

# Outils Mathématiques de la Physique

Peter Schlagheck  
Université de Liège

*Ces notes ont pour seule vocation d'être utilisées par les étudiants dans le cadre de leur cursus au sein de l'Université de Liège.*

*Aucun autre usage ni diffusion n'est autorisé, sous peine de constituer une violation de la Loi du 30 juin 1994 relative au droit d'auteur.*



# Chapitre 1

## L'analyse complexe

### 1.1 Les nombres complexes

Les nombres complexes sont tous les nombres de la forme  $z = x + iy$  où  $x$  et  $y$  sont des nombres réels. L'objet algébrique  $i$ , qui n'est pas un nombre réel, satisfait à l'équation

$$i^2 = i \cdot i = -1 \quad (1.1)$$

et on écrit parfois  $i = \sqrt{-1}$ .  $i$  est nommé *unité imaginaire* et on appelle  $x = \operatorname{Re}(z)$  la *partie réelle* et  $y = \operatorname{Im}(z)$  la *partie imaginaire* du nombre complexe  $z = x + iy$ . Si  $y = 0$ ,  $z = x \in \mathbb{R}$  sera un nombre réel, alors que si  $x = 0$ ,  $z = iy$  sera un nombre *imaginaire pur*.

L'ensemble des nombres complexes

$$\mathbb{C} = \{z = x + iy : x, y \in \mathbb{R}\} \quad (1.2)$$

forme un corps. En effet, on peut définir une addition de deux nombres complexes  $z_1 = x_1 + iy_1$  et  $z_2 = x_2 + iy_2$  par

$$z_1 + z_2 = z_2 + z_1 = (x_1 + x_2) + i(y_1 + y_2). \quad (1.3)$$

L'élément neutre de cette addition est évidemment le zéro complexe  $0 \equiv 0 + i0$  et l'élément inverse associé à  $z = x + iy$  est  $-z = -x + i(-y)$ . Pour définir la multiplication de deux nombres complexes  $z_1 = x_1 + iy_1$  et  $z_2 = x_2 + iy_2$ , on utilise la distributivité de la multiplication par rapport à l'addition ainsi que le fait que  $i^2 = -1$ . Ceci donne

$$z_1 z_2 = z_2 z_1 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + y_1 x_2). \quad (1.4)$$

L'élément neutre de cette multiplication est évidemment  $1 \equiv 1 + i0$ . Pour déterminer l'élément inverse  $z^{-1} \equiv 1/z = 1/(x + iy)$  associé au nombre complexe

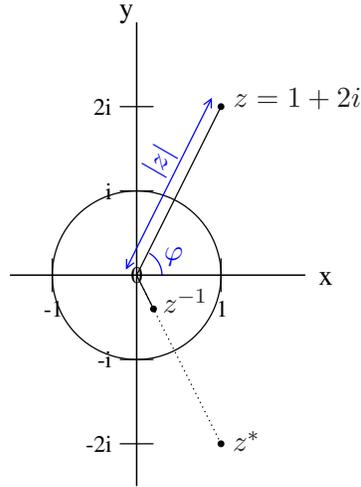


Figure 1.1: Visualisation du nombre complexe  $z = 1 + 2i$ , de son conjugué  $z^* = 1 - 2i$  ainsi que de son inverse  $z^{-1} = 0.2 - 0.4i = z^*/|z|^2$  qui se détermine par l'inversion de  $z^*$  au cercle de rayon unité dans le plan  $x-y$ . La représentation polaire de ce nombre complexe s'exprime par la distance  $|z|$  de  $z$  par rapport à l'origine du système des coordonnées ainsi que par l'angle géométrique  $\varphi$  de  $z$  par rapport à l'axe réel positif.

$z = x + iy$ , il faut multiplier le numérateur et le dénominateur de ce rapport par  $x - iy$  ce qui donne

$$\frac{1}{z} = \frac{1}{x + iy} = \frac{x - iy}{(x + iy)(x - iy)} = \frac{x}{x^2 + y^2} - i \frac{y}{x^2 + y^2}. \quad (1.5)$$

Il est utile de visualiser le nombre complexe  $z = x + iy$  dans le *plan complexe* où on représente la partie réelle  $x$  sur l'axe des abscisses et la partie imaginaire  $y$  sur l'axe des ordonnées, comme montré dans la figure 1.1. Cette représentation graphique donne naissance à certaines propriétés *géométriques* du nombre  $z$ . Notamment on peut introduire sa *norme* ou son *module*

$$|z| = \sqrt{x^2 + y^2}, \quad (1.6)$$

qui correspond à la distance de  $z$  par rapport à l'origine du système des coordonnées, ainsi que son *argument*  $\varphi$  qui correspond à l'angle que la ligne droite entre l'origine et  $z$  forme avec l'axe réel positif. On introduit aussi le *conjugué*<sup>1</sup> de  $z$  par  $z^* = x - iy$  ainsi que le *module carré*  $|z|^2 = x^2 + y^2 = |z^*|^2 = z^*z$ .

La norme  $|z|$  et l'angle  $\varphi$  constituent la *représentation polaire* du nombre complexe  $z$  (en opposition à la *représentation cartésienne*  $z = x + iy$ ). Notamment on peut écrire

$$z = |z| \cos(\varphi) + i|z| \sin(\varphi) \quad (1.7)$$

<sup>1</sup>Dans la littérature mathématique on écrit souvent  $\bar{z}$  au lieu de  $z^*$  pour le conjugué de  $z$ .

et

$$z^* = |z| \cos(\varphi) - i|z| \sin(\varphi) = |z| \cos(-\varphi) + i|z| \sin(-\varphi) \quad (1.8)$$

ce qui revient à dire que la conjugaison  $z \mapsto z^*$  est équivalente à l'inversion  $\varphi \mapsto -\varphi$  de l'angle  $\varphi$ . Alors que le calcul de la norme (1.6) est évident pour un nombre complexe  $z = x + iy$ , la détermination de l'angle  $\varphi$  associé demande un peu d'attention. On a évidemment  $x = |z| \cos \varphi$  et  $y = |z| \sin \varphi$  ce qui donne

$$\frac{x}{y} = \cot \varphi = \frac{\cos \varphi}{\sin \varphi}, \quad (1.9)$$

mais la simple solution  $\varphi = \operatorname{arccot}(x/y)$  de l'équation (1.9) ne fournit que des angles positifs contenus entre 0 et  $\pi$ . Une définition de l'angle  $\varphi$  en fonction de  $x$  et  $y$  nécessite donc une étude de cas et peut s'écrire

$$\varphi = \begin{cases} \operatorname{arccot}(x/y) & : y > 0 \\ \operatorname{arccot}(x/y) - \pi & : y < 0 \\ 0 & : y = 0 \text{ et } x > 0 \\ \pi & : y = 0 \text{ et } x < 0 \end{cases}, \quad (1.10)$$

ce qui fournit un angle unique  $\varphi$  contenu entre  $-\pi < \varphi \leq \pi$ .

## 1.2 Les fonctions complexes

On peut définir sur le plan complexe, où sur un sous-ensemble  $U$  de celui-ci, des fonctions

$$f : U \subseteq \mathbb{C} \rightarrow \mathbb{C}, z \mapsto f(z) \quad (1.11)$$

qui transforment un nombre complexe  $z \in U$  en un autre nombre complexe  $f(z) \in \mathbb{C}$ . Voici une liste de fonctions complexes qui peuvent être définies sur tout le plan complexe ( $U = \mathbb{C}$ ) :

$$f(z) = z, \quad (1.12)$$

$$f(z) = a \text{ avec } a \in \mathbb{C}, \quad (1.13)$$

$$f(z) = az + b \text{ avec } a, b \in \mathbb{C}, \quad (1.14)$$

$$f(z) = \operatorname{Re}(z), \quad (1.15)$$

$$f(z) = \operatorname{Im}(z), \quad (1.16)$$

$$f(z) = z^* = \operatorname{Re}(z) - i\operatorname{Im}(z), \quad (1.17)$$

$$f(z) = |z|^2 = z^*z, \quad (1.18)$$

$$f(z) = z^2 = z \cdot z, \quad (1.19)$$

$$f(z) = z^n = z \cdot z \cdot \dots \cdot z = \prod_{j=1}^n z \text{ avec } n \in \mathbb{N}, \quad (1.20)$$

$$f(z) = \sum_{n=0}^N a_n z^n \text{ avec } a_n \in \mathbb{C} \text{ et } N \in \mathbb{N}. \quad (1.21)$$

On peut se demander dans quelle mesure on a droit de laisser tendre le degré  $N$  du polynôme complexe (1.21) vers l'infini, c-à-d, de définir des *séries entières* sur le plan complexe. Cela dépend évidemment du choix des coefficients  $a_n$  et des propriétés de *convergence* de la série résultante, comme on le verra plus tard. Une série dont la convergence sur l'ensemble du plan complexe est connue avec certitude est la *série exponentielle*. Cette dernière donne naissance à la *fonction exponentielle complexe*

$$f : \mathbb{C} \rightarrow \mathbb{C}, z \mapsto f(z) = \exp(z) \equiv e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n. \quad (1.22)$$

En effet, comme on a  $|z_1 z_2| = |z_1| |z_2|$  pour tout produit de deux nombres complexes  $z_1$  et  $z_2$  (ce qui se vérifie facilement en utilisant les équations (1.4) et (1.6)), on a  $|z^n| = |z|^n$  pour tout  $n \in \mathbb{N}$  et on peut donc majorer

$$|e^z| = \left| \sum_{n=0}^{\infty} \frac{1}{n!} z^n \right| \leq \sum_{n=0}^{\infty} \frac{1}{n!} |z|^n = e^{|z|} \quad (1.23)$$

par la série exponentielle réelle  $e^x$  avec  $x = |z| \in \mathbb{R}$  dont on sait déjà qu'elle est convergente sur tout l'axe réel.

La fonction exponentielle satisfait à la propriété remarquable

$$e^{z_1+z_2} = e^{z_1} e^{z_2} \quad (1.24)$$

pour tous nombres complexes  $z_1, z_2 \in \mathbb{C}$ , car on a

$$\sum_{n=0}^{\infty} \frac{1}{n!} (z_1 + z_2)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^n \frac{n!}{k!(n-k)!} z_1^k z_2^{n-k} = \sum_{k=0}^{\infty} \frac{1}{k!} z_1^k \sum_{l=0}^{\infty} \frac{1}{l!} z_2^l \quad (1.25)$$

si on introduit  $l = n - k$ . La propriété (1.24) est également valable pour les parties réelle et imaginaire du nombre complexe  $z = x + iy$ . On a donc  $e^z = e^{x+iy} = e^x e^{iy}$  avec

$$e^{iy} = \sum_{n=0}^{\infty} \frac{1}{n!} (iy)^n = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} y^{2k} + i \sum_{l=0}^{\infty} \frac{(-1)^l}{(2l+1)!} y^{2l+1} = \cos(y) + i \sin(y) \quad (1.26)$$

où on utilise les séries entières des fonctions trigonométriques réelles  $\cos(x)$  et  $\sin(x)$ . L'équation (1.26) nous permet d'exprimer le cosinus et le sinus à l'aide de la fonction exponentielle complexe comme

$$\cos(x) = \frac{1}{2} (e^{ix} + e^{-ix}), \quad (1.27)$$

$$\sin(x) = \frac{1}{2i} (e^{ix} - e^{-ix}). \quad (1.28)$$

Si on rend complexes les relations (1.27) et (1.28), on obtient des fonctions trigonométriques complexes définies par

$$\cos(z) = \frac{1}{2} (e^{iz} + e^{-iz}), \quad (1.29)$$

$$\sin(z) = \frac{1}{2i} (e^{iz} - e^{-iz}) \quad (1.30)$$

pour tout nombre complexe  $z \in \mathbb{C}$  car on sait déjà que la fonction exponentielle  $e^{\pm iz}$  converge partout en  $\mathbb{C}$ . De même, on peut définir des fonctions hyperboliques complexes

$$\cosh(z) = \frac{1}{2} (e^z + e^{-z}) = \cos(iz), \quad (1.31)$$

$$\sinh(z) = \frac{1}{2} (e^z - e^{-z}) = -i \sin(iz). \quad (1.32)$$

Les relations (1.27) et (1.28) nous permettent de reformuler la représentation polaire (1.7) d'un nombre complexe  $z$ . Nous obtenons

$$z = r \cos(\varphi) + ir \sin(\varphi) = re^{i\varphi} \quad (1.33)$$

pour  $r > 0$  et  $-\pi < \varphi \leq \pi$ . Cette dernière représentation facilite le calcul du carré (1.19) d'un nombre complexe  $z = x + iy = r \cos(\varphi) + ir \sin(\varphi)$ . Par un calcul standard on obtient

$$\begin{aligned} z^2 &= x^2 - y^2 + 2ixy = r^2 (\cos^2(\varphi) - \sin^2(\varphi) + 2i \cos(\varphi) \sin(\varphi)) \\ &= r^2 (\cos(2\varphi) + i \sin(2\varphi)) \end{aligned} \quad (1.34)$$

où pour la dernière étape du calcul deux relations bien connues concernant les fonctions trigonométriques réelles ont été utilisées. Ce résultat est obtenu bien plus facilement à l'aide de la représentation  $z = re^{i\varphi}$  si on utilise la propriété (1.24) :

$$z^2 = (re^{i\varphi})^2 = r^2 e^{2i\varphi} = r^2 (\cos(2\varphi) + i \sin(2\varphi)). \quad (1.35)$$

Il en est de même pour les fonctions puissances (1.20) qui s'évaluent alors facilement selon

$$z^n = (re^{i\varphi})^n = r^n e^{in\varphi} = r^n (\cos(n\varphi) + i \sin(n\varphi)). \quad (1.36)$$

La nouvelle représentation polaire (1.33) nous permet également d'évaluer des fonctions avec des puissances *négatives*

$$f : \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C}, z \mapsto z^{-n} = \frac{1}{z^n} = \frac{1}{r^n} e^{-in\varphi} \quad (1.37)$$

pour  $n \in \mathbb{N}$  si on représente  $z = re^{i\varphi}$  avec  $r, \varphi \in \mathbb{R}$ . Evidemment,  $f$  ne peut pas être définie en  $z = 0$ . Plus généralement, on peut introduire des fonctions *rationnelles*

$$f(z) = \frac{\sum_{n=0}^N a_n z^n}{\sum_{m=0}^M b_m z^m} \quad (1.38)$$

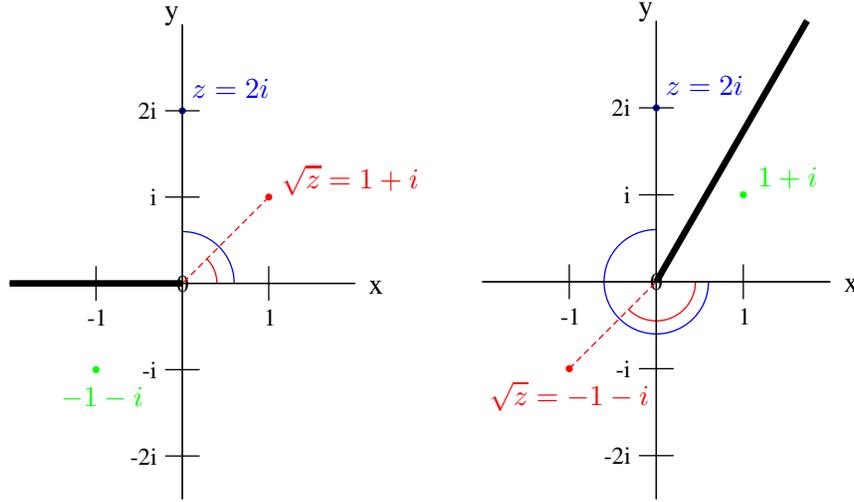


Figure 1.2: Deux différents résultats pour la racine de  $z = 2i$  qui dépendent de la définition de la représentation polaire. A gauche, on laisse varier l'angle polaire  $\varphi$  entre  $-\pi$  et  $\pi$ . La représentation polaire de  $z$  s'écrit donc  $z = 2e^{i\pi/2}$  et sa racine vaut  $\sqrt{z} = \sqrt{2}e^{i\pi/4} = 1 + i$ . A droite, l'angle polaire  $\varphi$  est censé varier entre  $-5\pi/3$  et  $\pi/3$ . On doit donc écrire  $z = 2e^{-3i\pi/2}$  dans cette représentation polaire et sa racine vaut  $\sqrt{z} = \sqrt{2}e^{-3i\pi/4} = -1 - i$ .

avec  $a_n, b_m \in \mathbb{C}$  et  $N, M \in \mathbb{N}$ , qui peuvent être définies partout sur  $\mathbb{C}$  à l'exception des zéros du polynôme dans le dénominateur de  $f$ .

La représentation polaire nous permet également de définir des fonctions de puissance *non entière* sur le plan complexe, mais cette dernière définition demande un peu de soin. Prenons, par exemple, la fonction racine. En représentant  $z = re^{i\varphi}$  avec  $r > 0$  et  $\varphi \in ]-\pi, \pi]$ , nous obtenons

$$\sqrt{z} = z^{1/2} = (re^{i\varphi})^{1/2} = \sqrt{r}e^{i\varphi/2} \quad (1.39)$$

ce qui immédiatement implique que la racine complexe peut être définie pour tout nombre complexe  $z \in \mathbb{C}$ , contrairement à ce qu'on peut dire de la racine réelle. En effet, en notant que  $-1 = e^{i\pi}$  dans cette représentation polaire, on a

$$\sqrt{-1} = \sqrt{e^{i\pi}} = e^{i\pi/2} = \cos(\pi/2) + i \sin(\pi/2) = i \quad (1.40)$$

ce qui est bien en accord avec la définition (1.1) du nombre  $i$ .

Le problème de la racine complexe est que sa définition (1.39) introduit une *discontinuité* le long de l'axe réel négatif. Considérons, par exemple, le nombre complexe  $z = -e^{i\delta}$  avec un petit angle réel  $\delta \rightarrow 0$  positif ou négatif. Evidemment on a

$$\lim_{\delta \rightarrow 0} (-e^{i\delta}) = -\lim_{\delta \rightarrow 0} (\cos(\delta) + i \sin(\delta)) = -1 \quad (1.41)$$

indépendamment du signe de  $\delta$ . Mais pour la racine de  $z$  nous obtenons deux différentes limites dépendant du signe de  $\delta$ , notamment

$$\lim_{\delta \rightarrow 0^-} \sqrt{-e^{i\delta}} = \lim_{\delta \rightarrow 0^-} \sqrt{e^{i(\pi+\delta)}} = \lim_{\delta \rightarrow 0^-} e^{i(\pi+\delta)/2} = i \lim_{\delta \rightarrow 0^-} e^{i\delta/2} = i, \quad (1.42)$$

$$\lim_{\delta \rightarrow 0^+} \sqrt{-e^{i\delta}} = \lim_{\delta \rightarrow 0^+} \sqrt{e^{i(-\pi+\delta)}} = \lim_{\delta \rightarrow 0^+} e^{i(-\pi+\delta)/2} = -i \lim_{\delta \rightarrow 0^+} e^{i\delta/2} = -i \quad (1.43)$$

ce qui est évidemment dû au fait que la représentation polaire définie ci-dessus nécessite l'introduction d'un angle  $\varphi$  qui est contenu entre  $-\pi$  (exclus) et  $\pi$  (inclus) avant de calculer la racine en l'équation (1.39).

