$0 = \int_A f(x) dx.$$

Ceci étant vrai pour tout borélien borné  $A$ , on peut en particulier prendre

$$A^+ = \{x \in \Omega, \text{t.q. } \|x\| \leq R, f(x) \geq 0\},$$

$$A^- = \{x \in \Omega, \text{t.q. } \|x\| \leq R, f(x) < 0\},$$

pour obtenir

$$\int_{\Omega \cap B(0, R)} |f(x)| dx = \int_{A^+} f(x) dx + \int_{A^-} (-f(x)) dx = 0 + 0 = 0.$$

Ceci prouve que  $f$  est nulle sur  $\Omega \cap B(0, R)$  pour tout  $R > 0$  et donc que  $f = 0$  sur  $\Omega$ .

■

1. un tel ouvert peut en effet s'écrire comme une union dénombrable de boules

2. Vérifiez que vous savez le faire !



### III Espace des fonctions test. Espace des distributions.

#### Notation

- Un  $d$ -uplet d'entiers positifs  $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$  est appelé un multi-index.
- L'entier  $|\alpha| = \alpha_1 + \dots + \alpha_d$  est la longueur du multi-index.
- Pour tout  $\alpha \in \mathbb{N}^d$  et toute fonction  $f$  suffisamment dérivable on définit la dérivée partielle

$$\partial^\alpha f = \frac{\partial^{|\alpha|}}{\partial_{x_1}^{\alpha_1} \dots \partial_{x_d}^{\alpha_d}} f.$$

Ces notations permettent souvent d'écrire de façon plus compacte des formules assez lourdes.<sup>3</sup>

#### III.1 Définitions, exemples

On a déjà utilisé l'ensemble des fonctions de classe  $\mathcal{C}^\infty$  à support compact dans le premier chapitre. Cet ensemble va jouer un rôle très important dans la suite et dorénavant, pour tout ouvert  $\Omega$  de  $\mathbb{R}^d$ , on va noter

$$\mathcal{D}(\Omega) = \mathcal{C}_c^\infty(\Omega) = \{f \in \mathcal{C}^\infty(\Omega), \text{ il existe un compact } K \subset \Omega \text{ tel que } f = 0 \text{ sur } K^c\}.$$

On prendra garde au fait que le compact  $K$  qui intervient dans cette définition dépend bien sûr de la fonction  $f$  considérée. On définira ainsi

$$\text{Supp } \varphi = \text{Le plus petit compact } K \text{ tel que } \varphi = 0 \text{ sur } K^c.$$

Il est utile de comprendre que de telles fonctions existent et qu'on peut en fabriquer aisément qui vérifient de bonnes propriétés.

#### Lemme A.6 (Construction de fonctions de classe $\mathcal{C}^\infty$ à support compact)

1. La fonction numérique définie par

$$\eta : s \in \mathbb{R} \mapsto \begin{cases} e^{-1/s^2}, & \text{si } s > 0, \\ 0, & \text{si } s \leq 0, \end{cases}$$

est de classe  $\mathcal{C}^\infty$  sur  $\mathbb{R}$ .

2. Soient  $a < b$ , la fonction  $\xi_{a,b}$  définie par

$$\xi_{a,b}(s) = \frac{\eta(s-a)}{\eta(s-a) + \eta(b-s)}, \quad \forall s \in \mathbb{R}$$

est de classe  $\mathcal{C}^\infty$  sur  $\mathbb{R}$  et vérifie

$$\xi_{a,b}(s) \begin{cases} = 0, & \text{si } s \leq a, \\ \in ]0, 1[, & \text{si } a < s < b, \\ = 1, & \text{si } s \geq b. \end{cases}$$

3. Soit  $a < \alpha < \beta < b$ , la fonction  $\varphi_{a,\alpha,\beta,b}$  définie par

$$\varphi_{a,\alpha,\beta,b}(s) = \eta_{a,\alpha}(s)(1 - \eta_{\beta,b}(s)), \quad \forall s \in \mathbb{R},$$

est de classe  $\mathcal{C}^\infty$  sur  $\mathbb{R}$  et vérifie

$$\varphi_{a,\alpha,\beta,b} = 0, \text{ en dehors de } [a, b],$$

$$\varphi_{a,\alpha,\beta,b} = 1, \text{ dans } [\alpha, \beta].$$

4. Soient  $U, V$  deux ouverts non vides de  $\mathbb{R}^d$  vérifiant  $\overline{U} \subset V$  et tel que  $U$  est borné. Alors, il existe une fonction  $\varphi$  de classe  $\mathcal{C}^\infty$  sur  $\mathbb{R}^d$  telle que

$$\varphi = 1, \text{ dans } U, \text{ et } \varphi = 0, \text{ en dehors de } V.$$

3. Voir par exemple : <https://fr.wikipedia.org/wiki/Multi-index>

**Preuve :**

1. Il est clair que  $\eta$  est de classe  $\mathcal{C}^\infty$  sur  $] -\infty, 0[$  (toutes ses dérivées y sont nulles !) et sur  $]0, +\infty[$ . Il reste donc à montrer que toutes les dérivées à droite en 0 se raccordent avec les dérivées à gauche (autrement dit : sont nulles). Pour cela, on établit par récurrence que

$$\forall k \geq 0, \exists P_k \in \mathbb{R}[X], \text{ tel que } \eta^{(k)}(s) = P_k(1/s)\eta(s), \quad \forall s > 0. \quad (\text{A.4})$$

Le cas  $k = 0$  est immédiat, et le cas  $k = 1$  donne

$$\eta'(s) = \frac{2}{s^3} e^{-1/s^2} = P_1(1/s)\eta(s),$$

avec  $P_1(X) = 2X^3$ .

Si on suppose maintenant le résultat vrai au rang  $k$ , on calcule la dérivée  $(k+1)$ -ième

$$\begin{aligned} \eta^{(k+1)}(s) &= (\eta^{(k)})'(s) \\ &= \frac{d}{ds} (P_k(1/s)\eta(s)) \\ &= \left( -\frac{1}{s^2} P_k'(s) + P_k(1/s) P_1(1/s) \right) \eta(s), \end{aligned}$$

ce qui donne le résultat avec le polynôme  $P_{k+1}(X) = -X^2 P_k'(X) + P_k(X) P_1(X)$ .

On déduit de (A.4) que

$$\lim_{s \rightarrow 0^+} \eta^{(k)}(s) = \lim_{x \rightarrow +\infty} (P_k(x) e^{-x^2}) = 0,$$

d'après les résultats de croissance comparée des puissances et des exponentielles.

2. On constate que le dénominateur ne s'annule jamais ce qui assure la régularité  $\mathcal{C}^\infty$  de la fonction. Le reste des propriétés est immédiate en utilisant que  $\eta = 0$  sur  $\mathbb{R}^-$ .
3. Idem que dans le point précédent.
4. Pour tout point  $x \in \overline{U}$ , il existe  $r_x > 0$  tel que  $B(x, r_x) \subset V$ . On écrit ensuite

$$\overline{U} \subset \bigcup_{x \in \overline{U}} B\left(x, \frac{r_x}{2}\right),$$

dont on extrait un sous-recouvrement fini, par compacité de  $\overline{U}$ . On a donc obtenu des  $x_i \in \overline{U}$  et des  $r_i > 0$ ,  $i = 1, \dots, N$ , tels que

$$\overline{U} \subset \bigcup_{i=1}^N B\left(x_i, \frac{r_i}{2}\right).$$

Pour chaque  $i \in \{1, \dots, N\}$ , on définit alors

$$\psi_i(x) = \xi_{r_i/2, r_i}(\|x - x_i\|),$$

de sorte que  $\psi_i \in \mathcal{C}^\infty$  (pas de problème avec la singularité de la norme en 0 ici, voyez-vous pourquoi ?) et vérifie

$$\psi_i = 0, \text{ dans } B(x_i, r_i/2),$$

$$\psi_i = 1, \text{ en dehors de } B(x_i, r_i).$$

On pose alors

$$\varphi(x) = 1 - \prod_{i=1}^N \psi_i(x),$$

qui est bien de classe  $\mathcal{C}^\infty$  et vérifie les propriétés attendues.

