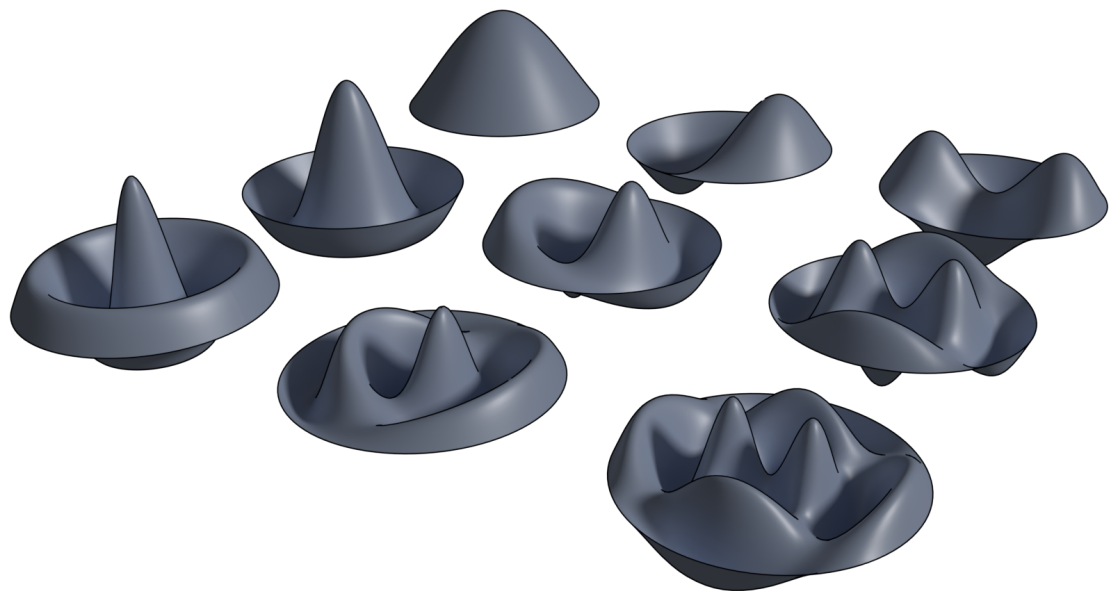


# MATHÉMATIQUES

2<sup>ÈME</sup> ANNÉE



Vincent Démery

2024–2025

Version du 18 août 2024



# Table des matières

<b>Table des matières</b>	<b>4</b>
<b>1 Équations aux dérivées partielles</b>	<b>5</b>
1.1 Introduction . . . . .	5
1.2 Exemples . . . . .	5
1.2.1 Équation de transport . . . . .	5
1.2.2 Équation de Poisson . . . . .	6
1.2.3 Équation de la chaleur (ou de diffusion) . . . . .	6
1.2.4 Équation des ondes . . . . .	6
1.2.5 Équations de Navier-Stokes . . . . .	6
1.3 Définitions et propriétés élémentaires . . . . .	6
1.3.1 Définitions générales . . . . .	6
1.3.2 EDP linéaires, fonction de Green . . . . .	7
1.3.3 EDP linéaires à coefficients constants . . . . .	8
1.3.4 Méthode des images . . . . .	8
1.4 EDP linéaires du 1 <sup>er</sup> ordre, méthode des caractéristiques . . . . .	9
1.4.1 Méthode des caractéristiques . . . . .	9
1.4.2 Exemple . . . . .	10
1.4.3 Résolution de l'équation de transport . . . . .	12
1.4.4 Remarques . . . . .	13
1.5 EDP linéaires du 2 <sup>nd</sup> ordre à coefficients constants . . . . .	13
1.5.1 Classification . . . . .	13
1.5.2 Équation de Poisson . . . . .	14
1.5.3 Équation de la chaleur . . . . .	17
1.5.4 Équation des ondes . . . . .	19
1.6 Analyse spectrale des EDP linéaires . . . . .	22
1.6.1 Exemples . . . . .	22
1.6.2 Généralités sur l'analyse spectrale . . . . .	22
1.6.3 Équation de Poisson sur une ligne . . . . .	23
1.6.4 Équation de Poisson sur un disque . . . . .	25

1.7	Pour aller plus loin . . . . .	26
<b>2</b>	<b>Calcul variationnel</b>	<b>29</b>
2.1	Introduction . . . . .	29
2.2	Calcul variationnel en dimension un . . . . .	29
2.2.1	Exemple : forme d'une corde pesante . . . . .	29
2.2.2	Équation d'Euler-Lagrange . . . . .	30
2.2.3	Minimisation sous contrainte . . . . .	33
2.2.4	Invariance par translation et intégrale du mouvement . . . . .	35
2.3	Calcul variationnel en dimension supérieure . . . . .	36
2.3.1	Exemples . . . . .	36
2.3.2	Équation d'Euler-Lagrange, conditions au bord . . . . .	37
2.4	Pour aller plus loin . . . . .	39
2.4.1	Calcul variationnel plus général . . . . .	39
2.4.2	Physique statistique . . . . .	39
2.4.3	Mécanique quantique . . . . .	39
<b>3</b>	<b>Probabilités</b>	<b>41</b>
3.1	Introduction . . . . .	41
3.2	Exemples . . . . .	41
3.3	Événements et probabilité . . . . .	42
3.3.1	Mesure de probabilité . . . . .	42
3.3.2	Conditionnement et indépendance . . . . .	42
3.4	Variables aléatoires . . . . .	43
3.4.1	Définition . . . . .	43
3.4.2	VA réelles . . . . .	44
3.4.3	VA discrètes et continues, lois classiques . . . . .	46
3.5	Suites de variables aléatoires . . . . .	49
3.5.1	Convergence de VA . . . . .	49
3.5.2	Loi des grands nombres . . . . .	50
3.5.3	Théorème de la limite centrale . . . . .	51
3.5.4	Intervalle de confiance . . . . .	52
3.6	Pour aller plus loin . . . . .	53
	<b>Index</b>	<b>55</b>

# Chapitre 1

## Équations aux dérivées partielles

### 1.1 Introduction

Les équations aux dérivées partielles (EDP), qui font intervenir les dérivées d'équations de plusieurs variables, sont omniprésentes en physique : électromagnétisme, mécanique des solides et des fluides, processus de transport, mécanique quantique, etc. Ces équations sont très difficiles à résoudre en général, et constituent un domaine de recherche actif des mathématiques.

Ici, après quelques généralités, nous nous concentrons sur les EDP linéaires. Nous voyons d'abord quelques propriétés communes, comme le principe de superposition, puis nous étudions « en détail » quelques EDP emblématiques (équations du transport, de Poisson, de la chaleur et des ondes). Pour chacune de ces équations, nous cherchons à répondre aux questions de base : existe-t-il une solution et comment la construire ? Est-elle unique ? Quelles sont ses propriétés ? Nous évoquons aussi l'analyse spectrale des EDP, qui donne un point de vue alternatif et très utile sur les EDP spatio-temporelles. Pour répondre aux différentes questions, nous introduisons divers outils comme la fonction de Green, la méthode des caractéristiques, la méthode de séparation des variables, le principe du maximum, etc. Ces outils sont rarement généraux mais il ne sont pas anecdotiques car ils constituent l'essentiel de ce qu'il faut savoir pour étudier des EDP.

### 1.2 Exemples

Pour commencer, nous donnons quelques exemples d'équations aux dérivées partielles qui nous serviront à discuter leurs propriétés.

#### 1.2.1 Équation de transport

On considère un élément présent en densité  $\rho(\mathbf{r}, t)$  dans un fluide en écoulement décrit par  $\mathbf{v}(\mathbf{r}, t)$  (qui est une donnée du problème), l'évolution de la densité est alors donnée par

$$\partial_t \rho(\mathbf{r}, t) = -\nabla \cdot [\rho(\mathbf{r}, t) \mathbf{v}(\mathbf{r}, t)]. \quad (1.1)$$

Si l'écoulement est uniforme et constant,  $\mathbf{v}(\mathbf{r}, t) = \mathbf{V}$  cette équation se réduit à

$$\partial_t \rho(\mathbf{r}, t) = -\mathbf{V} \cdot \nabla \rho(\mathbf{r}, t). \quad (1.2)$$

### 1.2.2 Équation de Poisson

L'équation de Poisson décrit par exemple le lien entre le potentiel électrostatique  $V(\mathbf{r})$  et une densité de charge  $\rho(\mathbf{r})$  :

$$\nabla^2 V(\mathbf{r}) = \rho(\mathbf{r}). \quad (1.3)$$

Nous avons enlevé le signe moins et pris  $\epsilon_0 = 1$ .

### 1.2.3 Équation de la chaleur (ou de diffusion)

La température  $\theta(\mathbf{r}, t)$  dans un fluide obéit en absence de convection à l'équation de la chaleur

$$\partial_t \theta(\mathbf{r}, t) = D \nabla^2 \theta(\mathbf{r}, t), \quad (1.4)$$

où  $D$  est le coefficient de diffusion.

### 1.2.4 Équation des ondes

L'équation des ondes décrit par exemple l'évolution d'un scalaire  $u(\mathbf{r}, t)$  associé au champ électrique dans le vide :

$$\partial_t^2 u(\mathbf{r}, t) = c^2 \nabla^2 u(\mathbf{r}, t), \quad (1.5)$$

où  $c$  est la vitesse de propagation des ondes.

### 1.2.5 Équations de Navier-Stokes

Les équations de Navier-Stokes décrivent l'évolution de la densité  $\rho(\mathbf{r}, t)$  et de la vitesse  $\mathbf{v}(\mathbf{r}, t)$  d'un fluide en écoulement

$$\rho [\partial_t \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v}] = -\nabla p + \mu \nabla^2 \mathbf{v}, \quad (1.6)$$

où  $p(\mathbf{r}, t)$  est la pression et  $\mu$  la viscosité dynamique du fluide. Cette équation doit être complétée par la conservation de la masse,  $\partial_t \rho = -\nabla \cdot [\rho \mathbf{v}]$  et l'équation d'état  $p = P(\rho)$ .

Si l'écoulement est en contact avec une surface solide décrite par la surface  $\Gamma$ , on peut prendre une condition de non-glissement au bord :

$$\forall \mathbf{r} \in \Gamma, \quad \mathbf{v}(\mathbf{r}, t) = 0. \quad (1.7)$$

## 1.3 Définitions et propriétés élémentaires

### 1.3.1 Définitions générales

Une *équation aux dérivées partielles* pour une fonction  $u(\mathbf{r})$  est une relation locale faisant intervenir des dérivées partielles de la fonction et qui doit être satisfaite sur un domaine  $\Omega \subset \mathbb{R}^d$ .

On appelle *ordre* d'une EDP l'ordre maximal des dérivées de  $u$  qui apparaissent dans l'équation. Dans les exemples ci-dessus, l'équation de transport est d'ordre 1 et toutes les autres équations sont d'ordre 2.

Une EDP est souvent complétée par des *conditions aux bords*, ou conditions *aux limites* satisfaites par l'équation sur le bord du domaine  $\Omega$ , noté  $\partial\Omega$ . Un exemple est donné pour les équations de Navier-Stokes en Éq. (1.7). Suivant la quantité qui est imposée sur le bord  $\partial\Omega$ , on distingue trois types de conditions aux limites :

- Condition de Dirichlet : on impose  $u(\mathbf{r})$ .
- Condition de Neumann : on impose  $\mathbf{n}(\mathbf{r}) \cdot \nabla u(\mathbf{r})$ , où  $\mathbf{n}(\mathbf{r})$  est le vecteur unitaire orthogonal à  $\partial\Omega$  en  $\mathbf{r}$ .
- Condition de Cauchy : on impose  $u(\mathbf{r})$  et  $\mathbf{n}(\mathbf{r}) \cdot \nabla u(\mathbf{r})$ .

Un *terme source* est un terme de l'EDP qui ne fait pas intervenir la fonction inconnue. C'est le cas de la densité de charge dans l'équation de Poisson (1.3).

Finalement, on dit qu'un problème d'EDP avec condition au bord est *bien posé* si la solution existe, est unique, et dépend continûment de la condition au bord.

### 1.3.2 EDP linéaires, fonction de Green

Une EDP est dite *linéaire* si tous les termes, hormis les termes source, sont linéaires en la fonction inconnue  $u(\mathbf{r})$ ; cette condition s'applique aussi aux conditions au bord. Dans les exemples, seule l'équation de Navier-Stokes n'est pas linéaire. On notera les termes linéaires  $\mathcal{L}u(\mathbf{r})$ , qui sont nécessairement de la forme

$$\mathcal{L}u(\mathbf{r}) = \sum_{i_1, \dots, i_d \geq 0} a_{i_1, \dots, i_d}(\mathbf{r}) \partial_1^{i_1} \cdots \partial_d^{i_d} u(\mathbf{r}). \quad (1.8)$$

Les fonctions  $a_{i_1, \dots, i_d}(\mathbf{r})$  sont les coefficients de l'équation. On notera  $\mathcal{C}u(\mathbf{r})$  les termes linéaires des conditions au bord, qui sont de la même forme. Une équation linéaire peut donc s'écrire

$$\mathcal{L}u(\mathbf{r}) = f(\mathbf{r}) \quad \forall \mathbf{r} \in \Omega, \quad (1.9)$$

$$\mathcal{C}u(\mathbf{r}) = g(\mathbf{r}) \quad \forall \mathbf{r} \in \partial\Omega, \quad (1.10)$$

$\mathcal{L}$  est l'*opérateur différentiel* de l'EDP, et les fonctions  $f(\mathbf{r})$  et  $g(\mathbf{r})$  sont les termes source. Dans la suite, on ne parle que de  $f(\mathbf{r})$ , mais il faut garder en tête que le terme source « de bord »  $g(\mathbf{r})$  peut être traité de la même façon.

Une EDP linéaire est *homogène* si elle ne contient pas de terme source. Une EDP linéaire et homogène peut donc s'écrire  $\mathcal{L}u(\mathbf{r}) = 0$ ,  $\mathcal{C}u(\mathbf{r}) = 0$ . Les équations (1.1, 1.4, 1.5) sont linéaires et homogènes. À une EDP linéaire avec des termes source, on peut associer une EDP homogène en prenant ces termes égaux à zéro.

Ces propriétés ont des conséquences importantes :

- toute combinaison linéaire de solutions d'une EDP linéaire homogène est aussi solution.
- si  $u_0(\mathbf{r})$  est solution d'une EDP linéaire avec un terme source  $f(\mathbf{r})$ ,  $\mathcal{L}u_0(\mathbf{r}) = f(\mathbf{r})$ , et  $v(\mathbf{r})$  est solution de l'EDP homogène associée,  $\mathcal{L}v(\mathbf{r}) = 0$ , alors  $u_0(\mathbf{r}) + v(\mathbf{r})$  est aussi solution de l'EDP avec un terme source.
- si  $u_1(\mathbf{r})$  est solution pour le terme source  $f_1(\mathbf{r})$ , et  $u_2(\mathbf{r})$  est solution pour le terme source  $f_2(\mathbf{r})$ , alors  $u(\mathbf{r}) = u_1(\mathbf{r}) + u_2(\mathbf{r})$  est solution pour le terme source  $f_1(\mathbf{r}) + f_2(\mathbf{r})$ ; c'est le *principe de superposition*.

Le principe de superposition permet d'écrire la solution pour un terme source  $f(\mathbf{r})$  quelconque à l'aide de la *fonction de Green*  $G(\mathbf{r}, \mathbf{r}')$  qui est solution de

$$\mathcal{L}G(\mathbf{r}, \mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}'), \quad (1.11)$$

où  $\delta$  est la distribution de Dirac, notée abusivement avec un argument  $\mathbf{r}$ .  $G(\mathbf{r}, \mathbf{r}')$  est donc la solution correspondant à un terme source qui est une distribution de Dirac centrée en  $\mathbf{r}'$ . La solution de  $\mathcal{L}u(\mathbf{r}) = f(\mathbf{r})$  est alors donnée par

$$u(\mathbf{r}) = \int G(\mathbf{r}, \mathbf{r}') f(\mathbf{r}') d\mathbf{r}'. \quad (1.12)$$

Cela se démontre en appliquant l'opérateur différentiel  $\mathcal{L}$  à cette égalité. On peut voir ça comme une réécriture du principe de superposition, en écrivant  $f(\mathbf{r})$  comme une « superposition » de distributions de Dirac,  $f(\mathbf{r}) = \int f(\mathbf{r}') \delta(\mathbf{r} - \mathbf{r}') d\mathbf{r}'$ .

### 1.3.3 EDP linéaires à coefficients constants

Les coefficients d'une EDP linéaire représentent les propriétés du milieu dans lequel on cherche à résoudre l'équation. Quand ce milieu est homogène, les coefficients sont donc constants. Si de plus le milieu est infini,  $\Omega = \mathbb{R}^d$ , alors le problème devient invariant par translation : pour une fonction  $u(\mathbf{r})$ , on définit la fonction translatée de  $\mathbf{a}$  par  $u_{\mathbf{a}}(\mathbf{r}) = u(\mathbf{r} - \mathbf{a})$ , alors  $\mathcal{L}u_{\mathbf{a}}(\mathbf{r}) = (\mathcal{L}u)_{\mathbf{a}}(\mathbf{r}) = (\mathcal{L}u)(\mathbf{r} - \mathbf{a})$ . Une conséquence est que si  $u(\mathbf{r})$  est solution avec le terme source  $f(\mathbf{r})$ , alors  $u_{\mathbf{a}}(\mathbf{r})$  est solution avec le terme source  $f_{\mathbf{a}}(\mathbf{r})$ .

Cette propriété permet de « simplifier » la fonction de Green : comme  $\delta(\mathbf{r} - \mathbf{r}')$  est aussi  $\delta(\mathbf{r})$  translaté de  $\mathbf{r}'$ ,  $G(\mathbf{r}, \mathbf{r}')$  est égal à la solution  $G(\mathbf{r}, 0)$  translatée de  $\mathbf{r}'$ , c'est à dire à  $G(\mathbf{r} - \mathbf{r}', 0)$ , c'est à dire

$$G(\mathbf{r}, \mathbf{r}') = G(\mathbf{r} - \mathbf{r}', 0) = G_0(\mathbf{r} - \mathbf{r}'), \quad (1.13)$$

où l'on a défini  $G_0(\mathbf{r})$ , qui est solution de

$$\mathcal{L}G_0(\mathbf{r}) = \delta(\mathbf{r}). \quad (1.14)$$

La fonction de Green d'un problème invariant par translation est donc plus simple que celle d'un problème qui n'est pas invariant, car elle ne dépend que d'une variable au lieu de deux. Dans la suite, on omet l'indice 0 ; quand la fonction de Green est écrite avec une seule variable, il est sous-entendu que c'est celle d'un problème invariant par translation.

La solution de  $\mathcal{L}u(\mathbf{r}) = f(\mathbf{r})$  peut alors s'écrire avec un produit de convolution :

$$u(\mathbf{r}) = \int G(\mathbf{r} - \mathbf{r}') f(\mathbf{r}') d\mathbf{r}' = G * f(\mathbf{r}). \quad (1.15)$$

On montre par exemple que la fonction de Green de l'équation de Poisson, Éq. (1.3), est donnée en dimension  $d = 3$  par  $G(\mathbf{r}) = -1/(4\pi r)$  (voir Sec. 1.5.2.1), donc la solution à l'équation de Poisson est donnée par

$$V(\mathbf{r}) = -\frac{1}{4\pi} \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'. \quad (1.16)$$

On peut voir ce potentiel comme la superposition des potentiels créés par des charges ponctuelles avec une densité  $\rho(\mathbf{r})$ .

### 1.3.4 Méthode des images

La fonction de Green d'un opérateur différentiel peut être utilisée pour certaines EDP à coefficients constants dans un milieu semi-infini, typiquement un demi-espace, si la condition au bord présente de bonnes propriétés de symétrie. Par exemple, considérons l'équation  $\mathcal{L}u(\mathbf{r}) = f(\mathbf{r})$  sur  $\Omega = \mathbb{R}^{d-1} \times \mathbb{R}_+$ , et notons  $\mathbf{r} = (\mathbf{r}_{\parallel}, z)$ . Il faut de plus que l'opérateur différentiel ne contienne que des dérivées d'ordre pair en  $z$ , ce qui implique que si on « retourne » la fonction  $u(\mathbf{r})$  en introduisant  $\bar{u}(\mathbf{r}) = u(\bar{\mathbf{r}}) = u(\mathbf{r}_{\parallel}, -z)$ , alors  $\mathcal{L}\bar{u}(\mathbf{r}) = \mathcal{L}u(\mathbf{r})$ . Cela implique que la fonction de Green est symétrique,  $G(\bar{\mathbf{r}}) = G(\mathbf{r})$ . On considère deux conditions aux limites sur le bord  $\partial\Omega = \mathbb{R}^{d-1} \times \{0\}$  :  $u(\mathbf{r}) = 0$  ou  $\partial_z u(\mathbf{r}) = 0$ .

L'idée de la *méthode des images* est de considérer le problème sur l'espace infini  $\mathbb{R}^d$  avec une source  $\hat{f}(\mathbf{r})$  choisie telle que la condition au bord soit satisfaite par symétrie. Par exemple, si la condition au bord est  $u(\mathbf{r}) = 0$ , on prendra pour la source  $\hat{f}(\mathbf{r}) = f(\mathbf{r})$  si  $\mathbf{r} \in \Omega$ , et  $\hat{f}(\mathbf{r}) = -f(\bar{\mathbf{r}})$  si  $\mathbf{r} \in \bar{\Omega}$  ; ce deuxième cas constitue la source image. On résout alors  $\mathcal{L}\hat{u}(\mathbf{r}) = \hat{f}(\mathbf{r})$  :  $\hat{u}(\mathbf{r}) = G * \hat{f}(\mathbf{r})$ . On peut alors montrer que



$\hat{u}(\mathbf{r}_{\parallel}, 0) = 0$ , ce qui montre que la restriction de  $\hat{u}(\mathbf{r})$  à  $\Omega$  est la solution recherchée :

$$\hat{u}(\mathbf{r}_{\parallel}, z) = G * \hat{f}(\mathbf{r}_{\parallel}, z) = \int G(\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel}, z - z') \hat{f}(\mathbf{r}'_{\parallel}, z') d\mathbf{r}'_{\parallel} dz' \quad (1.17)$$

$$= \int_{z' > 0} G(\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel}, z - z') f(\mathbf{r}'_{\parallel}, z') d\mathbf{r}'_{\parallel} dz' - \int_{z' < 0} G(\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel}, z - z') f(\mathbf{r}'_{\parallel}, -z') d\mathbf{r}'_{\parallel} dz' \quad (1.18)$$

$$= \int_{z' > 0} \left[ G(\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel}, z - z') - G(\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel}, z + z') \right] f(\mathbf{r}'_{\parallel}, z') d\mathbf{r}'_{\parallel} dz' \quad (1.19)$$

$$= \int_{z' > 0} \left[ G(\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel}, z - z') - G(\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel}, -z - z') \right] f(\mathbf{r}'_{\parallel}, z') d\mathbf{r}'_{\parallel} dz'. \quad (1.20)$$

Le terme entre crochets est nul si  $z = 0$ .

La source image est à adapter à la condition au bord. Par exemple, pour  $\partial_z u(\mathbf{r}) = 0$ , il faut prendre une source symétrique :  $\hat{f}(\mathbf{r}) = f(\bar{\mathbf{r}})$  si  $\mathbf{r} \in \bar{\Omega}$ .