Evidemment on peut changer la définition de la représentation polaire en imposant que l'angle  $\varphi$  doit être contenu par exemple dans l'intervalle  $-5\pi/3 < \varphi \leq \pi/3$  au lieu de l'intervalle  $]-\pi, \pi]$ , comme cela est montré dans la figure 1.2. Dans ce cas-là, la racine (1.39) est certainement continue le long de l'axe réel négatif, mais on introduit un autre axe de discontinuité dans le plan complexe, notamment le long de l'angle  $\varphi = \pi/3$ . Ceci fait qu'on évalue

$$\sqrt{i} = \sqrt{e^{-3i\pi/2}} = e^{-3i\pi/4} = -e^{i\pi/4} \quad (1.44)$$

vu que l'angle  $\pi/2 > \pi/3$  n'est pas contenu dans l'intervalle défini ci-dessus, alors qu'avec la définition standard (1.10) de l'angle  $\varphi$  on obtient bien  $\sqrt{i} = e^{i\pi/4}$ .

Cette espèce de discontinuité apparaît aussi pour d'autres fonctions de puissance non entière

$$f(z) = z^\alpha = (re^{i\varphi})^\alpha = r^\alpha e^{i\alpha\varphi} \quad (1.45)$$

avec  $\alpha \in \mathbb{R}$ , qui peuvent être définies sur tout le plan complexe si  $\alpha \geq 0$  et sur  $U = \mathbb{C} \setminus \{0\}$  si  $\alpha < 0$ . Si on définit l'angle  $\varphi$  de manière standard selon l'équation (1.10), on obtient également une discontinuité le long de l'axe réel négatif comme pour la racine.

Il est souvent utile de représenter des fonctions de puissance générale (1.45) à l'aide du *logarithme complexe*

$$f : \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C}, z \mapsto \ln(z). \quad (1.46)$$

Cette dernière fonction est définie de façon à vérifier que l'on a  $\exp[\ln(z)] = z$  pour tout  $z \neq 0$ , comme on le fait aussi pour le logarithme réel :  $\exp[\ln(r)] = r$  pour tout  $r > 0$ . En utilisant la représentation polaire on a

$$z = re^{i\varphi} = e^{\ln(r)} e^{i\varphi} = e^{\ln(r) + i\varphi}, \quad (1.47)$$

et en identifiant  $z = e^{\ln(z)}$  on obtient directement une définition du logarithme complexe

$$\ln(z) = \ln(r) + i\varphi \quad (1.48)$$

en représentation polaire. Il est facile de démontrer les relations  $\exp[\ln(z_1 z_2)] = \exp[\ln(z_1) + \ln(z_2)]$  ainsi que  $\exp[\ln(z^n)] = \exp[n \ln(z)]$ . Par la généralisation de

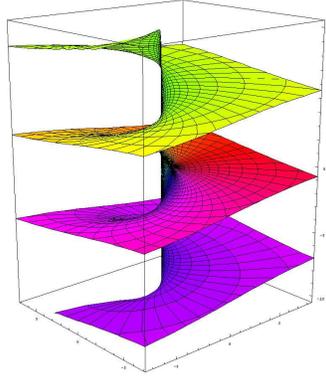


Figure 1.3: Représentation graphique des différentes branches du logarithme complexe (source de l'image : [http://en.wikipedia.org/wiki/Riemann\\_surface](http://en.wikipedia.org/wiki/Riemann_surface)). On montre sur l'axe vertical la partie imaginaire  $\varphi$  du logarithme complexe  $\ln(z)$  en fonction des parties réelle  $x$  et imaginaire  $y$  de  $z$ , c-à-d  $x = r \cos \varphi$  et  $y = r \sin \varphi$  avec  $r > 0$ , où  $\varphi$  n'est pas limité à un intervalle de longueur  $2\pi$ . Globalement on obtient comme ça une structure hélicoïdale avec un point de branchement à l'origine.

cette dernière relation pour des exposants négatifs et/ou non entiers, on obtient donc

$$z^\alpha = \exp[\ln(z^\alpha)] = \exp[\alpha \ln(z)] \quad (1.49)$$

si  $z \neq 0$ .

Aussi le logarithme complexe possède un axe de discontinuité le long de l'axe réel négatif si on utilise la convention standard (1.10) pour la définition de la représentation polaire, ou le long d'un autre axe si on utilise une autre convention. Si, d'un point de vue technique, on ne restreint pas l'angle  $\varphi \in \mathbb{R}$  à un intervalle particulier, la définition du logarithme complexe (1.48) fournit une "fonction" avec une infinité de valeurs  $\ln(r) + i(\varphi + 2\pi n)$  avec  $n \in \mathbb{Z}$  pour un nombre complexe  $z = re^{i\varphi}$  donné. Autrement dit, comme on le montre dans la figure 1.3, le logarithme complexe possède une infinité de feuillets différents dans le plan complexe qui sont connectés de manière hélicoïdale, et il faut donc "sauter" d'un feuillet à l'autre quelque part, par exemple le long de l'axe réel négatif, afin de définir une fonction qui fournit une seule valeur pour un nombre complexe donné. On appelle donc l'origine du plan complexe le *point de branchement* du logarithme complexe (et aussi des fonctions de puissance non entière) car c'est à ce point-là où commence cet axe de discontinuité.

### 1.3 La dérivée complexe

On définit la dérivée complexe de manière tout à fait analogue à la dérivée réelle. Cette dernière est définie pour une fonction continue  $f : U \subseteq \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x)$  par

$$f'(x) \equiv \frac{d}{dx}f(x) = \lim_{\delta \rightarrow 0} \frac{f(x + \delta) - f(x)}{\delta}. \quad (1.50)$$

La fonction  $f$  est appelée *dérivable* en  $x \in U$  si cette limite existe (ce qui n'est pas le cas, par exemple, pour la fonction  $f(x) = |x|$  en  $x = 0$ ). Elle est *continûment dérivable* en  $x$  si elle est dérivable en  $x$  et si sa dérivée  $f' : U' \subseteq U \rightarrow \mathbb{R}$  est continue en  $x$ .

Comme on sait calculer le quotient entre deux nombres complexes, on peut directement généraliser cette prescription (1.50) pour des fonctions complexes. Ceci donne la *dérivée complexe* qui est définie pour une fonction continue  $f : U \subseteq \mathbb{C} \rightarrow \mathbb{C}, z \mapsto f(z)$  par

$$f'(z) \equiv \frac{d}{dz}f(z) = \lim_{\delta \rightarrow 0} \frac{f(z + \delta) - f(z)}{\delta} \quad (1.51)$$

pour un  $\delta \in \mathbb{C}$  complexe. La fonction  $f$  est dite *dérivable* (de manière complexe) en  $z \in U$  si cette limite existe.

Il est facile de montrer que la fonction (1.20)  $f(z) = z^n$  est dérivable pour tout  $n \in \mathbb{N}$ . En effet, on a

$$f'(z) = \lim_{\delta \rightarrow 0} \frac{(z + \delta)^n - z^n}{\delta} = \lim_{\delta \rightarrow 0} \sum_{k=1}^n \frac{n!}{k!(n-k)!} z^{n-k} \delta^{k-1} = nz^{n-1} \quad (1.52)$$

pour tout  $z \in \mathbb{C}$ . Une combinaison linéaire  $f(z) = ag(z) + bh(z)$  de deux fonctions dérivables  $g, h$  (avec des constantes  $a, b \in \mathbb{C}$ ) est évidemment aussi dérivable avec pour dérivée  $f'(z) = ag'(z) + bh'(z)$ . Nous pouvons donc conclure que tout polynôme (1.21)  $f(z) = \sum_{n=0}^N a_n z^n$  est dérivable et sa dérivée est

$$f'(z) = \sum_{n=1}^N na_n z^{n-1} = \sum_{n=0}^{N-1} (n+1)a_{n+1} z^n. \quad (1.53)$$

Il en est de même pour une série entière

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \quad (1.54)$$

avec  $a_n \in \mathbb{C}$ , pourvu que la série converge en  $z \in \mathbb{C}$ . Dans ce cas-là, sa dérivée est obtenue par

$$f'(z) = \sum_{n=0}^{\infty} (n+1)a_{n+1} z^n. \quad (1.55)$$

Les fonctions exponentielle (1.22), trigonométriques (1.29,1.30) et hyperboliques (1.31,1.32) sont donc dérivables pour tout  $z \in \mathbb{C}$  avec pour dérivées

$$\frac{d}{dz} e^z = e^z, \quad (1.56)$$

$$\frac{d}{dz} \cos(z) = -\sin(z), \quad (1.57)$$

$$\frac{d}{dz} \sin(z) = \cos(z), \quad (1.58)$$

$$\frac{d}{dz} \cosh(z) = \sinh(z), \quad (1.59)$$

$$\frac{d}{dz} \sinh(z) = \cosh(z). \quad (1.60)$$

Parmi la liste de fonctions (1.13–1.21) fournie dans la section 1.2, il y a aussi des fonctions complexes qui ne sont pas dérivables. Ceci est notamment le cas pour la fonction (1.17)  $f(z) = z^*$ . Pour cette fonction la limite (1.51)

$$\lim_{\delta \rightarrow 0} \frac{(z + \delta)^* - z^*}{\delta} = \lim_{\delta \rightarrow 0} \frac{\delta^*}{\delta} \quad (1.61)$$

n'existe nulle part car le rapport  $\delta^*/\delta$  dépend de la direction dans le plan complexe par laquelle on s'approche de l'origine avec le paramètre complexe  $\delta$ . Il en est de même pour les parties réelle (1.15) et imaginaire (1.16) ainsi que pour le module carré (1.19) du nombre complexe  $z$ , qui ne sont dérivable nulle part dans le plan complexe.

On peut se demander sous quelles conditions une fonction complexe générale

$$f : U \subseteq \mathbb{C} \rightarrow \mathbb{C}, x + iy \mapsto f(x + iy) = u(x, y) + iv(x, y) \quad (1.62)$$

est dérivable si ses parties réelle et imaginaire sont représentées par les fonctions réelles et dérivables  $u, v : W \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}^2$ . En représentant  $\delta = \delta_x + i\delta_y$  avec  $\delta_x, \delta_y \in \mathbb{R}$  dans l'équation (1.51), nous obtenons

$$\begin{aligned} f'(z) &= \lim_{\delta \rightarrow 0} \frac{f(z + \delta) - f(z)}{\delta} \\ &= \lim_{\delta_x, \delta_y \rightarrow 0} \frac{u(x + \delta_x, y + \delta_y) + iv(x + \delta_x, y + \delta_y) - u(x, y) - iv(x, y)}{\delta_x + i\delta_y} \\ &= \lim_{\delta_x, \delta_y \rightarrow 0} \frac{\frac{\partial u}{\partial x}(x, y)\delta_x + \frac{\partial u}{\partial y}(x, y)\delta_y + i\frac{\partial v}{\partial x}(x, y)\delta_x + i\frac{\partial v}{\partial y}(x, y)\delta_y}{\delta_x + i\delta_y} \\ &= \begin{cases} \frac{\partial u}{\partial x}(x, y) + i\frac{\partial v}{\partial x}(x, y) : \delta_y = 0 \\ \frac{\partial u}{\partial y}(x, y) - i\frac{\partial v}{\partial y}(x, y) : \delta_x = 0 \end{cases} \quad (1.63) \end{aligned}$$

où, dans la dernière ligne du calcul (1.63), on considère particulièrement les deux cas spéciaux où l'on s'approche de l'origine du plan complexe le long des axes réel ( $\delta_y = 0$ ) et imaginaire ( $\delta_x = 0$ ). Ceci donne les *équations de Cauchy–Riemann*

$$\frac{\partial u}{\partial x}(x, y) = \frac{\partial v}{\partial y}(x, y), \quad (1.64)$$

$$\frac{\partial v}{\partial x}(x, y) = -\frac{\partial u}{\partial y}(x, y) \quad (1.65)$$

qui sont les conditions nécessaires — et, comme on peut également montrer, aussi suffisantes — pour que la dérivée complexe de la fonction  $f$  existe.

## 1.4 Les fonctions analytiques

Dans la suite, nous allons nous intéresser particulièrement aux fonctions qui sont dérivables partout ou presque partout sur le plan complexe. Ceci nous mène à la notion des fonctions *holomorphes*. Une fonction  $f : U \subseteq \mathbb{C} \rightarrow \mathbb{C}, z \mapsto f(z)$  est dite *holomorphe* en  $z_0 \in U$  s'il existe un voisinage ouvert  $V \subseteq U$  de  $z_0$  (avec  $z_0 \in V$ ) tel que  $f$  est dérivable pour tout  $z \in V$ . La fonction  $f$  est dite *holomorphe en  $U$*  si  $f$  est holomorphe en tout  $z \in U$ .

Un théorème important en analyse complexe dit que toute fonction  $f : U \subseteq \mathbb{C} \rightarrow \mathbb{C}$  qui est holomorphe en  $U$  est aussi *analytique* en  $U$ . En général, une fonction  $f : U \subseteq \mathbb{C} \rightarrow \mathbb{C}$  est dite *analytique* en  $z_0 \in U$  si on peut l'exprimer avec une série de Taylor autour de  $z_0$ . Plus précisément, pour que  $f$  soit analytique en  $z_0$ , il faut que  $f$  soit une infinité de fois dérivable en  $z_0$  et qu'il existe un voisinage  $V \subseteq U$  de  $z_0$  tel que la série de Taylor associée  $\sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(z_0)(z - z_0)^n$ , où  $f^{(n)}(z_0)$  est la  $n$ -ième dérivée de  $f$  en  $z_0$ , converge vers  $f(z)$  pour tout  $z \in V$ .  $f$  est dite *analytique en  $U$*  si  $f$  est analytique en tout  $z \in U$ .

Nous allons, dans la suite, utiliser les notions d'holomorphie et d'analyticité de manière synonyme, même si l'équivalence de ces deux notions nécessite une discussion plus approfondie de ce qui signifie une "fonction". De manière très technocrate, on peut considérer que la caractérisation d'une fonction  $f : U \rightarrow \mathbb{C}$  soit entièrement donnée par la spécification de sa région de définition  $U \subseteq \mathbb{C}$  ainsi que de chaque valeur  $f(z)$  que la fonction atteint pour tout argument  $z$  faisant partie de cette région de définition  $U$ . De ce point de vue, deux fonctions  $f : U \rightarrow \mathbb{C}$  et  $g : V \rightarrow \mathbb{C}$  qui sont spécifiées avec des régions de définition différentes,  $U \neq V$ , ne sont pas considérées "identiques",  $f \neq g$ , même si elles redonnent les mêmes valeurs sur l'intersection de ces régions de définition :  $f(z) = g(z)$  pour tout  $z \in U \cap V$ . Dans ce cas, on peut effectivement affirmer que toute fonction  $f$  qui est analytique en  $U \subseteq \mathbb{C}$  est également holomorphe en  $U$ . Cependant, la notion d'une fonction est très souvent utilisée d'une manière plus générale, impliquant, par exemple, qu'il existe une seule fonction de racine même s'il y a une multitude de manières de spécifier l'axe de discontinuité pour cette

dernière. Dans cette optique, on peut alors affirmer que “la racine”  $f : z \mapsto \sqrt{z}$  est analytique partout en  $\mathbb{C}$  à l’exception de l’origine car pour chaque  $z_0 \in \mathbb{C} \setminus \{0\}$  il est possible de représenter  $\sqrt{z}$  par une série entière développée autour de  $z_0$ , quitte à éventuellement modifier l’orientation de l’axe de discontinuité. Toutefois, à cause de la présence de l’axe de discontinuité, la racine n’est pas holomorphe en  $\mathbb{C} \setminus \{0\}$ , ce qui montre que cette dernière manière de considérer la notion d’une fonction n’implique pas que toute fonction  $f : U \rightarrow \mathbb{C}$  qui est analytique en  $U$  soit également holomorphe en  $U$ .

Les fonctions analytiques possèdent des propriétés remarquables. Considérons tout d’abord une fonction complexe  $f : U \subseteq \mathbb{C} \rightarrow \mathbb{C}, z \mapsto f(z)$  dont l’ensemble de ses zéros  $\{z \in U : f(z) = 0\}$  possède un *point d’accumulation* en  $U$ , c-à-d, un point  $z_0 \in U$  avec la propriété que pour tout  $\epsilon > 0$  il y a un zéro  $z \neq z_0$  de  $f$  dont la distance à  $z_0$  est plus petite que  $\epsilon : |z - z_0| < \epsilon$ . On peut alors montrer que si cette fonction est analytique en  $U$  (ce qui implique qu’on peut exprimer cette fonction par une série de Taylor autour de  $z_0$ ) elle est forcément *identique à zéro* partout en  $U : f(z) \equiv 0$  pour tout  $z \in U$ . Une conséquence immédiate de cette propriété est que deux fonctions  $f, g : U \subseteq \mathbb{C} \rightarrow \mathbb{C}$  qui sont toutes les deux analytiques en  $U$  doivent être identiques partout en  $U$ , c-à-d  $f(z) = g(z)$  pour tout  $z \in U$ , si l’ensemble des points où  $f$  est égal à  $g$ ,  $\{z \in U : f(z) = g(z)\}$ , possède un point d’accumulation en  $U$ .

Ceci nous mène à la notion de la *continuation analytique* d’une fonction réelle  $f : V \subseteq \mathbb{R} \rightarrow \mathbb{R}$  dans le plan complexe. Si  $f$  est analytique en  $V$  (c-à-d, si pour tout  $x \in V$  on peut exprimer  $f$  en série de Taylor réelle autour de  $x$  qui converge dans un voisinage ouvert de  $x$ ), il existe un sous-ensemble ouvert  $U \subseteq \mathbb{C}$  avec  $V \subset U$  dans lequel on peut définir de manière *unique* une fonction complexe  $g : U \rightarrow \mathbb{C}$  qui est analytique en  $U$  et qui fournit sur  $V$  les mêmes valeurs que  $f : f(x) = g(x + i0)$  pour tout  $x \in V$ . Par exemple, le logarithme complexe représente la continuation analytique du logarithme réel, ce dernier étant défini sur l’axe réel positif  $V = \mathbb{R}_+$ . La seule ambiguïté dans la définition de cette continuation analytique du logarithme réel réside dans le choix de l’axe de discontinuité (p.ex. le long de l’axe réel négatif) qui doit être exclu dans la définition de la région de définition  $U \subset \mathbb{C}$  du logarithme complexe.