■

On va avoir besoin de définir une notion de convergence pour les suites d'éléments de  $\mathcal{D}(\Omega)$  mais on ne traitera pas en détail la topologie de cet espace qui est assez complexe (en particulier ce n'est pas un espace vectoriel normé ...).

### Définition A.7 (Convergence dans $\mathcal{D}(\Omega)$ )

On dit qu'une suite  $(\varphi_n)_n \subset \mathcal{D}(\Omega)$  converge, au sens de  $\mathcal{D}(\Omega)$ , vers une autre fonction  $\varphi \in \mathcal{D}(\Omega)$  si et seulement si

1. Il existe un compact  $K \subset \Omega$  tel que  $\text{Supp}(\varphi_n) \subset K$  pour tout  $n$  et  $\text{Supp}(\varphi) \subset K$ .
2. Pour **tout** multi-indice  $\alpha \in \mathbb{N}^d$ , la suite des dérivées  $(\partial^\alpha \varphi_n)_n$  converge uniformément vers  $\partial^\alpha \varphi$  dans  $\Omega$ , c'est-à-dire

$$\|\partial^\alpha \varphi_n - \partial^\alpha \varphi\|_{L^\infty} \xrightarrow{n \rightarrow \infty} 0.$$

On notera  $\varphi_n \xrightarrow{n \rightarrow \infty} \varphi$  dans  $\mathcal{D}(\Omega)$ .

### Remarque A.8

Une remarque élémentaire : si  $(\varphi_n)_n \subset \mathcal{D}(\Omega)$  converge vers  $\varphi$ , alors pour tout multi-indice  $\beta \in \mathbb{N}^d$ , on a  $\partial^\beta \varphi_n \xrightarrow{n \rightarrow \infty} \partial^\beta \varphi$ .

On peut maintenant définir l'objet central de ce chapitre.

### Définition A.9 (Espace des distributions)

On dit qu'une forme linéaire  $T : \mathcal{D}(\Omega) \rightarrow \mathbb{R}$  est une **distribution sur  $\Omega$**  si elle est continue au sens suivant : pour toute suite  $(\varphi_n)_n \subset \mathcal{D}(\Omega)$  et  $\varphi \in \mathcal{D}(\Omega)$  on a

$$\left[ \varphi_n \xrightarrow{n \rightarrow \infty} \varphi \text{ dans } \mathcal{D}(\Omega) \right] \implies T(\varphi_n) \xrightarrow{n \rightarrow \infty} T(\varphi).$$

On notera  $\mathcal{D}'(\Omega)$  l'ensemble des distributions sur  $\Omega$ . C'est un espace vectoriel réel.

### Remarque A.10

L'espace  $\mathcal{D}'(\Omega)$  est donc, en un certain sens, le dual topologique de  $\mathcal{D}(\Omega)$ . En particulier, on pourra souvent utiliser la notation suivante

$$T(\varphi) = \langle T, \varphi \rangle_{\mathcal{D}', \mathcal{D}}, \quad \forall T \in \mathcal{D}'(\Omega), \forall \varphi \in \mathcal{D}(\Omega),$$

inspirée de la notation standard du produit scalaire dans un espace de Hilbert.

## III.1.a Les fonctions de $L^1_{loc}(\Omega)$

Commençons par un rappel

### Définition A.11 (Espace $L^1_{loc}(\Omega)$ )

Soit  $\Omega$  un ouvert de  $\mathbb{R}^d$ . On note

$$L^1_{loc}(\Omega) = \{f : \Omega \rightarrow \mathbb{R}, \text{ classe de fonctions mesurables telle que } 1_K f \in L^1(\Omega) \text{ pour tout compact } K \subset \Omega\}.$$

On peut munir cet espace d'une distance qui en fait un espace métrique complet et tel que  $(f_n)_n$  converge vers  $f$  dans  $L^1_{loc}(\Omega)$  si et seulement si

$$\|1_K(f_n - f)\|_{L^1} \xrightarrow{n \rightarrow \infty} 0, \quad \forall K \subset \Omega \text{ compact}.$$

**Remarque A.12**

Tous les espaces de fonctions usuels  $L^p(\Omega)$ ,  $1 \leq p \leq \infty$ ,  $\mathcal{C}^k(\Omega)$ ,  $k \geq 0$ , etc ... s'injectent naturellement et continument dans  $L^1_{loc}(\Omega)$ . On peut donc le voir comme un gros espace contenant tous les autres.

**Proposition A.13**

1. Pour toute fonction  $f \in L^1_{loc}(\Omega)$ , l'application  $T_f$  définie par

$$T_f : \varphi \in \mathcal{D}(\Omega) \mapsto \int_{\Omega} f \varphi \, dx,$$

définit une distribution sur  $\Omega$ .

2. L'application

$$f \in L^1_{loc}(\Omega) \mapsto T_f \in \mathcal{D}'(\Omega),$$

est injective.

**Preuve :**

1. Soit  $(\varphi_n)_n$  une suite de  $\mathcal{D}(\Omega)$  qui converge vers  $\varphi$ . On prend  $K$  un compact qui contient tous les supports des  $\varphi_n$  et de  $\varphi$ . On écrit alors

$$|T_f(\varphi_n) - T_f(\varphi)| \leq \int_K |f| |\varphi_n - \varphi| \, dx \leq \|f\|_{L^1(K)} \|\varphi_n - \varphi\|_{\infty} \xrightarrow{n \rightarrow \infty} 0.$$

2. Supposons que  $T_f = 0$ , cela signifie que

$$\int_{\Omega} f \varphi \, dx = 0, \quad \forall \varphi \in \mathcal{D}(\Omega).$$

Soit  $U$  n'importe quelle boule telle que  $\overline{U} \subset \Omega$ , on déduit de ce qui précède que

$$\int_U f \varphi \, dx = 0, \quad \forall \varphi \in \mathcal{C}_c^{\infty}(U),$$

et comme  $f \in L^1(U)$ , on peut alors appliquer le lemme A.4 pour déduire que  $f = 0$  sur  $U$ . Ceci étant vrai pour toute boule  $\overline{U}$  incluse dans  $\Omega$ , on a bien que  $f = 0$ . ■

En conséquence de cette proposition on dira, par léger abus de langage, que les fonctions de  $L^1_{loc}(\Omega)$  **sont des distributions** et on identifiera systématiquement  $f$  et  $T_f$ . Ainsi, on écrira

$$\langle f, \varphi \rangle_{\mathcal{D}', \mathcal{D}} = \int_{\Omega} f \varphi \, dx,$$

et

$$L^1_{loc}(\Omega) \subset \mathcal{D}'(\Omega).$$

Pour cette raison on ne s'étonnera pas que dans certains ouvrages les distributions sont appelées **fonctions généralisées**.