## 1.4 EDP linéaires du 1<sup>er</sup> ordre, méthode des caractéristiques

### 1.4.1 Méthode des caractéristiques

Nous allons voir que les EDP linéaires du premier ordre peuvent être résolues « facilement » en les transformant en équations différentielles ordinaires (EDO) par la *méthode des caractéristiques*. Par définition, une EDP linéaire du premier ordre est de la forme

$$\mathcal{L}u(\mathbf{r}) = \sum_{i=1}^d a_i(\mathbf{r}) \partial_i u(\mathbf{r}) + b(\mathbf{r}) u(\mathbf{r}) = f(\mathbf{r}). \quad (1.21)$$

En regroupant les coefficients  $a_i(\mathbf{r})$  dans un vecteur  $\mathbf{v}(\mathbf{r})$ , on peut réécrire cette équation comme

$$\mathbf{v}(\mathbf{r}) \cdot \nabla u(\mathbf{r}) + b(\mathbf{r}) u(\mathbf{r}) = f(\mathbf{r}). \quad (1.22)$$

L'idée de la méthode des caractéristiques est de suivre les lignes de courant du champ de vecteur  $\mathbf{v}(\mathbf{r})$ , car tout se passe le long de ces lignes.

Une ligne de courant du champ de vecteurs  $\mathbf{v}(\mathbf{r})$  est une courbe  $\hat{\mathbf{r}}(s)$  telle que  $\hat{\mathbf{r}}'(s) = \lambda(s) \mathbf{v}(\hat{\mathbf{r}}(s))$ , où  $\lambda(s)$  est une fonction et pour toute fonction  $g(s)$ ,  $g'(s)$  désigne sa dérivée par rapport à  $s$ . Nous prenons ici  $\lambda(s) = 1$ . Regardons ce qui se passe le long d'une ligne de courant : commençons par définir  $\hat{u}(s) = u(\hat{\mathbf{r}}(s))$ , alors

$$\hat{u}'(s) = \hat{\mathbf{r}}'(s) \cdot \nabla u(\hat{\mathbf{r}}(s)) = \mathbf{v}(\hat{\mathbf{r}}(s)) \cdot \nabla u(\hat{\mathbf{r}}(s)) = -b(\hat{\mathbf{r}}(s)) u(\hat{\mathbf{r}}(s)) + f(\hat{\mathbf{r}}(s)) = -\hat{b}(s) \hat{u}(s) + \hat{f}(s). \quad (1.23)$$

$\hat{u}(s)$  suit donc une EDO, qui est beaucoup plus facile à résoudre qu'une EDP. Pour déterminer complètement  $\hat{u}(s)$  il faut enfin connaître sa valeur en un point de la caractéristique. Pour pouvoir appliquer la méthode des caractéristiques et déterminer  $u(\mathbf{r})$ , il faut donc que la valeur de  $u(\mathbf{r})$  soit connue sur une surface  $\Gamma$  qui coupe toutes les caractéristiques une seule fois.

En pratique, la méthode des caractéristiques s'utilise de la façon suivante :

- À partir du champ vectoriel  $\mathbf{v}(\mathbf{r})$ , on détermine les caractéristiques  $\hat{\mathbf{r}}(s, \mathbf{r}_{\perp})$ , où  $s$  est la coordonnée le long de la caractéristique et  $\mathbf{r}_{\perp}$  une coordonnée dans la « direction transverse », qui indique sur quelle caractéristique on se trouve. On cherche la valeur de  $s$ , notée  $s^*(\mathbf{r}_{\perp})$ , pour laquelle chaque caractéristique coupe la surface  $\Gamma$  sur laquelle est donnée la condition aux limites :  $\hat{\mathbf{r}}(s^*(\mathbf{r}_{\perp}), \mathbf{r}_{\perp}) \in \Gamma$ .
- On résout l'EDO associée à l'EDP sur chaque caractéristique, ce qui donne  $\hat{u}(s, \mathbf{r}_{\perp})$ .
- On inverse l'équation des caractéristiques afin de retrouver, pour un point  $\mathbf{r}$ , à quelle caractéristique il appartient et où il se trouve sur cette caractéristique. Cela donne deux fonctions  $\tilde{\mathbf{r}}_{\perp}(\mathbf{r})$  et  $\tilde{s}(\mathbf{r})$ . Cela permet de déterminer la fonction  $u(\mathbf{r})$  :  $u(\mathbf{r}) = \hat{u}(\tilde{s}(\mathbf{r}), \tilde{\mathbf{r}}_{\perp}(\mathbf{r}))$ .

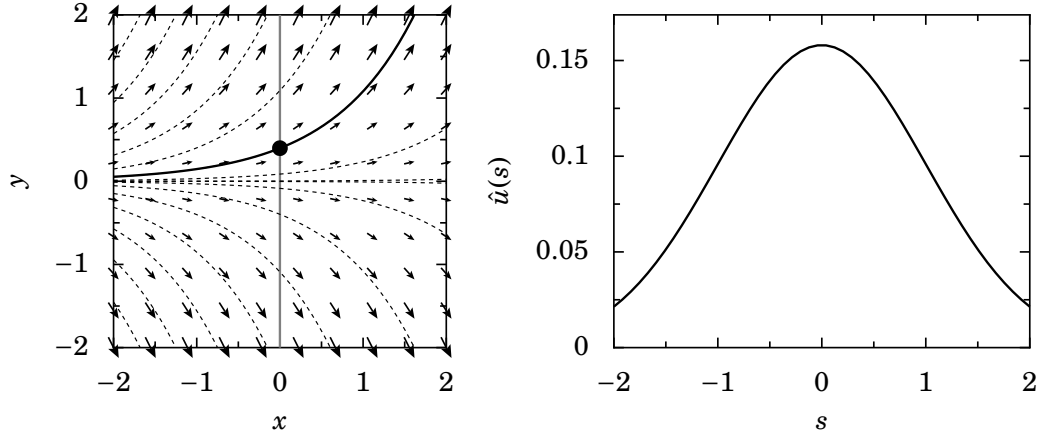


FIGURE 1.1 – Illustration de la méthode des caractéristiques. *Gauche* : champ de vecteurs défini par  $\mathbf{v}(\mathbf{r})$ , caractéristiques (lignes en pointillés et ligne continue), et surface  $\Gamma$  (ligne grise) où est définie la condition aux limites. *Droite* : valeur de  $u(\mathbf{r})$  le long de la caractéristique représentée par la ligne continue à gauche.

### 1.4.2 Exemple

On veut résoudre l'EDP suivante sur  $\mathbb{R}^2$  :

$$\partial_x u(x, y) + y \partial_y u(x, y) + x u(x, y) = 0, \quad (1.24)$$

avec la condition

$$u(0, y) = y^2 \quad \forall y \in \mathbb{R}. \quad (1.25)$$

La première étape est de déterminer les caractéristiques. Le champ de vecteurs est ici  $\mathbf{v}(x, y) = (1, y)$  (Fig. 1.1), il faut donc résoudre

$$\hat{x}'(s) = 1, \quad (1.26)$$

$$\hat{y}'(s) = \hat{y}(s). \quad (1.27)$$

On peut résoudre ce système avec

$$\hat{x}(s) = s, \quad (1.28)$$

$$\hat{y}(s) = y_0 e^s. \quad (1.29)$$

C'est ici  $y_0$  qui joue le rôle de la variable notée  $\mathbf{r}_\perp$  dans le cas général et qui indique sur quelle caractéristique on se trouve. Nous aurions pu prendre une constante d'intégration pour  $\hat{x}(s)$ , mais cela serait revenu à paramétrer différemment la même caractéristique. La valeur de  $u(x, y)$  est connue sur la surface  $\Gamma = \{(0, y), y \in \mathbb{R}\}$ . La caractéristique indexée par  $y_0$  coupe  $\Gamma$  en  $s = s^*(y_0) = 0$ .

Écrivons l'EDP satisfaite par  $\hat{u}(s, y_0)$  :

$$\hat{u}'(s, y_0) = \hat{x}'(s) \partial_x u(\hat{x}(s), \hat{y}(s)) + \hat{y}'(s) \partial_y u(\hat{x}(s), \hat{y}(s)) \quad (1.30)$$

$$= \partial_x u(\hat{x}(s), \hat{y}(s)) + \hat{y}(s) \partial_y u(\hat{x}(s), \hat{y}(s)) \quad (1.31)$$

$$= -\hat{x}(s) u(\hat{x}(s), \hat{y}(s)) \quad (1.32)$$

$$= -s \hat{u}(s, y_0). \quad (1.33)$$

La solution à cette équation est

$$\hat{u}(s, y_0) = \hat{u}(0, y_0)e^{-s^2/2}. \quad (1.34)$$

On peut maintenant utiliser la condition aux limites :

$$\hat{u}(s^*(y_0), y_0) = \hat{u}(0, y_0) \quad (1.35)$$

d'une part, et d'autre part

$$\hat{u}(s^*(y_0), y_0) = u(0, y_0) = y_0^2. \quad (1.36)$$

Finalement, on a donc

$$\hat{u}(s, y_0) = y_0^2 e^{-s^2/2}. \quad (1.37)$$

La dernière étape est d'inverser l'équation des caractéristiques pour savoir sur quelle caractéristique et à quelle endroit le long de cette caractéristique appartient un point  $(x, y) \in \mathbb{R}^2$ , on doit donc résoudre

$$x = \hat{x}(s) = s, \quad (1.38)$$

$$y = \hat{y}(s) = y_0 e^s. \quad (1.39)$$

On obtient donc

$$s = x, \quad (1.40)$$

$$y_0 = ye^{-x}. \quad (1.41)$$

On a ainsi la solution

$$u(x, y) = \hat{u}(x, ye^{-x}) = (ye^{-x})^2 e^{-x^2/2} = y^2 \exp\left(-\frac{x^2}{2} - 2x\right). \quad (1.42)$$

**Méthode de séparation des variables.** Le résultat final, de la forme  $u(x, y) = v(x)w(y)$ , suggère que ce problème aurait aussi pu être résolu avec la méthode de *séparation des variables*, très utilisée pour les EDP ; regardons comment la mettre en œuvre ici.

Cette méthode consiste à chercher la solution sous la forme  $u(x, y) = v(x)w(y)$ . En insérant cette décomposition dans l'équation (1.24), on trouve

$$v'(x)w(y) + yv(x)w'(y) + xv(x)w(y) = 0, \quad (1.43)$$

que l'on peut réécrire

$$[v'(x) + xv(x)]w(y) + v(x)yw'(y) = 0. \quad (1.44)$$

Si  $v(x)$  et  $w(y)$  sont non nuls, on peut écrire

$$\frac{v'(x) + xv(x)}{v(x)} + \frac{yw'(y)}{w(y)} = 0. \quad (1.45)$$

Le premier terme est indépendant de  $y$  et le deuxième est indépendant de  $x$ , comme ils sont opposés, ils sont tous les deux indépendants de  $x$  et  $y$ , c'est à dire qu'ils sont constants. On peut donc écrire

$$v'(x) + xv(x) = -Cv(x), \quad (1.46)$$

$$yw'(y) = Cw(y). \quad (1.47)$$

On obtient donc deux EDO qui ne sont couplées que par la valeur de  $C$ . La résolution de ces équations donne

$$v(x) = A \exp\left(-\frac{x^2}{2} - Cx\right), \quad (1.48)$$

$$w(y) = y^C. \quad (1.49)$$

Avec la condition  $u(0, y) = v(0)w(y) = y^2$ , on a  $A = 1$ ,  $C = 2$ , et on retrouve le résultat (1.42).

### 1.4.3 Résolution de l'équation de transport

Nous utilisons la méthode des caractéristiques pour résoudre l'équation de transport (1.2) dans un champ de vitesse uniforme en dimension d'espace 1 :

$$\partial_t u(x, t) + v \partial_x u(x, t) = 0. \quad (1.50)$$

#### 1.4.3.1 Résolution avec condition initiale

Commençons par résoudre cette équation avec des conditions initiales :

$$u(x, 0) = u_0(x). \quad (1.51)$$

La première étape est de déterminer les caractéristiques. Une caractéristique  $(\hat{x}(s), \hat{t}(s))$  vérifie  $\hat{t}'(s) = 1$ ,  $\hat{x}'(s) = v$ , on peut donc prendre  $\hat{t}(s) = s$ ,  $\hat{x}(s) = vs + x_0$ , où  $x_0$  est la coordonnée qui indique sur quelle caractéristique on se trouve. Ici les caractéristiques ont une forme très simple : ce sont des droites. Chaque caractéristique coupe la surface  $\Gamma$  où  $u(x, t)$  est donnée en  $s = 0$ .

Sur une caractéristique,  $\hat{u}(s, x_0) = u(\hat{x}(s), \hat{t}(s)) = u(vs + x_0, s)$  vérifie  $\hat{u}'(s, x_0) = 0$ , soit  $\hat{u}(s, x_0) = \hat{u}(0, x_0)$ . En utilisant la condition initiale, qui correspond à  $s = 0$ , on a  $\hat{u}(0, x_0) = u(x_0, 0) = u_0(x_0)$ . Finalement,  $\hat{u}(s, x_0) = u_0(x_0)$ .

Enfin, il faut inverser l'équation des caractéristiques, c'est à dire trouver  $s$  et  $x_0$  tels que  $t = s$  et  $x = vt + x_0$ ; cela donne  $s = t$  et  $x_0 = x - vt$ . On a donc

$$u(x, t) = u_0(x - vt). \quad (1.52)$$

#### Remarques :

- Le résultat (1.52) se généralise simplement en toute dimension d'espace et donne  $u(\mathbf{r}, t) = u_0(\mathbf{r} - \mathbf{v}t)$ .
- On a l'impression que la solution de l'équation de transport doit être dérivable dans les deux directions; cependant, on peut injecter n'importe quelle condition initiale dans la solution (1.52). On peut en fait montrer que l'équation de transport est toujours satisfaite au sens des distributions. Plus généralement, la question de la régularité des solutions, et de l'espace de fonctions dans lequel chercher des solutions est une problématique très importante des EDP que nous n'aborderons pas dans ce cours.

#### 1.4.3.2 Fonction de Green

On cherche maintenant la fonction de Green de l'équation de transport, c'est à dire la solution de

$$\partial_t G(x, t) + v \partial_x G(x, t) = \delta(x) \delta(t). \quad (1.53)$$

Nous allons d'abord transformer ce problème en problème avec condition initiale. Pour cela, on part de la solution  $g(x, t)$  de l'équation homogène, définie sur  $\mathbb{R} \times \mathbb{R}$ , avec la condition initiale  $g(x, 0) = \delta(x)$ ; ici,  $g(x, t)$  est donc donnée par  $g(x, t) = \delta(x - vt)$ . On pose ensuite

$$G(x, t) = H(t)g(x, t), \quad (1.54)$$

où  $H(t)$  est la fonction de Heaviside. Dérivons maintenant  $G(x, t)$  par rapport au temps :

$$\partial_t G(x, t) = [\partial_t H(t)]g(x, t) + H(t)\partial_t g(x, t) \quad (1.55)$$

$$= \delta(t)g(x, t) - vH(t)\partial_x g(x, t) \quad (1.56)$$

$$= \delta(t)g(x, 0) - v\partial_x G(x, t) \quad (1.57)$$

$$= \delta(t)\delta(x) - v\partial_x G(x, t), \quad (1.58)$$

donc  $G(x, t)$  est bien la fonction de Green de l'équation de transport, explicitement :

$$G(x, t) = H(t)\delta(x - vt). \quad (1.59)$$

#### 1.4.4 Remarques

- L'équation des caractéristiques n'est pas toujours facile à intégrer analytiquement, mais cela reste des EDO, plus faciles à intégrer qu'une EDP (par exemple numériquement).
- Certains problèmes peuvent survenir, par exemple si les caractéristiques se referment en des cercles. Dans ces cas, il peut ne pas exister de solution.

## 1.5 EDP linéaires du 2<sup>nd</sup> ordre à coefficients constants

### 1.5.1 Classification

Nous allons montrer que l'on peut « classer » les EDP linéaires d'ordre 2 à coefficients constants en s'inspirant de la classification des coniques. La classification a pour but de réduire le nombre d'EDP à étudier pour comprendre leur comportement.

Une telle EDP peut se mettre sous la forme

$$\sum_{i,j=1}^d a_{i,j}\partial_i\partial_j u(\mathbf{r}) + \sum_{i=1}^d b_i\partial_i u(\mathbf{r}) + cu(\mathbf{r}) = f(\mathbf{r}). \quad (1.60)$$

Comme on peut intervertir les dérivées partielles pour une fonction  $u(\mathbf{r})$  suffisamment régulière ( $\mathcal{C}^2(\mathbb{R}^d)$ ),  $\partial_i\partial_j u(\mathbf{r}) = \partial_j\partial_i u(\mathbf{r})$ , on peut toujours se retrouver avec une matrice  $a_{i,j}$  symétrique. La matrice  $a_{i,j}$  est donc diagonalisable en base orthonormée ; si on utilise cette nouvelle base pour les coordonnées de l'espace, l'EDP devient

$$\sum_i \lambda_i \partial_i^2 u(\mathbf{r}) + \sum_{i=1}^d \tilde{b}_i \partial_i u(\mathbf{r}) + cu(\mathbf{r}) = \tilde{f}(\mathbf{r}), \quad (1.61)$$

où  $\tilde{b}_i$  sont de nouveaux coefficients et  $\tilde{f}(\mathbf{r})$  un nouveau terme source.

Pour les directions dans lesquelles  $\lambda_i \neq 0$ , on peut se débarrasser des dérivées d'ordre 1 en posant :

$$u(\mathbf{r}) = \exp\left(-\sum_{i=1, \lambda_i \neq 0}^d \frac{\tilde{b}_i x_i}{2\lambda_i}\right) v(\mathbf{r}). \quad (1.62)$$

Alors la fonction  $v(\mathbf{r})$  est solution de

$$\sum_{i=1}^d \lambda_i \partial_i^2 v(\mathbf{r}) + \sum_{i=1, \lambda_i = 0}^d \tilde{b}_i \partial_i v(\mathbf{r}) + \left(c - \sum_{i=1, \lambda_i \neq 0}^d \frac{\tilde{b}_i^2}{4\lambda_i}\right) v(\mathbf{r}) = \exp\left(\sum_{i=1, \lambda_i \neq 0}^d \frac{\tilde{b}_i x_i}{2\lambda_i}\right) f(\mathbf{r}). \quad (1.63)$$

Cela montre que dans les directions où il y a une dérivée seconde, on peut se débarrasser de la dérivée première par un changement de variable.

Enfin, dans les directions où  $\lambda_i \neq 0$ , on peut changer d'échelle sur la coordonnée  $i$  pour ne conserver que le signe de  $\lambda_i$ ,  $\sigma(\lambda_i)$ ; la valeur absolue de  $\lambda_i$  n'a donc aucune importance.

Suivant les signes des valeurs propres  $\lambda_i$ , on distingue trois types d'EDP :

- Si tous les  $\lambda_i$  sont strictement positifs, l'EDP est *elliptique*.
- Si un  $\lambda_i$  est nul et tous les autres sont strictement positifs, l'EDP est *parabolique*.
- Si un  $\lambda_i$  est strictement négatif et tous les autres sont strictement positifs, l'EDP est *hyperbolique*.

Ces trois types d'EDP recouvrent les équations de Poisson (Éq. (1.3), elliptique), de la chaleur (Éq. (1.4), parabolique) et des ondes (Éq. (1.5), hyperbolique). Comme nous allons le voir, ces équations ont des propriétés très différentes.

#### Remarques :

- Les définitions ci-dessus sont données au signe près vu qu'on peut mettre un signe moins global devant l'équation.
- Ces trois classes ne recouvrent toutes les EDP du 2<sup>nd</sup> ordre à coefficients constants qu'en dimension  $d = 2$ ; en dimension supérieure il existe d'autres classes.

### 1.5.2 Équation de Poisson

On rappelle ici l'équation de Poisson, que l'on écrit sans terme source :

$$\nabla^2 u(\mathbf{r}) = 0. \quad (1.64)$$

#### 1.5.2.1 Fonction de Green

Calculons d'abord la fonction de Green en dimension  $d = 1$ , ce qui est facile, avant de calculer la fonction de Green en dimension quelconque. En dimension  $d = 1$ , il faut résoudre  $G''(x) = \delta(x)$ . Une première intégration donne  $G'(x) = H(x) + a$ , et une seconde intégration donne  $G(x) = \frac{|x|}{2} + (a + \frac{1}{2})x + b$ . L'ensemble des fonctions affines est l'ensemble des fonction harmoniques (de dérivée seconde nulle). On choisit la forme « la plus symétrique » comme fonction de Green :

$$G_1(x) = \frac{|x|}{2}. \quad (1.65)$$

On remarque qu'on ne peut pas imposer que  $G_1(x)$  s'annule à l'infini.

Le calcul en dimension quelconque peut se faire de façon générale. Le problème étant à symétrie radiale, on cherche une solution à symétrie radiale,  $G(\mathbf{r}) = f(r)$  telle que  $\nabla^2 G(\mathbf{r})$  s'annule pour tout  $r > 0$ . Le laplacien vaut

$$\nabla^2 f(r) = \nabla \cdot [\nabla f(r)] = \nabla \cdot [\mathbf{e}_r f'(r)] = (\nabla \cdot \mathbf{e}_r) f'(r) + f''(r) = \frac{d-1}{r} f'(r) + f''(r). \quad (1.66)$$

Nous avons utilisé que  $\nabla \cdot \mathbf{e}_r = (d-1)/r$ . Cela peut se montrer en calculant  $\nabla \cdot \mathbf{r}$  : d'une part,  $\nabla \cdot \mathbf{r} = d$ , et d'autre part,  $\nabla \cdot \mathbf{r} = \nabla \cdot (r \mathbf{e}_r) = (\nabla r) \cdot \mathbf{e}_r + r \nabla \cdot \mathbf{e}_r = 1 + r \nabla \cdot \mathbf{e}_r$ .