Comme l’analyticité d’une fonction complexe est étroitement liée à la convergence de sa série de Taylor, il est utile de discuter les propriétés de convergence des séries entières générales

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \quad (1.66)$$

avec des coefficients constants  $a_n \in \mathbb{C}$ . Toute série entière complexe (1.66) possède un *rayon de convergence*

$$\rho \in [0, \infty] \quad (1.67)$$

tel que la série (1.66) converge pour tout  $z$  avec  $|z| < \rho$  et diverge pour tout  $z$  avec  $|z| > \rho$ . La valeur  $\infty$  est, pour une fois, incluse dans l’intervalle (1.67) des

valeurs possibles pour  $\rho$ . En effet, un rayon de convergence  $\rho = \infty$  caractérise une série qui converge partout en  $\mathbb{C}$ , ce qui est le cas, par exemple, pour la série exponentielle (1.22). De l'autre côté, un rayon de convergence  $\rho = 0$  caractérise une série (plutôt pathologique) qui ne converge nulle part à l'exception de  $z = 0$ .

En pratique, le rayon de convergence associée à la série (1.66) peut être déterminé par le comportement des coefficients  $a_n$  pour  $n \rightarrow \infty$  au travers du critère de *Cauchy-Hadamard*

$$\frac{1}{\rho} = \limsup_{n \rightarrow \infty} \{ |a_k|^{1/k} \} \quad (1.68)$$

ou, de manière tout à fait équivalente, au travers du critère de *d'Alembert-Euler*

$$\frac{1}{\rho} = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|. \quad (1.69)$$

Ces deux critères confirment que le rayon de convergence de la série entière (1.66) est identique à celui de sa dérivée  $f'(z) = \sum_{n=0}^{\infty} (n+1)a_{n+1}z^n$ , car on obtient  $\lim_{n \rightarrow \infty} (n+1)/n = 1$  si, par exemple, on utilise le critère de d'Alembert et Euler. Pour la série exponentielle où  $a_n = 1/n!$  pour tout  $n \in \mathbb{N}_0$  on obtient  $1/\rho = 0$  par les deux critères, alors que la *série géométrique* où  $a_n = 1$  pour tout  $n \in \mathbb{N}_0$  possède évidemment le rayon de convergence  $\rho = 1$ .

## 1.5 Les singularités des fonctions analytiques

Ce dernier exemple de la série géométrique

$$f(z) = \sum_{n=0}^{\infty} z^n = \frac{1}{1-z} \quad \text{pour } |z| < 1 \quad (1.70)$$

nous indique qu'il existe aussi une autre méthode pour trouver le rayon de convergence de la série entière (1.66) qui représente une fonction complexe  $f(z)$  : il faut chercher les *singularités* de  $f$  dans le plan complexe. Le rayon de convergence est alors donné par la distance de l'origine (ou, plus généralement, du point  $z_0$  autour duquel on développe la série) à la singularité la plus proche à l'origine.

En général, les singularités d'une fonction complexe  $f$  sont tous les points  $z \in \mathbb{C}$  où  $f$  ne peut pas être définie de manière analytique, donc où  $f$  ne peut pas être représentée par une série entière développée autour de  $z$ . Ceci inclut tous les points  $z \in \mathbb{C}$  où  $f$  ne peut pas être défini du tout, mais aussi des points de branchement en cas des racines ou des logarithmes complexes. Par exemple, les

fonctions suivantes

$$f(z) = \frac{1}{z}, \quad (1.71)$$

$$f(z) = \frac{\sin(z)}{z}, \quad (1.72)$$

$$f(z) = e^{1/z}, \quad (1.73)$$

$$f(z) = \ln(z), \quad (1.74)$$

$$f(z) = \sqrt{z} \quad (1.75)$$

possèdent toutes une singularité à  $z = 0$ .

Par définition, une singularité  $z_0 \in \mathbb{C}$  d'une fonction complexe  $f$  est appelée *isolée* s'il existe un voisinage ouvert  $U \subseteq \mathbb{C}$  de  $z_0$  tel que  $f$  soit holomorphe sur  $U \setminus \{z_0\}$ . Ceci n'est ni le cas du logarithme complexe (1.74) ni de la racine (1.75) à cause de l'axe de discontinuité qui est connecté à l'origine. Par contre, les exemples (1.71–1.73) possèdent bien une singularité isolée en  $z = 0$ . Dans le cas d'une telle singularité isolée en  $z_0$ , on peut représenter la fonction complexe  $f$  de manière unique par une *série de Laurent*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n \quad \text{avec } a_n \in \mathbb{C} \quad (1.76)$$

autour de  $z_0$ . Cette série de Laurent peut être interprétée comme généralisation de la notion d'une série de Taylor, qui est applicable pour des fonctions qui ne sont pas définies au point  $z_0$ . Elle permet aussi de *classifier* les différents types des singularités isolées.

Par exemple, s'il s'avère que  $a_n = 0$  pour tout  $n < 0$  dans la série de Laurent (1.76), c-à-d, si la série de Laurent ne contient pas de termes avec des puissances négatives de  $(z - z_0)$  et est donc équivalente à une série entière, on dit que  $z_0$  représente une *singularité apparente*. Dans ce cas-là, on a  $\lim_{z \rightarrow z_0} |f(z)| = f_0$  pour un  $f_0 < \infty$ , même si formellement  $f$  ne peut pas être défini en  $z = z_0$ , et on peut donc bien définir une fonction analytique  $g : U \rightarrow \mathbb{C}$  tel que  $f(z) = g(z)$  pour tout  $z \in U \setminus \{z_0\}$ . Ceci est bien le cas pour la fonction (1.72) qui satisfait à la propriété  $\lim_{z \rightarrow 0} f(z) = 1$ . On peut alors dire que cette fonction (1.72) ne possède pas de "vraie" singularité à  $z = 0$ .

La fonction  $f$  possède un *pôle* au point  $z_0$  si on a  $\lim_{z \rightarrow z_0} |f(z)| = \infty$ . Ce pôle est dit *d'ordre  $N$*  (avec  $N \in \mathbb{N}$ ) si dans la série de Laurent (1.76) associée on a  $a_n = 0$  pour tout  $n < -N$  et  $a_{-N} \neq 0$ . Ceci est bien le cas pour la fonction (1.71) qui possède un pôle d'ordre 1 avec le coefficient associé  $a_{-1} = 1$ .

Finalement, une singularité isolée  $z_0$  d'une fonction complexe  $f$  est appelée *singularité essentielle* si la somme dans la série de Laurent (1.76) ne contient pas de terme minimal, c-à-d, si pour tout  $N \in \mathbb{N}$  on trouve un  $n < -N$  tel que

$a_n \neq 0$ . Ceci est le cas pour l'exemple (1.73) dont la série de Laurent s'écrit

$$e^{1/z} = \sum_{n=-\infty}^0 \frac{z^n}{(-n)!}. \quad (1.77)$$

Dans le cas d'une singularité essentielle en  $z_0$ , la limite de  $f(z)$  pour  $z \rightarrow z_0$  n'est pas définie. En effet, on peut montrer que pour tout  $\epsilon > 0$  et tout  $z \in \mathbb{C}$  il existe un  $z' \in \mathbb{C}$  avec  $|z' - z_0| < \epsilon$  tel que  $f(z') = z$ . Autrement dit,  $f$  assume toutes les valeurs complexes possibles dans tout voisinage infinitésimal autour de la singularité essentielle  $z_0$  de  $f$ .

### Exercices

1.1 Lesquelles des fonctions suivantes sont analytiques autour de l'origine  $z = 0$  ? Déterminer, en cas échéant, leurs rayons de convergence ainsi que les cinq premiers termes (jusqu'à l'ordre  $z^4$  inclus) de leurs séries entières.

(a)  $f(z) = \frac{1}{\ln(1-z)}$

(b)  $f(z) = \sqrt{\ln(1+z)}$

(c)  $f(z) = \ln(\sqrt{1+z})$

(d)  $f(z) = \cosh(\sqrt{z})$

(e)  $f(z) = \sqrt{\cosh(z)}$

(f)  $f(z) = \arctan(z)$

## 1.6 L'intégration dans le plan complexe

L'intégration dans le plan complexe est définie de la même manière que l'intégral de chemin dans un espace réel à deux dimensions. A cette fin, revoyons une fois encore l'intégration sur l'axe réel qui est définie par la *somme de Riemann*. On considère une fonction réelle  $f[a, b] \rightarrow \mathbb{R}, x \mapsto f(x)$  que l'on suppose continue, pour ne pas compliquer l'histoire<sup>2</sup>. On peut alors définir l'intégrale de  $f$  sur l'intervalle  $[a, b]$  par

$$\int_a^b f(x)dx = \lim_{\delta x \rightarrow 0, N \rightarrow \infty} \sum_{n=1}^N \frac{f(x_n) + f(x_{n-1})}{2} \delta x \quad (1.78)$$

où on introduit une *discrétisation* de l'intervalle  $[a, b]$  représentée par les points  $x_n = a + n\delta x$  ( $n = 0, 1, \dots, N$ ) tel que  $N\delta x = b - a$ , comme il est montré dans

<sup>2</sup>L'intégration d'une fonction  $f[a, b] \rightarrow \mathbb{R}$  qui possède une discontinuité à  $x = c \in ]a, b[$  peut être définie par la somme des intégrales de  $f$  sur les deux intervalles  $[a, c[$  et  $]c, b]$ .

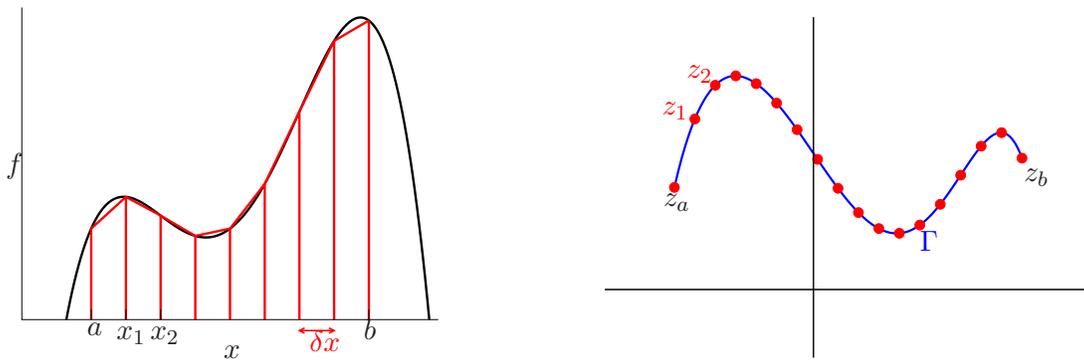


Figure 1.4: Intégration sur l'axe réel (à gauche) et dans le plan complexe (à droite). Sur l'axe réel, l'intégration est définie par la somme de Riemann qui correspond à la somme des aires des différents trapèzes obtenus par une discrétisation de l'intervalle  $[a, b]$ . Dans le plan complexe, l'intégration d'une fonction complexe est définie le long d'un chemin  $\Gamma$  qui relie un point  $z_a \in \mathbb{C}$  à un autre point  $z_b \in \mathbb{C}$ . Comme on fait pour l'intervalle  $[a, b]$  dans le cas réel, ce chemin est discrétisé, c-à-d représenté par de différents points de repère  $z_n \in \Gamma$  avec  $n = 1, \dots, N$ , et l'intégrale complexe correspond alors à la somme des valeurs de la fonction à  $z = z_n$  (ou, plus précisément, au juste milieu entre  $z_n$  et  $z_{n+1}$ ) multipliées par les distances  $z_n - z_{n+1}$ , comme on fait effectivement aussi pour l'intégrale de chemin dans un espace réel à deux dimensions.

la figure 1.4. Evidemment, la somme à droite de l'équation (1.78) correspond à l'aire géométrique située entre le graphe de la fonction et l'abscisse dans la limite où  $N \rightarrow \infty$  et  $\delta x \rightarrow 0$  en respectant  $N\delta x = b - a$ .

Pour définir l'intégrale d'une fonction complexe entre un point complexe  $z_a \in \mathbb{C}$  et un autre point complexe  $z_b \in \mathbb{C}$ , il faut spécifier un *chemin*  $\Gamma$  qui relie ces deux points dans le plan complexe. De manière tout à fait analogue au cas réel, on représente le chemin continu  $\Gamma$  par un ensemble discret des points  $\{z_0 = z_a, z_1, \dots, z_{N-1}, z_N = z_b\}$  avec  $z_n \in \Gamma$  pour tout  $n = 0, \dots, N$ , comme cela est montré dans la figure 1.4, et on définit l'intégrale de  $f$  le long du chemin  $\Gamma$  par

$$\int_{\Gamma} f(z) dz = \lim_{\delta z_n \rightarrow 0, N \rightarrow \infty} \sum_{n=1}^N \frac{f(z_n) + f(z_{n-1})}{2} \delta z_n \quad (1.79)$$

où  $\delta z_n = z_n - z_{n-1}$ . Cette intégrale complexe peut être exprimée par une intégrale réelle si on introduit une *paramétrisation* du chemin  $\Gamma$ , c-à-d, si on définit ce chemin par une transformation continue

$$\Gamma : [a, b] \in \mathbb{R} \rightarrow \mathbb{C}, t \mapsto \Gamma(t) \quad (1.80)$$

avec  $\Gamma(a) = z_a$  et  $\Gamma(b) = z_b$ . On obtient donc

$$\int_{\Gamma} f(z) dz = \int_a^b f[\Gamma(t)] \dot{\Gamma}(t) dt \quad (1.81)$$

où  $\dot{\Gamma}(t) \equiv \frac{d}{dt} \Gamma(t)$ .

Cette paramétrisation nous permet de calculer l'intégrale complexe pour différentes fonctions  $f$ . Considérons d'abord la fonction holomorphe  $f(z) = z^n$  avec  $n \in \mathbb{N}_0$ . Par l'application de la règle (1.81) avec

$$[\Gamma(t)]^n \dot{\Gamma}(t) = \frac{1}{n+1} \frac{d}{dt} [\Gamma(t)]^{n+1} \quad (1.82)$$

et par la règle  $\int_a^b \dot{F}(t) dt = F(b) - F(a)$  pour des intégrales réelles, nous obtenons

$$\int_{\Gamma} z^n dz = \frac{1}{n+1} (z_b^{n+1} - z_a^{n+1}) . \quad (1.83)$$

Evidemment, ce résultat est indépendant du choix du chemin  $\Gamma$  pourvu que celui-ci relie bien les points complexes  $z_a$  et  $z_b$ .

Par la linéarité de l'intégration, il en est de même pour un polynôme qui peut être représenté comme combinaison linéaire de différents monômes  $z^n$ . On a tendance à dire la même chose par rapport à une série entière

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \quad (1.84)$$

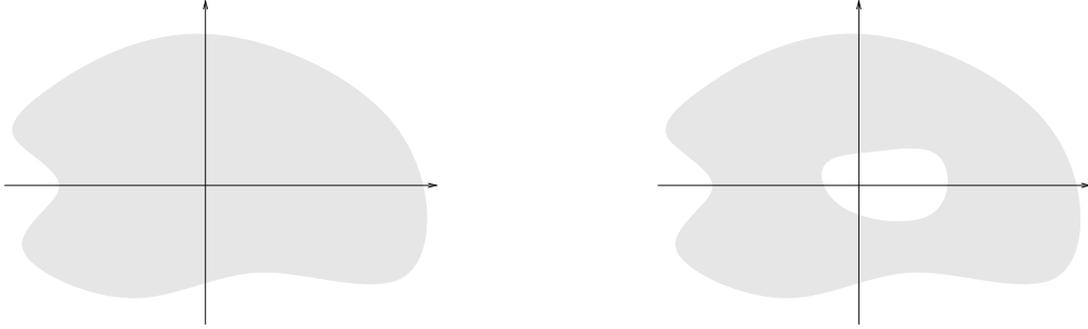


Figure 1.5: Exemple d'une région simplement connexe (à gauche) et non simplement connexe (à droite) dans le plan complexe.

et à calculer son intégrale le long de  $\Gamma$  par

$$\int_{\Gamma} \sum_{n=0}^{\infty} a_n z^n dz = \sum_{n=0}^{\infty} \frac{a_n}{n+1} (z_b^{n+1} - z_a^{n+1}) \quad (1.85)$$

en utilisant l'équation (1.83) et en échangeant la somme sur  $n$  avec l'intégration le long de  $\Gamma$ , mais cette dernière permutation de la somme avec l'intégrale nécessite que la série (1.84) converge en tout point  $z \in \Gamma$ . Il faut donc que non seulement les points  $z_a$  et  $z_b$  mais aussi la totalité du chemin  $\Gamma$  soient entièrement contenus à l'intérieur de la région de convergence de la série (1.84) ou, autrement dit, que  $|\Gamma(t)| < \rho$  pour tout  $t \in [a, b]$  avec  $\rho$  le rayon de convergence de la série (1.84).

L'ensemble des développements précédents nous permettent de conclure que l'intégrale complexe d'une fonction complexe  $f$  entre les points  $z_a$  et  $z_b$  est indépendante du chemin  $\Gamma$  parcouru à condition qu'il existe une région circulaire  $U \subset \mathbb{C}$  tel que  $z_a, z_b \in U$ , que l'intégralité du chemin  $\Gamma$  soit contenu dans  $U$  et que  $f$  soit holomorphe sur  $U$ . Dans ce cas-là, on peut évidemment représenter  $f$  par une série entière définie autour du centre  $z_0$  de  $U$  et ainsi obtenir le résultat (1.85) où il faut remplacer  $z$  par  $z - z_0$ . Ceci implique en particulier que pour un chemin  $\Gamma \subset U$  fermé, c-à-d un chemin (1.80) avec  $\Gamma(a) = \Gamma(b)$  et donc  $z_a = z_b$ , on obtient

$$\oint_{\Gamma} f(z) dz = 0 \quad (1.86)$$

où le symbole  $\oint$  généralement indique une intégrale le long d'un chemin fermé.

Le *théorème de Cauchy* généralise ces considérations et affirme essentiellement que l'équation (1.86) est valable non seulement pour des régions circulaires, mais pour toute région  $U \subseteq \mathbb{C}$  simplement connexe. Cette dernière notion de connectivité provient de la topologie. Une région  $U \subset \mathbb{C}$  est dite *connexe* si pour toute paire de points  $z_a, z_b \in \mathbb{C}$  il existe un chemin  $\Gamma : [a, b] \subset \mathbb{R} \rightarrow U$  avec  $\Gamma(a) = z_a$  et  $\Gamma(b) = z_b$  qui est entièrement contenu dans  $U$ .  $U \subset \mathbb{C}$  est *simplement connexe*

si tout chemin fermé  $\Gamma : [a, b] \rightarrow U$  avec  $\Gamma(a) = \Gamma(b)$  peut être contracté en un seul point dans  $U$  par des déformations continues de  $\Gamma$ . En gros, on peut dire que la région  $U \subset \mathbb{C}$  est simplement connexe si elle ne contient pas de trous, comme illustré dans la figure 1.5.