En particulier, si  $T \in \mathcal{D}'(\Omega)$  est une distribution quelconque, et  $E$  un espace de fonctions usuel ( $L^p$ ,  $\mathcal{C}^k$ , etc ...), on dira que  $T \in E$  s'il existe  $f \in E$  tel que  $T = T_f$  et alors identifiera  $T$  à la fonction  $f$  en question (qui est unique d'après la propriété d'injectivité).

**III.1.b Masses de Dirac**

Soit  $x_0 \in \Omega$ , alors la masse de Dirac  $\delta_{x_0}$  est une distribution sur  $\Omega$  définie par

$$\langle \delta_{x_0}, \varphi \rangle_{\mathcal{D}', \mathcal{D}} = \varphi(x_0), \quad \forall \varphi \in \mathcal{D}(\Omega).$$

**III.1.c Mesures**

Plus généralement toute mesure borélienne  $\mu$  localement finie sur  $\Omega$  est une distribution

$$\langle \mu, \varphi \rangle_{\mathcal{D}', \mathcal{D}} = \int_{\Omega} \varphi d\mu, \quad \forall \varphi \in \mathcal{D}(\Omega).$$

### III.1.d Autres exemples

L'ensemble des distributions contient des objets encore plus généraux comme par exemple

$$\delta'_{x_0} : \varphi \in \mathcal{D}(\Omega) \mapsto \langle \delta'_{x_0}, \varphi \rangle_{\mathcal{D}', \mathcal{D}} = -\varphi'(x_0),$$

pour  $x_0 \in \Omega$ ,

$$\delta_\Gamma : \varphi \in \mathcal{D}(\Omega) \mapsto \langle \delta_\Gamma, \varphi \rangle_{\mathcal{D}', \mathcal{D}} = \int_\Gamma \varphi d\sigma,$$

où  $\Gamma$  est une hypersurface régulière contenue dans  $\Omega$ .

## III.2 Convergence au sens des distributions

On a maintenant besoin de définir une topologie dans l'ensemble des distributions. Là encore on ne va pas entrer dans les détails abstraits de cette problématique mais se contenter d'un niveau pragmatique suffisant pour nos besoins.

### Définition A.14 (Convergence dans $\mathcal{D}'(\Omega)$ )

On dit qu'une suite de distributions  $(T_n)_n \subset \mathcal{D}'(\Omega)$  converge vers une distribution  $T \in \mathcal{D}'(\Omega)$  si et seulement si

$$\langle T_n, \varphi \rangle_{\mathcal{D}', \mathcal{D}} \xrightarrow{n \rightarrow \infty} \langle T, \varphi \rangle_{\mathcal{D}', \mathcal{D}}, \quad \forall \varphi \in \mathcal{D}(\Omega).$$

On notera  $T_n \rightarrow T$  dans  $\mathcal{D}'(\Omega)$ . Il est clair que, la limite d'une suite de distributions, si elle existe, est unique.

Cette convergence est assez aisée à manipuler car il s'agit simplement d'une convergence simple. En contre partie, on ne sera pas surpris d'apprendre que cette convergence est trop faible pour propager certaines bonnes propriétés. Ainsi si  $(f_n)_n$  est une suite de fonctions continues qui converge au sens des distributions vers une fonction  $f$ , alors il n'y a aucune raison que  $f$  soit continue. Voir les exemples plus bas.

### Remarque A.15

On peut démontrer, mais cela sort du cadre de ce cours, que si pour toute fonction test  $\varphi \in \mathcal{D}(\Omega)$ , les suites numériques  $(\langle T_n, \varphi \rangle_{\mathcal{D}', \mathcal{D}})_n$  sont convergentes, alors il existe une distribution  $T \in \mathcal{D}'(\Omega)$  telle que  $T_n \rightarrow T$  dans  $\mathcal{D}'(\Omega)$ .

Cette définition de la convergence au sens des distributions est compatible avec toutes les notions raisonnables de convergence auxquelles on peut penser. Donnons quelques exemples fondamentaux.

— Convergence  $L^1_{loc}$  et distributions

### Proposition A.16

Soit  $(f_n)_n$  une suite d'éléments de  $L^1_{loc}(\Omega)$  qui converge, dans cet espace, vers un certain  $f \in L^1_{loc}(\Omega)$ . Alors nous avons

$$f_n \xrightarrow{n \rightarrow \infty} f, \quad \text{dans } \mathcal{D}'(\Omega).$$

#### Preuve :

On se fixe un  $\varphi \in \mathcal{D}(\Omega)$  et on introduit le compact  $K = \text{Supp } \varphi$ . On écrit alors

$$|\langle f_n, \varphi \rangle_{\mathcal{D}', \mathcal{D}} - \langle f, \varphi \rangle_{\mathcal{D}', \mathcal{D}}| = \left| \int_\Omega (f_n - f) \varphi dx \right| = \left| \int_K (f_n - f) \varphi dx \right| \leq \|\varphi\|_\infty \|f_n - f\|_{L^1(K)} \xrightarrow{n \rightarrow \infty} 0,$$

ce qui montre bien le résultat. ■

— Attention, il est tout à faire possible (et c'est l'un des intérêts de cette théorie) que la suite  $(f_n)_n$  converge vers une distribution  $T$  qui n'est pas une fonction de  $L^1_{loc}(\Omega)$ . Etudions l'exemple suivant : on définit  $\Omega = B(0, 1)$  la boule unité de  $\mathbb{R}^d$  et on se fixe une fonction  $g \in L^1(\mathbb{R}^d)$  vérifiant  $\text{Supp } g \subset \Omega$  et  $\int_\Omega g dx = 1$ .

On pose alors  $f_n(x) = n^d g(nx)$ , qui est une fonction de  $L^1(\Omega)$ . On va montrer que

$$f_n \xrightarrow{n \rightarrow \infty} \delta_0, \quad \text{au sens des distributions sur } \Omega.$$

Pour cela, on choisit un  $\varphi \in \mathcal{D}(\Omega)$  et il s'agit de démontrer que

$$\langle f_n, \varphi \rangle_{\mathcal{D}', \mathcal{D}} \xrightarrow{n \rightarrow \infty} \langle \delta_0, \varphi \rangle_{\mathcal{D}', \mathcal{D}} = \varphi(0).$$

On constate que le support de  $f_n$  vérifie  $\text{Supp } f_n \subset B(0, 1/n)$  et aussi, par changement de variable, que

$$\int_{\Omega} f_n dx = \int_{\Omega} g dx = 1, \text{ et } \int_{\Omega} |f_n| dx = \int_{\Omega} |g| dx = \|g\|_{L^1}.$$

On écrit alors

$$\begin{aligned} |\langle f_n, \varphi \rangle_{\mathcal{D}', \mathcal{D}} - \varphi(0)| &= \left| \int_{\Omega} f_n \varphi dx - \varphi(0) \right| \\ &= \left| \int_{\Omega} f_n(x)(\varphi(x) - \varphi(0)) dx \right| \\ &= \left| \int_{B(0, 1/n)} f_n(x)(\varphi(x) - \varphi(0)) dx \right| \\ &\leq \left( \sup_{x \in B(0, 1/n)} |\varphi(x) - \varphi(0)| \right) \left| \int_{\Omega} |f_n|(x) dx \right| \\ &= \|g\|_{L^1} \left( \sup_{x \in B(0, 1/n)} |\varphi(x) - \varphi(0)| \right), \end{aligned}$$

et cette dernière quantité tend bien vers 0 par continuité de  $\varphi$  au point 0.