Cherchons  $f(r)$  sous la forme  $Ar^{-\alpha}$ , on doit donc avoir d'après l'équation (1.66) :

$$(d-1)\alpha = \alpha(\alpha+1). \quad (1.67)$$

On peut avoir  $\alpha = 0$  ou  $\alpha = d-2$ . Pour  $d = 1$  on retrouve  $\alpha = -1$ , qui correspond à la solution (1.65). Pour  $d = 2$ , il faut  $\alpha = 0$ , ce qui correspond à une constante ou un logarithme; la constante n'a pas la

nécessaire singularité en 0, on choisit donc le logarithme. En effet,  $f(r) = A \log(r)$  satisfait bien  $f''(r) + f'(r)/r = 0$ . On prend donc :  $f(r) = A g_d(r)$  avec

$$g_2(r) = \log(r) \quad (1.68)$$

$$g_{d \neq 2}(r) = r^{2-d}. \quad (1.69)$$

Ces fonctions ont une singularité en 0 ; il faut montrer qu'elle permet de retrouver une distribution en Dirac et calculer le facteur  $A$ . Il faut faire le calcul au sens des distributions. Prenons une fonction test  $\phi(\mathbf{r}) \in \mathcal{D}(\mathbb{R}^d)$ , et calculons

$$\langle \nabla^2 G(\mathbf{r}), \phi(\mathbf{r}) \rangle = \langle G(\mathbf{r}), \nabla^2 \phi(\mathbf{r}) \rangle \quad (1.70)$$

$$= A \int_{\mathbb{R}^d} \nabla^2 \phi(\mathbf{r}) g_d(r) d\mathbf{r} \quad (1.71)$$

$$= A \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^d \setminus \mathcal{B}_\epsilon} \nabla^2 \phi(\mathbf{r}) g_d(r) d\mathbf{r}, \quad (1.72)$$

où  $\mathcal{B}_\epsilon$  est la boule de rayon  $\epsilon$  centrée en 0 ; cette égalité est vraie car ce qui est dans l'intégrale est localement intégrable en 0. On peut ensuite procéder à l'intégration par parties, en utilisant le théorème de Stokes (ou de Green-Ostrogradski) :

$$\int_{\Omega} \nabla \cdot \mathbf{A}(\mathbf{r}) d\mathbf{r} = \int_{\partial\Omega} \mathbf{n}(\mathbf{r}) \cdot \mathbf{A}(\mathbf{r}) d\mathbf{r}, \quad (1.73)$$

où  $\Omega$  est un domaine de  $\mathbb{R}^d$ ,  $\partial\Omega$  son bord et  $\mathbf{A}(\mathbf{r})$  est un champ vectoriel sur ce domaine, et  $\mathbf{n}(\mathbf{r})$  est le vecteur unitaire normal au bord, dirigé vers l'extérieur de  $\Omega$ . En appliquant cette relation à  $\mathbf{A}(\mathbf{r}) = g_d(r) \nabla \phi(\mathbf{r})$ , on obtient :

$$\int_{\mathbb{R}^d \setminus \mathcal{B}_\epsilon} [g_d(r) \nabla^2 \phi(\mathbf{r}) + g'_d(r) \mathbf{e}_r \cdot \nabla \phi(\mathbf{r})] d\mathbf{r} = - \int_{\mathcal{S}_\epsilon} g_d(r) \mathbf{e}_r \cdot \nabla \phi(\mathbf{r}) d\mathbf{r}, \quad (1.74)$$

où  $\mathcal{S}_\epsilon$  est la sphère de rayon  $\epsilon$  et  $\mathbf{n}(\mathbf{r}) = -\mathbf{e}_r$  ; on a donc

$$\int_{\mathbb{R}^d \setminus \mathcal{B}_\epsilon} g_d(r) \nabla^2 \phi(\mathbf{r}) d\mathbf{r} = - \int_{\mathcal{S}_\epsilon} g_d(r) \mathbf{e}_r \cdot \nabla \phi(\mathbf{r}) d\mathbf{r} - \int_{\mathbb{R}^d \setminus \mathcal{B}_\epsilon} g'_d(r) \mathbf{e}_r \cdot \nabla \phi(\mathbf{r}) d\mathbf{r}. \quad (1.75)$$

Réutilisons la relation générale pour  $\mathbf{A}(\mathbf{r}) = \mathbf{e}_r g'_d(r) \phi(\mathbf{r})$  :

$$\int_{\mathbb{R}^d \setminus \mathcal{B}_\epsilon} [g'_d(r) \mathbf{e}_r \cdot \nabla \phi(\mathbf{r}) + \phi(\mathbf{r}) \nabla \cdot [g'_d(r) \mathbf{e}_r]] d\mathbf{r} = - \int_{\mathcal{S}_\epsilon} g'_d(r) \phi(\mathbf{r}) d\mathbf{r}. \quad (1.76)$$

Or  $\nabla \cdot [g'_d(r) \mathbf{e}_r] = \nabla^2 [g_d(r)] = 0$  par construction de  $g_d(r)$  pour  $r > 0$ , l'équation (1.76) devient donc

$$\int_{\mathbb{R}^d \setminus \mathcal{B}_\epsilon} g'_d(r) \mathbf{e}_r \cdot \nabla \phi(\mathbf{r}) d\mathbf{r} = -g'_d(\epsilon) \int_{\mathcal{S}_\epsilon} \phi(\mathbf{r}) d\mathbf{r}. \quad (1.77)$$

En utilisant cette relation avec la première intégration par parties, Éq. (1.75), on obtient

$$\int_{\mathbb{R}^d \setminus \mathcal{B}_\epsilon} g_d(r) \nabla^2 \phi(\mathbf{r}) d\mathbf{r} = -g_d(\epsilon) \int_{\mathcal{S}_\epsilon} \mathbf{e}_r \cdot \nabla \phi(\mathbf{r}) d\mathbf{r} + g'_d(\epsilon) \int_{\mathcal{S}_\epsilon} \phi(\mathbf{r}) d\mathbf{r}. \quad (1.78)$$

Les intégrales sur la sphère de rayon  $\epsilon$  sont d'ordre  $\epsilon^{d-1}$  (car l'aire de la sphère de dimension  $d-1$  et de rayon  $\epsilon$  est  $S_{d-1} \epsilon^{d-1}$ , et les fonctions à intégrer sont bornées),  $g_d(\epsilon) \propto \epsilon^{2-d}$  et  $g'_d(\epsilon) \propto \epsilon^{1-d}$ , ainsi seul le deuxième terme survit à la limite  $\epsilon \rightarrow 0$ , et par continuité

$$\lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}^d \setminus \mathcal{B}_\epsilon} g_d(r) \nabla^2 \phi(\mathbf{r}) d\mathbf{r} = S_{d-1} g'_d(1) \phi(0). \quad (1.79)$$

Nous avons utilisé que  $g'_d(\epsilon) = \epsilon^{1-d} g'_d(1)$ . On a donc

$$\langle \nabla^2 G(\mathbf{r}), \phi(\mathbf{r}) \rangle = A S_{d-1} g'_d(1) \phi(0). \quad (1.80)$$

En prenant  $A = [S_{d-1} g'_d(1)]^{-1}$ , on a donc  $\nabla^2 G(\mathbf{r}) = \delta(\mathbf{r})$ .

On trouve donc, pour  $1 \leq d \leq 4$  :

$$G_1(r) = \frac{r}{2}, \quad (1.81)$$

$$G_2(r) = \frac{\log(r)}{2\pi}, \quad (1.82)$$

$$G_3(r) = -\frac{1}{4\pi r}, \quad (1.83)$$

$$G_4(r) = -\frac{1}{4\pi^2 r^2}. \quad (1.84)$$

### 1.5.2.2 Principe du maximum

On montre ici le principe du maximum : si  $\nabla^2 u(\mathbf{r}) = 0$  sur un domaine  $\Omega$ , alors  $u(\mathbf{r})$  ne peut atteindre son maximum que sur le bord de  $\Omega$ , ou  $u(\mathbf{r})$  est constante sur chaque composante connexe de  $\Omega$ .

On commence par montrer le lemme suivant : soit  $\mathbf{r}_0 \in \Omega$  et  $R > 0$  tel que la boule de centre  $\mathbf{r}_0$  et de rayon  $R$ ,  $\mathcal{B}(\mathbf{r}_0, R)$ , soit dans  $\Omega$ , alors

$$u(\mathbf{r}_0) = \frac{1}{|\mathcal{S}(\mathbf{r}_0, R)|} \int_{\mathcal{S}(\mathbf{r}_0, R)} u(\mathbf{r}) d\mathbf{r}, \quad (1.85)$$

où  $\mathcal{S}(\mathbf{r}_0, R)$  est la sphère de centre  $\mathbf{r}_0$  et de rayon  $R$ . Pour démontrer cette propriété, posons

$$\phi(R) = \frac{1}{|\mathcal{S}(\mathbf{r}_0, R)|} \int_{\mathcal{S}(\mathbf{r}_0, R)} u(\mathbf{r}) d\mathbf{r}. \quad (1.86)$$

Nous allons montrer que  $\phi(R)$  est constante. Par un changement de variables,  $\mathbf{r} = \mathbf{r}_0 + R\mathbf{s}$ ,  $\mathbf{s} \in \mathcal{S}(0, 1)$ , on peut écrire

$$\phi(R) = \frac{1}{|\mathcal{S}(0, 1)|} \int_{\mathcal{S}(0, 1)} u(\mathbf{r}_0 + R\mathbf{s}) d\mathbf{s}. \quad (1.87)$$

La dérivée de  $\phi(R)$  vaut

$$\phi'(R) = \frac{1}{|\mathcal{S}(0, 1)|} \int_{\mathcal{S}(0, 1)} \mathbf{s} \cdot \nabla u(\mathbf{r}_0 + R\mathbf{s}) d\mathbf{s} \quad (1.88)$$

$$= \frac{1}{|\mathcal{S}(0, R)|} \int_{\mathcal{S}(0, R)} \mathbf{n}(\mathbf{r}) \cdot \nabla u(\mathbf{r}) d\mathbf{r} \quad (1.89)$$

$$= \frac{1}{|\mathcal{S}(0, R)|} \int_{\mathcal{B}(0, R)} \nabla^2 u(\mathbf{r}) d\mathbf{r} \quad (1.90)$$

$$= 0, \quad (1.91)$$

où  $\mathbf{n}(\mathbf{r})$  est le vecteur unitaire normal à  $\mathcal{S}(\mathbf{r}_0, R)$  en  $\mathbf{r}$ , pointant vers l'extérieur. Donc  $\phi(R)$  est constante et par continuité de  $u(\mathbf{r})$ ,  $\lim_{R \rightarrow 0} \phi(R) = u(\mathbf{r}_0)$ ; donc  $\phi(R) = u(\mathbf{r}_0)$  pour tout  $R$  tel que  $\mathcal{B}(\mathbf{r}_0, R) \subset \Omega$ . Enfin, la relation (1.85) est aussi vraie en remplaçant la sphère par la boule (ce qui se voit aisément en intégrant sur les sphères) :

$$u(\mathbf{r}_0) = \frac{1}{|\mathcal{B}(\mathbf{r}_0, R)|} \int_{\mathcal{B}(\mathbf{r}_0, R)} u(\mathbf{r}) d\mathbf{r}, \quad (1.92)$$

Il est maintenant facile de montrer le principe du maximum. Supposons que  $\Omega$  est un ouvert (il ne contient pas son bord) connexe et que  $u(\mathbf{r})$  atteigne son maximum en  $\mathbf{r}_0 \in \Omega$ , et montrons que  $u(\mathbf{r})$  est



uniforme sur  $\Omega$ . Soit  $R > 0$  tel que  $\mathcal{B}(\mathbf{r}_0, R) \subset \Omega$ , alors la relation (1.92) montre que la fonction  $u(\mathbf{r})$  est constante égale à  $u(\mathbf{r}_0)$  sur  $\mathcal{B}(\mathbf{r}_0, R)$ . En « itérant » cette relation, on montre que  $u(\mathbf{r})$  est égale à  $u(\mathbf{r}_0)$  sur  $\Omega$ .

On peut démontrer cette « itération ».  $\Omega$  est un ouvert d'un espace vectoriel normé, donc s'il est connexe il est connexe par arc. Soit  $\mathbf{r}_1 \in \Omega$ , on peut donc relier  $\mathbf{r}_0$  à  $\mathbf{r}_1$  par un arc  $\gamma$ , qui est compact. Cet arc peut être recouvert par des boules ouvertes centrées en tous les points de l'arc, et on peut extraire un sous-recouvrement fini de telles boules. Il suffit donc d'appliquer l'« itération » ci-dessus sur ces boules pour montrer que  $u(\mathbf{r}_0) = u(\mathbf{r}_1)$ .

### 1.5.2.3 Unicité d'une solution

On montre ici l'unicité de l'équation de Poisson avec certaines conditions au bord. On considère l'équation de Poisson avec second membre sur  $\Omega$ ,  $\nabla^2 u(\mathbf{r}) = f(\mathbf{r})$ , avec sur le bord  $\partial\Omega$  une des conditions suivantes :

- Dirichlet :  $u(\mathbf{r}) = g(\mathbf{r})$ ,
- Neumann :  $\mathbf{n}(\mathbf{r}) \cdot \nabla u(\mathbf{r}) = g(\mathbf{r})$ ,
- Mixte :  $\partial\Omega = \partial\Omega_1 \sqcup \partial\Omega_2$  avec une condition de Dirichlet sur  $\partial\Omega_1$  et une condition de Neumann sur  $\partial\Omega_2$ .

Soient  $u_1(\mathbf{r})$  et  $u_2(\mathbf{r})$  deux solutions à ce problème ; montrons que leur différence  $v(\mathbf{r}) = u_1(\mathbf{r}) - u_2(\mathbf{r})$  est nulle.  $v(\mathbf{r})$  est solution du problème homogène associé, donc

$$0 = \int_{\partial\Omega} v(\mathbf{r}) \mathbf{n}(\mathbf{r}) \cdot \nabla v(\mathbf{r}) d\mathbf{r} \quad (1.93)$$

$$= \int_{\Omega} \nabla \cdot [v(\mathbf{r}) \nabla v(\mathbf{r})] d\mathbf{r} \quad (1.94)$$

$$= \int_{\Omega} ([\nabla v(\mathbf{r})]^2 + v(\mathbf{r}) \nabla^2 v(\mathbf{r})) d\mathbf{r} \quad (1.95)$$

$$= \int_{\Omega} [\nabla v(\mathbf{r})]^2 d\mathbf{r}. \quad (1.96)$$

Cela implique que  $\nabla v(\mathbf{r}) = 0$  sur  $\Omega$ , donc que  $v(\mathbf{r})$  est constante. Pour des conditions au bord de Dirichlet ou mixtes,  $v(\mathbf{r}) = 0$  en un point de  $\partial\Omega$  donc  $v(\mathbf{r}) = 0$  partout. Pour des conditions de Neumann,  $v(\mathbf{r})$  est constante : la solution du problème est définie à une constante près.

**Remarque :** Nous avons montré l'unicité de la solution, mais pas l'existence d'une solution. En général, l'existence est beaucoup plus difficile à montrer que l'unicité et on le fait souvent au cas par cas.

## 1.5.3 Équation de la chaleur

### 1.5.3.1 Fonction de Green

Déterminons la fonction de Green de l'équation de la chaleur, qui satisfait

$$\partial_t G(\mathbf{r}, t) - D \nabla^2 G(\mathbf{r}, t) = \delta(\mathbf{r}) \delta(t), \quad (1.97)$$

pour une dimension d'espace  $d$  quelconque. Nous allons pour cela utiliser la transformée de Fourier (TF) : nous allons prendre la TF de cette équation (facile), résoudre l'équation dans l'espace de Fourier (facile) et calculer la TF inverse (plus difficile).

La TF de  $G(\mathbf{r}, t)$  et la TF inverse sont définies par

$$\tilde{G}(\mathbf{k}, \omega) = \int e^{-i(\mathbf{k} \cdot \mathbf{r} + \omega t)} G(\mathbf{r}, t) \frac{d\mathbf{r} dt}{(2\pi)^{(d+1)/2}}, \quad (1.98)$$

$$G(\mathbf{r}, t) = \int e^{i(\mathbf{k} \cdot \mathbf{r} + \omega t)} \tilde{G}(\mathbf{k}, \omega) \frac{d\mathbf{k} d\omega}{(2\pi)^{(d+1)/2}}. \quad (1.99)$$

En prenant la TF de l'équation (1.97) puis en utilisant les propriétés de la TF vis-à-vis de la dérivation, on obtient

$$(i\omega + Dk^2) \tilde{G}(\mathbf{k}) = \frac{1}{(2\pi)^{(d+1)/2}}. \quad (1.100)$$

L'inversion de cette équation est immédiat

$$\tilde{G}(\mathbf{k}) = \frac{1}{(2\pi)^{(d+1)/2}(i\omega + Dk^2)} \quad (1.101)$$

et il faut maintenant calculer la TF inverse pour avoir la solution :

$$G(\mathbf{r}, t) = \int \frac{e^{i(\mathbf{k} \cdot \mathbf{r} + \omega t)}}{i\omega + Dk^2} \frac{d\mathbf{k} d\omega}{(2\pi)^{d+1}} \quad (1.102)$$

L'intégration sur  $\omega$  peut se faire facilement avec le théorème des résidus. La fonction à intégrer est holomorphe sur le plan complexe privé du pôle en  $\omega^* = iDk^2$ , où le résidu est  $-ie^{i\mathbf{k} \cdot \mathbf{r} - Dk^2 t}$ . Pour utiliser le théorème des résidus, on ferme le segment  $[-R, R]$  par un demi-cercle. Quand  $t < 0$ , le demi-cercle doit être de partie imaginaire négative, le contour ne fait donc jamais le tour du pôle et le théorème des résidus donne  $G(\mathbf{r}, t) = 0$ . Quand  $t > 0$ , le demi-cercle doit être de partie imaginaire positive, le contour fait le tour du pôle et l'intégrale donne

$$G(\mathbf{r}, t > 0) = \int 2i\pi \times (-ie^{i\mathbf{k} \cdot \mathbf{r} - Dk^2 t}) \frac{d\mathbf{k}}{(2\pi)^{d+1}} = \int e^{i\mathbf{k} \cdot \mathbf{r} - Dk^2 t} \frac{d\mathbf{k}}{(2\pi)^d}. \quad (1.103)$$

On doit maintenant calculer l'intégrale sur  $\mathbf{k}$ . Il s'agit d'une intégrale gaussienne, qui se calcule facilement. On peut par exemple décomposer l'intégrale selon les directions d'espace :

$$G(\mathbf{r}, t > 0) = \prod_{i=1}^d \int e^{ik_i x_i - Dk_i^2 t} \frac{dk_i}{2\pi} = \prod_{i=1}^d \frac{e^{-x_i^2/(4Dt)}}{\sqrt{4\pi Dt}} = (4\pi Dt)^{-d/2} e^{-r^2/(4Dt)}. \quad (1.104)$$

La solution est donc finalement

$$G(\mathbf{r}, t) = \frac{H(t)}{(4\pi Dt)^{d/2}} \exp\left(-\frac{r^2}{4Dt}\right). \quad (1.105)$$

Réciproquement, on peut montrer que l'expression (1.105) satisfait (1.97). Pour  $t > 0$ , il s'agit simplement d'un calcul pénible de dérivées. Pour déterminer ce qui se passe en  $t = 0$ , on peut remplacer l'équation (1.97) par un problème avec condition initiale, comme en section 1.4.3.2 : il faut montrer que  $\lim_{t \rightarrow 0^+} G(\mathbf{r}, t) = \delta(\mathbf{r})$ . Cela se démontre facilement en remarquant la ressemblance avec une suite de Dirac :

$$G(\mathbf{r}, t > 0) = \frac{1}{\sqrt{2Dt}^d} f\left(\frac{\mathbf{r}}{\sqrt{2Dt}}\right), \quad (1.106)$$

avec

$$f(\mathbf{r}) = (2\pi)^{-d/2} e^{-r^2/2}. \quad (1.107)$$

Une fois écrit sous cette forme, il vient avec le théorème de convergence dominée que

$$\lim_{t \rightarrow 0^+} G(\mathbf{r}, t) = \delta(\mathbf{r}) \int f(\mathbf{r}') d\mathbf{r}' = \delta(\mathbf{r}). \quad (1.108)$$

**Remarque :** Nous avons obtenu la fonction de Green en toute dimension d'espace sans effort particulier ; nous allons voir qu'il n'en va pas toujours ainsi.

### 1.5.3.2 Unicité de la solution

On résout souvent l'équation de la chaleur pour  $(\mathbf{r}, t) \in \Omega \times \mathbb{R}_+$ , avec une condition initiale  $u(\mathbf{r}, 0)$  et des conditions au bord de Neumann ou de Dirichlet. Dans ce cas, on peut montrer l'unicité un peu comme pour l'équation de Poisson. Soit deux solutions  $u_1(\mathbf{r}, t)$  et  $u_2(\mathbf{r}, t)$  ; on introduit leur différence  $v(\mathbf{r}, t) = u_1(\mathbf{r}, t) - u_2(\mathbf{r}, t)$ .

Posons

$$V(t) = \int_{\Omega} v(\mathbf{r}, t)^2 d\mathbf{r}, \quad (1.109)$$

et calculons

$$\dot{V}(t) = 2 \int_{\Omega} v(\mathbf{r}, t) \partial_t v(\mathbf{r}, t) d\mathbf{r} \quad (1.110)$$

$$= 2D \int_{\Omega} v(\mathbf{r}, t) \nabla^2 v(\mathbf{r}, t) d\mathbf{r} \quad (1.111)$$

$$= 2D \left[ \int_{\partial\Omega} v(\mathbf{r}, t) \mathbf{n}(\mathbf{r}, t) \cdot \nabla v(\mathbf{r}, t) d\mathbf{r} - \int_{\Omega} [\nabla v(\mathbf{r}, t)]^2 d\mathbf{r} \right]. \quad (1.112)$$

Le premier terme est nul par hypothèse, le second est positif, donc  $V(t)$  est décroissante. Comme  $V(0) = 0$  et  $V(t) \geq 0$ ,  $V(t) = 0$  pour tout  $t$ , donc  $v(\mathbf{r}, t) = 0$  pour tout  $\mathbf{r}$  et  $t$ .

## 1.5.4 Équation des ondes

### 1.5.4.1 Fonction de Green en dimension $d = 3$

La fonction de Green de l'équation des ondes satisfait

$$c^{-2} \partial_t^2 G(\mathbf{r}, t) - \nabla^2 G(\mathbf{r}, t) = \delta(\mathbf{r}) \delta(t). \quad (1.113)$$

Nous allons la déterminer en dimension 3.

Commençons par prendre la TF de cette équation par rapport au temps, on obtient

$$-\frac{\omega^2}{c^2} \tilde{G}(\mathbf{r}, \omega) - \nabla^2 \tilde{G}(\mathbf{r}, \omega) = \frac{\delta(\mathbf{r})}{\sqrt{2\pi}}. \quad (1.114)$$

Il s'agit de l'équation de Helmholtz. On montre ensuite facilement que la solution à cette équation est, en dimension  $d = 3$ ,

$$\tilde{G}(\mathbf{r}, \omega) = \frac{e^{v i \omega r / c}}{4\pi \sqrt{2\pi} r}, \quad (1.115)$$

avec  $v \in \{-1, 1\}$ . Le laplacien de cette expression peut se prendre de façon usuelle avec les dérivées, en utilisant le résultat de la section 1.5.2.1 pour le laplacien de  $1/r$ .

Il faut ensuite prendre la TF inverse par rapport à  $\omega$ , ce qui est très rapide :

$$G(\mathbf{r}, t) = \int_{-\infty}^{\infty} \frac{e^{v i \omega r / c}}{4\pi \sqrt{2\pi} r} e^{i \omega t} \frac{d\omega}{\sqrt{2\pi}} = \frac{1}{4\pi r} \delta\left(t + \frac{vr}{c}\right). \quad (1.116)$$

Comme  $r$  est positif, cette fonction est nulle pour les temps positifs si  $v = 1$  et nulle pour les temps négatifs si  $v = -1$ . Cette dernière situation correspond à la fonction de Green *causale* ou *retardée* ; c'est celle qu'on utilise le plus, mais il est important de garder en tête qu'il y a plusieurs solutions et qu'il faut faire un choix. La fonction de Green retardée est donc

$$G_+(\mathbf{r}, t) = \frac{1}{4\pi r} \delta\left(t - \frac{r}{c}\right). \quad (1.117)$$

**Remarques :**

- L'équation des ondes présente une vitesse de propagation finie, qui est la vitesse  $c$  entrant dans sa définition. Une perturbation en  $\mathbf{r}$  n'a d'effet en  $\mathbf{r}'$  qu'après un temps  $|\mathbf{r} - \mathbf{r}'|/c$ .
- L'équation de Helmholtz est proche de l'équation de Poisson, et sa solution ressemble à celle de l'équation de Poisson « décorée » par l'exponentielle. Toutefois, un calcul simple montre que cela ne marche qu'en dimension  $d = 3$ .

**1.5.4.2 Solution générale en dimension d'espace  $d = 1$ , forme canonique**

On cherche la forme générale de la solution de l'équation des ondes en dimension  $d = 1$ , sans second membre :

$$\partial_t^2 u(x, t) - c^2 \partial_x^2 u(x, t) = 0. \quad (1.118)$$

On va montrer que la solution est de la forme

$$u(x, t) = f(x - ct) + g(x + ct). \quad (1.119)$$

La façon la plus rapide d'arriver au résultat est d'écrire

$$0 = (\partial_t^2 - c^2 \partial_x^2) u(x, t) = (\partial_t - c \partial_x)(\partial_t + c \partial_x) u(x, t). \quad (1.120)$$

Pour annuler  $\partial_t + c \partial_x$ , il faut que  $u(x, t) = f(x - ct)$  (voir Sec. 1.4.3); pour annuler  $\partial_t - c \partial_x$ , il faut que  $u(x, t) = g(x + ct)$ . Comme il suffit que l'un des deux termes différentiels s'annule,  $u(x, t)$  est une combinaison de ces deux solutions.