Le théorème de Cauchy dit alors que pour toute région  $U \subseteq \mathbb{C}$  simplement connexe et toute fonction complexe  $f : U \rightarrow \mathbb{C}$  qui est holomorphe en  $U$  on a  $\oint_{\Gamma} f(z)dz = 0$  pour tout chemin fermé  $\Gamma : [a, b] \rightarrow U$  qui est entièrement contenu dans  $U$ . Notons qu'il est possible d'avoir affaire à une fonction  $f$  qui peut être définie sur une région plus large  $V \subseteq \mathbb{C}$  qui n'est pas forcément simplement connexe. Par exemple, la fonction  $f(z) = 1/z$  peut être définie sur la région  $V = \mathbb{C} \setminus \{0\}$  qui n'est pas simplement connexe car elle contient un trou à l'origine. Le théorème de Cauchy est quand même valable pour cette fonction si le chemin  $\Gamma$  n'encercle pas cette singularité à l'origine. Dans ce cas-là, il sera possible de restreindre la définition de la fonction  $f$  à une région  $U \subset V$  simplement connexe avec  $0 \notin U$  qui contient la totalité du chemin  $\Gamma$ .

Pour illustrer la validité du théorème de Cauchy, nous considérons un chemin fermé  $\Gamma$  qui passe à côté d'une singularité  $z_0$  de la fonction  $f$  à intégrer, comme il est montré dans la figure 1.6. Tant que ce chemin n'encercle pas cette singularité, il est possible de définir une région ouverte  $U \subset \mathbb{C}$  qui inclut  $\Gamma$  et sur laquelle  $f$  est holomorphe. Mais cette région  $U$  ne peut pas forcément être définie avec une forme circulaire. Plus généralement encore, il n'est pas forcément possible de représenter  $f$  par une série entière unique qui converge partout sur  $U$ . En effet, comme il est montré dans la figure 1.6, la proximité de la singularité à  $z_0$  limite le rayon de convergence d'une telle série entière, de sorte que cette série ne converge pas sur la totalité du chemin  $\Gamma$ , mais uniquement sur une partie de  $\Gamma$ .

L'idée centrale de la démonstration du théorème de Cauchy est de sectionner le chemin  $\Gamma$  en différentes partitions  $\gamma_j$  avec  $j = 1, 2, \dots, J$  (pour un  $J \in \mathbb{N}$ ) dont chacune se trouve entièrement à l'intérieur du rayon de convergence de la série entière de  $f$  développée autour d'un point  $z_j \in U$  bien choisi. Ces différents chemins  $\gamma_j$  sont artificiellement fermés à l'intersection de deux régions de convergence voisines, par exemple en connectant leurs bouts par des lignes droites. Avec l'équation (1.86) on a donc

$$\oint_{\gamma_j} f(z)dz = 0 \quad (1.87)$$

pour tout  $j = 1, 2, \dots, J$ . Si on choisit de fermer deux chemins voisins  $\gamma_j$  et  $\gamma_{j'}$  le long des mêmes lignes droites, les deux intégrales de chemin le long de cette ligne droite s'annulent dans l'addition de  $\oint_{\gamma_j} f(z)dz$  et  $\oint_{\gamma_{j'}} f(z)dz$  car elles parcourent cette ligne droite dans des sens opposés. Globalement on obtient donc

$$\oint_{\Gamma} f(z)dz = \sum_{j=1}^J \oint_{\gamma_j} f(z)dz \quad (1.88)$$

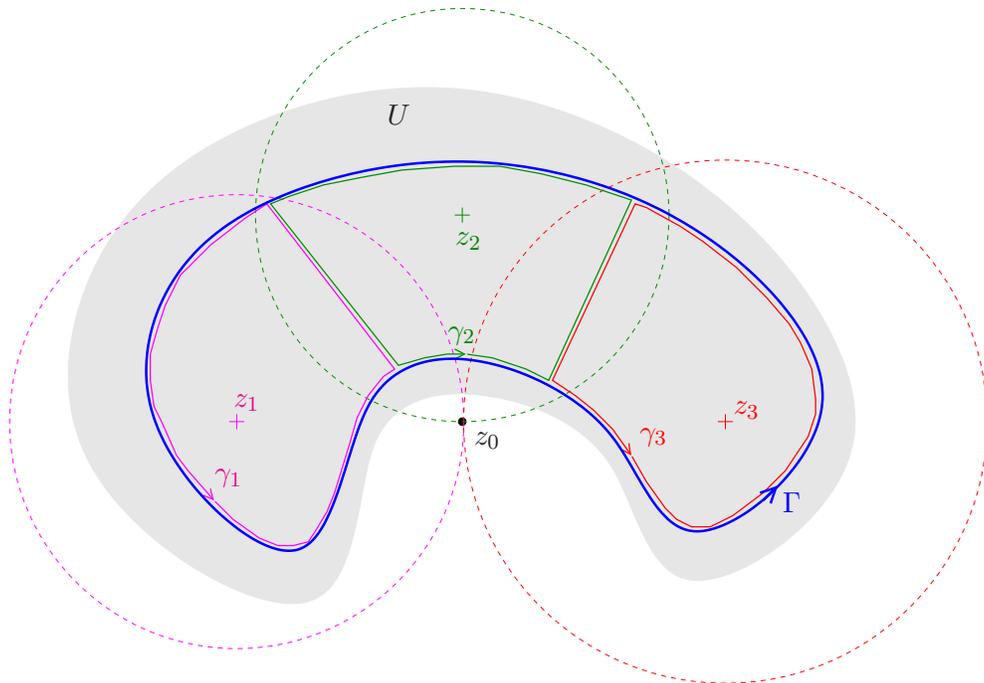


Figure 1.6: Illustration du théorème de Cauchy. On considère un chemin  $\Gamma$  qui passe à côté d'une singularité  $z_0$  de la fonction  $f$  à intégrer. Bien qu'il est possible de définir une région ouverte et simplement connexe  $U \subset \mathbb{C}$  (région grise) dans laquelle  $\Gamma$  est contenue et sur laquelle  $f$  est holomorphe, on ne peut pas représenter  $f$  par une série entière qui converge partout en  $U$ , à cause de la proximité de la singularité à  $z_0$  qui limite le rayon de convergence. On sectionne alors  $\Gamma$  en différentes partitions  $\gamma_j$  qui sont artificiellement fermées le long des lignes droites et dont chacune est entièrement contenue à l'intérieur du rayon de convergence de la série de Taylor de  $f$  développée autour d'un point  $z_j \in U$  bien choisi. On a donc  $\oint_{\gamma_j} f(z) dz = 0$  pour tout  $j$  et, par construction de ces chemins fermés,  $\oint_{\Gamma} f(z) dz = \sum_j \oint_{\gamma_j} f(z) dz$ , ce qui implique  $\oint_{\Gamma} f(z) dz = 0$ .

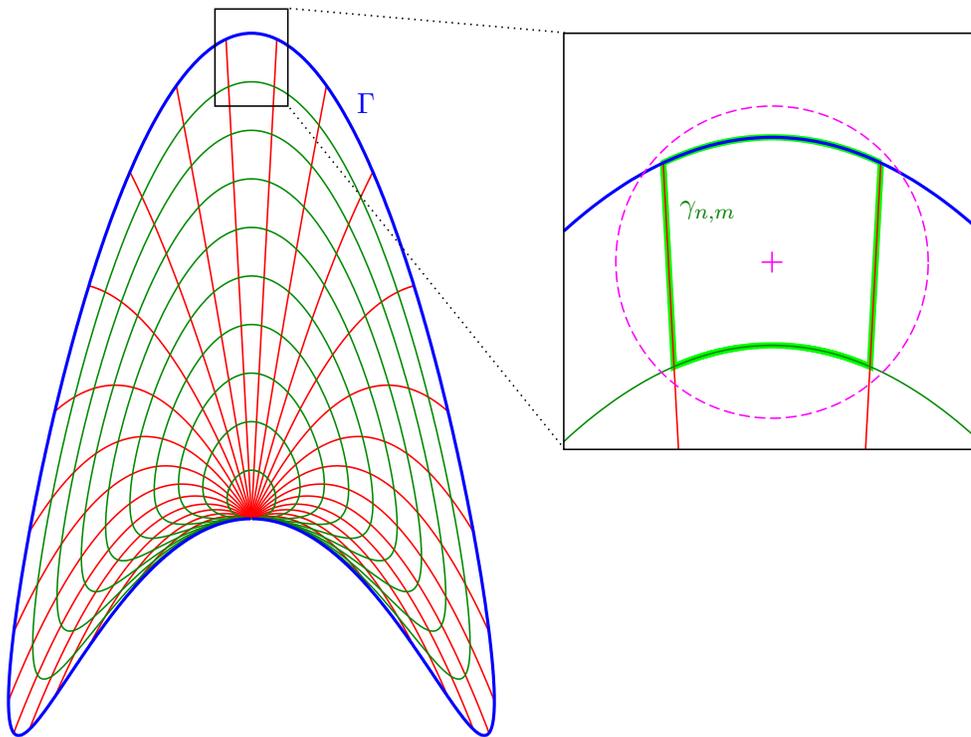


Figure 1.7: Illustration du maillage qui est introduit pour démontrer le théorème de Cauchy. Chaque maille est constitué d'un chemin fermé  $\gamma_{n,m}$ , avec  $n = 1, \dots, N$  et  $m = 1, \dots, M$ , qui, dans la limite  $N, M \rightarrow \infty$  est garanti se situer entièrement à l'intérieur du rayon de convergence de la série entière de  $f$  développée autour d'un point approprié, peu importe la proximité de singularités au voisinage de  $\Gamma$ .

ce qui avec l'équation (1.87) montre le théorème de Cauchy. Clairement, cette approche ne fonctionne pas pour une singularité présente à l'intérieur du chemin  $\Gamma$ . La condition de simple connexité de la région  $U$  est donc essentielle.

Une démonstration du théorème de Cauchy peut être effectuée en définissant une famille continue de chemins fermés qui permet de ramener le chemin  $\Gamma$  de manière continue en un seul point tout en restant à l'intérieur de  $U$ . Plus précisément, cette famille peut être définie via

$$\gamma : [a, b] \times ]0, 1[ \rightarrow U, (t, s) \mapsto \gamma(t, s) \quad (1.89)$$

avec  $\gamma$  étant continu en ses deux arguments  $t$  et  $s$ , où l'on exige les propriétés  $\gamma(t, 1) = \Gamma(t)$  pour tout  $t \in [a, b]$ ,  $\gamma(a, s) = \gamma(b, s) = z_a$  pour tout  $s \in ]0, 1[$ , ainsi que  $\lim_{s \rightarrow 0} \gamma(t, s) = z_a$  pour tout  $t \in [a, b]$ . Nous exigeons également<sup>3</sup> que la vitesse de paramétrisation soit bornée en  $t$  et  $s$ , c-à-d., il existe un  $L > 0$  tel que  $|\partial\gamma(t, s)/\partial t| < L$  et  $|\partial\gamma(t, s)/\partial s| < L$  pour tout  $t \in [a, b]$  et tout  $s \in ]0, 1[$ .

Comme il est illustré dans la Figure 1.7, nous pouvons ensuite introduire un maillage fin à l'intérieur du chemin  $\Gamma$ , qui est construit à partir de la famille de chemins (1.89) en sous-divisant les intervalles  $[a, b]$  et  $]0, 1[$  respectivement en  $N$  et  $M$  sous-intervalles équidistants, avec  $N, M \in \mathbb{N}$ . Ce maillage est constitué des chemins fermés  $\gamma_{n,m} : [0, 2\delta_N + 2\delta_M] \rightarrow U, \varphi \mapsto \gamma_{n,m}(\varphi)$  pour  $n = 1, \dots, N$  et  $m = 1, \dots, M$  avec  $\delta_N = (b - a)/N$  et  $\delta_M = 1/M$ , qui sont définis via

$$\gamma_{n,m}(\varphi) = \begin{cases} \gamma[(n-1)\delta_N + \varphi, m\delta_M] & : 0 \leq \varphi < \delta_N \\ \gamma[n\delta_N, m\delta_M - (\varphi - \delta_N)] & : 0 \leq \varphi - \delta_N < \delta_M \\ \gamma[n\delta_N - (\varphi - \delta_N - \delta_M), (m-1)\delta_M] & : 0 \leq \varphi - \delta_N - \delta_M < \delta_N \\ \gamma[(n-1)\delta_N, (m-1)\delta_M + (\varphi - 2\delta_N - \delta_M)] & : 0 \leq \varphi - 2\delta_N + \delta_M \leq \delta_M \end{cases} \quad (1.90)$$

pour  $m > 1$  ainsi que via

$$\gamma_{n,1}(\varphi) = \begin{cases} \gamma[(n-1)\delta_N + \varphi, \delta_M] & : 0 \leq \varphi < \delta_N \\ \gamma[n\delta_N, \delta_M - (\varphi - \delta_N)] & : 0 \leq \varphi - \delta_N < \delta_M \\ z_a & : 0 \leq \varphi - \delta_N - \delta_M \leq \delta_N \\ \gamma[(n-1)\delta_N, \varphi - 2\delta_N - \delta_M] & : 0 < \varphi - 2\delta_N + \delta_M \leq \delta_M \end{cases} \quad (1.91)$$

pour  $m = 1$ . Il est aisé de montrer

$$\oint_{\Gamma} f(z) dz = \sum_{n=1}^N \sum_{m=1}^M \oint_{\gamma_{n,m}} f(z) dz \quad (1.92)$$

pour tout choix du couple  $N, M$ , et il est également aisé de justifier que dans la limite  $N, M \rightarrow \infty$  nous avons  $\oint_{\gamma_{n,m}} f(z) dz = 0$  pour tout  $n = 1, \dots, N$  et tout

<sup>3</sup>Si cette condition n'est a priori pas satisfaite pour  $\Gamma : [a, b] \rightarrow \mathbb{C}$ , il convient de changer la paramétrisation de ce chemin afin d'avoir  $|\dot{\Gamma}(t)| < L$  pour tout  $t \in [a, b]$ .

$m = 1, \dots, M$ , peu importe la présence éventuelle d'une ou plusieurs singularités de  $f$  à côté du chemin  $\Gamma$ .<sup>4</sup>

Le théorème de Cauchy implique qu'on peut définir une *primitive* unique pour toute fonction  $f : U \rightarrow \mathbb{C}$  qui est holomorphe sur la région simplement connexe  $U$ , c-à-d une fonction holomorphe  $F : U \rightarrow \mathbb{C}$  qui satisfait à la relation  $F'(z) = f(z)$  pour tout  $z \in U$ . On obtient donc

$$\int_{\gamma} f(z) dz = F(z_b) - F(z_a) \quad (1.93)$$

pour tout chemin  $\gamma : [a, b] \subset \mathbb{R} \rightarrow U$  avec  $z_a = \gamma(a)$  et  $z_b = \gamma(b)$ . Evidemment, l'intégrale (1.93) est indépendante du choix de chemin  $\gamma$  pourvu que ce dernier soit entièrement contenu dans  $U$ . En effet, s'il y avait deux chemins  $\gamma, \gamma' \subset U$  connectant  $z_a$  et  $z_b$  dont les intégrales (1.93) associées fournissent des résultats différents, on pourrait construire une contradiction au théorème de Cauchy, notamment en combinant le premier chemin  $\gamma$  avec le deuxième chemin parcouru dans l'autre sens, ce dernier étant défini par  $\bar{\gamma}' : [a, b] \rightarrow U, t \mapsto \bar{\gamma}'(t) = \gamma'(b+a-t)$ , parce que l'intégrale de  $f$  le long du chemin fermé résultant

$$\oint_{\gamma \circ \bar{\gamma}'} f(z) dz = \int_{\gamma} f(z) dz + \int_{\bar{\gamma}'} f(z) dz = \int_{\gamma} f(z) dz - \int_{\gamma'} f(z) dz \quad (1.94)$$

serait non nulle dans ce cas-là.

## 1.7 La formule intégrale de Cauchy

Il nous reste à examiner ce qui se passe si le chemin fermé  $\Gamma$  encercle une singularité isolée de la fonction à intégrer. Considérons d'abord le cas spécial d'une singularité isolée à l'origine et d'un chemin

$$\Gamma : [0, 2\pi] \rightarrow \mathbb{C}, \varphi \mapsto \Gamma(\varphi) = r e^{i\varphi} \quad \text{avec } r > 0 \quad (1.95)$$

qui entoure cette dernière de manière circulaire. Comme il est possible de représenter la fonction en question par une série de Laurent (1.76), il suffit de calculer l'intégrale de chaque terme de cette série, c-à-d, l'intégrale de la fonction  $1/z^n$  pour tout  $n \in \mathbb{N}$ . En utilisant la paramétrisation (1.95) nous obtenons

$$\oint_{\Gamma} \frac{1}{z^n} dz = \int_0^{2\pi} \frac{\dot{\Gamma}(\varphi)}{[\Gamma(\varphi)]^n} d\varphi = \frac{i}{r^{n-1}} \int_0^{2\pi} e^{-i(n-1)\varphi} d\varphi = \begin{cases} 0 & : n > 1 \\ 2\pi i & : n = 1 \end{cases} \quad (1.96)$$

<sup>4</sup>Cette façon de démontrer le théorème de Cauchy nous indique également que le champ d'application de ce théorème n'est pas restreint à des fonctions et chemins définis par rapport à une région simplement connexe, mais peut être élargi pour inclure aussi des chemins définis dans des régions non simplement connexes, pourvu qu'il soit possible de continûment contracter le chemin en un seul point sans toucher une singularité de la fonction  $f$  à intégrer.

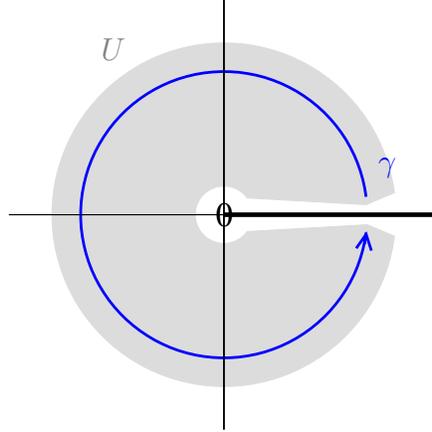


Figure 1.8: Visualisation de l'intégration de la fonction  $f(z) = 1/z$  le long du chemin non fermé  $\gamma$  (1.97). Afin de calculer l'intégrale par l'évaluation de la primitive  $F(z) = \ln(z)$  de  $f$ , il faut définir une région simplement connexe  $U$  qui inclut  $\gamma$  et exclut la singularité de  $f$  à l'origine. Comme on peut définir cette région  $U$  tel qu'elle exclut l'axe réel positif, ce dernier est alors choisi comme axe de discontinuité du logarithme complexe afin d'assurer que la primitive  $F$  soit holomorphe sur  $U$ .

où la dernière intégrale réelle sur  $\varphi$  peut être calculée en séparant les parties réelle et imaginaire de l'intégrand. L'intégrale de la fonction  $1/z^n$  sur ce chemin fermé vaut donc zéro pour tout entier  $n \in \mathbb{Z} \setminus \{1\}$  et elle est indépendante du choix du rayon  $r$  aussi dans le cas  $n = 1$ .