- On peut aussi avoir  $f_n \xrightarrow[n \rightarrow \infty]{} f$  au sens des distributions sans avoir convergence de  $(f_n)_n$  vers  $f$  dans  $L^1_{loc}(\Omega)$ .

Si on reprend le calcul précédent mais avec  $\int_{\Omega} g dx = 0$  cette fois (et  $g$  non identiquement nulle), alors on peut montrer que  $(f_n)_n$  tend vers 0 au sens des distributions et pourtant on a

$$\|f_n\|_{L^1} = \|g\|_{L^1} \neq 0, \quad \forall n,$$

ce qui montre qu'elle ne peut pas converger vers 0 dans  $L^1$ .

- La convergence au sens des distributions n'est pas compatible avec la multiplication des fonctions ! Prenons un exemple :

Soit  $f_n : \mathbb{R} \rightarrow \mathbb{R}$  la fonction définie par

$$f_n(x) = \cos(nx), \quad \forall x \in \mathbb{R}.$$

Alors nous avons

$$f_n \xrightarrow[n \rightarrow \infty]{} 0, \text{ dans } \mathcal{D}'(\mathbb{R}),$$

$$f_n^2 \xrightarrow[n \rightarrow \infty]{} \frac{1}{2}, \text{ dans } \mathcal{D}'(\mathbb{R}).$$

En effet, fixons une fonction test  $\varphi \in \mathcal{D}(\mathbb{R})$  et effectuons l'intégration par parties (en primitivant  $f_n$  et dérivant  $\varphi$ )

$$\int_{\mathbb{R}} f_n \varphi dx = - \int_{\mathbb{R}} \frac{\sin(nx)}{n} \varphi'(x) dx.$$

On a donc

$$\left| \int_{\mathbb{R}} f_n \varphi dx \right| \leq \frac{1}{n} \|\varphi'\|_{L^1} \xrightarrow[n \rightarrow \infty]{} 0.$$

Si maintenant on calcule  $f_n^2$ , on trouve grâce aux formules de trigonométrie,

$$f_n^2(x) = \frac{1}{2} + \frac{\cos(2nx)}{2} = \frac{1}{2}(1 + f_{2n}),$$

et donc d'après ce qui précède on a bien la convergence de  $f_n^2$  vers  $1/2$ .

## IV Dérivation au sens des distributions.

Comme les distributions sont des *fonctions généralisées* destinées à aider à la résolution des équations aux dérivées partielles, il est nécessaire de savoir les dériver.

L'un des énormes avantages de cette théorie est que **toutes les distributions sont dérivables** ce qui simplifie énormément l'analyse, au moins sur ce point.

### Définition et Proposition A.17 (Dérivée d'une distribution)

Soit  $T \in \mathcal{D}'(\Omega)$  une distribution et  $\alpha \in \mathbb{N}^d$  un multi-indice. L'application  $\partial^\alpha T : \mathcal{D}(\Omega) \rightarrow \mathbb{R}$  définie par

$$\partial^\alpha T : \varphi \mapsto (-1)^{|\alpha|} \langle T, \partial^\alpha \varphi \rangle_{\mathcal{D}', \mathcal{D}},$$

est une distribution sur  $\Omega$  appelée *distribution dérivée  $\alpha$ -fois de  $T$* .

#### Preuve :

On remarque tout d'abord que si  $\varphi \in \mathcal{D}(\Omega)$ , alors  $\partial^\alpha \varphi$  est aussi dans  $\mathcal{D}(\Omega)$  et donc l'application  $\partial^\alpha T$  est bien définie. Il reste à voir qu'elle est continue. Pour cela, on constate que si  $\varphi_n \rightarrow \varphi$  dans  $\mathcal{D}(\Omega)$  alors  $\partial^\alpha \varphi_n \rightarrow \partial^\alpha \varphi$  dans  $\mathcal{D}(\Omega)$ .

■

Bien entendu, cette notion ne fait que généraliser la notion usuelle d'après la proposition suivante

### Proposition A.18

Si  $f \in \mathcal{C}^k(\Omega, \mathbb{R})$ , alors nous avons pour tout  $\alpha \in \mathbb{N}^d$  tel que  $|\alpha| \leq k$

$$\partial^\alpha (T_f) = T_{\partial^\alpha f}.$$

Autrement dit, la notion de dérivée au sens des distributions et celle de dérivée usuelle coïncident.

#### Preuve :

Il s'agit simplement d'une utilisation répétée de la formule d'intégration par parties pour les fonctions à support compact. ■

Une dernière chose très agréable avec les distributions est la possibilité d'intervertir la dérivation et la limite d'une suite "sans réfléchir".

### Théorème A.19

Soit  $(T_n)_n$  une suite de distributions qui converge vers une distribution  $T$  et soit  $\alpha \in \mathbb{N}^d$  alors on a

$$\partial^\alpha T_n \xrightarrow[n \rightarrow \infty]{} \partial^\alpha T, \text{ au sens des distributions.}$$

#### Preuve :

On écrit juste les définitions

$$\langle \partial^\alpha T_n, \varphi \rangle_{\mathcal{D}', \mathcal{D}} = (-1)^{|\alpha|} \langle T_n, \partial^\alpha \varphi \rangle_{\mathcal{D}', \mathcal{D}} \xrightarrow[n \rightarrow \infty]{} (-1)^{|\alpha|} \langle T, \partial^\alpha \varphi \rangle_{\mathcal{D}', \mathcal{D}} = \langle \partial^\alpha T, \varphi \rangle_{\mathcal{D}', \mathcal{D}}.$$

■

Exemples :

— Dérivée d'une masse de Dirac en 1D :

On calcule

$$\langle \delta'_{x_0}, \varphi \rangle_{\mathcal{D}', \mathcal{D}} = -\langle \delta_{x_0}, \varphi' \rangle_{\mathcal{D}', \mathcal{D}} = -\varphi'(x_0).$$

— Dérivée d'une fonction  $\mathcal{C}^1$  par morceaux en 1D :

Soit  $f : [0, 1] \rightarrow \mathbb{R}$  de classe  $\mathcal{C}^1$  par morceaux. On note  $x_1 < \dots < x_n$  les points de discontinuité potentielle de  $f$  et  $x_0 = 0, x_{n+1} = 1$ . Un simple calcul d'intégration par parties permet de montrer la **formule des sauts**

$$\begin{aligned} \langle \partial_x f, \varphi \rangle_{\mathcal{D}', \mathcal{D}} &= -\langle f, \partial_x \varphi \rangle_{\mathcal{D}', \mathcal{D}} \\ &= -\sum_{i=0}^n \int_{x_i}^{x_{i+1}} f \varphi' dx \\ &= -\sum_{i=0}^n \left[ f(x_{i+1}^-) \varphi(x_{i+1}) - f(x_i^+) \varphi(x_i) - \int_{x_i}^{x_{i+1}} f' \varphi dx \right] \\ &= \sum_{i=1}^n \varphi(x_i) [f(x_i^+) - f(x_i^-)] + \sum_{i=0}^n \int_{x_i}^{x_{i+1}} f' \varphi dx. \end{aligned}$$

Ainsi, au sens des distributions nous avons

$$\partial_x f = f' + \sum_{i=1}^n [f(x_i^+) - f(x_i^-)] \delta_{x_i},$$

où on a noté  $f'$  la fonction continue par morceaux qui coïncide avec la dérivée de  $f$  sur chaque intervalle  $]x_i, x_{i+1}[$ . Les termes de Dirac montrent l'influence des discontinuités de  $f$  dans la dérivée.