Plus rigoureusement, on peut définir  $u(x, t) = \hat{u}(x - ct, x + ct)$ ; on note  $\partial_1 \hat{u}$  et  $\partial_2 \hat{u}$  les dérivées de  $\hat{u}$  par rapport à sa première et à sa deuxième variable, respectivement. On peut donc calculer les dérivées :

$$\partial_t u(x, t) = -c \partial_1 \hat{u}(x - ct, x + ct) + c \partial_2 \hat{u}(x - ct, x + ct), \quad (1.121)$$

$$\partial_t^2 u(x, t) = c^2 (\partial_1^2 - 2 \partial_1 \partial_2 + \partial_2^2) \hat{u}(x - ct, x + ct), \quad (1.122)$$

$$\partial_x u(x, t) = \partial_1 \hat{u}(x - ct, x + ct) + \partial_2 \hat{u}(x - ct, x + ct), \quad (1.123)$$

$$\partial_x^2 u(x, t) = (\partial_1^2 + 2 \partial_1 \partial_2 + \partial_2^2) \hat{u}(x - ct, x + ct). \quad (1.124)$$

Ainsi, l'équation (1.118) s'écrit

$$0 = (\partial_t^2 - c^2 \partial_x^2) u(x, t) = -4c^2 \partial_1 \partial_2 \hat{u}(x - ct, x + ct), \quad (1.125)$$

c'est à dire que  $\hat{u}(p, q)$  vérifie

$$\partial_p \partial_q \hat{u}(p, q) = 0. \quad (1.126)$$

C'est la *forme canonique* de l'EDP hyperbolique (1.118).

La forme canonique peut s'intégrer rigoureusement. En intégrant par rapport à  $p$  pour chaque valeur de  $q$ , on obtient

$$\partial_q \hat{u}(p, q) = f(q). \quad (1.127)$$

En intégrant maintenant par rapport à  $q$ , il vient

$$\hat{u}(p, q) = F(q) + g(p), \quad (1.128)$$

où  $F(q)$  est une primitive de  $f(q)$ . Il s'agit du résultat recherché.

On peut maintenant utiliser la forme générale pour obtenir la solution d'un problème avec conditions initiales et la fonction de Green. La condition initiale  $u(x, 0) = \phi(x)$  se traduit par

$$f(x) + g(x) = \phi(x), \quad (1.129)$$

ce qui n'est pas suffisant pour déterminer les fonctions  $f(x)$  et  $g(x)$ . Il faut donc une condition supplémentaire, par exemple  $\partial_t u(x, 0) = \psi(x)$ , qui donne

$$-cf'(x) + cg'(x) = \psi(x). \quad (1.130)$$

En introduisant une primitive  $\Psi(x)$  de  $\psi(x)$ , on peut écrire

$$f(x) = \frac{\phi(x)}{2} - \frac{\Psi(x)}{2c}, \quad (1.131)$$

$$g(x) = \frac{\phi(x)}{2} + \frac{\Psi(x)}{2c}. \quad (1.132)$$

La solution est donc

$$u(x, t) = f(x - ct) + g(x + ct) \quad (1.133)$$

$$= \frac{\phi(x - ct) + \phi(x + ct)}{2} + \frac{\Psi(x + ct) - \Psi(x - ct)}{2c} \quad (1.134)$$

$$= \frac{\phi(x - ct) + \phi(x + ct)}{2} + \frac{1}{2c} \int_{x-ct}^{x+ct} \psi(x') dx'. \quad (1.135)$$

On retrouve bien la propriété de causalité observée avec la fonction de Green en dimension 3, que ce soit pour la condition sur  $u(x, 0)$  ou sur  $\partial_t u(x, 0)$ .

**Remarque :** Il n'est pas surprenant qu'il faille définir la valeur initiale de la fonction mais aussi celle de sa dérivée temporelle, car l'EDP est du second ordre en temps : il se passe la même chose pour les EDO.

#### 1.5.4.3 Conservation de l'énergie, unicité

On montre ici qu'on peut définir l'énergie d'une solution  $u(\mathbf{r}, t)$  et que cette énergie est conservée pour des conditions au bord assez générales. On utilise ensuite cette propriété pour montrer l'unicité de la solution de l'équation des ondes dans certains cas.

Définissons l'énergie d'une fonction  $u(\mathbf{r}, t)$  définie sur  $\Omega \times I$ ,  $I \subset \mathbb{R}$  par

$$E_u(t) = \frac{1}{2} \int_{\Omega} ([\partial_t u(\mathbf{r}, t)]^2 + c^2 [\nabla u(\mathbf{r}, t)]^2) d\mathbf{r}. \quad (1.136)$$

Calculons maintenant la dérivée de cette énergie quand  $u(\mathbf{r}, t)$  est solution de l'équation des ondes homogène :

$$\dot{E}_u(t) = \int_{\Omega} [\partial_t u(\mathbf{r}, t) \partial_t^2 u(\mathbf{r}, t) + c^2 \nabla u(\mathbf{r}, t) \cdot \nabla \partial_t u(\mathbf{r}, t)] d\mathbf{r} \quad (1.137)$$

$$= c^2 \int_{\Omega} [\partial_t u(\mathbf{r}, t) \nabla^2 u(\mathbf{r}, t) + \nabla u(\mathbf{r}, t) \cdot \nabla \partial_t u(\mathbf{r}, t)] d\mathbf{r} \quad (1.138)$$

$$= c^2 \int_{\partial\Omega} \partial_t u(\mathbf{r}, t) \mathbf{n}(\mathbf{r}) \cdot \nabla u(\mathbf{r}, t) d\mathbf{r}. \quad (1.139)$$

Cette intégrale s'annule si  $u(\mathbf{r}, t)$  vérifie la condition aux limites de Dirichlet homogène ( $u(\mathbf{r}, t) = 0$ , donc  $\partial_t u(\mathbf{r}, t) = 0$ ) ou de Neumann homogène ( $\mathbf{n}(\mathbf{r}) \cdot \nabla u(\mathbf{r}, t) = 0$ ) en chaque point du bord (condition mixte).

On peut maintenant prouver l'unicité de la solution avec une condition au bord mixte. Le calcul nous renseignera sur les conditions initiales à prendre pour avoir l'unicité. Comme d'habitude, on considère la différence  $v(\mathbf{r}, t)$  entre deux solutions, qui satisfait la même équation avec des conditions au bord homogènes. Alors l'énergie de  $v(\mathbf{r}, t)$ ,  $E_v(t)$  (Éq. (1.136)), est constante. D'après l'expression de l'énergie, on voit qu'elle est nulle à l'instant initial  $t = 0$ , et donc pour tout  $t$ , si les valeurs de  $u(\mathbf{r}, 0)$  et  $\partial_t u(\mathbf{r}, 0)$  sont imposées. Si l'énergie est nulle, alors  $\partial_t v(\mathbf{r}, t) = 0$  et  $\nabla v(\mathbf{r}, t) = 0 \forall \mathbf{r}, t$ , donc comme  $v(\mathbf{r}, 0) = 0$ ,  $v(\mathbf{r}, t) = 0$ .

**Remarque :** En dimension  $d = 1$  en domaine infini, on montre après un court calcul que l'énergie de la solution  $u(x, t) = f(x - ct) + g(x + ct)$  vaut

$$E_u(t) = c^2 \int_{-\infty}^{\infty} [f'(x - ct)^2 + g'(x + ct)^2] dx = c^2 \int_{-\infty}^{\infty} [f'(x)^2 + g'(x)^2] dx; \quad (1.140)$$

nous avons effectué un changement de variables pour parvenir à la deuxième intégrale. Cette expression est clairement constante.

## 1.6 Analyse spectrale des EDP linéaires

### 1.6.1 Exemples

#### 1.6.1.1 Vibrations d'une corde ou d'un tambour

On décrit la position d'une corde ou de la surface d'un tambour par une fonction  $u(\mathbf{r}, t)$ , avec  $\mathbf{r} \in \Omega = [0, L]$  pour une corde et  $\mathbf{r} \in \Omega \subset \mathbb{R}^2$  pour un tambour. La position  $u(\mathbf{r}, t)$  satisfait l'équation des ondes (Éq. (1.5)),

$$\partial_t^2 u(\mathbf{r}, t) = c^2 \nabla^2 u(\mathbf{r}, t), \quad (1.141)$$

et la condition  $u(\mathbf{r}, t) = 0$  sur le bord  $\partial\Omega$ .

On s'intéresse souvent aux *modes propres* de la corde ou du tambour, qui sont des solutions particulières de la forme

$$u(\mathbf{r}, t) = \cos(\omega t) u_\omega(\mathbf{r}). \quad (1.142)$$

En mettant cette forme dans l'équation des ondes, on trouve

$$-\omega^2 u_\omega(\mathbf{r}) = c^2 \nabla^2 u_\omega(\mathbf{r}). \quad (1.143)$$

On voit que  $u_\omega(\mathbf{r})$  est vecteur propre du laplacien, avec la condition au bord  $u_\omega(\mathbf{r}) = 0$  sur  $\partial\Omega$ , pour la valeur propre  $-\omega^2/c^2$ .

#### 1.6.1.2 Équation de Schrödinger

L'équation de Schrödinger pour la fonction d'onde  $\psi(\mathbf{r}, t)$  d'une particule de masse  $m$  dans un potentiel  $V(\mathbf{r})$  s'écrit

$$i\hbar \partial_t \psi(\mathbf{r}, t) = -\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}, t) + V(\mathbf{r}) \psi(\mathbf{r}, t). \quad (1.144)$$

Le terme de droite correspond à l'opérateur hamiltonien appliqué à la fonction d'onde, souvent noté  $\hat{H}\psi(\mathbf{r}, t)$ . On cherche souvent à calculer les états propres de cet opérateur, dont les valeurs propres donnent les niveaux d'énergie du système. Un état propre  $\psi_i(\mathbf{r})$  et d'énergie  $E_i$  vérifie donc

$$\hat{H}\psi_i(\mathbf{r}) = -\frac{\hbar^2}{2m} \nabla^2 \psi_i(\mathbf{r}) + V(\mathbf{r}) \psi_i(\mathbf{r}) = E_i \psi_i(\mathbf{r}, t). \quad (1.145)$$

### 1.6.2 Généralités sur l'analyse spectrale

Comme on le voit sur les exemples, de nombreuses EDP faisant intervenir le temps peuvent être mises sous la forme

$$\partial_t^n u(\mathbf{r}, t) = \mathcal{L}u(\mathbf{r}, t), \quad (1.146)$$

où  $n \in \{1, 2\}$  le plus souvent, et  $\mathcal{L}$  est un opérateur différentiel agissant uniquement sur les variables spatiales. Ces équations sont définies sur  $\Omega \times I$ ,  $\Omega \subset \mathbb{R}^d$ ,  $I \subset \mathbb{R}$ , avec des conditions sur le bord  $\partial\Omega$  homogènes et indépendantes du temps.

On cherche alors à identifier les *modes propres* de l'opérateur  $\mathcal{L}$ , qui vérifient sur  $\Omega$

$$\mathcal{L}u(\mathbf{r}) = \lambda u(\mathbf{r}), \quad (1.147)$$

et les bonnes conditions sur  $\partial\Omega$ .

Si l'opérateur  $\mathcal{L}$  est auto-adjoint (pour un certain produit scalaire), alors les modes propres forment une base orthogonale de l'espace de Hilbert des fonctions avec ce produit scalaire, c'est le *théorème spectral*. En dimension finie, ce théorème signifie que toute matrice réelle symétrique est diagonalisable en base orthogonale. En dimension infinie, même l'énoncé précis de ce théorème dépasse le cadre de ce cours.

L'identification des modes propres a plusieurs avantages :

- Les modes propres ont souvent une signification « physique » importante et leur connaissance permet de mieux comprendre le système.
- Les modes propres sont plus faciles à calculer que la solution de l'équation de départ : il y a une dimension de moins, et il s'agit souvent d'un problème elliptique pour des EDP d'ordre 2, qui a de meilleures propriétés.
- Grâce au théorème spectral, les solutions de l'équation de départ peuvent être décomposées sur les modes propres.

### 1.6.3 Équation de Poisson sur une ligne

Nous allons faire l'analyse spectrale d'un problème très simple,

$$\mathcal{L}u(x) = u''(x) \quad (1.148)$$

sur  $\Omega = [0, L]$ , avec  $u(0) = u(L) = 0$ . Nous verrons ensuite comment utiliser ces résultats dans l'équation de la chaleur et l'équation des ondes.

#### 1.6.3.1 Analyse spectrale

On cherche une solution à

$$u''(x) = \lambda u(x). \quad (1.149)$$

Si  $\lambda > 0$ ,  $u(x) = a \exp(\pm \sqrt{\lambda}x)$ , donc pour satisfaire les conditions au bord il faut  $a = 0$  : il n'y a pas de solution non nulle. Si  $\lambda = 0$ ,  $u(x) = ax + b$ , et les conditions au bord imposent aussi  $a = b = 0$ . Si  $\lambda < 0$ ,  $u(x) = a \cos(\sqrt{|\lambda|x}) + b \sin(\sqrt{|\lambda|x})$ . La condition en 0 impose  $a = 0$ , et la condition en  $L$  impose  $b = 0$  ou  $\sqrt{|\lambda|}L = k\pi$ ,  $k \in \mathbb{N}^*$  (si  $k$  est négatif, il s'agit de la même solution que  $-k$ , et 0 donne  $u(x) = 0$ ).

Les modes propres  $u_k(x)$  et les valeurs propres  $\lambda_k$  associées sont donc, pour  $k \in \mathbb{N}^*$ ,

$$u_k(x) = \sin\left(\frac{k\pi x}{L}\right), \quad (1.150)$$

$$\lambda_k = -\left(\frac{k\pi}{L}\right)^2. \quad (1.151)$$

L'opérateur différentiel  $\mathcal{L} : u(x) \mapsto u''(x)$  est auto-adjoint pour le produit scalaire  $\langle f, g \rangle = \int_0^L f(x)g(x)dx$  :

$$\langle \mathcal{L}f, g \rangle = \int_0^L f''(x)g(x)dx = \int_0^L f(x)g''(x)dx = \langle f, \mathcal{L}g \rangle, \quad (1.152)$$

où nous avons utilisé deux intégrations par parties et la condition au bord  $f(0) = f(L) = 0$ . On vérifie que les modes propres sont orthogonaux pour ce produit scalaire :

$$\langle u_k, u_l \rangle = \int_0^L \sin\left(\frac{k\pi x}{L}\right) \sin\left(\frac{l\pi x}{L}\right) dx \quad (1.153)$$

$$= \int_0^L \frac{1}{2} \left[ \cos\left(\frac{[k-l]\pi x}{L}\right) - \cos\left(\frac{[k+l]\pi x}{L}\right) \right] dx \quad (1.154)$$

$$= \frac{L}{2} \delta_{k,l}. \quad (1.155)$$

On note au passage que ces modes ne sont pas normalisés, leur norme est  $\|u_k\| = \sqrt{L/2}$ .

Les coefficients  $a_k$  de la décomposition d'une fonction  $f(x)$  sur les modes propres,

$$f(x) = \sum_{k=1}^{\infty} a_k u_k(x), \quad (1.156)$$

sont donnés par

$$a_k = \frac{\langle f, u_k \rangle}{\|u_k\|^2} = \frac{2}{L} \int_0^L f(x) \sin\left(\frac{k\pi x}{L}\right) dx. \quad (1.157)$$

**Remarque :** On aurait pu prévoir certaines propriétés du spectre à partir de l'opérateur différentiel. En effet, l'opérateur  $\mathcal{L}$  est négatif car

$$\langle \mathcal{L}u, u \rangle = \int_0^L u''(x)u(x)dx = - \int_0^L u'(x)^2 dx \leq 0, \quad (1.158)$$

avec égalité si et seulement si  $u(x) = 0$ . Cela montre que

- 0 n'est pas valeur propre, car la relation ci-dessus montre que  $\mathcal{L}u = 0 \Rightarrow u = 0$ .
- les valeurs propres sont strictement négatives.

Cette remarque s'applique aussi pour l'équation de Poisson sur un domaine  $\Omega \subset \mathbb{R}^d$  avec une condition au bord de Dirichlet homogène.

### 1.6.3.2 Application à l'équation de la chaleur

On cherche ici à résoudre l'équation de la chaleur,

$$\partial_t u(x, t) = D \partial_x^2 u(x, t), \quad (1.159)$$

avec  $u(x, 0) = u_0(x)$ ,  $u(0, t) = u(L, t) = 0$ . Cette équation peut représenter la diffusion d'un élément de densité initiale  $u_0(x)$  avec des bords absorbants.

La décomposition de  $u(x, t)$  en modes propres s'écrit

$$u(x, t) = \sum_{k=1}^{\infty} a_k(t) u_k(x). \quad (1.160)$$

En injectant cette décomposition dans l'équation, on obtient d'une part

$$\partial_t u(x, t) = \sum_{k=1}^{\infty} \dot{a}_k(t) u_k(x), \quad (1.161)$$

et d'autre part

$$\partial_t u(x, t) = D \partial_x^2 u(x, t) = D \sum_{k=1}^{\infty} a_k(t) u_k''(x) = D \sum_{k=1}^{\infty} \lambda_k a_k(t) u_k(x). \quad (1.162)$$



En identifiant les coefficients de  $u_k(x)$  dans ces deux expressions (c'est possible car les modes sont orthogonaux), on obtient l'évolution des coefficients  $a_k(t)$  :  $\dot{a}_k(t) = D\lambda_k a_k(t)$ , et donc

$$a_k(t) = a_k(0)e^{D\lambda_k t}. \quad (1.163)$$

La valeur initiale des coefficients est donnée par la décomposition de la condition initiale :

$$u_0(x) = \sum_{k=1}^{\infty} a_k(0)u_k(x). \quad (1.164)$$

Finalement,

$$u(x, t) = \sum_{k=1}^{\infty} a_k(0) \exp\left(-\frac{D\pi^2 k^2 t}{L^2}\right) \sin\left(\frac{k\pi x}{L}\right). \quad (1.165)$$

On peut par exemple déduire de ce résultat la limite de temps long :

$$u(x, t) \underset{t \rightarrow \infty}{\sim} a_1(0) \exp\left(-\frac{D\pi^2 t}{L^2}\right) \sin\left(\frac{\pi x}{L}\right), \quad (1.166)$$

avec

$$a_1(0) = \frac{2}{L} \int_0^L u_0(x) \sin\left(\frac{\pi x}{L}\right) dx. \quad (1.167)$$

### 1.6.3.3 Application à l'équation des ondes

Appliquons maintenant l'étude spectrale à l'équation des ondes,

$$\partial_t^2 u(x, t) = c^2 \partial_x^2 u(x, t), \quad (1.168)$$

avec  $u(x, 0) = u_0(x)$ ,  $\partial_t u(x, 0) = v_0(x)$ ,  $u(0, t) = u(L, t) = 0$ . Cette équation représente par exemple, comme nous l'avons vu plus haut, la vibration d'une corde fixée à une de ses extrémités.

On peut utiliser la même décomposition que pour l'équation de la chaleur, Éq. (1.160); en appliquant le même raisonnement, on trouve que l'évolution des coefficients est donnée par  $\ddot{a}_k(t) = c^2 \lambda_k a_k(t)$  ( $\lambda_k < 0$ ), et donc

$$a_k(t) = a_k(0) \cos\left(\sqrt{|\lambda_k|} ct\right) + \frac{\dot{a}_k(0)}{\sqrt{|\lambda_k|} c} \sin\left(\sqrt{|\lambda_k|} ct\right). \quad (1.169)$$

Finalement,

$$u(x, t) = \sum_{k=1}^{\infty} \left[ a_k(0) \cos\left(\frac{k\pi ct}{L}\right) + \frac{L \dot{a}_k(0)}{k\pi c} \sin\left(\frac{k\pi ct}{L}\right) \right] \sin\left(\frac{k\pi x}{L}\right). \quad (1.170)$$

### 1.6.4 Équation de Poisson sur un disque

Nous faisons ici l'étude spectrale de l'équation de Poisson sur un disque de rayon  $R$  avec une condition au bord de Dirichlet : on cherche à déterminer les modes propres d'un tambour circulaire.

Comme dans le cas de l'équation de Poisson sur une ligne, nous savons que les valeurs propres seront strictement négatives ; on cherche donc les solutions non nulles de

$$\nabla^2 u(\mathbf{r}) = -\lambda u(\mathbf{r}), \quad (1.171)$$

avec  $\lambda > 0$ . Nous utilisons les coordonnées sphériques et écrivons  $u(\mathbf{r}) = u(r, \theta)$ ; l'équation de Poisson devient

$$\partial_r^2 u(r, \theta) + \frac{1}{r} \partial_r u(r, \theta) + \frac{1}{r^2} \partial_\theta^2 u(r, \theta) = -\lambda u(r, \theta), \quad (1.172)$$

et la condition au bord s'écrit  $u(R, \theta) = 0$ .

On utilise la méthode de séparation des variables (voir aussi Sec. 1.4.2) en cherchant la solution sous la forme  $u(r, \theta) = P(r)Q(\theta)$ . L'équation devient alors

$$P''(r)Q(\theta) + \frac{P'(r)}{r}Q(\theta) + \frac{P(r)}{r^2}Q''(\theta) = -\lambda P(r)Q(\theta), \quad (1.173)$$

et la condition au bord est simplement  $P(R) = 0$ . En multipliant tous les termes de l'équation par  $r^2/[P(r)Q(\theta)]$ , on arrive à

$$\frac{r^2 P''(r) + r P'(r)}{P(r)} + \lambda r^2 + \frac{Q''(\theta)}{Q(\theta)} = 0. \quad (1.174)$$

Comme nous l'avons vu plus haut, les termes dépendants uniquement de  $r$  (les deux premiers) et celui dépendant uniquement de  $\theta$  sont constants :

$$\frac{r^2 P''(r) + r P'(r)}{P(r)} + \lambda r^2 = -\frac{Q''(\theta)}{Q(\theta)} = C. \quad (1.175)$$

L'équation la plus facile à résoudre est  $Q''(\theta) = -CQ(\theta)$  :  $Q(\theta) = e^{in\theta}$ ,  $n \in \mathbb{Z}$ , et  $C = n^2$ . Si on préfère des solutions réelles, on peut prendre  $Q(\theta) = \cos(n\theta)$  ou  $\sin(n\theta)$ . En injectant cela dans l'équation pour  $P(r)$ , on trouve

$$r^2 P''(r) + r P'(r) + (\lambda r^2 - n^2) P(r) = 0. \quad (1.176)$$

$\lambda$  peut être absorbé dans un « changement d'échelle » de  $r$ , c'est à dire en définissant  $P(r) = \tilde{P}(\sqrt{\lambda}r)$ , de sorte que  $P'(r) = \sqrt{\lambda}\tilde{P}'(\sqrt{\lambda}r)$ , etc. Alors l'équation pour  $\tilde{P}(x)$  devient

$$x^2 \tilde{P}''(x) + x \tilde{P}'(x) + (x^2 - n^2) \tilde{P}(x) = 0 : \quad (1.177)$$

il s'agit de l'équation qui définit la fonction de Bessel d'ordre  $n$ ,  $J_n(x)$ ; ainsi,  $P(r) = J_n(\sqrt{\lambda}r)$ .