On arrive à la même conclusion en calculant cette intégrale par une primitive, comme nous venons de le discuter dans la section précédente. On a besoin, pour cela, de la définition d'une région  $U$  simplement connexe dans laquelle la fonction  $1/z^n$  est holomorphe et qui contient l'ensemble du chemin d'intégration. Ceci nécessite de "couper" le chemin fermé (1.95) tel qu'on le définit, par exemple, par

$$\gamma : [\epsilon, 2\pi - \epsilon] \rightarrow \mathbb{C}, \varphi \mapsto \Gamma(\varphi) = r e^{i\varphi} \quad (1.97)$$

avec  $\epsilon \rightarrow 0_+$ , comme il est indiqué dans la figure 1.8. On obtient alors

$$\int_{\gamma} \frac{1}{z^n} dz = \left[ \frac{-1}{(n-1)z^{n-1}} \right]_{\gamma(\epsilon)}^{\gamma(2\pi-\epsilon)} = -\frac{2i \sin[(n-1)\epsilon]}{(n-1)r^{n-1}} \xrightarrow{\epsilon \rightarrow 0} 0 \quad (1.98)$$

pour  $n \neq 1$ , tandis que pour  $n = 1$  on a

$$\int_{\gamma} \frac{1}{z} dz = [\ln(z)]_{\gamma(\epsilon)}^{\gamma(2\pi-\epsilon)} = [\ln(r) + i\varphi]_{\varphi=\epsilon}^{\varphi=2\pi-\epsilon} = 2i(\pi - \epsilon) \xrightarrow{\epsilon \rightarrow 0} 2\pi i \quad (1.99)$$

où on a, pour cette fois, choisi la représentation polaire  $z = r e^{i\varphi}$  avec  $0 < \varphi < 2\pi$

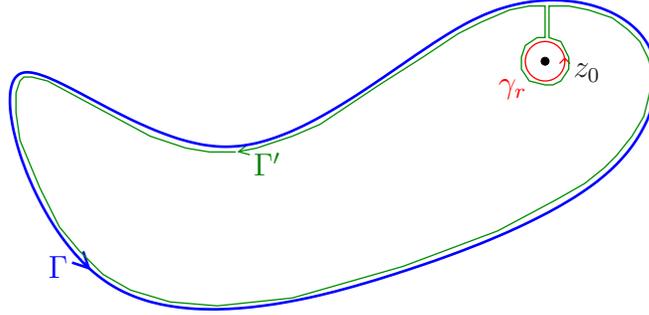


Figure 1.9: Illustration des chemins fermés  $\Gamma$ ,  $\Gamma'$  et  $\gamma_r$  utilisés dans l'équation (1.100). Comme le chemin  $\Gamma'$  n'encercle pas la singularité à  $z_0$ , le théorème de Cauchy nous permet de réduire l'intégrale de  $1/(z - z_0)^n$  le long de  $\Gamma$  à une intégrale le long du chemin circulaire  $\gamma_r$  pour laquelle le résultat est connu par l'équation (1.96).

dans l'évaluation du logarithme complexe, afin d'assurer que cette primitive de  $1/z$  soit holomorphe le long du chemin  $\gamma$ .

On peut généraliser ces considérations pour des chemins fermés  $\Gamma : [a, b] \rightarrow \mathbb{C}$  de forme arbitraire. Considérons, comme montré dans la figure 1.9, un chemin qui encercle une singularité isolée de la fonction  $f$  à intégrer, située à  $z = z_0$ , de sorte que le chemin effectue dans le sens trigonométrique un tour unique autour de la singularité. Dans ce cas-là, il est possible de décomposer l'intégrale de  $f$  le long de  $\Gamma$  selon

$$\oint_{\Gamma} f(z)dz = \oint_{\Gamma'} f(z)dz + \oint_{\gamma_r} f(z)dz \quad (1.100)$$

où on définit

$$\gamma_r : [0, 2\pi] \rightarrow \mathbb{C}, \varphi \mapsto \gamma_r(\varphi) = z_0 + re^{i\varphi} \quad (1.101)$$

pour un rayon  $r > 0$ . Le chemin  $\Gamma'$  est défini de façon à ce qu'il parcoure  $\Gamma$ , hormis en un point où il rejoint, suivant une ligne droite, le cercle  $\gamma_r$  qu'il parcourt dans le sens opposé, pour finir par rejoindre  $\Gamma$  le long de la même ligne droite, comme illustré dans la figure 1.10. Visiblement,  $\Gamma'$  représente un chemin fermé qui n'encercle pas la singularité  $z_0$  de  $f$ . En supposant que  $z_0$  soit la seule singularité de  $f$  se trouvant à l'intérieur du chemin  $\Gamma$ , on a donc  $\oint_{\Gamma'} f(z)dz = 0$  par le théorème de Cauchy. Par l'équation (1.100), l'intégrale le long du chemin fermé  $\Gamma$  peut donc être "réduite" à une intégrale le long du cercle  $\gamma_r$  pour un rayon  $r > 0$  qui peut être arbitrairement petit.

Plus spécifiquement, on obtient donc

$$\oint_{\Gamma} \frac{1}{(z - z_0)^n} dz = \oint_{\gamma_r} \frac{1}{(z - z_0)^n} dz = 2\pi i \delta_{n1} = \begin{cases} 0 & : n > 1 \\ 2\pi i & : n = 1 \end{cases} \quad (1.102)$$

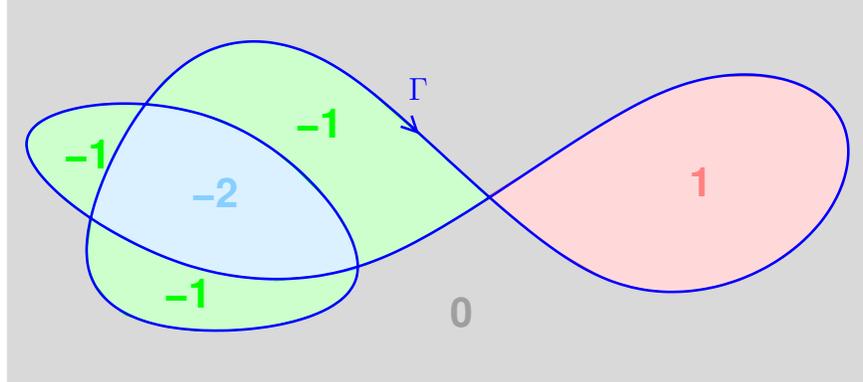


Figure 1.10: Exemple d'un chemin  $\Gamma$  non simple. Les différentes couleurs indiquent les différentes régions dans le plan complexe où l'indice  $\text{Ind}_\Gamma(z)$  du chemin vaut 0 (gris), 1 (rouge),  $-1$  (vert) ou  $-2$  (bleu).

comme généralisation de l'équation (1.96) pour ce chemin  $\Gamma$ . La procédure est également claire dans le cas où le chemin  $\Gamma$  encercle la singularité  $z_0$  plusieurs fois, avec éventuellement un sens de parcours anti-trigonométrique. On a toujours  $\oint_\Gamma (z - z_0)^{-n} dz = 0$  pour tout entier  $n > 1$ , mais pour  $n = 1$  on obtient

$$\oint_\Gamma \frac{1}{z - z_0} dz = 2\pi i k \quad (1.103)$$

où  $k \in \mathbb{Z}$  représente le *nombre de tours* du chemin  $\Gamma$  autour de la singularité à  $z_0$  comptés dans le sens trigonométrique (des tours dans l'autre sens sont comptés de manière négative). Ceci nécessite évidemment que  $\Gamma$  ne touche pas la singularité directement, c-à-d, que  $z_0 \neq \Gamma(t)$  pour tout  $t \in [a, b]$ .

D'un point de vu mathématique, on peut voir les choses autrement et *définir* le nombre de tours d'un chemin autour d'un point  $z_0 \in \mathbb{C}$  par justement la relation (1.103). Ce "nombre de tours" est appelé *indice* du chemin  $\Gamma$  en  $z_0$  et est défini par

$$\text{Ind}_\Gamma(z_0) = \frac{1}{2\pi i} \oint_\Gamma \frac{1}{z - z_0} dz. \quad (1.104)$$

Pour un chemin  $\Gamma$  *simple* qui ne se croise pas — c-à-d, un chemin fermé  $\Gamma : [a, b] \rightarrow \mathbb{C}$  pour lequel  $\Gamma(t_1) = \Gamma(t_2)$  implique  $t_1 = t_2$  pour tous  $t_1, t_2 \in [a, b[$  — on obtient deux possibilités pour l'indice de  $\Gamma$ , notamment  $\text{Ind}_\Gamma(z) = 0$  si  $z$  se trouve à l'*extérieur* de  $\Gamma$  et  $\text{Ind}_\Gamma(z) = \pm 1$  si  $z$  se trouve à l'*intérieur* de  $\Gamma$ , le signe dépendant de l'orientation (trigonométrique ou anti-trigonométrique) de  $\Gamma$ . Evidemment, la situation est plus compliquée pour le chemin non simple montré dans la figure 1.10.

Le cheminement qui nous a mené à l'équation (1.104) peut facilement être étendu au cas plus général où le terme  $1/(z - z_0)$  dans l'intégrale est multiplié

par une fonction  $f(z)$  holomorphe sur  $\Gamma$  et en son intérieur. Aussi dans ce cas-là, l'intégrale le long de  $\Gamma$  peut être réduite à une ou plusieurs intégrales (dépendant de l'indice de  $\Gamma$  en  $z_0$ ) le long d'un chemin circulaire entourant le pôle en  $z_0$ , comme illustré dans la figure 1.9. En définissant ce cercle avec un rayon qui est suffisamment petit, on peut justifier que  $f(z)$  peut être remplacé par  $f(z_0)$  dans cette dernière intégrale. Ceci nous conduit à la *formule intégrale* de Cauchy : Si on peut définir une région  $U \subseteq \mathbb{C}$  simplement connexe tel que le chemin fermé  $\Gamma : [a, b] \subset \mathbb{R} \rightarrow U$  soit entièrement contenu en  $U$  et  $f : U \rightarrow \mathbb{C}$  soit holomorphe en  $U$ , on obtient la relation

$$\frac{1}{2\pi i} \oint_{\Gamma} \frac{f(z)}{z - z_0} dz = \text{Ind}_{\Gamma}(z_0) f(z_0) \quad (1.105)$$

pour tout  $z_0 \in U$ . La valeur d'une fonction holomorphe à un point  $z_0 \in \mathbb{C}$  est donc liée à ces valeurs dans les alentours de ce point.

## 1.8 Le théorème des résidus

Nous sommes donc maintenant en position de discuter ce qui se passe si on intègre une fonction holomorphe le long d'un chemin fermé. Supposons d'abord que cette fonction possède une seule singularité isolée à l'intérieur du chemin. Plus spécifiquement, on suppose qu'il existe une région  $U \subseteq \mathbb{C}$  ouverte et simplement connexe dans laquelle on définit un chemin fermé  $\Gamma : [a, b] \subset \mathbb{R} \rightarrow U$  et une fonction complexe  $f : U \setminus \{z_0\} \rightarrow \mathbb{C}, z \mapsto f(z)$  qui est holomorphe en  $U \setminus \{z_0\}$  et qui éventuellement possède une singularité isolée en  $z_0$ . Dans ce cas-là, on peut représenter  $f$  par une série de Laurent

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n \quad (1.106)$$

dans un voisinage ouvert de  $z_0$  (avec un rayon de convergence qui est, comme pour une série entière, limité par la distance minimale à une autre singularité éventuelle de  $f$ ). En adaptant la procédure développée dans la section précédente à ce cas-là et en utilisant la représentation (1.106) dans l'évaluation de l'intégrale le long du chemin

$$\gamma_r : [0, 2\pi] \rightarrow \mathbb{C}, \varphi \mapsto \gamma_r(\varphi) = z_0 + r e^{i\varphi} \quad (1.107)$$

qui entoure cette singularité sur un cercle avec un petit rayon  $r > 0$ , nous obtenons

$$\oint_{\Gamma} f(z) dz = 2\pi i \text{Ind}_{\Gamma}(z_0) a_{-1} \quad (1.108)$$

où  $\text{Ind}_{\Gamma}(z_0)$  est l'indice du chemin  $\Gamma$  en  $z_0$ , défini par l'équation (1.104), et  $a_{-1}$  est le coefficient associé au terme  $(z - z_0)^{-1}$  de la série de Laurent (1.106). On appelle

$$a_{-1} = \text{Res}_f(z_0) \quad (1.109)$$

le *résidu* de  $f$  à  $z_0$ .

De manière plus générale, nous obtenons

$$\oint_{\Gamma} (z - z_0)^n f(z) dz = 2\pi i \operatorname{Ind}_{\Gamma}(z_0) a_{-1-n} \quad (1.110)$$

pour tout  $n \in \mathbb{Z}$ . Encore une fois, on peut voir les choses autrement et exprimer les coefficients de la série de Laurent de  $f$  autour de  $z_0$  par

$$a_n = \frac{1}{2\pi i} \oint_{\gamma_r} \frac{f(z)}{(z - z_0)^{n+1}} dz \quad (1.111)$$

pour tout  $n \in \mathbb{Z}$ , où  $\gamma_r$  est le chemin circulaire défini par l'équation (1.107) pour un rayon  $r > 0$  qui est suffisamment petit. Cette dernière relation (1.111) serait également valable pour des fonctions  $f$  qui sont holomorphes en  $z_0$  ou qui y possèdent une singularité apparente. Dans ce cas-là, on obtient l'expression

$$f^{(n)}(z_0) \equiv \frac{d^n f}{dz^n}(z_0) = \frac{n!}{2\pi i} \oint_{\gamma_r} \frac{f(z)}{(z - z_0)^{n+1}} dz \quad (1.112)$$

pour la  $n$ -ième dérivée de  $f$  à  $z = z_0$ .

Plus généralement encore, supposons que la fonction complexe  $f$  en question possède plusieurs singularités isolées  $z_0, z_1, z_2, \dots \in U$  dans la région simplement connexe  $U \subseteq \mathbb{C}$ , de telle manière que  $f : U \setminus S \rightarrow \mathbb{C}$  soit holomorphe en  $U \setminus S$  où  $S = \{z_0, z_1, z_2, \dots\}$  représente l'ensemble des singularités isolées de  $f$  en  $U$ . On parle, dans ce contexte-là, d'une fonction qui est *méromorphe* en  $U$ , pourvu que les singularités  $z_k$  de  $f$  ne soient pas essentielles. Par une généralisation évidente de l'approche développée ci-dessus (qui évidemment nécessite d'encercler plusieurs singularités à l'intérieur  $\Gamma$  par des chemins circulaires), nous obtenons

$$\oint_{\Gamma} f(z) dz = 2\pi i \sum_k \operatorname{Ind}_{\Gamma}(z_k) \operatorname{Res}_f(z_k) \quad (1.113)$$

pour tout chemin fermé  $\Gamma : [a, b] \subset \mathbb{R} \rightarrow U$ , où  $\operatorname{Res}_f(z_k)$  est le résidu de  $f$  à la singularité  $z_k$  défini par l'équation (1.109).

De manière assez remarquable, ce théorème des résidus (1.113) permet de calculer certaines intégrales définies sur l'axe réel. L'idée principale est de transformer une telle intégrale réelle en une intégrale *complexe* qui est définie le long d'un *chemin fermé* dans le plan complexe. L'évaluation de cette intégrale se fait donc par l'identification des *singularités isolées* de la fonction à intégrer qui se trouvent à l'intérieur de ce chemin. Selon l'équation (1.113) il faut donc déterminer les indices du chemin en de telles singularités ainsi que les *résidus* associés de la fonction à intégrer.

Considérons, par exemple, une intégrale de la forme

$$J = \int_0^{2\pi} \frac{F[\cos(\varphi), \sin(\varphi)]}{G[\cos(\varphi), \sin(\varphi)]} d\varphi \quad (1.114)$$

où  $F(x, y)$  et  $G(x, y)$  sont des polynômes (réels ou complexes) en  $x$  et  $y$  avec la propriété que  $G[\cos(\varphi), \sin(\varphi)] \neq 0$  pour tout  $\varphi \in [0, 2\pi]$ . Cette intégrale peut être transformée en une intégrale complexe le long du chemin fermé

$$\Gamma : [0, 2\pi] \rightarrow \mathbb{C}, \varphi \mapsto \Gamma(\varphi) = e^{i\varphi}. \quad (1.115)$$

en identifiant les égalités<sup>5</sup>

$$\cos(\varphi) = \frac{1}{2} \left( \Gamma(\varphi) + \frac{1}{\Gamma(\varphi)} \right), \quad (1.116)$$

$$\sin(\varphi) = \frac{1}{2i} \left( \Gamma(\varphi) - \frac{1}{\Gamma(\varphi)} \right) \quad (1.117)$$

le long de ce chemin  $\Gamma$ . En prenant en compte  $\dot{\Gamma}(\varphi) = i\Gamma(\varphi)$  pour le chemin (1.115), nous obtenons  $J = \oint_{\Gamma} f(z)dz$  avec la fonction méromorphe

$$f(z) = \frac{1}{iz} \frac{F\left[\frac{1}{2}\left(z + \frac{1}{z}\right), \frac{1}{2i}\left(z - \frac{1}{z}\right)\right]}{G\left[\frac{1}{2}\left(z + \frac{1}{z}\right), \frac{1}{2i}\left(z - \frac{1}{z}\right)\right]}. \quad (1.118)$$

Il faut alors identifier les pôles  $z_0, z_1, \dots$  de cette fonction rationnelle  $f$  avec  $|z_k| < 1$ , qui se trouvent donc à l'intérieur du chemin  $\Gamma$ , et en déterminer les résidus associés de  $f$ . Comme  $\text{Ind}_{\Gamma}(z_k) = 1$  pour tous ces pôles, on obtient finalement

$$\int_0^{2\pi} \frac{F[\cos(\varphi), \sin(\varphi)]}{G[\cos(\varphi), \sin(\varphi)]} d\varphi = 2\pi i \sum_k \text{Res}_f(z_k). \quad (1.119)$$

Une autre application importante du théorème des résidus concerne les *intégrales de Fourier*

$$\hat{f} : \mathbb{R} \rightarrow \mathbb{C}, k \mapsto \hat{f}(k) = \int_{-\infty}^{\infty} f(x)e^{ikx} dx \quad (1.120)$$

pour une fonction rationnelle

$$f : \mathbb{R} \rightarrow \mathbb{C}, x \mapsto f(x) = \frac{\sum_{n=0}^N p_n z^n}{\sum_{m=0}^M q_m z^m} \quad (1.121)$$

avec  $M > N$  pour laquelle les zéros du dénominateur ne se situent pas sur l'axe réel. Afin d'évaluer cette intégrale (1.120), nous définissons un chemin complexe  $\Gamma = \Gamma_1 \circ \Gamma_2$  qui consiste en la *concaténation* des deux chemins

$$\Gamma_1 : [-R, R] \rightarrow \mathbb{C}, x \mapsto x, \quad (1.122)$$

$$\Gamma_2 : [0, \pi] \rightarrow \mathbb{C}, \varphi \mapsto Re^{\pm i\varphi} \quad (1.123)$$

---

<sup>5</sup>L'identification  $\cos(\varphi) = [\Gamma(\varphi) + \Gamma^*(\varphi)]/2$  et  $\sin(\varphi) = [\Gamma(\varphi) - \Gamma^*(\varphi)]/(2i)$  est, en principe, aussi correcte, mais elle ne mène pas à une fonction holomorphe pour laquelle on peut appliquer le théorème des résidus.