On retrouve le fait que  $f \in H^1(]0, 1[)$  si et seulement si les sauts sont nuls, autrement dit, si  $f$  est continue sur  $[0, 1]$ .

- Piège classique : le produit de deux distributions n'est pas bien défini en général (ni même le produit d'une distribution par une fonction insuffisamment régulière).

Ainsi, si  $u \in L^2(]0, 1[)$ , la quantité  $u(\partial_x u)$  n'est pas bien définie en général (elle l'est si  $u \in H^1(]0, 1[)$ ). Par contre, on sait que (au moins formellement) on a  $u(\partial_x u) = \frac{1}{2} \partial_x(u^2)$  et cette deuxième écriture est, elle, parfaitement définie car  $u^2 \in L^1(]0, 1[)$  est une distribution qui admet bien une dérivée.

On conclut cette annexe par un résultat très naturel et utile mais aussi trivial que ce qu'il y paraît au premier abord.

**Proposition A.20 (Distributions à gradient nul)**

Soit  $T \in \mathcal{D}'(\Omega)$  une distribution à gradient nul, c'est-à-dire que

$$\partial_{x_i} T = 0, \text{ dans } \mathcal{D}'(\Omega), \text{ pour tout } 1 \leq i \leq d,$$

alors  $T$  est une constante, c'est-à-dire qu'il existe  $\alpha \in \mathbb{R}$  vérifiant

$$\langle T, \varphi \rangle_{\mathcal{D}', \mathcal{D}} = \alpha \int_{\Omega} \varphi dx.$$

**Preuve :**

On va se contenter de la preuve en dimension 1,  $\Omega$  étant alors un intervalle  $I$ , car le cas général étant un peu plus délicat (voir par exemple [3]).

Soit donc  $T \in \mathcal{D}'(I)$  telle que pour toute fonction test  $\varphi \in \mathcal{D}(I)$  on a :

$$\langle T, \varphi' \rangle_{\mathcal{D}', \mathcal{D}} = 0. \quad (\text{A.5})$$

On fixe une fois pour toute une fonction test  $\theta \in \mathcal{D}(I)$  telle que  $\int_I \theta(t) dt = 1$ . Maintenant pour toute autre fonction test  $\psi \in \mathcal{D}(I)$  on pose

$$\varphi(x) = \int_a^x \psi(t) dt - \left( \int_I \psi \right) \left( \int_a^x \theta(t) dt \right).$$

Cette fonction est bien sûr de classe  $\mathcal{C}^\infty$ , et on vérifie qu'elle est également à support compact. De plus on a

$$\varphi'(x) = \psi(x) - \left( \int_I \psi \right) \theta(x).$$

Appliquons (A.5) à la fonction  $\varphi$  ainsi construite, on trouve

$$\langle T, \psi \rangle_{\mathcal{D}', \mathcal{D}} = \left( \int_I \psi(x) dx \right) \underbrace{\langle T, \theta \rangle_{\mathcal{D}', \mathcal{D}}}_{\stackrel{\text{def}}{=} m_{T, \theta}}.$$

On a donc obtenu que, pour toute fonction test  $\psi$ , on a

$$\langle T, \psi \rangle_{\mathcal{D}', \mathcal{D}} = m_{T, \theta} \int_I \psi(x) dx = \langle m_{T, \theta}, \psi \rangle_{\mathcal{D}', \mathcal{D}},$$

et donc que  $T$  est égale à la constante  $m_{T, \theta}$ . ■



## Annexe B

# La formule de Stokes

Le but de la suite de cette annexe est de comprendre ce qu'il convient de faire quand on ne travaille plus sur  $\mathbb{R}^d$  tout entier avec des fonctions à support compact mais dans un domaine quelconque et des fonctions potentiellement non nulles au bord. Comme on le sait d'après notre expérience de l'intégration par parties en 1D, des termes de bord doivent nécessairement apparaître dans la formule.

Notre but va donc être de prouver le résultat suivant

### Théorème B.1 (Formule de la divergence / Formule de Stokes)

Soit  $\Omega$  un ouvert borné de  $\mathbb{R}^d$  satisfaisant des hypothèses détaillées plus loin, alors pour tout champ de vecteurs  $F \in C^1(\mathbb{R}^d, \mathbb{R}^d)$  on a l'égalité

$$\int_{\Omega} \operatorname{div} F \, dx = \int_{\partial\Omega} F \cdot n \, d\sigma,$$

où  $n$  désigne la **normale unitaire sortante** au domaine  $\Omega$ .

Pour montrer ce théorème, il est donc nécessaire de définir ce qu'est la normale unitaire sortante d'un domaine  $\Omega$  d'une part, et ce qu'est l'intégrale d'une quantité scalaire sur le bord  $\partial\Omega$  de ce domaine (il ne peut s'agir de l'intégrale de Lebesgue usuelle car selon toute probabilité  $\partial\Omega$  est de mesure nulle dans  $\mathbb{R}^d$ ). Ce sera l'objet des sections suivantes.

A l'aide de la formule ci-dessus, on peut en particulier généraliser les résultats de la Proposition A.3 de la façon suivante pour des fonctions non nécessairement à support compact

### Proposition B.2

Soit  $\Omega$  comme dans le théorème précédent.

1. Soient  $u \in C^1(\mathbb{R}^d, \mathbb{R})$  et  $V \in C^1(\mathbb{R}^d, \mathbb{R}^d)$  alors on a

$$\int_{\Omega} u(\operatorname{div} V) \, dx = - \int_{\Omega} (\nabla u) \cdot V \, dx + \int_{\partial\Omega} u(V \cdot n) \, d\sigma.$$

2. Soient  $u, v \in C^2(\mathbb{R}^d, \mathbb{R})$  tels que  $u$  ou  $v$  soit à support compact, alors on a

$$\begin{aligned} \int_{\Omega} u(-\Delta v) \, dx &= \int_{\Omega} (\nabla u) \cdot (\nabla v) \, dx - \int_{\partial\Omega} u(\nabla v \cdot n) \, d\sigma, \\ \int_{\Omega} (\Delta u)v \, dx - \int_{\Omega} u(\Delta v) \, dx &= \int_{\partial\Omega} v(\nabla u \cdot n) \, d\sigma - \int_{\partial\Omega} u(\nabla v \cdot n) \, d\sigma. \end{aligned}$$

## I Hypersurfaces de $\mathbb{R}^d$ . Intégrale de surface

### I.1 Courbes planes

#### Définition B.3

Soit  $J \subset \mathbb{R}$  un intervalle compact de  $\mathbb{R}$  et  $\gamma : J \rightarrow \mathbb{R}^2$  une fonction de classe  $C^1$  telle que  $\gamma'(t) \neq 0$  pour tout  $t \in J$  et telle que  $\gamma$  est injective. On dit alors que l'image  $\Gamma = \gamma(J) \subset \mathbb{R}^2$  est une courbe simple de  $\mathbb{R}^2$  et que le couple  $(J, \gamma)$  est une paramétrisation de  $\Gamma$ .

**Définition B.4**

Pour tout  $t \in J$ , le vecteur  $\gamma'(t) \in \mathbb{R}^2$  est un **vecteur tangent** à la courbe au point  $\gamma(t)$ .

On note  $R_{\pi/2}$  la rotation d'angle  $\pi/2$  (une orientation étant préalablement choisie). Pour tout  $t \in J$ , le vecteur

$$n = \frac{R_{\pi/2}\gamma'(t)}{\|\gamma'(t)\|},$$

est un **vecteur normal unitaire** à la courbe au point  $\gamma(t)$ . Celui-ci ne dépend pas de la paramétrisation, à part peut-être son orientation.