Pour l'instant, il n'y a aucune contrainte sur  $\lambda$ . Il faut utiliser la condition au bord,  $0 = P(R) = J_n(\sqrt{\lambda}R)$  :  $\sqrt{\lambda}R$  est un zéro de  $J_n$ . Les fonctions de Bessel ont une infinité de zéros; notons  $\mu_{n,p}$  le  $p$ -ème zéro de la fonction de Bessel d'ordre  $n$ ,  $p \in \mathbb{N}$ . Nous avons donc  $\lambda = \mu_{n,p}^2/R^2$  : les valeurs propres sont donc indexées par deux entiers  $n$  et  $p$ , de même que les fonctions propres :

$$\lambda_{n,p} = \frac{\mu_{n,p}^2}{R^2}, \quad (1.178)$$

$$u_{n,p}(\mathbf{r}) = J_n\left(\frac{\mu_{n,p}r}{R}\right) e^{in\theta}. \quad (1.179)$$

Il faut noter que les espaces propres sont de dimension 2, car  $n$  et  $-n$  correspondent à la même valeur propre. Une autre façon de le voir est de ne prendre que  $n \geq 0$  et de remplacer  $e^{in\theta}$  par  $\cos(n\theta)$  et  $\sin(n\theta)$  pour le mode propre. Les premiers modes sont représentés sur la figure 1.2.

## 1.7 Pour aller plus loin

Nous n'avons vu que des EDP d'ordre un ou deux, mais on peut bien sûr rencontrer des EDP d'ordre supérieur comme l'équation des plaques élastiques. Les EDP d'ordre supérieur à deux peuvent s'étudier avec les techniques présentées ici.

Nous avons ici complètement occulté la question des espaces fonctionnels associés aux EDP. D'une part, ces espaces définissent où chercher des solutions, où prendre les termes sources, etc. Nous avons aussi vu en section 1.3.1 qu'un problème était bien posé si sa solution dépendait continûment de la

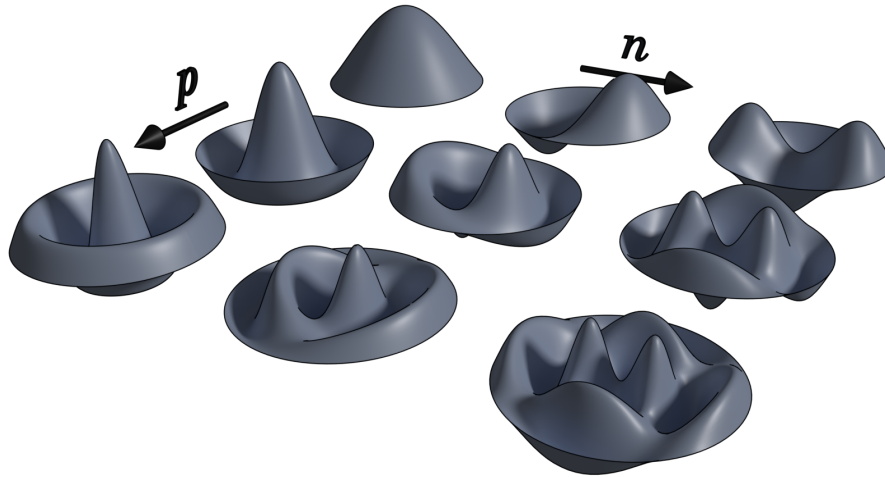


FIGURE 1.2 – Modes propres d'un tambour pour  $0 \leq n \leq 2$ ,  $1 \leq p \leq 3$ .

condition au bord : la définition même de la continuité requiert la définition d'une norme sur l'espace des conditions au bord et sur l'espace des solutions.

Enfin, à cause de l'absence de solution générale, les EDP doivent souvent être résolues numériquement, ce qui représente un domaine des mathématiques à part entière. Il faut noter que pour les EDP linéaires, la résolution numérique associée est un problème linéaire, revenant en bref à inverser une matrice une fois que la procédure de discrétisation a été choisie. La résolution numérique des EDP linéaires est abordée dans le cours de Méthodes Numériques de 2<sup>ème</sup> année.



## Chapitre 2

# Calcul variationnel

### 2.1 Introduction

Dans de nombreuses situations en physique, le problème n'est pas posé en terme d'équation à résoudre mais d'énergie à minimiser. Un exemple classique est celui de la corde pesante attachée entre deux points, mais il y a aussi la forme d'un ménisque ou d'un film de savon, etc. Dans ces cas, l'inconnue est une fonction et l'énergie une fonction de cette fonction. Minimiser l'énergie donne des équations satisfaites par la fonction inconnue, qui peuvent être des EDO ou des EDP, suivant la nature et la dimension du problème. Voir comment déterminer ces équations à partir de l'énergie est l'objet principal de ce chapitre.

Il se trouve que la formulation du problème en terme d'« énergie » à minimiser présente aussi des avantages, notamment car elle permet d'utiliser les invariances du problème pour simplifier les équations obtenues. Il peut ainsi être intéressant de transformer un problème classique posé sous forme d'équation en un problème de minimisation; cela a notamment été fait en mécanique du point ou en optique.

### 2.2 Calcul variationnel en dimension un

#### 2.2.1 Exemple : forme d'une corde pesante

On cherche à déterminer la forme d'une corde de masse linéique  $\rho$  dans le champ de pesanteur. On se place dans un plan vertical et on utilise le système de coordonnées «  $(x, z)$  ». Une extrémité de la corde est attachée en  $(0, 0)$ ; pour la deuxième extrémité, on considère trois possibilités :

- (a) La corde passe par une poulie en  $(L, 0)$  et est attachée à un poids de masse  $m$ .
- (b) La corde est de longueur  $L_0 > L$  et la deuxième extrémité est attachée en  $(L, 0)$ .
- (c) La corde est de longueur  $L_0 > L$  et la deuxième extrémité est libre de coulisser le long de l'axe  $x = L$ .

On paramètre la hauteur de la corde à l'abscisse  $x \in [0, L]$  par  $h(x)$ . L'énergie potentielle de la corde est une fonction de  $h(x)$  :

$$E[h] = \rho g \int_0^L h(x) \sqrt{1 + h'(x)^2} dx. \quad (2.1)$$

Dans le cas (a), l'énergie totale du système contient aussi l'énergie potentielle du poids (on néglige l'énergie potentielle du bout de corde situé entre la poulie et le poids)

$$E_p[h] = mg \int_0^L \sqrt{1 + h'(x)^2} dx. \quad (2.2)$$

Nous avons choisi par convention une énergie nulle quand  $L = 0$ .

Dans tous les cas, la forme de la corde est celle qui minimise l'énergie. Dans le cas (a), il faut minimiser la somme des énergies (2.1) et (2.2). Dans les cas (b) et (c), il faut minimiser l'énergie (2.1) avec la contrainte que la longueur de la corde est  $L_0$ .

Le problème mathématique est le suivant : on doit minimiser une *fonctionnelle*, c'est à dire une fonction d'une fonction, avec des conditions au bord et éventuellement avec une contrainte globale sur la fonction.

**Remarque :** Les cas (b) et (c) font appel à une contrainte globale sur la longueur de la corde, qui n'est pas évidente à traiter. Ce n'est pas le cas des contraintes au bord, qui se rencontrent dans tous les cas et qui sont faciles à traiter.

## 2.2.2 Équation d'Euler-Lagrange

Nous allons voir comment trouver la fonction  $f(x)$  qui minimise une fonctionnelle de la forme

$$I[f] = \int_a^b L(x, f(x), f'(x)) dx, \quad (2.3)$$

où  $L(x, y, y')$  est une fonction réelle de trois variables, que l'on appelle *lagrangien*, avec des conditions aux limites pour  $f(x)$  en  $a$  et  $b$ .

Pour minimiser la fonctionnelle  $I[f]$ , on utilise la démarche classique utilisée pour minimiser une fonction : on cherche les points où la dérivée de la fonction s'annule. La difficulté réside dans le fait qu'il faille ici calculer la dérivée par rapport à une fonction. Pour calculer la dérivée d'une fonction  $g(x)$  en  $x$ , il s'agit de calculer  $g(x+\epsilon) - g(x)$  au premier ordre en  $\epsilon$ . C'est ce que nous allons faire ici, avec la fonction  $\epsilon(x)$  : à l'ordre 1 en  $\epsilon(x)$ ,

$$I[f + \epsilon] = \int_a^b L(x, f(x) + \epsilon(x), f'(x) + \epsilon'(x)) dx \quad (2.4)$$

$$= \int_a^b [L(x, f(x), f'(x)) + \epsilon(x) \partial_y L(x, f(x), f'(x)) + \epsilon'(x) \partial_{y'} L(x, f(x), f'(x))] dx \quad (2.5)$$

$$= I[f] + [\epsilon(x) \partial_{y'} L(x, f(x), f'(x))]_a^b + \int_a^b \epsilon(x) \left( \partial_y L(x, f(x), f'(x)) - \frac{d}{dx} [\partial_{y'} L(x, f(x), f'(x))] \right) dx. \quad (2.6)$$

Nous avons intégré par parties pour arriver à la dernière équation.

La fonction  $f(x)$  minimise (ou maximise, en fait) la fonctionnelle  $I[f]$  si la dérivée fonctionnelle de  $I[f]$  en  $f(x)$  est nulle, c'est à dire si les termes d'ordre  $\epsilon(x)$  s'annulent dans l'équation (2.6). Il y a deux termes d'ordre  $\epsilon(x)$  : le deuxième qui est un terme de bord, et le troisième qui est un terme de cœur. Annuler le terme de cœur donne l'équation d'Euler-Lagrange :

$$\partial_y L(x, f(x), f'(x)) = \frac{d}{dx} [\partial_{y'} L(x, f(x), f'(x))] \quad (2.7)$$

Annuler le terme de bord conduit à annuler  $\epsilon(x) \partial_{y'} L(x, f(x), f'(x))$  aux extrémités, il faut donc que l'un des deux termes soit nul :

- Avec une condition au bord de Dirichlet, où  $f(x)$  est fixé au bord,  $\epsilon(x) = 0$  au bord et on n'obtient aucune nouvelle condition. C'est le cas des conditions (a), (b), et de l'extrémité en  $x = 0$  du cas (c) de l'exemple (Sec. 2.2.1).
- Si la valeur de  $f(x)$  n'est pas fixée au bord, on obtient la condition

$$\partial_{y'} L(x, f(x), f'(x)) = 0. \quad (2.8)$$

C'est le cas du bord en  $x = L$  de la condition (c) de l'exemple.

### 2.2.2.1 Application à la corde tendue par un poids

Avec l'équation d'Euler-Lagrange, nous pouvons calculer la forme de la corde tendue par un poids, ce qui correspond au cas (a) de l'exemple. Il faut donc minimiser l'énergie totale

$$I[h] = \rho g \int_0^L h(x) \sqrt{1 + h'(x)^2} dx + m g \int_0^L \sqrt{1 + h'(x)^2} dx = \int_0^L L(x, h(x), h'(x)) dx, \quad (2.9)$$

avec

$$L(x, y, y') = \rho g (y + \lambda) \sqrt{1 + y'^2}, \quad (2.10)$$

où  $\lambda = m/\rho$ . La condition au bord est  $h(0) = h(L) = 0$ .

Pour écrire l'équation d'Euler-Lagrange, il faut calculer

$$\partial_y L(x, y, y') = \rho g \sqrt{1 + y'^2}, \quad (2.11)$$

$$\partial_{y'} L(x, y, y') = \rho g \frac{(y + \lambda) y'}{\sqrt{1 + y'^2}}. \quad (2.12)$$

L'équation d'Euler-Lagrange (Éq. (2.7)) donne donc

$$\sqrt{1 + h'(x)^2} = \frac{d}{dx} \left( \frac{[h(x) + \lambda] h'(x)}{\sqrt{1 + h'(x)^2}} \right). \quad (2.13)$$

Après développement du terme de droite et simplification, on arrive à

$$1 + h'(x)^2 = [h(x) + \lambda] h''(x), \quad (2.14)$$

qui est une EDO non linéaire du deuxième ordre. En multipliant les deux côtés par  $h'(x)/([h(x) + \lambda][1 + h'(x)^2])$ , on obtient

$$\frac{h'(x)}{h(x) + \lambda} = \frac{h'(x) h''(x)}{1 + h'(x)^2}, \quad (2.15)$$

ou

$$\frac{d}{dx} \log(h(x) + \lambda) = \frac{1}{2} \frac{d}{dx} \log(1 + h'(x)^2). \quad (2.16)$$

On en déduit

$$\frac{d}{dx} \log \left( \frac{h(x) + \lambda}{\sqrt{1 + h'(x)^2}} \right) = 0, \quad (2.17)$$

donc

$$\frac{h(x) + \lambda}{\sqrt{1 + h'(x)^2}} = A, \quad (2.18)$$

où  $A$  est une constante. Introduisons  $g(x) = h(x) + \lambda$  qui vérifie  $g'(x)^2 = A^{-2} g(x)^2 - 1$ . On remarque que le cosinus hyperbolique vérifie  $\text{ch}'(x)^2 = \text{sh}(x)^2 = \text{ch}(x)^2 - 1$ , donc  $g(x) = A \text{ch}([x - x_0]/A)$  est solution, d'où  $h(x) = A \text{ch}([x - x_0]/A) - \lambda$ . La symétrie impose  $x_0 = L/2$ , d'où

$$h(x) = A \text{ch} \left( \frac{x - \frac{L}{2}}{A} \right) - \lambda, \quad (2.19)$$

et les conditions au bord donnent  $A \text{ch}(L/(2A)) = \lambda$ , ce qui est une équation implicite pour  $A$ . Suivant la valeur de  $\lambda$ , cette équation peut avoir deux solutions ou aucune solution. Quand il n'y a pas de solution, cela veut dire que la corde est trop lourde pour être retenue par la masse; quand il y a deux solutions, on peut se convaincre « physiquement » que c'est la plus grande solution qui est stable.

Nous avons pu déterminer la forme de la corde dans le cas (a), mais les cas (b) et (c) nécessitent une minimisation sous contrainte, que nous ne savons pas faire pour l'instant.

### 2.2.2.2 Bord mobile

On rencontre parfois une condition au bord plus « exotique », où la valeur de la fonction au bord est fixée mais où la position du bord n'est pas fixée. Considérons le cas général où le bord  $b$  n'est pas fixé, mais où la valeur de  $f(b)$  est connue. Il faut déterminer  $b$  en minimisant la fonctionnelle par rapport à  $b$ . On considère alors une variation  $\eta$  de  $b$  et on compare  $I[f + \epsilon, b + \eta]$  à  $I[f, b]$ . Au premier ordre en  $\epsilon(x)$  et  $\eta$ , on obtient un terme de bord supplémentaire par rapport au résultat (2.6), qui est  $\eta L(b, f(b), f'(b))$ . Les termes de bord en  $b$  doivent s'annuler, c'est à dire

$$\epsilon(b) \partial_{y'} L(b, f(b), f'(b)) + \eta L(b, f(b), f'(b)) = 0. \quad (2.20)$$

Mais comme  $f(b)$  est fixée, le déplacement du bord  $\eta$  et la variation  $\epsilon(x)$  doivent satisfaire  $f(b) = f(b + \eta) + \epsilon(b + \eta) = f(b) + \eta f'(b) + \epsilon(b)$  au premier ordre en  $\epsilon(x)$  et  $\eta$ , ce qui donne  $\epsilon(b) = -\eta f'(b)$ . En utilisant cette relation dans l'équation (2.20), il vient

$$f'(b) \partial_{y'} L(b, f(b), f'(b)) = L(b, f(b), f'(b)), \quad (2.21)$$

ce qui est la nouvelle condition au bord.

**Application à la corde pesante.** Si la poulie du cas (a) peut glisser le long l'axe  $z = 0$ ,  $L$  est une inconnue. La condition au bord (2.21) s'écrit

$$\frac{h'(L)^2}{\sqrt{1 + h'(L)^2}} = \sqrt{1 + h'(L)^2}, \quad (2.22)$$

où on a utilisé  $h(L) = 0$ . Cette équation semble conduire à  $1 = 0$ , mais elle est satisfaite dans la limite  $|h'(L)| \rightarrow \infty$ , elle impose donc  $|h'(L)| = \infty$ . Avec la forme de la corde, cela conduit à  $L = 0$ , ce qui n'est pas très intéressant.

### 2.2.2.3 Équations linéaires et lagrangien quadratique

Les équations différentielles linéaires, quelles soient en dimension un (EDO) ou plus (EDP) sont les plus faciles à traiter. Les équations d'Euler-Lagrange étant obtenues par une dérivées fonctionnelle, les équations linéaires proviennent de lagrangiens quadratiques, qui constituent donc la classe la plus simple de lagrangiens.

Illustrons cela pour la corde pesante dans le cas (a) où elle est soumise à une tension  $mg$ . Le lagrangien (2.10) est proportionnel à  $L(x, y, y') = (y + \lambda) \sqrt{1 + y'^2}$ . Développons ce lagrangien à l'ordre 2 en  $y$  et  $y'$  :

$$L(x, y, y') = \lambda + y + \frac{1}{2} \lambda y'^2 + \mathcal{O}(y^3). \quad (2.23)$$

On en déduit l'équation d'Euler-Lagrange

$$1 = \lambda h''(x), \quad (2.24)$$

qui est une équation linéaire, et qui correspond à la linéarisation de l'équation (2.14). La solution de cette équation avec les conditions aux limites  $h(0) = h(L) = 0$  donne

$$h(x) = \frac{1}{2\lambda} \left[ \left( x - \frac{L}{2} \right)^2 - \frac{L^2}{4} \right]. \quad (2.25)$$

Comparons cette solution à la solution générale (Éq. (2.19)). L'approximation de lagrangien quadratique correspond à la limite de faibles variations de hauteur, soit  $|h'(x)| \ll 1$ ; dans l'équation générale, cela correspond à la limite  $A \gg L$  où la solution devient

$$h(x) \simeq A - \lambda + \frac{1}{2A} \left( x - \frac{L}{2} \right)^2. \quad (2.26)$$



Les conditions aux limites utilisées permettent enfin de retrouver la solution de l'équation linéarisée.

Face à un problème variationnel, il est souvent utile de commencer par étudier le problème linéarisé correspondant au lagrangien quadratique pour se faire une idée du comportement du système et éventuellement étudier certains cas limites quand le problème général n'a pas de solution explicite simple.

### 2.2.3 Minimisation sous contrainte

On cherche ici à minimiser une fonctionnelle  $I[f]$  de la forme (2.3) avec une contrainte globale de la forme

$$J[f] = \int_a^b G(x, f(x), f'(x)) dx = J_0. \quad (2.27)$$

Nous allons voir que cela n'est pas difficile, mais nous allons commencer par un exemple en dimension finie.

#### 2.2.3.1 Minimisation sous contrainte en dimension finie

Plaçons-nous en dimension finie et cherchons à minimiser la fonction  $f : \mathbf{x} \in \mathbb{R}^d \mapsto f(\mathbf{x}) \in \mathbb{R}$  sous la contrainte  $g(\mathbf{x}) = g_0 \in \mathbb{R}$ . Sans contrainte, il suffit de chercher les points où le gradient de  $f(\mathbf{x})$  s'annule :  $\nabla f(\mathbf{x}) = 0$ . Avec la contrainte,  $\mathbf{x} \in \Omega = g^{-1}(g_0) \subset \mathbb{R}^d$ , qu'il faut se représenter comme une surface. Dire que  $f(\mathbf{x})$  est maximale en  $\mathbf{x}_0 \in \Omega$  revient à dire que  $f(\mathbf{x})$  ne varie pas autour de  $\mathbf{x}_0$  le long de cette surface, autrement dit que  $\nabla f(\mathbf{x}_0)$  est perpendiculaire à la surface. Or,  $\nabla g(\mathbf{x}_0)$  est par définition perpendiculaire à la surface, il existe donc  $\lambda$  tel que  $\nabla f(\mathbf{x}_0) = \lambda \nabla g(\mathbf{x}_0)$ . Cette relation peut se réécrire  $\nabla[f(\mathbf{x}_0) - \lambda g(\mathbf{x}_0)] = 0$ .

On en déduit que pour minimiser  $f(\mathbf{x})$  sous la contrainte  $g(\mathbf{x}) = g_0$ , il faut minimiser  $f(\mathbf{x}) - \lambda g(\mathbf{x})$  pour un certain  $\lambda$ . Dans le cadre du calcul variationnel,  $\lambda$  est appelé *multiplicateur de Lagrange*, il doit être déterminé à la fin du calcul.

**Remarque :** Pour voir plus précisément pourquoi  $\nabla g(\mathbf{x}_0)$  et  $\nabla f(\mathbf{x}_0)$  doivent être colinéaires, calculons  $f(\mathbf{x}_0 + \boldsymbol{\epsilon})$  au premier ordre en  $\boldsymbol{\epsilon}$  :  $f(\mathbf{x}_0 + \boldsymbol{\epsilon}) = f(\mathbf{x}_0) + \boldsymbol{\epsilon} \cdot \nabla f(\mathbf{x}_0)$ . Pour que cette variation s'annule, il faut  $\boldsymbol{\epsilon} \cdot \nabla f(\mathbf{x}_0) = 0$ . Mais  $\boldsymbol{\epsilon}$  n'est pas quelconque, car  $\mathbf{x}_0 + \boldsymbol{\epsilon}$  doit satisfaire la contrainte, c'est à dire  $g_0 = g(\mathbf{x}_0) + \boldsymbol{\epsilon} \cdot \nabla g(\mathbf{x}_0)$ ; or  $g(\mathbf{x}_0) = g_0$ , ce qui implique  $\boldsymbol{\epsilon} \cdot \nabla g(\mathbf{x}_0) = 0$ . On en déduit que  $\mathbf{x}_0$  maximise  $f(\mathbf{x})$  sous la contrainte si

$$\forall \boldsymbol{\epsilon} \in \mathbb{R}^d, \quad \boldsymbol{\epsilon} \cdot \nabla g(\mathbf{x}_0) = 0 \quad \Rightarrow \quad \boldsymbol{\epsilon} \cdot \nabla f(\mathbf{x}_0) = 0. \quad (2.28)$$

Cela n'est possible que si les deux vecteurs  $\nabla g(\mathbf{x}_0)$  et  $\nabla f(\mathbf{x}_0)$  sont colinéaires.

#### 2.2.3.2 Minimisation sous contrainte en calcul variationnel

D'après la section précédente, pour minimiser la fonctionnelle  $I[f]$  sous la contrainte  $J[f] = J_0$ , on doit minimiser la fonctionnelle

$$I'[f] = I[f] - \lambda J[f]. \quad (2.29)$$

Le multiplicateur de Lagrange  $\lambda$  doit être traité comme une constante pour la minimisation et doit ensuite être déterminé de façon à satisfaire la contrainte.

On peut souvent attribuer une signification physique au multiplicateur de Lagrange. Quand  $I[f]$  est une énergie,  $\lambda$  est la quantité *conjuguée* à la grandeur  $J$ , comme nous allons le voir dans l'exemple.

**Remarque :** Il peut y avoir plusieurs contraintes,  $J_i[f] = j_i$ , il faut dans ce cas introduire un multiplicateur de Lagrange  $\lambda_i$  par contrainte et minimiser

$$I'[f] = I[f] - \sum_{i=1}^n \lambda_i J_i[f]. \quad (2.30)$$

### 2.2.3.3 Application à la corde pesante

Nous pouvons maintenant résoudre le problème de la corde pesante dans les cas (b) et (c) : il faut minimiser l'énergie gravitationnelle de la corde (2.1) sous la contrainte

$$J[h] = \int_0^L \sqrt{1 + h'(x)^2} dx = L_0. \quad (2.31)$$

En introduisant un multiplicateur de Lagrange  $\sigma$ , il faut donc minimiser

$$I[h] = E[h] - \sigma J[h] = \rho g \int_0^L h(x) \sqrt{1 + h'(x)^2} dx - \sigma \int_0^L \sqrt{1 + h'(x)^2} dx. \quad (2.32)$$

En remplaçant  $-\sigma$  par  $mg$ , on retrouve exactement la fonctionnelle à minimiser par une corde tendue par un poids de masse  $m$  (Éq. (2.9)) : le multiplicateur de Lagrange utilisé pour contraindre la longueur est la tension dans la corde.