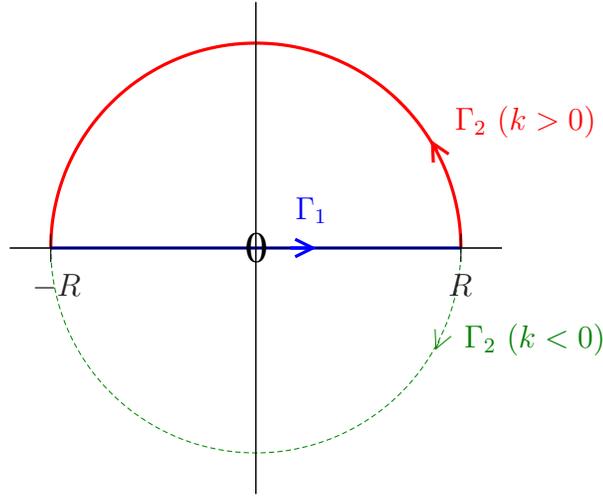


Figure 1.11: Visualisation des deux chemins  $\Gamma_1$  et  $\Gamma_2$  définis selon les équations (1.122) et (1.123).

où le signe dans l'exposant dans l'équation (1.123) correspond au signe de  $k$  et vaut donc  $+1$  si  $k > 0$  et  $-1$  si  $k < 0$ <sup>6</sup>. Comme il est illustré dans la figure 1.11, le chemin  $\Gamma$  résultant est bien fermé car  $\Gamma_1(R) = \Gamma_2(0)$  et  $\Gamma_1(-R) = \Gamma_2(\pi)$ , et nous obtenons

$$\oint_{\Gamma} f(z)e^{ikz} dz = \int_{\Gamma_1} f(z)e^{ikz} dz + \int_{\Gamma_2} f(z)e^{ikz} dz. \quad (1.124)$$

Nous allons maintenant discuter l'équation (1.124) dans la limite  $R \rightarrow \infty$ . Dans cette limite-là nous obtenons évidemment

$$\lim_{R \rightarrow \infty} \int_{\Gamma_1} f(z)e^{ikz} dz = \int_{-\infty}^{\infty} f(x)e^{ikx} dx = \hat{f}(k) \quad (1.125)$$

par la paramétrisation (1.122). L'autre intégrale *disparaît* dans cette limite, c-à-d

$$\lim_{R \rightarrow \infty} \int_{\Gamma_2} f(z)e^{ikz} dz = 0, \quad (1.126)$$

comme on peut montrer pour la fonction rationnelle (1.121). Nous obtenons donc

$$\hat{f}(k) = \lim_{R \rightarrow \infty} \oint_{\Gamma} f(z)e^{ikz} dz = \pm 2\pi i \sum_k \text{Res}_F(z_k) \quad (1.127)$$

avec  $F(z) \equiv f(z)e^{ikz}$ , où  $z_0, z_1, \dots$  représentent les pôles complexes de la fonction  $f$  qui se trouvent au-dessus de l'axe réel (c-à-d  $\text{Im}(z_k) > 0$ ) si on a choisi un

<sup>6</sup>Si  $k = 0$ , on a la liberté de choisir un signe positif ou négatif dans l'exposant. Notons que dans ce cas spécial il faut imposer  $M > N + 1$  pour être sûr que l'intégrale (1.120) converge.

signe positif dans l'exposant dans l'équation (1.123) et au-dessous de l'axe réel (c-à-d  $\text{Im}(z_k) < 0$ ) dans le cas d'un signe négatif dans l'exposant dans l'équation (1.123). Cette méthode permet, par exemple, de vérifier que la transformée de Fourier d'une fonction lorentzienne  $f(x) = a/(x^2 + a^2)$  avec  $a > 0$  correspond à une fonction exponentielle de la forme  $\hat{f}(k) = \pi \exp(-a|k|)$ .

### Exercices

1.2 (a) Exprimer les fonctions complexes

$$(i) f(z) = \frac{z^2 - z - 2}{z^3 + 3z^2 + 3z + 1}$$

$$(ii) f(z) = \frac{z + i}{z^4 - 2z^3 + z^2 - 2z}$$

$$(iii) f(z) = \frac{z^2 + 1}{z^3 - 2z^2 + z}$$

sous la forme

$$f(z) = \sum_k \sum_{n=-\infty}^{-1} a_n^{(k)} (z - z_k)^n$$

avec  $a_n^{(k)} \in \mathbb{C}$ , où  $z_k$  ( $k \in \mathbb{N}$ ) sont les pôles de  $f(z)$ .

(b) Pour chacune de ces fonctions, calculer l'intégrale  $\oint_{\Gamma} f(z) dz$  avec

$$\Gamma : [0, 2\pi] \rightarrow \mathbb{C}, \varphi \mapsto \left( \frac{3}{4} + \frac{1}{2} \sin \varphi \right) e^{i\varphi}$$

1.3 On considère les fonctions complexes

$$(i) f(z) = \frac{1}{(z^2 + 1)\sqrt{z - i}} + \frac{2z + i}{\cosh(\pi z)} + z \cos\left(\frac{\pi}{z}\right)$$

$$(ii) f(z) = \frac{z^2 + \pi^2}{e^z + 1} \ln(z + 5) + \frac{e^z}{z(z - 1)}$$

$$(iii) f(z) = ze^{1/z} + \frac{z}{e^z - 1} + \frac{1}{3z - 1} + \frac{1}{z + 2}$$

Pour chacune d'entre elles :

(a) Déterminer les singularités de cette fonction ainsi que leurs types (singularité apparente, pôle, singularité essentielle ou point de branchement).

(b) Calculer l'intégrale  $\oint_{\Gamma} f(z) dz$  pour

$$\Gamma : [0, 2\pi] \rightarrow \mathbb{C}, \varphi \mapsto \left( 1 - \frac{2}{3} \sin \varphi \right) e^{i\varphi}$$

1.4 Calculer avec le théorème des résidus

(a)  $\frac{1}{2\pi} \int_0^{2\pi} \frac{e^{in\varphi}}{1 + \epsilon \sin \varphi} d\varphi$  pour  $0 < \epsilon < 1$  avec  $n \in \mathbb{Z}$ .

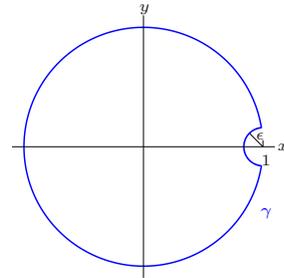
(b)  $\int_{-\infty}^{\infty} \frac{1}{z^4 + a^4} dz$  pour  $a > 0$ .

(c)  $\int_{-\infty}^{\infty} \frac{1}{z^3 + ia^3} dz$  pour  $a > 0$ .

1.5 Calculer avec le théorème des résidus

(a)  $\frac{1}{2\pi} \int_0^{2\pi} \frac{1}{\sqrt{1 - e^{i\varphi}}} d\varphi$

pour  $a > 0$ , en utilisant le chemin ci-contre et en plaçant la discontinuité de la racine d'une manière appropriée.

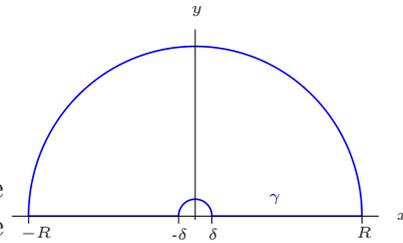


1.6 Calculer avec le théorème des résidus

(a)  $\int_0^{\infty} \frac{\ln x}{x^2 + a^2} dx$

(b)  $\int_0^{\infty} \frac{\ln x}{x^4 + a^4} dx$

pour  $a > 0$ , en utilisant le chemin ci-contre et en plaçant la discontinuité du logarithme d'une manière appropriée.

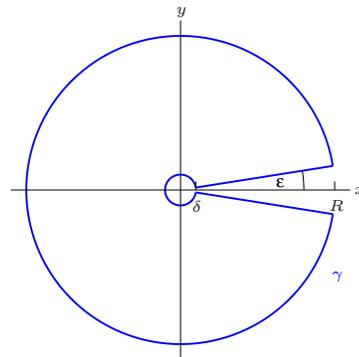


1.7 Calculer avec le théorème des résidus

(a)  $\int_0^{\infty} \frac{\sqrt{x}}{x^2 + a^2} dx$

(b)  $\int_0^{\infty} \frac{1}{\sqrt{x}(x^2 + a^2)} dx$

pour  $a > 0$ , en utilisant le chemin ci-contre et en plaçant la discontinuité de la racine d'une manière appropriée.



# Chapitre 2

## Les transformations intégrales

### 2.1 La fonction delta

Nous cherchons une fonction  $\delta : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \delta(x)$  qui satisfait aux exigences suivantes : elle est symétrique, c.à.d.

$$\delta(x) = \delta(-x) \quad (2.1)$$

pour tout  $x \in \mathbb{R}$ , et on exige

$$\int_{-\infty}^{\infty} f(x)\delta(x)dx = f(0) \quad (2.2)$$

pour tout  $f : \mathbb{R} \rightarrow \mathbb{C}$  qui est continu en  $x = 0$ , où nous identifions de manière générale

$$\int_{-\infty}^{\infty} g(x)dx \equiv \lim_{R \rightarrow \infty} \int_{-R}^R g(x)dx \quad (2.3)$$

pour une fonction  $g : \mathbb{R} \rightarrow \mathbb{C}$  quelconque. Cette dernière exigence (2.2) nous indique que la fonction recherchée possède des propriétés assez pathologiques, vu que l'intégrale dans l'équation (2.2) est censée toujours redonner la valeur de  $f$  à l'origine, indépendamment de ce que  $f$  vaut ailleurs. Nous sommes donc amenés à conclure (de manière prématurée à ce stade-là) que la fonction en question doit valoir

$$\delta(x) = \begin{cases} \infty & : x = 0 \\ 0 & : x \neq 0 \end{cases} \quad (2.4)$$

Evidemment, une fonction avec de telles propriétés ne peut pas être analytique. D'un point de vue mathématique, il ne s'agit même pas d'une fonction, mais plutôt d'une *distribution*, c-à-d, d'une fonctionnelle linéaire et continue. Le terme *fonctionnelle* signifie une transformation d'un espace de fonctions à un espace de nombres réels ou complexes. Définissons, dans notre cas, l'espace

$$C = \{f : U \subseteq \mathbb{R} \rightarrow \mathbb{C}, x \mapsto f(x) \text{ continu en } x = 0\} \quad (2.5)$$

des fonctions  $f$  qui sont définies sur un sous-ensemble  $U$  de l'axe réel et qui sont continues à l'origine. La “fonction” delta, qui est aussi nommée *distribution de Dirac*, peut alors être définie de la manière suivante<sup>1</sup> :

$$\bar{\delta} : C \rightarrow \mathbb{C}, f \mapsto \bar{\delta}(f) = f(0). \quad (2.6)$$

Elle correspond à une fonctionnelle *linéaire*, car

$$\bar{\delta}(\alpha f + \beta g) = \alpha \bar{\delta}(f) + \beta \bar{\delta}(g) \quad (2.7)$$

pour tout  $\alpha, \beta \in \mathbb{C}$  et tout  $f, g \in C$ , et elle représente une transformation *continue*, car pour toute suite convergente  $\{f_n\}_{n \in \mathbb{N}}$  de fonctions  $f_n \in C$  on a

$$\bar{\delta} \left( \lim_{n \rightarrow \infty} f_n \right) = \lim_{n \rightarrow \infty} \bar{\delta}(f_n). \quad (2.8)$$

Bien que le concept des distributions représente le bon cadre mathématique pour une introduction rigoureuse de la fonction delta (que nous continuons à appeler “fonction” dans la suite, même si elle ne correspond pas vraiment à une fonction à proprement parler), il peut toutefois être utile de l'imaginer en tant que fonction hypothétique qui satisfait aux propriétés (2.1) et (2.2), surtout si on souhaite déterminer des règles de calcul associées à la distribution de Dirac. Dans ce but, nous allons maintenant discuter comment cette fonction pathologique  $\delta : \mathbb{R} \rightarrow \mathbb{R}$  peut être approximée par des fonctions tout à fait régulières. De manière générale, une telle *régularisation* de la fonction delta peut être générée par une fonction  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ ,  $\xi \mapsto \varphi(\xi)$  qui satisfait aux propriétés  $\varphi(-\xi) = \varphi(\xi)$  et

$$\int_{-\infty}^{\infty} \varphi(\xi) d\xi = 1. \quad (2.9)$$

Nous pouvons ainsi “construire” la fonction delta par

$$\delta(x) = \lim_{\epsilon \rightarrow 0_+} \frac{1}{\epsilon} \varphi \left( \frac{x}{\epsilon} \right). \quad (2.10)$$

En effet, un calcul rapide,

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) \lim_{\epsilon \rightarrow 0_+} \frac{1}{\epsilon} \varphi \left( \frac{x}{\epsilon} \right) dx &= \lim_{\epsilon \rightarrow 0_+} \int_{-\infty}^{\infty} f(x) \varphi \left( \frac{x}{\epsilon} \right) \frac{1}{\epsilon} dx \\ &= \lim_{\epsilon \rightarrow 0_+} \int_{-\infty}^{\infty} f(\epsilon y) \varphi(y) dy \\ &= \int_{-\infty}^{\infty} \lim_{\epsilon \rightarrow 0_+} f(\epsilon y) \varphi(y) dy = f(0), \end{aligned} \quad (2.11)$$

---

<sup>1</sup>Nous ne souhaitons pas confondre la *fonctionnelle*  $\bar{\delta} : C \rightarrow \mathbb{C}$  avec la *fonction* hypothétique  $\delta : \mathbb{R} \rightarrow \mathbb{C}$  qui est reliée à la première par l'équation  $\bar{\delta}(f) = \int_{-\infty}^{\infty} f(x) \delta(x) dx$ .

mais pas tout à fait rigoureux (nous avons échangé deux fois la limite  $\epsilon \rightarrow 0_+$  et l'intégration sur l'axe réel) nous indique que la propriété (2.2) est bien satisfaite par une telle construction.

Considérons, par exemple, la fonction

$$\varphi(\xi) = \begin{cases} 1 & : -\frac{1}{2} < \xi < \frac{1}{2} \\ 0 & : \text{sinon} \end{cases} \quad (2.12)$$

qui est localement constante et qui satisfait par construction à la condition (2.9). L'application de la règle (2.10) nous mène à la régularisation la plus répandue de la fonction delta, notamment

$$\delta(x) = \lim_{\epsilon \rightarrow 0_+} \begin{cases} \frac{1}{\epsilon} & : -\frac{\epsilon}{2} < x < \frac{\epsilon}{2} \\ 0 & : \text{sinon} \end{cases} . \quad (2.13)$$

Si nous préférons approximer la fonction delta par des fonctions analytiques (p.ex. afin d'introduire sa dérivée), nous pouvons choisir une gaussienne,  $\varphi(\xi) = \exp(-\xi^2)/\sqrt{\pi}$ , ou une lorentzienne,  $\varphi(\xi) = 1/[\pi(1 + \xi^2)]$ , ce qui mène respectivement aux régularisations suivantes

$$\delta(x) = \lim_{\epsilon \rightarrow 0_+} \frac{1}{\epsilon\sqrt{\pi}} e^{-(x/\epsilon)^2}, \quad (2.14)$$

$$\delta(x) = \lim_{\epsilon \rightarrow 0_+} \frac{1}{\pi} \frac{\epsilon}{x^2 + \epsilon^2}. \quad (2.15)$$

Aussi les *transformées de Fourier* des régularisations définies ci-dessus (à discuter dans la section suivante 2.2) peuvent être utilisées afin d'approximer la fonction delta. Nous obtenons ainsi les régularisations

$$\delta(x) = \lim_{\epsilon \rightarrow 0_+} \frac{1}{2\epsilon} e^{-|x|/\epsilon}, \quad (2.16)$$

$$\delta(x) = \lim_{\epsilon \rightarrow 0_+} \frac{\sin(x/\epsilon)}{\pi x}, \quad (2.17)$$

qui respectivement proviennent de la fonction exponentielle,  $\varphi(\xi) = \exp(-|\xi|)/2$ , et de la fonction *sinc*,  $\varphi(\xi) = \sin(\xi)/(\pi\xi)$ .

Notons que ce dernier choix de régularisation ne satisfait pas au comportement intuitif (2.4) de la fonction delta, car nous ne pouvons pas déduire de l'équation (2.17) la propriété  $\delta(x) = 0$  pour tout  $x \neq 0$ . Cette dernière propriété n'est donc pas une conséquence nécessaire de l'équation (2.2). Il convient donc d'ajouter l'équation (2.4) aux propriétés constitutives de la fonction delta, c-à-d de considérer que cette fonction soit définie par les trois propriétés (2.1), (2.2) et (2.4), surtout si on souhaite utiliser la distribution de Dirac pour décrire p.ex. la densité de charge d'une particule ponctuelle qui est censée être strictement nulle hors de la position précise de la particule.

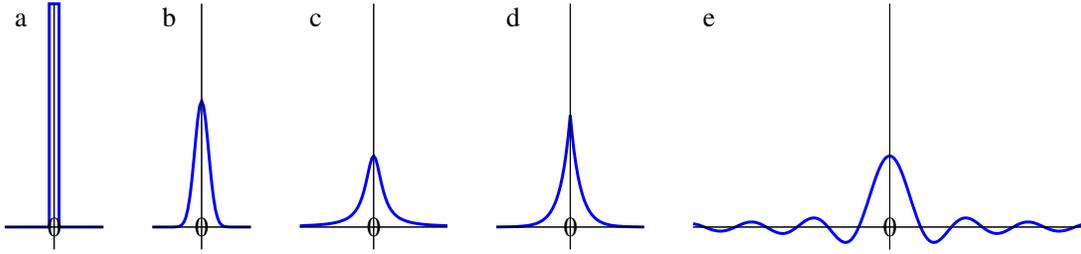


Figure 2.1: Exemples des fonctions régulières approximant la fonction delta, correspondant à (a) la fonction localement constante (2.13), (b) la fonction gaussienne (2.14), (c) la fonction lorentzienne (2.15), (d) la fonction exponentielle (2.16) ainsi que (e) la fonction *sinc* (2.17).