**Définition et Proposition B.5**

Soit  $\Gamma = \gamma(J)$  une courbe simple de  $\mathbb{R}^2$  et  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction continue, on définit alors la quantité

$$I_\gamma(f) = \int_J f(\gamma(t)) \|\gamma'(t)\| dt,$$

où  $\|\cdot\|$  désigne ici la norme euclidienne dans  $\mathbb{R}^2$ .

Alors nous avons :

- $I_\gamma(f)$  ne dépend que des valeurs de  $f$  sur  $\Gamma$ .
- $I_\gamma(f)$  ne dépend pas de la paramétrisation de  $\Gamma$ , c'est-à-dire que si  $(\tilde{J}, \tilde{\gamma})$  est une autre paramétrisation vérifiant  $\tilde{\gamma}(\tilde{J}) = \Gamma$ , alors on a

$$I_{\tilde{\gamma}}(f) = I_\gamma(f).$$

Ainsi, la quantité  $I_\gamma(f)$  sera appelée l'intégrale de  $f$  sur  $\Gamma$  et notée

$$\int_\Gamma f d\sigma, \text{ ou } \int_\Gamma f(x) d\sigma(x).$$

**Preuve :**

On ne va pas faire la preuve complète mais en expliquer rapidement le principe, le point-clé étant le théorème de changement de variable.

Soit  $(J, \gamma)$  est une paramétrisation de  $\Gamma$  (on supposera sans perte de généralité que  $J = [0, T]$ ) et soit  $S : t \in [0, T] \mapsto [0, S(T)]$  une application de classe  $\mathcal{C}^1$  bijective et telle que  $S'(t) > 0$  pour tout  $t \in [0, T]$ . Alors on peut construire une nouvelle paramétrisation de  $\Gamma$  donnée par  $(\tilde{J} = [0, S(T)], \tilde{\gamma})$  et

$$\tilde{\gamma}(s) = \gamma(S^{-1}(s)), \quad \forall s \in [0, S(T)].$$

Comme  $S$  est bijective, cette application est bien définie et on a bien  $\gamma(J) = \tilde{\gamma}(\tilde{J}) = \Gamma$ . Enfin, nous avons

$$\gamma(t) = \tilde{\gamma}(S(t)), \quad \forall t \in [0, T], \tag{B.1}$$

et donc

$$\gamma'(t) = S'(t)\tilde{\gamma}'(S(t)), \quad \forall t \in [0, T], \tag{B.2}$$

ce qui prouve en particulier que  $\tilde{\gamma}'$  ne peut pas s'annuler. On a donc bien affaire à une nouvelle paramétrisation du même ensemble  $\Gamma$ .

On effectue maintenant le changement de variable  $s = S(t)$  dans l'intégrale qui définit  $I_\gamma(f)$

$$\begin{aligned} I_\gamma(f) &= \int_0^T f(\gamma(t)) \|\gamma'(t)\| dt \\ &= \int_0^T f(\tilde{\gamma}(S(t))) \|S'(t)\tilde{\gamma}'(S(t))\| dt \\ &= \int_0^T f(\tilde{\gamma}(S(t))) \|\tilde{\gamma}'(S(t))\| S'(t) dt \\ &= \int_0^{S(T)} f(\tilde{\gamma}(s)) \|\tilde{\gamma}'(s)\| ds \\ &= I_{\tilde{\gamma}}(f), \end{aligned}$$

ce qui est bien ce qu'on voulait démontrer.

Pour clore complètement la preuve, il faudrait montrer que **toutes** les paramétrisations de  $\Gamma$  ont bien, plus ou moins, la forme  $\tilde{\gamma} = \gamma \circ S^{-1}$  pour une application  $S$  bien choisie vérifiant les hypothèses plus haut. Nous laissons cette partie aux soins du lecteur. ■

### Remarque B.6

Avec cette définition, la **longueur** de la courbe  $\Gamma$  est l'intégrale de la fonction constante égale à 1

$$|\Gamma| = \int_{\Gamma} 1 \, d\sigma.$$

Donnons quelques exemples :

- Intégrale sur un segment : Soient  $A = (a_1, a_2)$  et  $B = (b_1, b_2)$  deux points de  $\mathbb{R}^2$  et  $\Gamma = [A, B]$  le segment joignant  $A$  et  $B$ . On choisit la paramétrisation naturelle de ce segment

$$\gamma : t \in [0, 1] \mapsto tA + (1-t)B = (ta_1 + (1-t)b_1, ta_2 + (1-t)b_2) \in \mathbb{R}^2.$$

Pour une fonction  $f$  quelconque on a donc

$$\int_{\Gamma} f \, d\sigma = \int_0^1 f(tA + (1-t)B) \|A - B\| \, dt = \|A - B\| \int_0^1 f(tA + (1-t)B) \, dt.$$

Pour  $f = 1$ , on retrouve que la longueur du segment vaut  $\|A - B\|$ .

- Intégrale sur un cercle : On considère  $\Gamma = C(0, R)$  le cercle de centre 0 et de rayon  $R$  et une paramétrisation donnée par

$$\gamma : t \in [0, 2\pi] \mapsto \gamma(t) = R(\cos(t), \sin(t)).$$

Remarquons que  $\gamma(0) = \gamma(2\pi)$  ici et donc que  $\gamma$  n'est pas injective, néanmoins on se convainc aisément que les propriétés établies plus haut s'adaptent sans difficulté au cas des paramétrisations périodiques (courbes fermées).

Dans ces conditions, comme  $\|\gamma'(t)\| = R$  pour tout  $t$ , on a

$$\int_{C(0, R)} f \, d\sigma = R \int_0^{2\pi} f(R \cos(t), R \sin(t)) \, dt,$$

et pour  $f = 1$  on retrouve bien le périmètre du cercle égal à  $2\pi R$ .

- Intégrale sur un graphe : Soit  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  une fonction de classe  $\mathcal{C}^1$  et  $[\alpha, \beta] \subset \mathbb{R}$ . On s'intéresse à la courbe du plan formée de la portion de graphe de  $\varphi$  et définie par

$$\Gamma = \{(x, \varphi(x)), x \in [\alpha, \beta]\} \subset \mathbb{R}^2.$$

Cette courbe possède une paramétrisation “naturelle” donnée par

$$\gamma : x \in [\alpha, \beta] \mapsto (x, \varphi(x)) \in \Gamma.$$

On a alors pour une fonction  $f$  quelconque

$$\int_{\Gamma} f \, d\sigma = \int_{\alpha}^{\beta} f(x, \varphi(x)) \sqrt{1 + \varphi'(x)^2} \, dx.$$

## I.2 Intégrales sur des hypersurfaces de $\mathbb{R}^d$

On se contente ici du cas  $d = 3$  par simplicité mais toute la théorie peut s'adapter à la dimension quelconque.

On va admettre que l'on peut définir, comme précédemment, l'intégrale de n'importe quelle fonction (disons continue) sur une hypersurface  $\Gamma$  de  $\mathbb{R}^3$  par la formule

$$\int_{\Gamma} f \, d\sigma = \int_U f(\gamma(u)) \left\| \frac{\partial \gamma}{\partial u_1} \wedge \frac{\partial \gamma}{\partial u_2} \right\| \, du,$$

où

$$\gamma : u \in U \subset \mathbb{R}^2 \mapsto \gamma(u) \in \mathbb{R}^3,$$

est une paramétrisation de  $\Gamma$  (i.e. une bijection régulière telle que  $\frac{\partial \gamma}{\partial u_1} \wedge \frac{\partial \gamma}{\partial u_2} \neq 0$  en tout point  $u$ ).