Dans le cas (b), la condition au bord est maintenant la même que dans le cas (a), et on peut utiliser la solution (2.19) :

$$h(x) = A \left[ \operatorname{ch} \left( \frac{x - \frac{L}{2}}{A} \right) - \operatorname{ch} \left( \frac{L}{2A} \right) \right], \quad (2.33)$$

où  $A$  doit être déterminé par la contrainte sur la longueur, qui s'écrit

$$L_0 = \int_0^L \sqrt{1 + h'(x)^2} dx = \int_0^L \operatorname{ch} \left( \frac{x - \frac{L}{2}}{A} \right) dx = 2A \operatorname{sh} \left( \frac{L}{2A} \right). \quad (2.34)$$

On peut ensuite utiliser  $\lambda = \operatorname{ch}(L/[2A])$  puis  $\sigma = -\rho g \lambda$  ; la valeur de  $\sigma$  n'est toutefois pas ce qui nous intéresse ici. La forme de la corde pour  $L = 1$  et différentes valeurs de  $L_0$  est représentée en figure 2.1.

Dans le cas (c), la condition en  $x = L$  est différente : la hauteur n'est pas imposée, il faut donc utiliser la condition au bord (2.8), qui s'écrit

$$0 = \partial_{y'} L(L, h(L), h'(L)) = \rho g \frac{[h(L) + \lambda] h'(L)}{\sqrt{1 + h'(L)^2}}. \quad (2.35)$$

Cette relation impose  $h'(L) = 0$ . En reprenant le calcul à tension imposée, on voit que cela conduit à

$$h(x) = A \operatorname{ch} \left( \frac{x - L}{A} \right) - \lambda, \quad (2.36)$$

avec la condition en  $x = 0$  qui donne  $\lambda = A \operatorname{ch}(L/A)$  et la contrainte sur la longueur

$$L_0 = A \operatorname{sh} \left( \frac{L}{A} \right). \quad (2.37)$$

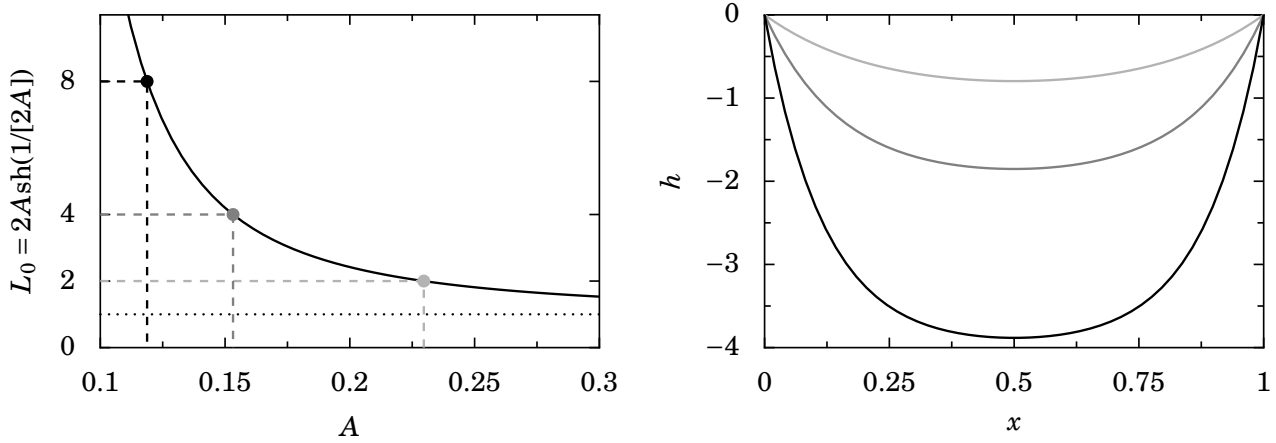


FIGURE 2.1 – Forme d’une corde pesante de longueur donnée attachée en  $x = 0$  et  $x = 1$ , pour différentes longueurs de la corde  $L_0 = 2, 4, 8$  (du plus clair au plus foncé). *Gauche* : équation (2.34) à résoudre pour déterminer le paramètre  $A$  qui caractérise la forme de la corde. *Droite* : forme de la corde (Éq. (2.33)).

### 2.2.4 Invariance par translation et intégrale du mouvement

On montre ici qu’une invariance du problème par translation est reliée à l’existence d’une *intégrale du mouvement*. Le système est invariant par translation si le lagrangien qui entre dans la fonctionnelle à minimiser ne dépend pas explicitement de la variable d’espace :

$$L(x, y, y') = L(y, y'). \quad (2.38)$$

Dans ce cas, pour une fonction  $f(x)$  qui minimise la fonctionnelle associée au lagrangien  $L$ , calculons la dérivée totale

$$\frac{d}{dx} [L(x, f(x), f'(x))] = f'(x) \partial_y L(x, f(x), f'(x)) + f''(x) \partial_{y'} L(x, f(x), f'(x)) \quad (2.39)$$

$$= f'(x) \frac{d}{dx} [\partial_{y'} L(x, f(x), f'(x))] + f''(x) \partial_{y'} L(x, f(x), f'(x)) \quad (2.40)$$

$$= \frac{d}{dx} [f'(x) \partial_{y'} L(x, f(x), f'(x))]. \quad (2.41)$$

Nous avons utilisé le fait que  $f(x)$  vérifie l’équations d’Euler-Lagrange. On obtient donc

$$\frac{d}{dx} [f'(x) \partial_{y'} L(x, f(x), f'(x)) - L(x, f(x), f'(x))] = 0 : \quad (2.42)$$

le terme entre crochets,

$$y' \partial_{y'} L(x, y, y') - L(x, y, y'), \quad (2.43)$$

est une intégrale du mouvement, sa valeur est constante sur l’intervalle  $[a, b]$ .

#### 2.2.4.1 Application à la corde pesante

Le lagrangien de la corde pesante (Éq. (2.10)) ne dépend pas explicitement de la variable  $x$ , on peut donc utiliser l’intégrale du mouvement (2.42) :

$$y' \partial_{y'} L(y, y') - L(y, y') = \rho g \frac{(y + \lambda) y'^2}{\sqrt{1 + y'^2}} - \rho g (y + \lambda) \sqrt{1 + y'^2} = \rho g \frac{y + \lambda}{\sqrt{1 + y'^2}}. \quad (2.44)$$

Cette quantité est constante, c'est à dire

$$\frac{h(x) + \lambda}{\sqrt{1 + h'(x)^2}} = A ; \quad (2.45)$$

on retrouve le résultat (2.18), qui n'a pas été facile à obtenir à partir de l'EDO initiale.

On voit que l'intégrale du mouvement donne une EDO non linéaire du premier ordre, qui est plus facile à résoudre que l'équation d'Euler-Lagrange ; il faut en général privilégier l'intégrale du mouvement. Utiliser l'intégrale du mouvement, c'est utiliser l'invariance par translation, qui est une symétrie du problème ; il est bien connu qu'utiliser les symétries d'un problème permet de le simplifier par rapport à l'approche « frontale ».

## 2.3 Calcul variationnel en dimension supérieure

### 2.3.1 Exemples

Nous allons voir quelques exemples pour lesquels l'espace de départ ou d'arrivée de la fonction inconnue  $f(x)$  sont de dimension supérieure à un.

#### 2.3.1.1 Mécanique du point

Considérons une particule de masse  $m$  qui évolue dans un potentiel  $V(\mathbf{r})$ . L'équation de Newton pour l'évolution de sa position est  $m\ddot{\mathbf{r}}(t) = -\nabla V(\mathbf{r}(t))$ . Nous allons voir que cette équation peut être dérivée du Lagrangien

$$L(\mathbf{r}, \dot{\mathbf{r}}) = \frac{m}{2} \dot{\mathbf{r}}^2 - V(\mathbf{r}). \quad (2.46)$$

Dans ce contexte, on appelle *action* la fonctionnelle  $S[\mathbf{r}] = \int_{t_1}^{t_2} L(\mathbf{r}(t), \dot{\mathbf{r}}(t)) dt$ . Selon le *principe de moindre action*, la trajectoire de la particule minimise l'action.

#### 2.3.1.2 Optique géométrique

En optique géométrique, on peut décrire le cheminement d'un rayon lumineux par une fonction  $\mathbf{r}(s)$ , où  $s$  est l'abscisse curviligne, qui est régie par les lois de Snell-Descartes. On peut montrer que ces lois reviennent à minimiser la longueur optique

$$L[\mathbf{r}] = \int n(\mathbf{r}(s)) ds, \quad (2.47)$$

où  $n(\mathbf{r})$  est l'indice optique du milieu.

#### 2.3.1.3 Forme d'un film de savon

Un film de savon est un objet bidimensionnel décrit par une fonction  $\mathbf{r}(u, v)$ , qui cherche à minimiser son énergie de surface, qui est proportionnelle à son aire

$$A[\mathbf{r}] = \int |\partial_u \mathbf{r}(u, v) \times \partial_v \mathbf{r}(u, v)| du dv. \quad (2.48)$$

On peut parfois décrire le film par la fonction de hauteur  $h(x, y)$ , ce qui correspond à  $\mathbf{r}(x, y) = (x, y, h(x, y))$ . L'aire est alors donnée par

$$A[h] = \int \sqrt{1 + [\nabla h(x, y)]^2} dx dy. \quad (2.49)$$

Enfin, si la pente est faible,  $|\nabla h(x, y)| \ll 1$ , on peut faire l'approximation  $\sqrt{1 + [\nabla h(x, y)]^2} \simeq 1 + \frac{1}{2}[\nabla h(x, y)]^2$  et il faut minimiser

$$A'[h] = \frac{1}{2} \int [\nabla h(x, y)]^2 dx dy. \quad (2.50)$$

### 2.3.2 Équation d'Euler-Lagrange, conditions au bord

#### 2.3.2.1 Fonction inconnue $f : \mathbb{R} \rightarrow \mathbb{R}^d$

On cherche ici à déterminer la fonction  $f : \mathbb{R} \rightarrow \mathbb{R}^d$  qui minimise la fonctionnelle

$$I[f] = \int_a^b L(x, f(x), f'(x)) dx, \quad (2.51)$$

avec cette fois  $L : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ . L'approche est la même qu'en section 2.2.2, mis à part que les dérivées partielles de  $L(x, y, y')$  par rapport à  $y$  et  $y'$  sont remplacées par des gradients, on obtient donc

$$\nabla_y L(x, f(x), f'(x)) = \frac{d}{dx} [\nabla_{y'} L(x, f(x), f'(x))]. \quad (2.52)$$

On obtient aussi la même chose pour les conditions au bord : si  $f(x)$  n'est pas fixée, il faut  $\nabla_{y'} L(x, f(x), f'(x)) = 0$ .

L'existence d'une intégrale du mouvement pour un problème invariant par translation vue en Sec. 2.2.4 se généralise immédiatement à une fonction vectorielle : si le lagrangien ne dépend pas de la variable d'espace,  $L(x, y, y') = L(y, y')$ , alors

$$y' \cdot \nabla_{y'} L(y, y') - L(y, y') \quad (2.53)$$

est une intégrale du mouvement.

#### 2.3.2.2 Application à la mécanique du point

On peut maintenant écrire l'équation d'Euler-Lagrange associée au lagrangien (2.46) :

$$\nabla_y L(x, y, y') = -\nabla V(y) \quad (2.54)$$

$$\nabla_{y'} L(x, y, y') = m y'. \quad (2.55)$$

L'équation d'Euler-Lagrange (2.52) s'écrit donc

$$-\nabla V(r(t)) = \frac{d}{dt} [m \dot{r}(t)] = m \ddot{r}(t) : \quad (2.56)$$

il s'agit bien de l'équation de Newton.

Si le problème est indépendant du temps, le lagrangien ne dépend pas explicitement du temps : le problème admet l'intégrale du mouvement (2.53) qui s'écrit

$$\dot{r} \cdot \nabla_r L(r, \dot{r}) - L(r, \dot{r}) = \frac{m}{2} \dot{r}^2 + V(r). \quad (2.57)$$

On reconnaît l'énergie mécanique du système : sa conservation peut être déduite de l'invariance du problème dans le temps de façon très générale.

### 2.3.2.3 Fonction inconnue $f : \mathbb{R}^d \rightarrow \mathbb{R}$

On considère maintenant une fonction inconnue  $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}$ , la fonctionnelle s'écrit donc

$$I[f] = \int_{\Omega} L(\mathbf{r}, f(\mathbf{r}), \nabla f(\mathbf{r})) d\mathbf{r}, \quad (2.58)$$

avec  $L : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ . Calculons la variation de cette fonctionnelle pour une perturbation  $\epsilon(\mathbf{r})$  (toujours à l'ordre  $\epsilon$ ) :

$$I[f + \epsilon] = I[f] + \int_{\Omega} [\epsilon(\mathbf{r}) \partial_y L(\mathbf{r}, f(\mathbf{r}), \nabla f(\mathbf{r})) + \nabla \epsilon(\mathbf{r}) \cdot \nabla_{\nabla y} L(\mathbf{r}, f(\mathbf{r}), \nabla f(\mathbf{r}))] d\mathbf{r} \quad (2.59)$$

$$= I[f] + \int_{\partial\Omega} \epsilon(\mathbf{r}) \mathbf{n}(\mathbf{r}) \cdot \nabla_{\nabla y} L(\mathbf{r}, f(\mathbf{r}), \nabla f(\mathbf{r})) d\mathbf{r} \quad (2.60)$$

$$+ \int_{\Omega} \epsilon(\mathbf{r}) [\partial_y L(\mathbf{r}, f(\mathbf{r}), \nabla f(\mathbf{r})) - \nabla \cdot \nabla_{\nabla y} L(\mathbf{r}, f(\mathbf{r}), \nabla f(\mathbf{r}))] d\mathbf{r}. \quad (2.61)$$

L'équation d'Euler-Lagrange s'écrit donc

$$\partial_y L(\mathbf{r}, f(\mathbf{r}), \nabla f(\mathbf{r})) = \nabla \cdot \nabla_{\nabla y} L(\mathbf{r}, f(\mathbf{r}), \nabla f(\mathbf{r})), \quad (2.62)$$

et la condition sur le bord  $\partial\Omega$  si  $f(\mathbf{r})$  n'est pas imposée est

$$\mathbf{n}(\mathbf{r}) \cdot \nabla_{\nabla y} L(\mathbf{r}, f(\mathbf{r}), \nabla f(\mathbf{r})) = 0. \quad (2.63)$$

**Remarque :** Le cas le plus général est celui d'une fonction inconnue  $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ . La méthode pour obtenir l'équation d'Euler-Lagrange est la même, mais il faut manipuler des tenseurs. Une notation avec des indices est alors plus adaptée que la notation vectorielle utilisée ici.

### 2.3.2.4 Application à la forme d'un film de savon

On veut ici minimiser l'aire d'un film de savon décrit par une fonction de hauteur  $h(x, y)$ , Éq. (2.49), qui correspond au lagrangien  $L(\mathbf{r}, y, \nabla y) = \sqrt{1 + (\nabla y)^2}$ . On a

$$\partial_y L(\mathbf{r}, y, \nabla y) = 0, \quad (2.64)$$

$$\nabla_{\nabla y} L(\mathbf{r}, y, \nabla y) = \frac{\nabla y}{\sqrt{1 + (\nabla y)^2}}. \quad (2.65)$$

L'équation d'Euler-Lagrange (2.62) s'écrit donc

$$0 = \nabla \cdot \left[ \frac{\nabla h(\mathbf{r})}{\sqrt{1 + [\nabla h(\mathbf{r})]^2}} \right]. \quad (2.66)$$

On obtient une EDP non linéaire.

Dans l'approximation de faible pente, avec la fonctionnelle (2.50), l'équation d'Euler-Lagrange donne simplement

$$\nabla^2 h(\mathbf{r}) = 0. \quad (2.67)$$

On obtient cette fois une EDP linéaire, qui est l'équation de Poisson homogène. À partir de la forme de l'équation d'Euler-Lagrange, on voit que si le laplacien est quadratique en la fonction inconnue, on obtient une équation (EDO ou EDP) linéaire.

## 2.4 Pour aller plus loin

### 2.4.1 Calcul variationnel plus général

Nous avons d'abord exposé le calcul variationnel pour des fonctions inconnues  $\mathbb{R} \rightarrow \mathbb{R}$ , puis nous l'avons étendu à des dimensions plus grandes de l'espace de départ ou de l'espace d'arrivée. Il est bien sûr possible de l'étendre à des dimensions quelconques des deux espaces. L'intégrale du mouvement déduite de l'invariance par translation dans l'espace de départ peut être étendue à des invariances par symétries continues (translations, rotations) de l'espace d'arrivée : c'est le *théorème de Noether*. En mécanique, ce théorème montre la conservation de la quantité de mouvement pour un problème invariant par translation, et du moment cinétique pour un problème invariant par rotation. Les symétries continues peuvent aussi combiner les deux espaces, on peut par exemple faire de la mécanique du point avec un potentiel tournant.

Dans ce que nous avons traité, l'énergie ne dépend que de la fonction inconnue et de sa dérivée première (ou gradient). Il se peut toutefois que l'énergie fasse intervenir des dérivées secondes, comme par exemple l'énergie élastique de courbure d'une plaque. Il faut alors généraliser le calcul variationnel pour incorporer ces dérivées d'ordre supérieur. L'équation d'Euler-Lagrange fait alors intervenir des dérivées (éventuellement partielles) d'ordre supérieur à deux, de manière générale deux fois l'ordre maximal des dérivées présentes dans le Lagrangien.

### 2.4.2 Physique statistique

En physique statistique, la densité d'une grandeur  $x$  d'énergie  $E(x)$  à l'équilibre dans l'ensemble canonique est donnée par la distribution de Gibbs-Boltzmann

$$f(x) \propto e^{-\beta E(x)}, \quad (2.68)$$

où  $\beta = (k_B T)^{-1}$ ,  $T$  étant la température.

Si l'état du système est caractérisé non par une variable réelle  $x$  mais par une fonction  $y(x)$ , d'énergie  $E[y]$ , la densité « dans l'espace des fonctions » est donnée par

$$f[y] \propto e^{-\beta E[y]}. \quad (2.69)$$

À température nulle, c'est à dire quand  $\beta \rightarrow \infty$ , la densité est dominée par la configuration  $y_0(x)$  qui minimise  $E[y]$ .

La formulation du problème classique en terme d'énergie à minimiser permet ainsi de donner la probabilité d'occupation des états quand la température est positive.

### 2.4.3 Mécanique quantique

En mécanique quantique, l'équation du mouvement de Newton pour la position  $\mathbf{r}(t)$  d'un point matériel est remplacée par l'équation de Schrödinger pour la fonction d'onde  $\psi(\mathbf{r}, t)$  (Éq. (1.144)). En dehors de la « dérivation » de l'équation de Schrödinger avec les opérateurs position et impulsion, le lien entre ces deux équations n'est pas évident et il n'est pas facile de montrer que l'équation de Newton peut être obtenue comme limite classique de celle de Schrödinger.

La formulation de la mécanique quantique à partir de l'intégrale de chemin, introduite par Feynman, permet de résoudre ce problème et d'établir un lien clair entre la mécanique quantique et la mécanique classique. En effet, Feynman a montré que la fonction suivante, définie par une somme sur

les « chemins »

$$K(\mathbf{r}_1, t_1; \mathbf{r}_0, t_0) = \int_{\mathbf{r}(t_0)=\mathbf{r}_0}^{\mathbf{r}(t_1)=\mathbf{r}_1} \exp\left(\frac{i}{\hbar} S[\mathbf{r}]\right) [d\mathbf{r}(t)], \quad (2.70)$$

où  $S[\mathbf{r}]$  est l'action de la mécanique du point (voir Sec. 2.3.1.1), obéit à l'équation de Schrödinger avec la condition initiale  $\psi(\mathbf{r}, t_0) = \delta(\mathbf{r} - \mathbf{r}_0)$ .

Quand  $\hbar \rightarrow 0$ , seule la trajectoire  $\mathbf{r}(t)$  qui rend l'action  $S[\mathbf{r}]$  stationnaire contribue, c'est à dire la trajectoire qui minimise l'action : on retrouve le principe de moindre action, et donc la mécanique classique comme limite de la mécanique quantique.



## Chapitre 3

# Probabilités

### 3.1 Introduction

Il est fréquent en physique ou en biologie que le résultat d'une expérience contienne un part d'aléa. Le plus souvent, l'aléa provient d'une connaissance incomplète du système (par exemple la position et l'impulsion des molécules d'un liquide, qui déterminent le mouvement d'un colloïde immergé); parfois, comme en mécanique quantique, il est dû à la nature du système et de l'observable considérée (le trou par lequel passe le photon dans l'expérience des fentes d'Young). Il est donc important de savoir formaliser le caractère aléatoire du résultat et de savoir manipuler les résultats aléatoires : c'est l'enjeu de la théorie des probabilités.

Dans ce chapitre, nous introduisons les bases de la théorie des probabilités. Nous commençons avec les notions ensemblistes d'événement et de mesure de probabilité, qui empruntent beaucoup à la théorie de la mesure vue en première année. Puis nous introduisons les variables aléatoires, qui constituent l'élément central de ce cours, et comment les manipuler; nous voyons notamment comment la théorie de la mesure permet de décrire les variables aléatoires continues et discrètes de façon unifiée. Nous voyons enfin les propriétés des suites de variables aléatoires indépendantes, qui obéissent à la loi des grands nombres et au théorème de la limite centrale.

### 3.2 Exemples

Commençons par donner quelques questions auxquelles ce cours permettra de répondre :

- Une maladie touche une personne sur  $10^4$ . Un test qui permet de la dépister est positif avec probabilité 99 % si le patient est malade et 1 % s'il est sain. Si un patient a un test positif, qu'elle est la probabilité qu'il soit malade ?
- Un atome radioactif a une probabilité  $\delta t/\tau$  de se désintégrer pendant un temps  $\delta t$ . Quelle est la distribution du temps de survie  $T$  de cet atome ? Comment « simuler » numériquement ce temps de survie ?
- Sur un échantillon de 2000 personnes, 48 % voteront pour le candidat  $A$  et 52 % pour le candidat  $B$ . Que peut-on dire sur le résultat final du vote ?

### 3.3 Événements et probabilité

#### 3.3.1 Mesure de probabilité

On considère une expérience aléatoire, dont on note  $\Omega$  l'ensemble des résultats possibles, aussi appelé *espace des états*. Un *événement* est un sous-ensemble  $A \subset \Omega$ . On veut maintenant associer une probabilité  $P(A) \in [0, 1]$  à l'événement  $A$ . La théorie de la mesure fournit le cadre approprié pour définir une probabilité. Nous rappelons rapidement les notions de *tribu* et de *mesure* qui ont été vues en première année.

Une tribu  $\mathcal{A}$  sur  $\Omega$  est un sous-ensemble des parties de  $\Omega$ ,  $\mathcal{A} \subset \mathcal{P}(\Omega)$ , qui vérifie les propriétés suivantes

- $\emptyset \in \mathcal{A}$ ,
- Si  $A \in \mathcal{A}$ , alors  $A^c \in \mathcal{A}$ .
- Si  $(A_n)_{n \in \mathbb{N}}$  est une suite d'éléments de  $\mathcal{A}$ ,  $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$  et  $\bigcap_{n \in \mathbb{N}} A_n \in \mathcal{A}$ .

Une mesure  $\mu$  est une application  $\mathcal{A} \rightarrow \mathbb{R}_+$  telle que  $\mu(\emptyset) = 0$  et pour toute suite  $(A_n)_{n \in \mathbb{N}}$  d'éléments de  $\mathcal{A}$  deux-à-deux disjoints ( $\forall i, j \in \mathbb{N}, i \neq j, A_i \cap A_j = \emptyset$ ),

$$\mu\left(\bigsqcup_{n \in \mathbb{N}} A_n\right) = \sum_{n=0}^{\infty} \mu(A_n), \quad (3.1)$$

où  $\sqcup$  désigne la réunion d'ensembles disjoints.