Notons également que l'introduction d'une certaine régularisation de la fonction delta peut éventuellement restreindre la classe des "fonctions test"  $f$  admissibles dans l'intégrale de l'équation (2.2), ce qui évidemment traduit la faiblesse de la "démonstration" (2.11). Ceci n'est pas le cas pour la régularisation (2.13). Par contre, les régularisations *sinc* (2.17) et lorentzienne (2.15) ne fournissent pas  $f(0)$  mais  $\infty$  dans l'équation (2.11) si p.ex. on considère une fonction  $f$  avec un comportement asymptotique  $f(x) \sim x^2$  pour  $x \rightarrow \pm\infty$ . La régularisation exponentielle (2.16) échoue pour toute fonction  $f$  avec le comportement asymptotique  $f(x) \sim \exp(x^2)$  pour  $x \rightarrow \pm\infty$ , et la régularisation gaussienne (2.14) ne marche pas p.ex. pour la fonction  $f(x) = \exp(x^4)$ . Il est possible d'approximer la fonction delta par une fonction qui possède des propriétés analytiques tout à fait satisfaisantes sur l'axe réel et ne connaît pas de telles limitations en ce qui concerne des fonctions test admissibles, notamment si on choisit

$$\varphi(\xi) = \begin{cases} \frac{1}{C} \exp\left(-\frac{1}{1-\xi^2}\right) & : -1 < \xi < 1 \\ 0 & : \text{sinon} \end{cases} \quad (2.18)$$

dans la définition (2.10) de la fonction delta, avec

$$C = \int_{-1}^1 \exp\left(-\frac{1}{1-\xi^2}\right) d\xi \simeq 0.444. \quad (2.19)$$

Le concept de régularisation de la fonction delta étant donc à traiter avec une certaine précaution, il s'avère néanmoins très utile si on veut se convaincre de la pertinence de certaines règles de calcul associées à la distribution de Dirac. Remplaçons p.ex. dans l'équation (2.2) le domaine d'intégration par un intervalle fini  $[a, b] \subset \mathbb{R}$  avec  $a < b$ . Nous obtenons

$$\int_a^b f(x)\delta(x)dx = f(0) [\theta(b) - \theta(a)] \quad (2.20)$$

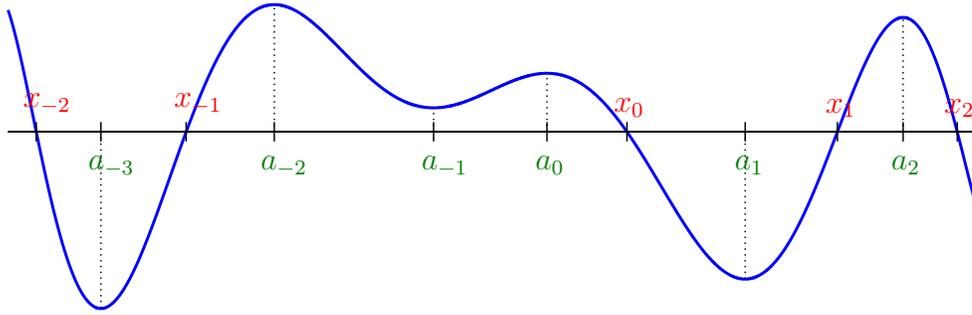


Figure 2.2: Illustration des notations des zéros ( $x_n$ ) et extrema ( $a_\nu$ ) de la fonction  $g$  dont le graphique est montré par la courbe bleue. La ligne noire horizontale correspond au niveau  $g = 0$ .

où nous avons introduit la *fonction de Heaviside*

$$\theta(x) = \int_{-\infty}^x \delta(x') dx' = \begin{cases} 0 & : x < 0 \\ \frac{1}{2} & : x = 0 \\ 1 & : x > 0 \end{cases} \quad (2.21)$$

correspondant à la primitive de la fonction delta. L'intégrale (2.20) vaut donc zéro si l'origine n'est pas contenu dans l'intervalle  $[a, b]$ , et on obtient  $f(0)$ , comme dans l'équation (2.2), si  $a < 0 < b$ . Par la propriété de symétrie (2.1) de la fonction delta, l'intégrale (2.20) vaut  $f(0)/2$  si  $a = 0$  ou  $b = 0$ .

En supposant que  $\delta$  soit une fonction comme une autre, il est aisé d'évaluer ce qui se passe si on remplace l'argument  $x$  de la fonction delta par une expression plus compliquée. En utilisant la technique de l'intégration par changement de variable, nous obtenons

$$\int_{-\infty}^{\infty} f(x) \delta(x - a) dx = \int_{-\infty}^{\infty} f(y + a) \delta(y) dy = f(a), \quad (2.22)$$

$$\int_{-\infty}^{\infty} f(x) \delta(ax) dx = \int_{-\infty}^{\infty} f\left(\frac{y}{a}\right) \delta(y) \frac{1}{a} dx = \frac{f(0)}{a} \quad (2.23)$$

pour tout  $a \in \mathbb{R} \setminus \{0\}$ . Cette dernière équation (2.23) peut aussi s'exprimer par la relation

$$\delta(ax) = \frac{1}{a} \delta(x). \quad (2.24)$$

Déterminons maintenant ce que vaut  $\delta[g(x)]$  pour une fonction générale  $g : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto g(x)$ . Nous supposons pour cela que  $g$  soit continûment dérivable sur  $\mathbb{R}$  et ne possède que de zéros simples, ces derniers se trouvant aux points  $x = x_n$  où l'entier  $n \in I$  fait partie d'un ensemble d'entiers  $I \subseteq \mathbb{Z}$  fini (p.ex. dans le cas d'un polynôme) ou infini (p.ex. dans le cas d'une fonction trigonométrique). Afin d'appliquer la méthode de changement de variable, nous dénotons par  $a_\nu \in \mathbb{R}$

avec  $\nu \in J$  les points où  $g$  acquiert un extremum, avec la propriété  $a_\nu < a_{\nu+1}$  pour tout  $\nu \in J$  comme illustré dans la figure 2.2, où  $J \subseteq \mathbb{Z}$  est un autre ensemble d'entiers fini ou infini. Nous pouvons donc écrire

$$\int_{-\infty}^{\infty} f(x)g[\delta(x)]dx = \sum_{\nu \in J} \int_{a_{\nu-1}}^{a_\nu} f(x)g[\delta(x)]dx, \quad (2.25)$$

où le premier et dernier point dans l'ensemble  $\{a_\nu\}_{\nu \in J}$  sont respectivement identifié à  $-\infty$  et  $+\infty$  dans le cas d'un nombre d'extrema fini. Grâce à la continuité de la fonction  $g$ , nous pouvons sans perte de généralité supposer que  $a_\nu$  corresponde à un maximum local de  $g$  pour tout  $\nu$  pair et à un minimum local de  $g$  pour tout  $\nu$  impair. Nous pouvons donc introduire le changement de variable  $x \mapsto y = \pm g(x)$  dans l'intervalle  $x \in ]a_{\nu-1}, a_\nu[$  pour un  $\nu \in J$  pair (+) ou impair (-), ce qui donne le jacobien

$$dx = \frac{dy}{\pm g'[g^{-1}(y)]} \quad (2.26)$$

où  $g^{-1}$  signifie l'inverse de la fonction  $g$  restreinte à l'intervalle  $]a_{\nu-1}, a_\nu[$ . L'application de la règle de calcul (2.20) fait intervenir  $g^{-1}(0)$  ce qui vaut par définition  $x_n$  pour un certain  $n \in I$ , pourvu que  $g(a_{\nu-1})g(a_\nu) < 0$ , c-à-d un zéro de  $g$  se trouve dans l'intervalle  $]a_{\nu-1}, a_\nu[$  ; sinon, l'intégration de  $a_{\nu-1}$  à  $a_\nu$  ne fournit pas de contribution à l'équation (2.25). En notant que  $\pm g'[g^{-1}(0)] = \pm g'(x_n) = |g'(x_n)|$  dans le jacobien (2.26), nous obtenons finalement le résultat

$$\int_{-\infty}^{\infty} f(x)\delta[g(x)]dx = \sum_{n \in I} \frac{f(x_n)}{|g'(x_n)|} \quad (2.27)$$

de l'intégrale (2.25), ce qui peut s'exprimer de manière alternative par la relation

$$\delta[g(x)] = \sum_{n \in I} \frac{\delta(x - x_n)}{|g'(x_n)|}. \quad (2.28)$$

Nous notons a posteriori que la simplicité des zéros  $x_n$  de  $g$ , impliquant  $g'(x_n) \neq 0$ , est une condition nécessaire pour que l'expression  $\delta[g(x)]$  représente une distribution admissible. En effet,  $\delta[g(x)]$  ne peut pas être proprement défini si  $g$  possède des zéros multiples. Considérons le cas particulier  $g(x) = x^2$ . Par la méthode de changement de variable nous obtenons

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\delta(x^2)dx &= \int_{-\infty}^0 f(x)\delta(x^2)dx + \int_0^{\infty} f(x)\delta(x^2)dx \\ &= \int_0^{\infty} \frac{f(-\sqrt{y}) + f(\sqrt{y})}{2\sqrt{y}} \delta(y)dy \end{aligned} \quad (2.29)$$

ce qui tend vers l'infini pour tout  $f$  qui ne s'annule pas à l'origine.

L'approximation par une fonction analytique (gaussienne ou lorentzienne p.ex.) selon l'équation (2.10) s'avère aussi utile pour la définition de la *dérivée* de la fonction delta :

$$\delta'(x) \equiv \frac{d}{dx}\delta(x) = \lim_{\epsilon \rightarrow 0^+} \frac{1}{\epsilon^2} \varphi'(\xi)|_{\xi=x/\epsilon}. \quad (2.30)$$

Cette dérivée représente aussi une distribution, notamment celle qui transforme une fonction  $f : U \subseteq \mathbb{R} \rightarrow \mathbb{C}$  (qui est continûment dérivable à l'origine) en la valeur  $-f'(0)$ . En effet, par l'intégration par parties nous obtenons

$$\int_{-\infty}^{\infty} f(x)\delta'(x)dx = \int_{-\infty}^{\infty} \left( \frac{d}{dx}[f(x)\delta(x)] - f'(x)\delta(x) \right) dx = -f'(0) \quad (2.31)$$

où nous avons utilisé la propriété  $\delta(x)|_{x \rightarrow \pm\infty} = 0$  qui est une conséquence évidente de l'équation (2.4). De la même manière, nous pouvons aussi définir des dérivées supérieures de la fonction delta. Sa dérivée seconde  $\delta''$  p.ex. fait transformer la fonction  $f$  en la valeur  $f''(0)$  car nous obtenons

$$\int_{-\infty}^{\infty} f(x)\delta''(x)dx = - \int_{-\infty}^{\infty} f'(x)\delta'(x)dx = f''(0) \quad (2.32)$$

après deux applications de l'intégration par parties.

Le concept de la fonction delta peut aussi être généralisé pour des espaces vectoriels réels à  $n > 1$  dimensions. Dans ce cas-là, il convient de rédéfinir, dans la définition (2.6) de la distribution de Dirac, l'espace  $C$  par

$$C = \{f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{C}, \mathbf{r} \mapsto f(\mathbf{r}) \text{ continu en } \mathbf{r} = 0\}, \quad (2.33)$$

c-à-d, la distribution de Dirac à  $n$  dimensions transforme un champ scalaire (ou vectoriel)  $f$  défini sur un sous-ensemble  $U$  de  $\mathbb{R}^n$  en sa valeur  $f(0)$  à l'origine  $\mathbf{r} = 0$ . La "fonction" delta associée peut s'écrire comme produit des fonctions delta à une dimension,

$$\delta(\mathbf{r}) = \delta(r_1) \cdots \delta(r_n), \quad (2.34)$$

chacune d'elles associée à une coordonnée du vecteur  $\mathbf{r} \equiv (r_1, \dots, r_n)$ .

Les règles générales pour évaluer des intégrales contenant des produits des fonctions delta peuvent aussi être obtenues en supposant que  $\delta$  soit approximée par une fonction régulière. En définissant  $g(x) = f(x)\delta(x - b)$  (ou, de manière alternative,  $g(x) = f(x)\delta(x - a)$ ) et en appliquant la règle de calcul (2.22) où on remplace  $f$  par  $g$  (et éventuellement aussi  $a$  par  $b$ ), nous obtenons

$$\int_{-\infty}^{\infty} f(x)\delta(x - a)\delta(x - b)dx = f(b)\delta(b - a) = f(a)\delta(a - b) \quad (2.35)$$

ce qui s'exprime de manière alternative par la loi

$$\delta(x - a)\delta(x - b) = \delta(x - a)\delta(a - b) = \delta(x - b)\delta(a - b). \quad (2.36)$$

Bien évidemment, l'identification directe  $b = a$  n'est pas admissible dans cette dernière expression car  $\delta(0) \rightarrow \infty$ . Autrement dit, le carré  $\delta^2$  d'une fonction delta ne peut pas être défini.

Evaluons finalement comment on peut exprimer la fonction delta en coordonnées sphériques  $(r, \theta, \varphi)$  dans un espace vectoriel à  $n = 3$  dimensions. Si on suppose que le vecteur  $\mathbf{r}_0$  est caractérisé, de manière unique, par le rayon  $r_0 > 0$ , l'angle polaire  $\theta_0 \in ]0, \pi[$  et l'angle azimutale  $\varphi_0 \in [0, 2\pi[$ , nous obtenons

$$\delta(\mathbf{r} - \mathbf{r}_0) = \frac{1}{r_0^2 \sin \theta_0} \delta(r - r_0) \delta(\theta - \theta_0) \delta(\varphi - \varphi_0). \quad (2.37)$$

Dans la limite  $r_0 \rightarrow 0$  nous pouvons donc simplement écrire

$$\delta(\mathbf{r}) = \frac{1}{2\pi r^2} \delta(r) \quad (2.38)$$

où, pour la détermination du préfacteur, nous avons utilisé la propriété  $\int_0^\infty f(r) \delta(r) dr = f(0)/2$  qui est une conséquence de la symétrie (2.1) de la fonction delta.

## 2.2 Les transformations de Fourier

Nous considérons une fonction  $f : \mathbb{R} \rightarrow \mathbb{C}$ ,  $x \mapsto f(x)$  qui est continue et *intégrable*, c-à-d, l'expression

$$\int_{-\infty}^{\infty} |f(x)| dx \equiv \lim_{R \rightarrow \infty} \int_{-R}^R |f(x)| dx < \infty \quad (2.39)$$

existe et fournit une valeur finie, où nous nous servons de la définition (2.3) pour une intégrale qui couvre l'axe réel dans sa totalité. Pour une telle fonction  $f$  on définit la *transformée de Fourier* par

$$\hat{f} : \mathbb{R} \rightarrow \mathbb{C}, k \mapsto \hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx. \quad (2.40)$$

On peut alors montrer que  $f(x)$  peut être déterminé par les valeurs de la fonction  $\hat{f}$  au travers de la *transformation inverse*

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(k) e^{ikx} dk, \quad (2.41)$$

ce qui relie les fonctions  $f$  et  $\hat{f}$  l'une à l'autre de manière unique.

La relation (2.41) peut être montrée en utilisant le concept de la fonction delta développé dans la section 2.1. En effet, nous pouvons calculer

$$\begin{aligned} \int_{-\infty}^{\infty} e^{ikx} dx &= \lim_{R \rightarrow \infty} \int_{-R}^R e^{ikx} dx = \lim_{R \rightarrow \infty} \frac{e^{ikR} - e^{-ikR}}{ik} \\ &= \lim_{R \rightarrow \infty} \frac{2}{k} \sin(kR) = 2\pi \delta(k), \end{aligned} \quad (2.42)$$

où pour la dernière équivalence dans l'équation (2.42) nous avons utilisé le fait que la fonction *sinc* représente une régularisation de la fonction delta suivant l'expression (2.17) qui est compatible avec des fonctions test intégrables. D'une manière alternative, nous pouvons aussi démontrer la relation (2.41) via

$$\begin{aligned} \int_{-\infty}^{\infty} e^{ikx} dx &= \lim_{\epsilon \rightarrow 0^+} \int_{-\infty}^{\infty} e^{ikx - \epsilon|x|} dx = \lim_{\epsilon \rightarrow 0^+} \left( \int_{-\infty}^0 e^{(\epsilon+ik)x} dx + \int_0^{\infty} e^{-(\epsilon-ik)x} dx \right) \\ &= \lim_{\epsilon \rightarrow 0^+} \left( \frac{1}{\epsilon + ik} + \frac{1}{\epsilon - ik} \right) = \lim_{\epsilon \rightarrow 0^+} \frac{2\epsilon}{\epsilon^2 + k^2} = 2\pi\delta(k), \end{aligned} \quad (2.43)$$

où nous utilisons la régularisation lorentzienne (2.15) de la fonction delta, qui est également compatible avec des fonctions test intégrables. Nous pouvons alors évaluer la transformation inverse

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(k) e^{ikx} dk &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx' f(x') e^{-ikx'} e^{ikx} \\ &= \int_{-\infty}^{\infty} dx' f(x') \underbrace{\frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(x-x')}}_{\delta(x-x')} = f(x), \end{aligned} \quad (2.44)$$

où l'échange d'ordre des deux intégrations sur  $k$  et  $x'$  peut être justifiée si  $f$  satisfait à la condition d'intégrabilité (2.39). Dans ce cas-là,  $\hat{f}$  satisfait également à cette condition (2.39) et est donc admissible en tant que "fonction test" pour la régularisation (2.17) du type *sinc*.

Notons que l'exigence (2.39) peut s'avérer trop restrictive en pratique si on souhaite déterminer si une fonction particulière peut être soumise à une transformation de Fourier ou non. En effet, des fonctions  $f : \mathbb{R} \rightarrow \mathbb{C}$  qui satisfont à la "simple" intégrabilité

$$\int_{-\infty}^{\infty} f(x) dx \equiv \lim_{R \rightarrow \infty} \int_{-R}^R f(x) dx < \infty \quad (2.45)$$

mais non pas à la condition (2.39) ont toutefois des transformées de Fourier bien définies, éventuellement sous une forme de distribution comme il est le cas pour la fonction exponentielle complexe. Nous notons cependant que cette notion de "simple intégrabilité" ne peut pas être proprement définie d'un point de vue mathématique pour une fonction  $f$  qui n'est pas intégrable selon la condition (2.39), car pour une telle fonction la valeur de l'intégrale  $\int_{-\infty}^{\infty} f(x) dx$  dépend généralement de l'ordre dans lequel les aires des différents trapèzes sont sommées dans l'expression de Riemann (1.78) de cette intégrale (cf. figure 1.4). Par exemple, la fonction  $\text{sinc}(x) = \sin(x)/x$  est bien "intégrable" selon le critère (2.45), avec  $\int_{-\infty}^{\infty} \text{sinc}(x) dx = \pi$  suivant la prescription (2.2), mais une autre manière d'intégrer cette fonction, p.ex. en sommant d'abord les intégrales de *sinc* sur les

intervalles  $[0, \pi]$ ,  $[2\pi, 3\pi]$ ,  $[4\pi, 5\pi]$ , ... et ensuite les intégrales sur les autres intervalles  $[\pi, 2\pi]$ ,  $[3\pi, 4\pi]$ , ... en ce qui concerne l'axe réel positif, peut mener à un autre résultat ou à une expression divergente. Il en est de même pour la somme  $\sum_{n=1}^{\infty} (-1)^{n-1}/n = 1 - 1/2 + 1/3 - \dots$  qui fournit la valeur  $\ln(2)$  si on l'évalue suivant la prescription  $\lim_{N \rightarrow \infty} \sum_{n=1}^N (-1)^{n-1}/n$ , mais qui peut atteindre toute autre valeur réelle si on décide de sommer les différents termes de cette somme dans un autre ordre.