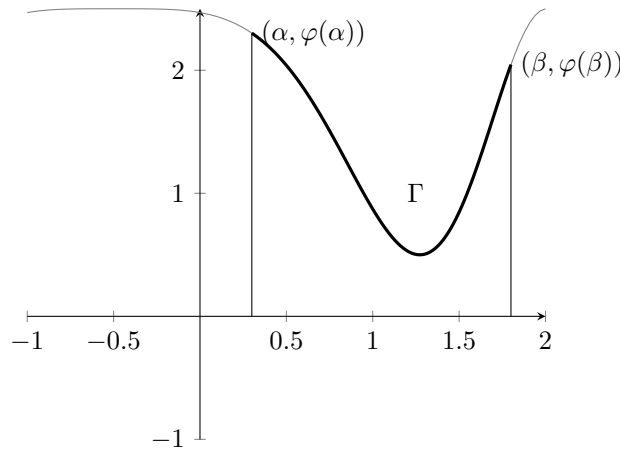


FIGURE B.1 – Intégrale sur le graphe d'une fonction

En pratique, il n'est pas toujours vrai qu'une telle paramétrisation globale de  $\Gamma$  existe et on doit raisonner par des paramétrisations locales et une technique de recollement mais on ne va pas détailler ces points.

### Définition B.7

Pour tout  $u \in U$ , les vecteurs  $\frac{\partial \gamma}{\partial u_1}(u), \frac{\partial \gamma}{\partial u_2}(u) \in \mathbb{R}^3$  sont des **vecteurs tangents** à la surface au point  $\gamma(u)$ . De plus, ils sont linéairement indépendants et le plan qu'ils engendrent ne dépend pas de la paramétrisation choisie, on l'appelle **le plan tangent** à la surface en ce point.

Pour tout  $u \in U$ , le vecteur

$$n = \frac{1}{\left\| \frac{\partial \gamma}{\partial u_1} \wedge \frac{\partial \gamma}{\partial u_2} \right\|} \frac{\partial \gamma}{\partial u_1} \wedge \frac{\partial \gamma}{\partial u_2},$$

est un **vecteur normal unitaire** à la surface au point  $\gamma(u)$ . Celui-ci ne dépend pas de la paramétrisation, à part peut-être son orientation.

Le cas des graphes est toujours intéressant à étudier de plus près. Soit donc  $\varphi : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction régulière, le graphe de  $\varphi$  est défini par

$$\Gamma = \{(u, \varphi(u)), u \in U\} \subset \mathbb{R}^3,$$

et il est naturellement paramétrisé par l'application

$$\gamma : u = (u_1, u_2) \in U \mapsto (u_1, u_2, \varphi(u)) \in \mathbb{R}^3.$$

Un calcul immédiat montre que

$$\frac{\partial \gamma}{\partial u_1} = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial \varphi}{\partial u_1} \end{pmatrix}, \quad \text{et} \quad \frac{\partial \gamma}{\partial u_2} = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial \varphi}{\partial u_2} \end{pmatrix},$$

de sorte que

$$\frac{\partial \gamma}{\partial u_1} \wedge \frac{\partial \gamma}{\partial u_2} = \begin{pmatrix} -\frac{\partial \varphi}{\partial u_1} \\ -\frac{\partial \varphi}{\partial u_2} \\ 1 \end{pmatrix},$$

et en particulier, on a

$$\left\| \frac{\partial \gamma}{\partial u_1} \wedge \frac{\partial \gamma}{\partial u_2} \right\| = \sqrt{1 + \|\nabla \varphi(u)\|^2} > 0,$$

ce qui prouve que la paramétrisation est licite. L'intégrale sur  $\Gamma$  s'exprime alors par la formule

$$\int_{\Gamma} f \, d\sigma = \int_U f(u_1, u_2, \varphi(u_1, u_2)) \sqrt{1 + \|\nabla \varphi(u)\|^2} \, du. \quad (\text{B.3})$$

Par ailleurs, la normale unitaire sur ce graphe est donnée par

$$n = \frac{1}{\sqrt{1 + \|\nabla \varphi(u)\|^2}} \begin{pmatrix} -\frac{\partial \varphi}{\partial u_1} \\ -\frac{\partial \varphi}{\partial u_2} \\ 1 \end{pmatrix}, \quad (\text{B.4})$$

et on observe que ce vecteur est orienté “vers le haut”, c’est-à-dire dans le sens des  $x_3$  croissants, et ce indépendamment de  $\varphi$ .

## II Domaines réguliers de $\mathbb{R}^d$

### Définition B.8

On dit qu’un ouvert  $\Omega$  de  $\mathbb{R}^d$  est un **domaine régulier** (ou par abus de langage **un ouvert régulier**) si, localement,  $\partial\Omega$  est une hypersurface régulière de  $\mathbb{R}^d$  et que  $\Omega$  se situe d’un seul côté de  $\partial\Omega$ . Ainsi, en tout point de  $\partial\Omega$ , on peut définir l’unique vecteur normal unitaire orienté de l’intérieur vers l’extérieur de  $\Omega$ .

Quelques exemples d’ouverts qui **ne sont pas** des domaines réguliers :

—  $\Omega = \mathbb{R}^d \setminus \{0\}$

En effet, le bord de  $\Omega$  est réduit à un point et n’est donc pas une hypersurface de  $\mathbb{R}^d$ .

—  $\Omega = \mathbb{R}^d \setminus (\{0\} \times \mathbb{R}^{d-1})$  est une union de deux demi-espaces.

Son bord est bien une hypersurface de  $\mathbb{R}^d$  mais  $\Omega$  se situe localement des deux côtés de sa frontière, on ne peut donc pas définir de normale *sortante*.

Remarquons que, dans ce type de situation, on peut souvent (mais pas toujours) décomposer  $\Omega$  en deux ouverts disjoints qui, eux, sont des domaines réguliers.

—  $\Omega = ]0, 1[ \times ]0, 1[ \subset \mathbb{R}^2$

Son bord n’est pas une hypersurface régulière à cause des coins du carré. En pratique, on peut quand même travailler avec ce type de domaines à condition de relaxer un peu (mais pas trop) les hypothèses de régularité sur sa frontière.

— Domaine à cusp :

L’ouvert

$$\Omega = \{(x, y), 0 < x < 1, 0 < y < x^2\},$$

dessiné dans la figure B.2 n’est pas un domaine régulier à cause de la singularité “pointue” dans le coin inférieur gauche. Ce type de domaine présente un certain nombre de particularités que nous ne détaillerons pas plus avant ici.

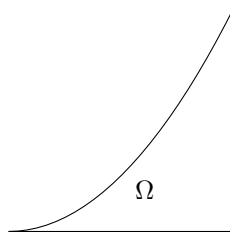


FIGURE B.2 – Domaine à cusp

## III Formule de Stokes

On souhaite ici démontrer le théorème B.1 au moins dans des cas particuliers.

### III.1 Le cas du demi-espace $\mathbb{R}_+^d$

On se place sur  $\Omega = \mathbb{R}_+^d = \mathbb{R}^{d-1} \times \mathbb{R}_+^1$ . Dans le cas  $d = 3$ , on a donc  $\mathbb{R}_+^3 = \{(x_1, x_2, x_3), x_3 > 0\}$ . Le bord de cet ouvert est le plan  $\Gamma = \partial\Omega = \mathbb{R}^2 \times \{0\}$ .