On définit alors une *probabilité*  $P$  comme une mesure de masse totale 1, c'est à dire  $P(\Omega) = 1$ . Le triplet  $(\Omega, \mathcal{A}, P)$  est un *espace probabilisé*.

#### Remarques :

- Quand  $\Omega$  est dénombrable, on prend souvent  $\mathcal{A} = \mathcal{P}(\Omega)$ , et on définit la probabilité pour tous les événements possibles. Quand  $\Omega$  n'est pas dénombrable, il est pratique d'avoir une tribu plus petite.
- Les propriétés naturelles de la probabilité découlent naturellement des propriétés de la mesure et de  $P(\Omega) = 1$ . Par exemple  $P(\emptyset) = 0$ ,  $P(\Omega) = 1$ ,  $P(A \sqcup B) = P(A) + P(B)$ ,  $P(A) + P(A^c) = 1$ , etc.

#### 3.3.2 Conditionnement et indépendance

On définit la *probabilité conditionnelle* de  $A$  sachant  $B \neq \emptyset$ ,  $P(A|B)$ , par

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \quad (3.2)$$

On note que la probabilité « sachant  $B$  » vérifie toutes les propriétés d'une probabilité.

On a immédiatement la *formule des probabilités totales* : si  $(B_n)_{n \in \mathbb{N}}$  est une partition de  $\Omega$  (ou *système complet d'événements*), alors

$$P(A) = \sum_{n \in \mathbb{N}} P(A \cap B_n) = \sum_{n \in \mathbb{N}} P(A|B_n)P(B_n). \quad (3.3)$$

De cette formule, on peut déduire la *formule de Bayes* :

$$P(B_n|A) = \frac{P(A \cap B_n)}{P(A)} = \frac{P(A|B_n)P(B_n)}{\sum_k P(A|B_k)P(B_k)}. \quad (3.4)$$

Cette formule est intéressante car elle permet « d'inverser » le conditionnement : si on connaît  $P(B_n)$  et  $P(A|B_n)$  pour tout  $n$ , alors on peut tout calculer, et en particulier  $P(B_n|A)$ .

On dit que deux événements  $A$  et  $B$  sont *indépendants* si  $P(A \cap B) = P(A)P(B)$ . Cette définition s'explique par le fait qu'alors  $P(A|B) = P(A \cap B)/P(B) = P(A)$ , c'est à dire que la connaissance de  $B$  ne modifie pas la probabilité de  $A$ .

On dit que des événements  $(A_n)_{n \in \mathbb{N}}$  sont indépendants si pour tout  $I \subset \mathbb{N}$  fini,

$$P\left(\bigcap_{n \in I} A_n\right) = \prod_{n \in I} P(A_n). \quad (3.5)$$

Attention : l'indépendance deux à deux n'entraîne pas l'indépendance.

**Exemple :** Un exemple classique d'application de la formule de Bayes est celui du test médical donné en exemple. On considère une maladie qui touche une fraction  $f$  de la population. Un test qui permet de la dépister est juste avec probabilité  $j_s$  si le patient est sain et  $j_m$  si le patient est malade. La question que l'on se pose est : si le test est positif, qu'elle est la probabilité d'être malade ?

En notant  $M$  l'événement « l'individu est malade »,  $S$  « l'individu est sain »,  $P$  « le test est positif » et  $N$  « le test est négatif », les données sont  $P(M) = f$ ,  $P(P|M) = j_m$ ,  $P(N|S) = j_s$ . On connaît donc toutes les probabilités conditionnées par la famille complète  $(M, S)$ , et on cherche  $P(M|P)$ . On a donc

$$P(M|P) = \frac{P(M \cap P)}{P(P)} = \frac{P(P|M)P(M)}{P(P|M)P(M) + P(P|S)P(S)} = \frac{j_m f}{j_m f + (1 - j_s)(1 - f)}. \quad (3.6)$$

En général,  $f, 1 - j_m, 1 - j_s \ll 1$ , et le dénominateur est  $j_m f + (1 - j_s)(1 - f) \simeq f + (1 - j_s)$  : un test positif peut venir d'un individu malade ou d'un faux positif. Si  $f \gg 1 - j_s$ , l'individu a de grandes chances d'être malade alors que si  $f \ll 1 - j_s$  il a peu de chances de l'être.

## 3.4 Variables aléatoires

### 3.4.1 Définition

Une *variable aléatoire* (VA) est une grandeur qui dépend du résultat de l'expérience aléatoire, c'est donc une application  $X : \Omega \rightarrow F$ , où  $F$  peut être  $\mathbb{N}$ ,  $\mathbb{R}$ ,  $\mathbb{R}^d$ , etc.

Une variable aléatoire permet de « transporter » sur  $F$  la structure d'espace probabilisé sur  $\Omega$ ,  $(\Omega, \mathcal{A}, P)$ . En effet,

$$\mathcal{F} = \{B \subset F | X^{-1}(B) \in \mathcal{A}\} \quad (3.7)$$

définit bien une tribu sur  $F$ , et l'application

$$P_X : \begin{cases} \mathcal{F} & \rightarrow \mathbb{R} \\ B & \mapsto P(X^{-1}(B)) \end{cases} \quad (3.8)$$

définit une probabilité sur  $(F, \mathcal{F})$  ; c'est la *loi* de la variable aléatoire  $X$ . En général, on note  $(X \in B) = X^{-1}(B) = \{\omega \in \Omega | X(\omega) \in B\}$  ; on écrit ainsi  $P(X \in B)$ . La probabilité  $P_X$  est la probabilité *image* (ou *induite*) de  $P$  par  $X$ . Cela s'applique à toutes les mesures, et pas seulement aux probabilités.

On travaille plus souvent sur des VA que sur des événements sur  $\Omega$ , car l'espace d'arrivée est plus facile à décrire. Il est toutefois bon de garder en tête qu'une VA est en fait une fonction.

De même que pour une seule variable aléatoire, la loi d'un couple de VA  $(X, Y)$  sur  $F \times G$  est une probabilité  $P_{(X, Y)}$  sur  $F \times G$ . La *loi marginale* de la variable  $X$  est définie par  $P_X(A) = P_{(X, Y)}(A \times G)$ .

Deux variables aléatoires  $X$  et  $Y$  sur  $F$  et  $G$  sont indépendantes si pour tous  $A \subset F$ ,  $B \subset G$ , les événements  $X \in A$  et  $Y \in B$  sont indépendants. Cela revient à dire que la mesure  $P_{(X, Y)}$  induite par la

VA  $(X, Y)$  sur  $F \times G$  est égale au produit des mesures induites  $P_X$  et  $P_Y$  : la loi de  $(X, Y)$  est le produit des lois de  $X$  et  $Y$ . On définit de même que pour les événements l'indépendance d'une famille de VA  $(X_n)_{n \in \mathbb{N}}$ .

Soit  $f : F \rightarrow G$  une fonction, alors  $f(X) = f \circ X$  est une variable aléatoire sur  $G$ .

### 3.4.2 VA réelles

#### 3.4.2.1 Fonction de répartition, tirage numérique de VA

Pour une VA  $X$  à valeur dans  $\mathbb{R}$ , on définit sa *fonction de répartition*  $F_X(x)$  par

$$F_X(x) = P(X \leq x) = P_X([-\infty, x]). \quad (3.9)$$

La fonction de répartition est croissante,  $\lim_{x \rightarrow -\infty} F_X(x) = 0$  et  $\lim_{x \rightarrow \infty} F_X(x) = 1$ .

On a souvent besoin de générer numériquement des valeurs de VA suivant une loi donnée. Or, les générateurs aléatoires génèrent des nombres uniformément répartis sur l'intervalle  $[0, 1[$ . Comment utiliser ces nombres pour obtenir des nombres aléatoires suivant la loi recherchée? Notons  $F(x)$  la fonction de répartition de la loi recherchée, et  $R$  une variable aléatoire de loi uniforme sur  $[0, 1]$  :

$$P_R = \chi_{[0,1]} \lambda, \quad (3.10)$$

où  $\lambda$  est la mesure de Lebesgue.

On cherche une fonction  $g(r)$  telle que la VA  $X = g(R)$  ait la fonction de répartition  $F(x)$ , c'est à dire

$$F(x) = P(g(R) \leq x) = P(R \leq g^{-1}(x)) = g^{-1}(x), \quad (3.11)$$

où nous avons supposé que la fonction  $g(r)$  était croissante et admettait une fonction réciproque  $g^{-1}(x)$ , et  $g^{-1}(x) \in [0, 1]$ . Il faut donc prendre  $g(r) = F^{-1}(r)$ . On peut maintenant déterminer quand il est possible de faire ça.  $F$  est une fonction croissante  $\mathbb{R} \rightarrow [0, 1]$ , sa fonction réciproque est donc  $F^{-1} : [0, 1] \rightarrow \mathbb{R}$ , ce qui est bien ce que nous recherchons ; elle est de plus croissante. Enfin, on peut se convaincre que les difficultés qui peuvent apparaître pour définir la fonction réciproque (discontinuité de  $F(x)$ , intervalle où  $F(x)$  est constante) ne concernent que des intervalles de mesure nulle et n'ont donc aucun effet sur la loi de  $X$ . Pour résumer, nous avons montré que si  $R$  était une VA de loi uniforme sur  $[0, 1]$ , alors  $X = F^{-1}(R)$  admettait la fonction de répartition  $F(x)$ .

#### 3.4.2.2 Espérance, variance

La probabilité  $P_X$  est une mesure sur  $\mathbb{R}$ , on peut donc l'utiliser pour intégrer. L'espérance est définie par

$$E(X) = \int_{\Omega} X(\omega) dP(\omega) = \int_{\mathbb{R}} x dP_X(x) \quad (3.12)$$

si cette intégrale existe. L'égalité de ces deux intégrales vient de la définition de la mesure image.

On peut calculer l'espérance de  $f(X)$  par

$$E(f(X)) = \int_{\Omega} f(X(\omega)) dP(\omega) = \int_{\mathbb{R}} f(x) dP_X(x). \quad (3.13)$$

On définit le *moment* d'ordre  $n$  de  $X$  par  $E(X^n)$ . À l'ordre 2, on introduit aussi la variance

$$\text{Var}(X) = E([X - E(X)]^2) = E(X^2) - E(X)^2. \quad (3.14)$$

L'espérance contient l'information sur les probabilités des événements  $X \in A$  : en effet, si  $\chi_A$  est la fonction indicatrice de  $A$ ,

$$E(\chi_A(X)) = \int_{\mathbb{R}} \chi_A(x) dP_X(x) = \int_A dP_X(x) = P_X(A) = P(X \in A). \quad (3.15)$$

Dans le cas où  $A = ]-\infty, x]$ , on retrouve la fonction de répartition

$$E(\chi_{]-\infty, x]}(X)) = P(X \in ]-\infty, x]) = P(X \leq x) = F_X(x). \quad (3.16)$$

On peut caractériser l'indépendance de deux VA avec l'espérance :  $X$  et  $Y$  sont indépendantes si et seulement si pour toutes fonctions  $f(x)$  et  $g(y)$ ,

$$E(f(X)g(Y)) = E(f(X)) \times E(g(Y)). \quad (3.17)$$

Cela se montre facilement :

- Si cette propriété est vraie, alors en utilisant la relation (3.15), on montre que  $X$  et  $Y$  sont indépendantes.
- Réciproquement, si  $X$  et  $Y$  sont indépendantes, alors cette propriété est vraie pour les fonctions indicatrices, puis pour les fonctions étagées, puis pour toutes les fonctions par définition de l'intégrale de Lebesgue.

On définit la *covariance* d'un couple de VA  $(X, Y)$  par

$$\text{Cov}(X, Y) = E([X - E(X)][Y - E(Y)]) = E(XY) - E(X)E(Y). \quad (3.18)$$

D'après la propriété (3.17), la covariance de deux VA indépendantes est nulle (la réciproque est bien sûr fausse). Le coefficient de corrélation linéaire des VA  $X$  et  $Y$  est défini par

$$f(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}. \quad (3.19)$$

En général,  $|f(X, Y)| \leq 1$  ; le cas d'égalité correspond à  $|f(X, aX + b)| = 1$ . La variance de la somme  $X + Y$  est donnée par

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y). \quad (3.20)$$

Ainsi, si  $X$  et  $Y$  sont indépendantes, la variance de la somme est égale à la somme des variances.

### 3.4.2.3 Inégalités

On montre ici quelques inégalités utiles faisant intervenir l'espérance ou la variance d'une VA  $X$ . Pour  $a > 0$ , on peut écrire  $|X| \geq a \chi_{|X| \geq a}$ . En prenant la puissance  $p$  et l'espérance, on obtient

$$E(|X|^p) \geq a^p E(\chi_{|X| \geq a}) = a^p P(|X| \geq a), \quad (3.21)$$

et on obtient l'*inégalité de Markov* :

$$P(|X| \geq a) \leq \frac{E(|X|^p)}{a^p}. \quad (3.22)$$

En appliquant le même raisonnement pour l'écart à la moyenne,  $X - E(X)$ , avec  $p = 2$ , on obtient l'*inégalité de Bienaymé-Tchebychev* :

$$P(|X - E(X)| \geq a) \leq \frac{E([X - E(X)]^2)}{a^2} = \frac{\text{Var}(X)}{a^2}. \quad (3.23)$$

Enfin, on sait que pour une fonction convexe  $f(x)$  et deux réels  $x_1$  et  $x_2$ , on a  $[f(x_1) + f(x_2)]/2 \geq f([x_1 + x_2]/2)$ . Cette inégalité se généralise à l'espérance : si  $f(x)$  est convexe, pour tout  $a$  il existe  $\lambda$  tel que pour tout  $x$ ,  $f(x) \geq f(a) + \lambda(x - a)$ . Avec  $a = E(X)$ , on obtient donc  $f(X) \geq f(E(X)) + \lambda[X - E(X)]$ . En prenant l'espérance, on obtient l'*inégalité de Jensen* :

$$E(f(X)) \geq f(E(X)). \quad (3.24)$$

### 3.4.2.4 Fonction caractéristique

La fonction caractéristique de la VA réelle  $X$  est définie par :

$$\phi_X(k) = \mathbb{E}\left(e^{ikX}\right). \quad (3.25)$$

Comme  $|e^{ikX}| \leq 1$ , la fonction caractéristique est définie pour tout  $k \in \mathbb{R}$ . La fonction caractéristique est la transformée de Fourier de la loi de  $X$  et caractérise donc complètement la loi de  $X$ .

Par dérivation, la fonction caractéristique permet de retrouver les moments, en effet

$$\phi_X^{(n)}(k) = \mathbb{E}\left([iX]^n e^{ikX}\right), \quad (3.26)$$

donc

$$\phi_X^{(n)}(0) = i^n \mathbb{E}(X^n), \quad (3.27)$$

La fonction caractéristique vérifie aussi  $\phi_X(0) = 1$ .

Si deux VA  $X$  et  $Y$  sont indépendantes, la fonction caractéristique de leur somme vérifie

$$\phi_{X+Y}(k) = \mathbb{E}\left(e^{ik[X+Y]}\right) = \mathbb{E}\left(e^{ikX} e^{ikY}\right) = \mathbb{E}\left(e^{ikX}\right) \mathbb{E}\left(e^{ikY}\right) = \phi_X(k) \phi_Y(k). \quad (3.28)$$

Nous avons utilisé la propriété (3.17).

### 3.4.3 VA discrètes et continues, lois classiques

#### 3.4.3.1 VA discrètes

Une VA *discrète*  $X$  a des valeurs dans un ensemble fini ou dénombrable  $F = \{x_i\}_{i \in I} \subset \mathbb{R}$ , avec  $I \subset \mathbb{N}$ ; par exemple  $F = \{1, \dots, N\}$  ou  $F = \mathbb{N}$ . Dans ce cas, sa loi est caractérisée par les probabilités  $p_i = \mathbb{P}(X = x_i)$  et on peut écrire

$$\mathbb{P}_X = \sum_{i \in I} p_i \delta_{x_i}, \quad (3.29)$$

où  $\delta_{x_i}$  est la mesure de Dirac en  $x_i$  :  $\delta_{x_i}(B) = 1$  si  $x_i \in B$ , et vaut 0 sinon. Alors l'espérance de  $f(X)$  s'écrit

$$\mathbb{E}(f(X)) = \sum_{i \in I} p_i f(x_i). \quad (3.30)$$

Quand la VA est à valeurs dans  $\mathbb{N}$ , on définit plutôt sa fonction caractéristique par

$$G_X(s) = \mathbb{E}\left(s^X\right) = \sum_{n \in \mathbb{N}} p_n s^n, \quad (3.31)$$

à la place de la définition donnée en section 3.4.2.4. Elle est liée à la définition précédente par  $\phi_X(k) = G_X(e^{ik})$ , et elle vérifie des propriétés analogues :

- $G_X(1) = 1$ ,
- $G_X^{(n)}(1) = \mathbb{E}(X[X-1] \dots [X-n+1])$ ,
- si  $X$  et  $Y$  sont indépendantes,  $G_{X+Y}(s) = G_X(s)G_Y(s)$ .

#### 3.4.3.2 VA continues

Une VA est continue, ou « à densité », s'il existe une fonction  $f_X : \mathbb{R} \rightarrow \mathbb{R}$  telle que la fonction de répartition de  $X$  s'écrive

$$F_X(x) = \int_{-\infty}^x f_X(y) dy. \quad (3.32)$$

La fonction  $f_X$  est la *densité*, ou densité de probabilité, de  $X$ . Cela revient à dire que  $P_X$  admet la densité  $f$  par rapport à la mesure de Lebesgue  $\lambda$  :  $dP_X(x) = f(x)d\lambda(x)$ . Alors, pour une fonction  $g(x)$ , l'espérance de  $g(X)$  s'écrit

$$E(g(X)) = \int_{\mathbb{R}} g(x)f_X(x)dx. \quad (3.33)$$

**Remarque :** Il existe bien sûr des VA qui ne sont ni discrètes ni continues. Par exemple, notons  $T$  le temps passé par un élève dans la salle pour un examen de durée  $t_e$ .  $T$  est une VA à valeurs dans  $[0, t_e]$ , mais  $T = t_e$  avec une probabilité finie. On s'attend donc à une loi de la forme

$$P_T = f\lambda + p_{t_e}\delta_{t_e}, \quad (3.34)$$

où  $f(t)$  est une fonction continue.  $T$  est donc une VA « mixte », dont la loi contient une partie continue et une partie discrète. La normalisation impose  $\int_0^{t_e} f(t')dt' + p_{t_e} = 1$ , et la fonction de répartition est

$$F_T(t) = \begin{cases} 0 & t < 0 \\ \int_0^t f(t')dt' & 0 < t < t_e \\ 1 & t \geq t_e \end{cases}. \quad (3.35)$$

### 3.4.3.3 Lois classiques

On donne ici des lois qui apparaissent souvent, avec la situation dans laquelle elles apparaissent et leurs principales propriétés. On commence avec des lois discrètes et on termine avec des lois continues.

Une VA qui peut prendre  $n$  valeurs différentes  $\{1, \dots, n\}$  suit une *loi uniforme* si

$$P(X = i) = 1/n, \quad (3.36)$$

pour tout  $i$ . On parle aussi de probabilité uniforme sur un espace d'états  $\Omega$  fini si  $P(\{\omega\}) = 1/\text{Card}(\Omega)$  pour tout  $\omega \in \Omega$ . Des lois uniformes apparaissent souvent dans des problèmes de dénombrement : le résultat d'un lancer de dé non biaisé suit une loi uniforme sur  $\{1, \dots, 6\}$ , la main que l'on obtient aux cartes suit une loi uniforme sur toutes les mains possibles, etc.

Une VA  $X$  suit une *loi de Bernouilli* si elle est à valeurs dans  $\{0, 1\}$ . Elle est alors caractérisée par la probabilité  $P(X = 1) = p$ . Ses moments sont donnés par  $E(X^n) = p$  pour  $n > 0$  et sa variance est  $\text{Var}(X) = p(1 - p)$ . Cette loi apparaît dès que seules deux issues sont possibles : lancer d'une pièce à pile ou face, loi du vote pour une élection à deux candidats, succès ou échec à un jeu.

Une VA  $X$  à valeurs dans  $\{0, \dots, n\}$  suit une *loi binomiale* de paramètres  $n$  et  $p$  si

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad (3.37)$$

On la note  $\mathcal{B}(n, p)$ . C'est la somme de  $n$  VA de Bernouilli indépendantes. Son espérance et sa variance valent donc  $E(X) = np$ ,  $\text{Var}(X) = np(1 - p)$ . Elle représente le nombre de succès à un jeu après  $n$  tentatives quand la probabilité de succès est  $p$ .

Une VA  $X$  à valeurs dans  $\mathbb{N}_+$  suit une *loi géométrique* si

$$P(X = n) = (1 - p)p^{n-1}, \quad (3.38)$$

avec  $0 < p < 1$ . Son espérance et sa variance valent  $E(X) = 1/(1 - p)$  et  $\text{Var}(X) = p/(1 - p)^2$ . La loi géométrique est la loi du moment du premier succès pour une suite d'expériences indépendantes de probabilité de succès  $1 - p$ .

Une VA  $X$  à valeurs dans  $\mathbb{N}$  suit une *loi de Poisson* de paramètre  $\lambda > 0$  si

$$P(X = n) = e^{-\lambda} \frac{\lambda^n}{n!}. \quad (3.39)$$

Son espérance vaut  $E(X) = \lambda$  et sa variance est  $\text{Var}(X) = \lambda$ . Si  $\rho L$  points sont pris uniformément (voir ci-dessous) dans l'intervalle  $[0, L]$ , alors la loi de Poisson de paramètre  $\rho$  décrit le nombre de points dans l'intervalle  $[0, 1]$  dans la limite  $L \rightarrow \infty$ . C'est par exemple la loi du nombre de noisettes dans un carré de chocolat.

Une VA continue  $X$  à valeurs dans l'intervalle  $[a, b]$  suit une loi uniforme si pour tous  $x < y \in [a, b]$ ,

$$P(X \in [x, y]) = \frac{y - x}{b - a}. \quad (3.40)$$

La probabilité est la restriction de la mesure de Lebesgue à l'intervalle  $[a, b]$ , avec un facteur  $1/(b - a)$ . Elle se généralise à des domaines du plan, à la sphère, etc. On utilise souvent la loi uniforme sur l'intervalle  $[0, 1]$ , d'espérance  $1/2$  et de variance  $1/12$ . Elle représente par exemple la loi du point d'impact sur la cible d'une fléchette lancée par un joueur ivre (conditionnée par le fait que la cible soit atteinte).

Une VA continue  $X$  à valeurs dans  $\mathbb{R}_+$  suit une loi exponentielle de paramètre  $\lambda$  si

$$P(X \geq x) = e^{-\lambda x}. \quad (3.41)$$

Son espérance vaut  $E(X) = 1/\lambda$  et sa variance est  $\text{Var}(X) = 1/\lambda^2$ . Elle représente le temps de vie d'un atome radioactif, ou de façon générale le temps de fin d'un processus qui « ne vieillit pas ». Cela se traduit par le fait que

$$P(X \geq x + x_0 | X \geq x_0) = P(X \geq x) : \quad (3.42)$$

on ne peut pas voir que le processus a déjà « survécu » jusqu'à  $x_0$ .

Une VA continue  $X$  à valeurs dans  $\mathbb{R}$  suit une *loi normale* d'espérance  $\mu$  et de variance  $\sigma^2$  si elle admet la densité

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{[x - \mu]^2}{2\sigma^2}\right). \quad (3.43)$$

On note cette loi  $\mathcal{N}(\mu, \sigma^2)$ ;  $\mathcal{N}(0, 1)$  est la loi normale *centrée réduite*, on note sa densité  $f_{\mathcal{N}}(x)$ . Cette loi apparaît comme une loi limite dans la « moyenne » d'un grand nombre de processus aléatoires indépendants dans le théorème de la limite centrale (Sec. 3.5.3). La fonction de répartition de la loi normale centrée réduite est donnée par la fonction erreur  $\text{erf}(x)$  :

$$F_{\mathcal{N}}(x) = \int_{-\infty}^x f_{\mathcal{N}}(x') dx' = \frac{1}{2} \left[ 1 + \text{erf}\left(\frac{x}{\sqrt{2}}\right) \right] = \Pi(x). \quad (3.44)$$

Cette fonction est très utile pour estimer des intervalles de confiance, comme nous le verrons en section 3.5.4. On utilise aussi sa fonction inverse,

$$\Pi^{-1}(x) = \sqrt{2} \text{erf}^{-1}(2x - 1); \quad (3.45)$$

on peut aussi utiliser des tables de valeurs.

**Exemple :** On montre comment estimer la valeur de  $\pi$  à partir d'une suite de réalisations d'une VA uniforme sur  $[0, 1]$ . Prenons  $X$  et  $Y$  deux variables aléatoires uniformes sur  $[0, 1]$ , alors le couple  $(X, Y)$  est une VA uniforme sur  $[0, 1]^2$ . La probabilité que  $X^2 + Y^2 \leq 1$  est donnée par la mesure de Lebesgue du domaine sur lequel cette inégalité est vérifiée, qui vaut  $\pi/4$  divisée par la mesure de Lebesgue du carré  $[0, 1]^2$ , qui vaut 1 :

$$P(X^2 + Y^2 \leq 1) = \frac{\pi}{4}. \quad (3.46)$$



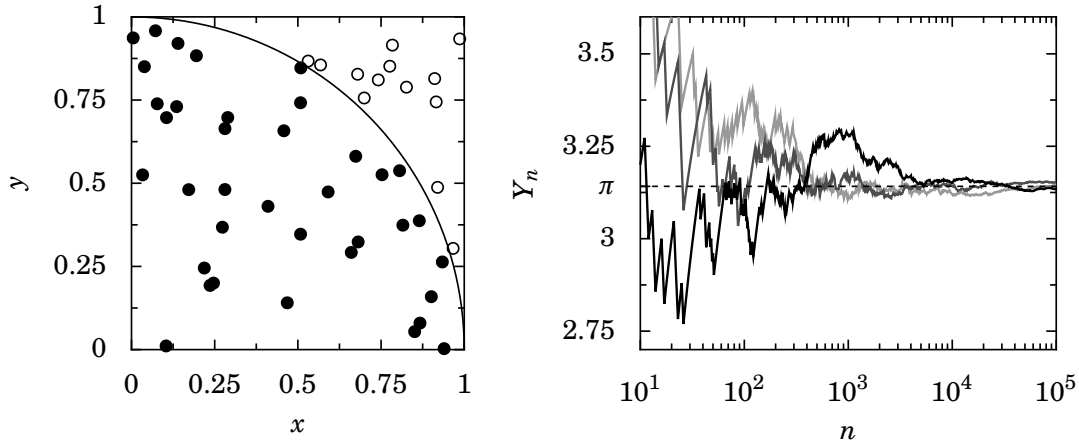


FIGURE 3.1 – Estimation de  $\pi$  à partir de nombres aléatoires de loi uniforme sur  $[0,1]$ . *Gauche* : 50 points aléatoires de loi uniforme sur le carré  $[0,1]^2$  sont tirés, on distingue ceux qui sont à l'intérieur du quart de cercle de ceux qui sont à l'extérieur. *Droite* : en prenant le nombre de points à l'intérieur du cercle divisé par le nombre total de points, on obtient une estimation de  $\pi$ , notée  $Y_n$ . Trois « trajectoires » de l'estimation sont tracées en différents niveaux de gris.

On tire maintenant  $2n$  nombres aléatoires de loi uniforme sur  $[0,1]$ , que l'on note  $((x_k, y_k))_{1 \leq k \leq n}$  (voir Fig. 3.1). Notons  $N_n$  le nombre d'indices  $k$  pour lesquels  $x_k^2 + y_k^2 \leq 1$ , alors  $4N_n/n$  fournit une estimation de  $\pi$ , qui devient meilleure quand  $n$  augmente.

Pour formaliser cela, on introduit

$$Z = 4H(1 - X^2 - Y^2), \quad (3.47)$$

où  $H(x)$  est la fonction de Heaviside, dont la loi est  $P(Z = 0) = 1 - \pi/4$ ,  $P(Z = 4) = \pi/4$ , donc son espérance est  $E(Z) = \pi$ .

## 3.5 Suites de variables aléatoires

### 3.5.1 Convergence de VA

#### 3.5.1.1 Égalité de VA

Avant de discuter des types de convergence des VA, commençons par définir les types d'égalité de VA les plus simples. On dit que deux VA  $X$  et  $Y$  sont égales

- presque sûrement si  $P(X \neq Y) = 0$ ,
- en moyenne si  $E(|X - Y|) = 0$ ,
- en loi si  $P_X = P_Y$ .

Les deux premières égalités sont équivalentes, mais elles ne sont pas équivalentes à la troisième : deux VA égales presque sûrement ont même loi, mais la réciproque n'est pas vraie.  $X$  et  $Y$  sont égales presque sûrement si elles ont la même valeur,  $X(\omega) = Y(\omega)$ , pour toute réalisation  $\omega$  de l'expérience aléatoire, sauf pour un ensemble de mesure nulle de réalisations.  $X$  et  $Y$  peuvent avoir même loi mais être différentes pour toute réalisation. Par exemple, si  $X$  est une VA uniforme sur  $[0,1]$ ,  $1 - X$  a même loi que  $X$ , mais  $P(X \neq 1 - X) = 1$ .

### 3.5.1.2 Types de convergence de VA

On dit qu'une suite de VA  $(X_n)_{n \in \mathbb{N}}$  converge vers la VA  $X$

- presque sûrement si  $\forall \omega \in \Omega, X_n(\omega) \rightarrow X(\omega)$ , on écrit  $X_n \xrightarrow{\text{p.s.}} X$  ;
- en probabilité si  $\forall \epsilon > 0, P(|X_n - X| > \epsilon) \rightarrow 0$ , on écrit  $X_n \xrightarrow{P} X$  ;
- en moyenne si  $E(|X_n - X|) \rightarrow 0$ , on écrit  $X_n \xrightarrow{L^1} X$  ;
- en loi si pour toute fonction  $f(x)$  continue bornée,  $E(f(X_n)) \rightarrow E(f(X))$ , on écrit  $X_n \xrightarrow{\mathcal{L}} X$ .

La convergence presque sûre est l'analogue de la convergence simple vue en première année. Elle peut notamment permettre d'utiliser le théorème de convergence dominée. La convergence en moyenne correspond à la convergence en norme  $L^1$  car  $E(|X_n - X|) = \|X_n - X\|_1$ .

On a les relations suivantes entre ces types de convergence :

- $X_n \xrightarrow{\text{p.s.}} X$  et  $|X_n| \leq Y \ \forall n, Y$  intégrable  $\Rightarrow X_n \xrightarrow{L^1} X$ .
- $X_n \xrightarrow{\text{p.s.}} X \Rightarrow X_n \xrightarrow{P} X$ .
- $X_n \xrightarrow{L^1} X \Rightarrow X_n \xrightarrow{P} X$ .
- $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{\mathcal{L}} X$ .

Démontrons ces relations. La convergence presque sûre  $X_n \xrightarrow{\text{p.s.}} X$  correspond à la convergence simple. Si de plus la suite  $X_n$  est dominée par une VA  $Y$ , on peut alors utiliser le théorème de convergence dominée qui donne  $E(|X_n - X|) \rightarrow 0 : X_n \xrightarrow{L^1} X$ .

Supposons la convergence presque sûre et prenons  $\epsilon > 0$ . On introduit les ensembles  $A_{n,\epsilon} = (|X_n - X| > \epsilon)$ , et  $Y_{n,\epsilon} = \chi_{A_{n,\epsilon}}$  leur fonction indicatrice. Pour presque tout  $\omega$ ,  $|X_n(\omega) - X(\omega)| \leq \epsilon$  à partir d'un certain  $n$ , c'est à dire  $Y_{n,\epsilon}(\omega) \rightarrow 0$ , donc  $Y_{n,\epsilon} \xrightarrow{\text{p.s.}} 0$ . De plus,  $Y_{n,\epsilon}$  est dominée ( $0 \leq Y_{n,\epsilon} \leq 1$ ), donc  $Y_{n,\epsilon} \xrightarrow{L^1} 0$ . Cela s'écrit

$$0 = \lim E(Y_{n,\epsilon}) = \lim P(A_{n,\epsilon}) = \lim P(|X_n - X| > \epsilon) : \quad (3.48)$$

$X_n \xrightarrow{P} X$ .

Si  $X_n \xrightarrow{L^1} X$ , soit  $\epsilon > 0$ . Par l'inégalité de Markov (Éq. (3.22)), on a

$$P(|X - X_n| > \epsilon) \leq \frac{E(|X_n - X|)}{\epsilon} \rightarrow 0. \quad (3.49)$$

Enfin, si  $X_n \xrightarrow{P} X$  on utilise sans les démontrer que pour  $f(x)$  continue bornée  $f(X_n) \xrightarrow{P} f(X)$  et que si  $(Y_n)_{n \in \mathbb{N}}$  est bornée presque sûrement,  $Y_n \xrightarrow{P} Y \Rightarrow Y_n \xrightarrow{L^1} Y$ , que l'on applique à  $Y_n = f(X_n)$ .

### 3.5.1.3 Caractérisation de la convergence en loi

On utilise, sans les démontrer, les propriétés suivantes :

- La convergence en loi est équivalente à la convergence des fonctions de répartition, en tout point où la fonction de répartition limite est continue.
- La convergence en loi est équivalente à la convergence simple des fonctions caractéristiques (théorème de Lévy).

## 3.5.2 Loi des grands nombres

On considère ici une suite de VA  $(X_n)_{n \in \mathbb{N}^*}$  indépendantes et de même loi. On note  $E(X_n) = \mu$  et  $\text{Var}(X_n) = \sigma^2$ . On s'intéresse au comportement de la somme

$$S_n = \sum_{k=1}^n X_k \quad (3.50)$$

à grand  $n$ . On peut facilement en calculer l'espérance et la variance :  $E(S_n) = n\mu$ ,  $\text{Var}(S_n) = n\sigma^2$ . La moyenne  $M_n = S_n/n$  vérifie donc  $E(M_n) = \mu$  et  $\text{Var}(M_n) = \sigma^2/n$ . Cela signifie que la VA  $M_n$  tend vers la VA certaine égale à  $\mu$ . Voyons de quel type de convergence il s'agit.

D'après l'inégalité de Bienaymé-Tchebychev (Éq. (3.23)),

$$P(|M_n - \mu| > \epsilon) \leq \frac{\text{Var}(M_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}. \quad (3.51)$$

Ainsi  $M_n \xrightarrow{P} \mu$ . On peut en fait faire un peu mieux, en utilisant que  $E(|M_n - \mu|)^2 \leq E([M_n - \mu]^2)$  (inégalité de Jensen, Éq. (3.24)) et

$$E([M_n - \mu]^2) = \text{Var}(M_n) = \frac{\sigma^2}{n} \rightarrow 0, \quad (3.52)$$

donc

$$E(|M_n - \mu|) \rightarrow 0, \quad (3.53)$$

ce qui signifie que

$$M_n \xrightarrow{L^1} \mu. \quad (3.54)$$

Ce résultat est la *loi des grands nombres*.

Plus précisément, le résultat (3.54) est la loi *faible* des grands nombres. La loi forte des grands nombres, que nous ne démontrons pas ici, indique que  $M_n \xrightarrow{\text{p.s.}} \mu$ . De plus, la loi forte des grands nombres n'a pas besoin de l'hypothèse de variance finie des VA  $X_n$ .

### 3.5.3 Théorème de la limite centrale

Nous nous plaçons dans le cadre de la section précédente. La loi des grands nombres montre que la moyenne  $M_n$  converge vers l'espérance  $E(X_n) = \mu$ ; nous allons maintenant chercher à caractériser les fluctuations autour de  $\mu$ . La variance de  $M_n$  nous renseigne sur l'échelle de ces fluctuations, et il apparaît que  $\text{Var}(\sqrt{n}M_n) = \sigma^2$ , mais maintenant  $E(\sqrt{n}M_n) = \sqrt{n}\mu$  n'est plus constante. La bonne VA à considérer est en fait

$$Y_n = \frac{S_n - n\mu}{\sqrt{n}\sigma}, \quad (3.55)$$

qui vérifie

$$E(Y_n) = 0, \quad (3.56)$$

$$\text{Var}(Y_n) = \frac{\text{Var}(S_n)}{n\sigma^2} = 1. \quad (3.57)$$

L'espérance et la variance de  $Y_n$  sont constantes, mais peut-on en dire plus sur la forme limite de  $Y_n$  quand  $n \rightarrow \infty$ ?

Considérons pour cela la fonction caractéristique de  $Y_n$ ,  $\phi_{Y_n}(k) = \Phi_n(k)$ , que nous allons exprimer avec la fonction caractéristique de  $X_n$ ,  $\phi_{X_n}(k) = \phi(k)$ . Nous avons besoin des propriétés suivantes de la fonction caractéristique :

$$\phi_{aX}(k) = E(e^{ik a X}) = \phi_X(ak), \quad (3.58)$$

$$\phi_{X+a}(k) = E(e^{ik(X+a)}) = e^{ia k} \phi_X(k). \quad (3.59)$$

Ainsi,

$$\Phi_n(k) = \phi_{S_n - n\mu}\left(\frac{k}{\sqrt{n}\sigma}\right) = e^{-i\frac{\sqrt{n}\mu k}{\sigma}} \phi_{S_n}\left(\frac{k}{\sqrt{n}\sigma}\right) = e^{-i\frac{\sqrt{n}\mu k}{\sigma}} \left[\phi\left(\frac{k}{\sqrt{n}\sigma}\right)\right]^n, \quad (3.60)$$

où nous avons utilisé la fonction caractéristique d'une somme de VA indépendantes (Éq. (3.28)). Pour  $k$  fixé, cherchons la limite de  $\Phi_n(k)$ . Nous avons besoin du développement de Taylor de  $\phi(k)$  en  $k = 0$ , qui est

$$\phi(k) = 1 + i\mu k - \frac{E(X_n^2)}{2}k^2 + o(k^2) = 1 + i\mu k - \frac{\mu^2 + \sigma^2}{2}k^2 + o(k^2). \quad (3.61)$$

Ainsi,

$$\Phi_n(k) = \exp\left(-i\frac{\sqrt{n}\mu k}{\sigma} + n \log\left(1 + i\frac{\mu k}{\sqrt{n}\sigma} - \frac{[\mu^2 + \sigma^2]}{2n\sigma^2}k^2 + o\left(\frac{1}{n}\right)\right)\right). \quad (3.62)$$

En utilisant le développement de Taylor de  $\log(x)$  à l'ordre 2 autour de  $x = 1$ ,  $\log(x) = x - \frac{x^2}{2} + o(x^2)$ , on obtient

$$\Phi_n(k) = \exp\left(-\frac{k^2}{2} + o(1)\right) \rightarrow \exp\left(-\frac{k^2}{2}\right). \quad (3.63)$$

Il s'agit de la fonction caractéristique d'une variable normale centrée réduite  $\mathcal{N}(0, 1)$ . D'après le théorème de Lévy, il y a convergence en loi :

$$Y_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1). \quad (3.64)$$

Ce résultat est le *théorème de la limite centrale*. Il est aussi appelé théorème central limite, et souvent noté TCL.

Ce théorème est très fort, car il montre que les fluctuations autour de l'espérance de la moyenne statistique  $M_n$  ont une forme universelle qui ne dépend pas de la loi des variables de départ, qu'elles soient discrètes ou continues.

**Exemple :** On peut utiliser le théorème de la limite centrale pour estimer une probabilité d'écart à la moyenne. On lance 1000 fois une pièce et on cherche la probabilité d'obtenir plus de  $p = 550$  fois « pile ». On note  $S_n$  le nombre de pile pour  $n$  lancers, alors on sait que

$$Y_n = \frac{S_n - n/2}{\sqrt{n}(1/2)} \sim \mathcal{N}(0, 1), \quad (3.65)$$

donc

$$P(S_n \geq p) = P\left(Y_n \geq \frac{2p - n}{\sqrt{n}}\right) \simeq \int_{\frac{2p-n}{\sqrt{n}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 1 - \Pi\left(\frac{2p - n}{\sqrt{n}}\right). \quad (3.66)$$

Pour  $n = 1000$  et  $p = 550$ , on trouve

$$P(S_{1000} \geq 550) \simeq 1 - \Pi(3.16) \simeq 7.8 \cdot 10^{-4}. \quad (3.67)$$

### 3.5.4 Intervalle de confiance

On effectue un sondage pour estimer le vainqueur d'une élection à deux candidats  $A$  et  $B$ . Sur un échantillon de  $n = 2000$  personnes, on trouve qu'une proportion  $\hat{p}_n = 48\%$  des électeurs interrogés souhaitent voter pour le candidat  $A$ . Quelle confiance peut-on accorder à ce résultat? Plus précisément, si  $p$  est la proportion réelle des électeurs votant pour  $A$ , pour un risque  $\alpha \in ]0, 1[$  donné, quel est l'écart  $\epsilon$  tel que

$$P(|\hat{p}_n - p| \leq \epsilon) = \alpha ? \quad (3.68)$$

On cherche donc un intervalle autour de  $\hat{p}_n$  dans lequel se trouve  $p$  avec une probabilité  $\alpha$ .

Chaque personne  $i$  interrogée,  $1 \leq i \leq n$ , correspond à une VA  $X_i$  qui vaut 1 si elle vote pour  $A$ , et 0 si elle vote pour  $B$ .  $X_i$  est donc une VA de Bernouilli avec  $P(X_i = 1) = p$ . Le sondage fournit

$$\hat{p}_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad (3.69)$$

qui est une VA qui fournit une estimation de  $p$ . Son espérance vaut  $E(\hat{p}_n) = p$ , et sa variance est  $\text{Var}(\hat{p}_n) = p(1-p)/n$ . Comme l'espérance de  $\hat{p}_n$  vaut exactement  $p$ , on dit que  $\hat{p}_n$  est un *estimateur sans biais* de  $p$ . La taille  $n$  de l'échantillon étant assez grande ( $n \geq 50$ ), on peut utiliser que  $(\hat{p}_n - p)/\sqrt{\text{Var}(\hat{p}_n)}$  suit une loi normale centrée réduite. Un point embêtant est que la variance de  $\hat{p}_n$  dépend de  $p$  qui n'est pas connu ; heureusement, on peut la majorer :  $\text{Var}(\hat{p}_n) \leq 1/(4n)$ . Dans le cas général, il faut utiliser un estimateur de la variance, ce qui rend le problème plus compliqué.

On obtient donc

$$\begin{aligned} P(|\hat{p}_n - p| \leq \epsilon) &= P\left(\frac{|\hat{p}_n - p|}{\sqrt{p(1-p)/n}} \leq \frac{\epsilon}{\sqrt{p(1-p)/n}}\right) \geq P\left(\frac{|\hat{p}_n - p|}{\sqrt{p(1-p)/n}} \leq 2\sqrt{n}\epsilon\right) \\ &= \int_{-2\sqrt{n}\epsilon}^{2\sqrt{n}\epsilon} f_{\mathcal{N}}(x)dx = \Pi(2\sqrt{n}\epsilon) - \Pi(-2\sqrt{n}\epsilon) = 2\Pi(2\sqrt{n}\epsilon) - 1 ; \quad (3.70) \end{aligned}$$

nous avons utilisé la symétrie de la loi normale centrée,  $\Pi(x) = 1 - \Pi(-x)$ . Il faut donc résoudre  $2\Pi(2\sqrt{n}\epsilon) - 1 = \alpha$ , ce qui donne

$$\epsilon = \frac{1}{2\sqrt{n}} \Pi^{-1}\left(\frac{1+\alpha}{2}\right) = \frac{1}{\sqrt{2n}} \text{erf}^{-1}(\alpha). \quad (3.71)$$

Ainsi, on dit que

$$I_{n,\alpha} = \left[ \hat{p}_n - \frac{\text{erf}^{-1}(\alpha)}{\sqrt{2n}}, \hat{p}_n + \frac{\text{erf}^{-1}(\alpha)}{\sqrt{2n}} \right] \quad (3.72)$$

est l'intervalle de confiance de risque  $\alpha$  pour  $p$ .

Avec  $n = 2000$ , on obtient

$$I_{2000,0.95} = [0.458, 0.502], \quad (3.73)$$

$$I_{2000,0.99} = [0.451, 0.509]. \quad (3.74)$$

Les instituts de sondage donnent leurs résultats sous cette forme.

### 3.6 Pour aller plus loin

L'intervalle de confiance vu en section 3.5.4 est en fait une introduction à un vaste domaine : nous avons considéré le cas très simple où nous cherchons à estimer un paramètre d'une VA de Bernoulli à partir de réalisations de cette VA. Les choses se compliquent si on change de loi (loi exponentielle, loi gaussienne), où il faut parfois aussi estimer la variance. L'étape suivante consiste à tester l'adéquation à une loi : une expérience fournit des résultats aléatoires, et on cherche à savoir si ces résultats suivent une distribution uniforme, exponentielle, etc. Des exemples simples sont de savoir si un dé est pipé à partir d'un nombre fini de lancers, ou si les notes aux examens sont corrélées au temps passé au foyer. Pour répondre à ces questions, il faut construire des *estimateurs statistiques* et mettre en place des *tests d'hypothèses* ; ceux-ci ont été abordés dans le cours de Statistique Appliquée de 1<sup>ère</sup> année.

D'autre part, notre étude des suites de VA s'est limitée au VA indépendantes et identiquement distribuées. Un prolongement intéressant est le cas des modèles dynamiques aléatoires, où les VA sont reliées entre elles par des relations de récurrence. Ces modèles sont utilisés pour décrire des processus dynamiques, en écologie pour l'évolution des espèces, en finance pour l'évolution du cours d'une action, et en physique pour décrire des systèmes hors équilibre, ce qui fait partie du cours de Concepts Clés en Physique Statistique de 3<sup>ème</sup> année.



# Index

- Condition aux bord, 6
  - de Cauchy, 7
  - de Dirichlet, 7
  - de Neumann, 7
- EDP
  - elliptique, 14
  - homogène, 7
  - hyperbolique, 14
  - linéaire, 7
  - parabolique, 14
- Équation
  - d'Euler-Lagrange, 30, 37, 38
  - de la chaleur, 6, 17
  - de Poisson, 6, 14
  - de transport, 5, 12
  - des ondes, 6, 19
- Espace des états, 42
- Espace probabilisé, 42
- Espérance, 44
- Événement, 42
- Événements indépendants, 42
- Fonction
  - caractéristique, 46
  - de Green, 7
  - de répartition, 44
- Formule
  - de Bayes, 42
- Inégalité
  - de Bienaymé-Tchebychev, 45
  - de Jensen, 45
  - de Markov, 45
- Intégrale du mouvement, 35
- Intervalle de confiance, 52
- Lagrangien, 30
- Loi
  - binomiale, 47
  - d'une variable aléatoire, 43
  - de Bernouilli, 47
  - de Poisson, 47
  - des grands nombres, 50
  - exponentielle, 48
  - géométrique, 47
  - marginale, 43
  - normale, 48
  - uniforme, 47, 48
- Mesure, 42
- Méthode
  - de séparation des variables, 11, 26
  - des caractéristiques, 9
  - des images, 8
- Modes propres, 23
- Multiplicateur de Lagrange, 33
- Principe
  - de moindre action, 36, 37
  - de superposition, 7
  - du maximum, 16
- Probabilité, 42
  - conditionnelle, 42
- Théorème
  - de la limite centrale, 51
  - de Lévy, 50
  - spectral, 23
- Tribu, 42
- Variable aléatoire, 43
  - continue, 46
  - discrète, 46