On prend maintenant un champ de vecteurs  $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$  de classe  $C^1$  et à support compact dans  $\mathbb{R}^d$ . On intègre la divergence de  $F$  sur  $\Omega$ . La contribution des termes  $\partial_{x_1} F_1$  et  $\partial_{x_2} F_2$  est toujours nulle d’après le théorème de Fubini. En revanche celle du troisième terme devient

$$\int_{\Omega} \partial_{x_3} F_3 dx = \int_{\mathbb{R}} \int_{\mathbb{R}} \left( \int_0^{+\infty} \partial_{x_3} F_3(x_1, x_2, x_3) dx_3 \right) dx_1 dx_2 = - \int_{\mathbb{R}^2} F_3(x_1, x_2, 0) dx_1 dx_2.$$

En remarquant que la normale unitaire **sortante** à  $\Omega$  sur  $\Gamma$  est donnée par  $n = {}^t(0, 0, -1)$ , on a établi

$$\int_{\Omega} \operatorname{div} F dx = \int_{\mathbb{R}^2} F \cdot n(x_1, x_2, 0) dx_1 dx_2.$$

Cette dernière est bien égale à l'intégrale sur l'hypersurface (plate)  $\partial\Omega$  que l'on a défini plus haut.

### III.2 Le cas du demi-espace à frontière non plane

On se donne maintenant une fonction  $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$  et on considère l'ouvert (appelé épigraphe de  $\varphi$ ) défini par

$$\Omega = \{x = (x_1, x_2, x_3) \mid x_3 > \varphi(x_1, x_2)\},$$

dont le bord  $\Gamma = \partial\Omega$  n'est rien d'autre que le graphe de  $\varphi$ . L'ensemble  $\Omega$  est donc la partie de l'espace qui se trouve *au-dessus* du graphe  $\Gamma$ .

On prend à nouveau un champ de vecteurs  $F$  régulier et à support compact. On introduit le changement de variable

$$\Psi : \tilde{x} \in \mathbb{R}^3 \mapsto (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3 + \varphi(\tilde{x}_1, \tilde{x}_2)) \in \mathbb{R}^3.$$

On vérifie aisément que celui-ci est de classe  $\mathcal{C}^1$  et bijectif de  $\mathbb{R}^3$  dans  $\mathbb{R}^3$ . De plus, sa jacobienne vaut

$$\text{Jac}\Psi(\tilde{x}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \partial_{x_1}\varphi & \partial_{x_2}\varphi & 1 \end{pmatrix},$$

dont le déterminant vaut 1.

Enfin, on observe que l'image du demi-espace droit  $\mathbb{R}_+^3$  par  $\Psi$  est exactement le domaine  $\Omega$  qui nous intéresse. On définit donc un nouveau champ de vecteur  $\tilde{F}$  par la formule

$$\tilde{F} = F \circ \Psi.$$

On calcule immédiatement

$$(\text{div } \tilde{F})(\tilde{x}) = (\text{div } F)(\Psi(\tilde{x})) + \partial_{x_1}\varphi(\tilde{x})(\partial_{x_3}F_1)(\Psi(\tilde{x})) + \partial_{x_2}\varphi(\tilde{x})(\partial_{x_3}F_2)(\Psi(\tilde{x})).$$

On remarque les deux derniers termes s'écrivent comme des dérivées par rapport à  $\tilde{x}_3$ . Autrement dit si on définit un nouveau champ de vecteur

$$G(\tilde{x}) = \tilde{F}(\tilde{x}) + \begin{pmatrix} 0 \\ 0 \\ -(\partial_{x_1}\varphi)\tilde{F}_1 - (\partial_{x_2}\varphi)\tilde{F}_2 \end{pmatrix},$$

alors on a montré que

$$(\text{div } G)(\tilde{x}) = (\text{div } F)(\Psi(\tilde{x})).$$

On intègre alors cette égalité sur le demi-espace  $\mathbb{R}_+^3$  pour obtenir

$$\int_{\mathbb{R}_+^3} (\text{div } G)(\tilde{x}) d\tilde{x} = \int_{\mathbb{R}_+^3} (\text{div } F)(\Psi(\tilde{x})) d\tilde{x}.$$

Comme le Jacobien de  $\Psi$  vaut 1 partout et que  $\Psi$  envoie  $\mathbb{R}_+^3$  sur  $\Omega$ , on voit que le second terme vaut exactement  $\int_{\Omega} \text{div } F dx$ . Quant au premier terme on peut lui appliquer la formule de Stokes sur un demi-espace que nous avons prouvée dans le paragraphe précédent. On a donc finalement établi

$$\int_{\Omega} \text{div } F dx = - \int_{\mathbb{R}^2 \times \{0\}} G_3 d\sigma = - \int_{\mathbb{R}^2} G_3(\tilde{x}_1, \tilde{x}_2, 0) d\tilde{x}_1 d\tilde{x}_2.$$

Par définition de  $G$  ceci s'écrit finalement

$$\begin{aligned} \int_{\Omega} \text{div } F dx &= \int_{\mathbb{R}^2} \left[ (\partial_{x_1}\varphi)\tilde{F}_1(\tilde{x}_1, \tilde{x}_2, 0) + (\partial_{x_2}\varphi)\tilde{F}_2(\tilde{x}_1, \tilde{x}_2, 0) - \tilde{F}_3(\tilde{x}_1, \tilde{x}_2, 0) \right] d\tilde{x}_1 d\tilde{x}_2 \\ &= \int_{\mathbb{R}^2} \left[ (\partial_{x_1}\varphi)F_1(\tilde{x}_1, \tilde{x}_2, \varphi(\tilde{x}_1, \tilde{x}_2)) + (\partial_{x_2}\varphi)F_2(\tilde{x}_1, \tilde{x}_2, \varphi(\tilde{x}_1, \tilde{x}_2)) - F_3(\tilde{x}_1, \tilde{x}_2, \varphi(\tilde{x}_1, \tilde{x}_2)) \right] d\tilde{x}_1 d\tilde{x}_2. \end{aligned}$$

Par les définitions rappelées plus haut (en particulier les formules (B.3), (B.4)), et en prenant garde à l'orientation de la normale, on a bien obtenu que

$$\int_{\Omega} \text{div } F dx = \int_{\partial\Omega} (F \cdot n) d\sigma.$$

# Bibliographie

- [1] Sylvie Benzoni-Gavage. *Calcul différentiel et équations différentielles*. Dunod, 2010.
- [2] Florent Berthelin. *Equations différentielles*. Cassini, 2017.
- [3] Franck Boyer and Pierre Fabrie. *Mathematical Tools for the Study of the Incompressible Navier-Stokes Equations and Related Models*. Springer New York, 2013.
- [4] Philippe G. Ciarlet. *Linear and nonlinear functional analysis with applications*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2013.
- [5] Jean-Pierre Demailly. *Analyse Numérique et équations différentielles*. Collection Grenoble Sciences. Presses Universitaires de Grenoble, 1991.
- [6] Françoise Demengel and Gilbert Demengel. *Espaces fonctionnels*. Savoirs Actuels (Les Ulis). [Current Scholarship (Les Ulis)]. EDP Sciences, Les Ulis ; CNRS Éditions, Paris, 2007. Utilisation dans la résolution des équations aux dérivées partielles. [Application to the solution of partial differential equations].
- [7] Benoît Perthame. *Transport Equations in Biology*. Birkhäuser Basel, 2007.