En pratique, la transformation de Fourier peut souvent être calculée en utilisant le théorème des résidus comme nous avons déjà vu à la fin de la section 1.8. Les transformées de Fourier des fonctions suivantes

$$f(x) = \frac{1}{\sqrt{a}} \exp \left[ -\frac{1}{2} \left( \frac{x}{a} \right)^2 \right], \quad (2.46)$$

$$g(x) = \frac{1}{\sqrt{a}} e^{-|x|/a}, \quad (2.47)$$

$$h(x) = \frac{1}{\sqrt{a}} \frac{1}{1 + (x/a)^2}, \quad (2.48)$$

définies pour un  $a > 0$ , sont ainsi évaluées comme

$$\hat{f}(k) = \sqrt{a} \exp \left[ -\frac{1}{2} (ak)^2 \right], \quad (2.49)$$

$$\hat{g}(k) = \sqrt{\frac{2a}{\pi}} \frac{1}{1 + (ak)^2}, \quad (2.50)$$

$$\hat{h}(k) = \sqrt{\frac{\pi a}{2}} e^{-a|k|}, \quad (2.51)$$

respectivement.

Nous notons bien une complémentarité entre la taille du pic central des fonctions  $f$ ,  $g$  et  $h$ , qui est caractérisée par la constante  $a$ , et la taille du pic central des transformées de Fourier associées  $\hat{f}$ ,  $\hat{g}$  et  $\hat{h}$ , qui est caractérisée par  $1/a$ . D'une manière générale, si

$$\hat{\varphi} : \mathbb{R} \rightarrow \mathbb{C}, \kappa \mapsto \hat{\varphi}(\kappa) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(\xi) e^{-i\kappa\xi} d\xi \quad (2.52)$$

représente la transformée de Fourier de la fonction  $\varphi : \mathbb{R} \rightarrow \mathbb{C}, \xi \mapsto \varphi(\xi)$  qui est supposée être intégrable, nous obtenons pour la fonction

$$f : \mathbb{R} \rightarrow \mathbb{C}, x \mapsto f(x) = \frac{1}{\sqrt{a}} \varphi \left( \frac{x}{a} \right) \quad (2.53)$$

définie pour un  $a > 0$  la transformée de Fourier associée

$$\hat{f} : \mathbb{R} \rightarrow \mathbb{C}, k \mapsto \hat{f}(k) = \sqrt{a} \hat{\varphi}(ak). \quad (2.54)$$

Par conséquent, une fonction  $f$  qui possède un maximum assez étroit et pointu près de l'origine, ce qui correspond à la limite  $a \rightarrow 0$  dans l'équation (2.53), mène à une transformée de Fourier  $\hat{f}$  dont le maximum est très large et peu prononcé comme on peut le voir dans l'équation (2.54). Cette dernière propriété générique de la transformation de Fourier est au cœur de la relation d'incertitude de Heisenberg en mécanique quantique, qui stipule que le produit des incertitudes de la position d'une particule quantique (qui est de l'ordre  $a$  si  $f$  représente la fonction d'onde de la particule définie dans l'espace de configuration) et de son impulsion (qui est de l'ordre  $\hbar/a$  car  $\hat{f}$  représente la fonction d'onde de la particule dans l'espace réciproque dans ce cas-là) vaut au moins  $\hbar/2$  où  $\hbar = h/(2\pi)$  est la constante de Planck réduite.

Les transformations de Fourier s'avèrent très utiles pour la résolution des équations différentielles linéaires car elles transforment des opérations de dérivée en des simples produits avec la variable  $k$  dans l'espace réciproque. Considérons une fonction  $f : \mathbb{R} \rightarrow \mathbb{C}$  qui est dérivable. La transformation de Fourier de sa dérivée  $f' : \mathbb{R} \rightarrow \mathbb{C}, x \mapsto f'(x) = \frac{d}{dx}f(x)$  est calculée en effectuant une intégration par parties suivant

$$\begin{aligned}\widehat{f'}(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{df}{dx}(x) e^{-ikx} dx = \frac{1}{\sqrt{2\pi}} [f(x)e^{-ikx}]_{-\infty}^{\infty} + \frac{ik}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx \\ &= ik\widehat{f}(k),\end{aligned}\tag{2.55}$$

où nous avons utilisé la propriété  $\lim_{x \rightarrow \pm\infty} f(x) = 0$  pour une fonction intégrable. D'une manière tout à fait analogue nous obtenons

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{d\widehat{f}}{dk}(k) e^{ikx} dk = -ixf(x)\tag{2.56}$$

pour la transformation inverse de la dérivée de  $\widehat{f}$ . En appliquant la règle (2.55) d'une manière récursive, nous pouvons aussi calculer les transformations de Fourier des dérivées supérieures de la fonction  $f$ . Par exemple, la transformée de Fourier de la dérivée seconde  $f'' : \mathbb{R} \rightarrow \mathbb{C}, x \mapsto f''(x) = \frac{d^2}{dx^2}f(x) = \frac{d}{dx}f'(x)$  est évaluée suivant

$$\widehat{f''}(k) = ik\widehat{f'}(k) = -k^2\widehat{f}(k).\tag{2.57}$$

La transformation de Fourier peut aussi être définie pour des *champs*, c-à-d, des fonctions qui sont définies sur des espaces vectoriels pluridimensionnels  $f : \mathbb{R}^n \rightarrow \mathbb{C}, \mathbf{r} \mapsto f(\mathbf{r})$  avec  $n \in \mathbb{N}$ . Pourvu qu'une telle fonction  $f$  soit intégrable sur  $\mathbb{R}^n$ , nous pouvons en définir sa transformée de Fourier suivant

$$\widehat{f} : \mathbb{R}^n \rightarrow \mathbb{C}, \mathbf{k} \mapsto \widehat{f}(\mathbf{k}) = \frac{1}{\sqrt{2\pi}^n} \int d^n r f(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}},\tag{2.58}$$

ce qui revient à appliquer la transformation de Fourier unidimensionnelle (2.40) pour chacune des coordonnées de  $\mathbf{r}$ . La transformation inverse associée s'écrit

alors

$$f(\mathbf{r}) = \frac{1}{\sqrt{2\pi}^n} \int d^n k \hat{f}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \quad (2.59)$$

en parfaite analogie avec l'expression unidimensionnelle (2.41).

Si la fonction  $f : \mathbb{R}^n \rightarrow \mathbb{C}$  est dérivable, nous pouvons définir un champ vectoriel

$$\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{C}, \mathbf{r} \mapsto \mathbf{g}(\mathbf{r}) = \nabla f(\mathbf{r}) \equiv \frac{\partial}{\partial \mathbf{r}} f(\mathbf{r}) \quad (2.60)$$

qui correspond au gradient de la fonction  $f$ . Pourvu que  $\mathbf{g}$  soit intégrable, sa transformée de Fourier est évaluée suivant

$$\hat{\mathbf{g}}(\mathbf{k}) = \frac{1}{\sqrt{2\pi}^n} \int d^n r e^{-i\mathbf{k}\cdot\mathbf{r}} \nabla f(\mathbf{r}) = i\hat{f}(\mathbf{k})\mathbf{k} \quad (2.61)$$

en parfaite analogie avec le calcul unidimensionnel (2.55), où nous avons encore une fois utilisé l'intégration par parties. De la même manière, la transformation de Fourier du laplacien  $\Delta f(\mathbf{r}) \equiv \frac{\partial^2}{\partial \mathbf{r}^2} f(\mathbf{r})$  d'une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{C}$  qui est deux fois dérivable est calculée suivant

$$\widehat{\Delta f}(\mathbf{k}) = \frac{1}{\sqrt{2\pi}^n} \int d^n r e^{-i\mathbf{k}\cdot\mathbf{r}} \frac{\partial^2}{\partial \mathbf{r}^2} f(\mathbf{r}) = -k^2 \hat{f}(\mathbf{k}). \quad (2.62)$$

Cette dernière relation s'avère très utile pour la résolution de certaines équations aux dérivées partielles linéaires où le laplacien apparaît, à savoir l'équation d'ondes, l'équation de diffusion ou alors l'équation de Schrödinger (cette dernière en l'absence d'une énergie potentielle de la particule quantique en question).

## Exercices

2.1 Calculer pour  $\alpha > 0$  la transformée de Fourier de la fonction

(a)  $f(x) = e^{-\alpha|x|}$

(b)  $f(x) = e^{-\alpha x^2}$

(c)  $f(x) = e^{i\alpha x^2}$

(d)  $f(x) = \frac{1}{x^2 + \alpha^2}$

(e)  $f(x) = \frac{1}{x^3 + i\alpha^3}$

(f)  $f(x) = \frac{1}{x^4 + \alpha^4}$

en utilisant le théorème des résidus.

2.2 Calculer la transformée de Fourier de la fonction  $f(x) = \frac{\sin(qx)}{x}$  pour  $q > 0$ .

*Indice* : Calculer d'abord  $\oint_{\gamma} \frac{e^{ikz}}{z}$  pour  $k > 0$  le long du chemin  $\gamma$  de l'exercice 1.6.

## 2.3 Les séries de Fourier

Nous considérons une fonction  $f : \mathbb{R} \rightarrow \mathbb{C}, x \mapsto f(x)$  qui est continue et *périodique* en  $x$  avec la période  $a > 0$ , c-à-d,  $f(x) = f(x + a)$  pour tout  $x \in \mathbb{R}$ . Dans ce cas-là nous pouvons exprimer cette fonction par sa *série de Fourier*

$$f(x) = \sum_{k=-\infty}^{\infty} \hat{f}_k e^{2\pi i k x / a} \quad (2.63)$$

où les *coefficients de Fourier*  $\hat{f}_k \in \mathbb{C}$  sont définis par

$$\hat{f}_k = \frac{1}{a} \int_{-a/2}^{a/2} f(x) e^{-2\pi i k x / a} dx. \quad (2.64)$$

La fonction  $f$  est donc entièrement représentée et caractérisée par l'ensemble  $\{\hat{f}_k : k \in \mathbb{Z}\}$  de ses coefficients de Fourier.

La validité de l'expression (2.64) peut être facilement montrée en utilisant la relation  $\int_{-\pi}^{\pi} e^{i\nu\xi} d\xi = 2\pi\delta_{\nu 0}$  pour tout  $\nu \in \mathbb{Z}$ . Nous calculons ainsi

$$\frac{1}{a} \int_{-a/2}^{a/2} f(x) e^{-2\pi i k x / a} dx = \sum_{k'=-\infty}^{\infty} \hat{f}_{k'} \frac{1}{a} \int_{-a/2}^{a/2} e^{2\pi i (k' - k)x / a} dx = \hat{f}_k \quad (2.65)$$

pour  $f(x)$  étant donné par l'expression (2.63). D'une manière tout à fait analogue, nous obtenons

$$\begin{aligned} \sum_{k=-\infty}^{\infty} \hat{f}_k e^{2\pi i k x / a} &= \sum_{k=-\infty}^{\infty} \frac{1}{a} \int_{-a/2}^{a/2} f(x') e^{2\pi i k (x - x') / a} dx' \\ &= \int_{-a/2}^{a/2} f(x') \frac{1}{a} \sum_{n=-\infty}^{\infty} \delta\left(\frac{x - x'}{a} - n\right) dx' = f(x) \end{aligned} \quad (2.66)$$

pour toute fonction périodique  $f : \mathbb{R} \rightarrow \mathbb{C}, x \mapsto f(x) = f(x + a)$  dont les coefficients  $\hat{f}_k$  sont définis par l'expression (2.64). Dans ce calcul nous avons utilisé la *formule de Poisson*

$$\sum_{k=-\infty}^{\infty} e^{2\pi i k \xi} = \sum_{n=-\infty}^{\infty} \delta(\xi - n) \quad (2.67)$$

pour tout  $\xi \in \mathbb{R}$  où  $\delta(\xi - n)$  représente bien la fonction delta introduite dans la section 2.1.

**Exercices**

2.3 (a) Calculer pour  $\epsilon > 0$  et  $n \in \mathbb{Z}$

$$f_n = \frac{1}{2\pi} \int_0^{2\pi} \frac{e^{in\varphi}}{1 + \epsilon \cos \varphi} d\varphi$$

en utilisant le théorème des résidus.

(b) Calculer pour  $\epsilon > 0$  la série de Fourier de la fonction  $f(x) = \frac{1}{1 + \epsilon \cos x}$ .

**2.4 Les transformations de Laplace**

Dans beaucoup de contextes physiques on a affaire à des équations différentielles linéaires qui sont définies en fonction du temps, comme nous allons détailler dans le chapitre suivant. Ces équations différentielles sont généralement accompagnées par la spécification des conditions initiales à un certain temps  $t = t_0$ , qui typiquement caractérisent la préparation initiale de l'expérience en question. La méthode de la transformation de Fourier n'est pas tout à fait adaptée pour calculer la solution d'un tel problème car la fonction recherchée n'est généralement pas définie avant ce temps initial  $t_0$ . De plus, dépendant des coefficients qui constituent l'équation différentielle en question, elle risque de subir une croissance exponentielle en fonction du temps, ce qui la rendrait non intégrable et ainsi non admissible pour une transformation de Fourier.

Dans une telle situation il est conseillé d'effectuer une *transformation de Laplace* au lieu. Cette dernière peut être calculée pour toute fonction  $f : \mathbb{R}_+ \rightarrow \mathbb{C}, t \mapsto f(t)$  qui est définie sur l'axe réel positif (en supposant  $t_0 = 0$  sans perte de généralité) et qui peut être majorée par une fonction exponentielle, c-à-d, il existe des constantes  $C, \gamma \in \mathbb{R}$  avec  $C > 0$  tel que  $|f(t)| \leq Ce^{\gamma t}$  pour tout  $t \in \mathbb{R}_+$ . La *transformée de Laplace* d'une telle fonction  $f$  est alors définie en tant que fonction complexe suivant

$$\tilde{f} : U \subseteq \mathbb{C} \rightarrow \mathbb{C}, z \mapsto \tilde{f}(z) = \int_0^\infty f(t)e^{-zt} dt, \quad (2.68)$$

où  $U$  est un sous-ensemble ouvert du plan complexe. Si nous introduisons par

$$\gamma_0 = \inf \{ \gamma \in \mathbb{R} : \exists C > 0 \text{ tel que } |f(t)| \leq Ce^{\gamma t} \text{ pour tout } t > 0 \} \quad (2.69)$$

l'exposant effectif du comportement exponentiel asymptotique de la fonction  $f$ , il paraît évident que nous devons définir  $U = \{z \in \mathbb{C} : \operatorname{Re}(z) > \gamma_0\}$  car c'est dans cette région-là que l'intégrale dans l'équation (2.68) converge vers une valeur finie.

Notons cependant que la région de définition  $U$  de la fonction résultante  $\tilde{f}$  peut généralement être élargie dans le plan complexe, grâce au principe de la continuation analytique. Considérons p.ex. la fonction

$$f(t) = f_0 e^{-(\lambda + i\omega)t} \quad (2.70)$$

où  $f_0, \lambda, \omega \in \mathbb{R}$  sont des constantes réelles. Sa transformée de Laplace

$$\tilde{f}(z) = \frac{f_0}{z + \lambda + i\omega} \quad (2.71)$$

peut être évaluée pour  $\operatorname{Re}(z) > -\lambda$  suivant la prescription (2.68). Cependant, comme elle ne possède qu'une seule singularité à  $z = -\lambda - i\omega$  qui est isolée, la fonction  $\tilde{f}$  peut bien être définie sur  $U = \mathbb{C} \setminus \{-\lambda - i\omega\}$ . Cette situation est tout à fait analogue au phénomène du rayon de convergence qui limite la région  $U \subseteq \mathbb{C}$  dans laquelle une série entière peut être évaluée. Comme nous avons vu dans la section 1.4, ce rayon de convergence est effectivement donné par le module de la singularité “la plus petite” — c-à-d, qui est la plus proche de l'origine — de la fonction complexe envers laquelle série entière en question converge. De la même manière, l'exposant effectif  $\gamma_0$  qui limite la convergence de l'intégrale (2.68) est donné par la partie réelle de la singularité “la plus à droite” de  $f$  — c-à-d, dont la partie réelle est la plus grande. Ce critère nous fournit bien  $\gamma_0 = -\lambda$  pour la fonction (2.70), en parfait accord avec les considérations développées ci-dessus.

L'inversion de la transformation de Laplace (2.68) peut être effectuée par l'intégrale de  $\tilde{f}$  le long du chemin

$$\Gamma : [-R, R] \rightarrow \mathbb{C}, \omega \mapsto \Gamma(\omega) = \gamma + i\omega \quad (2.72)$$

qui est défini pour un  $\gamma > \gamma_0$  dans la limite  $R \rightarrow \infty$ , comme illustré dans la figure 2.3 : nous obtenons

$$f(t) = \lim_{R \rightarrow \infty} \frac{1}{2\pi i} \int_{\Gamma} \tilde{f}(z) e^{zt} dz \quad (2.73)$$

pour la transformation inverse. Cette relation peut facilement être montrée en utilisant le fait que l'intégrale sur  $t$  dans l'expression (2.68) pour  $\tilde{f}$  converge tout au long du chemin  $\Gamma$ , vu que  $\operatorname{Re}[\Gamma(\omega)] = \gamma > \gamma_0$  pour tout  $\omega \in [-R, R]$ . Nous pouvons ainsi échanger l'intégrale sur  $t$  avec l'intégrale de chemin le long de  $\Gamma$ , ce qui nous fournit

$$\begin{aligned} \lim_{R \rightarrow \infty} \frac{1}{2\pi i} \int_{\Gamma} \tilde{f}(z) e^{zt} dz &= \int_0^{\infty} dt' f(t') \lim_{R \rightarrow \infty} \frac{1}{2\pi i} \int_{-R}^R i d\omega e^{i\omega(t-t')} e^{\gamma(t-t')} \\ &= \begin{cases} f(t) & : t > 0 \\ 0 & : t < 0 \end{cases} \end{aligned} \quad (2.74)$$

après la mise en évidence de la paramétrisation (2.72) du chemin  $\Gamma$ . Nous avons ici utilisé la relation (2.42) stipulant que l'expression  $\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega(t-t')} d\omega = \delta(t-t')$  représente une régularisation de la fonction delta pour des fonctions test qui tendent vers zéro pour  $t' \rightarrow \infty$ , ce qui est bien le cas pour  $f(t')e^{\gamma(t-t')}$ .

En passant, nous avons aussi démontré que la transformation de Laplace est une opération unique. C-à-d, si deux fonctions  $f_1, f_2 : \mathbb{R}_+ \rightarrow \mathbb{C}$  fournissent la même transformée de Laplace  $\tilde{f} : U \subseteq \mathbb{C} \rightarrow \mathbb{C}$  suivant la définition (2.68), le

raisonnement (2.74) développé ci-dessus nous confirme que ces deux fonctions doivent être identiques :  $f_1 \equiv f_2$ . Cette unicité peut aussi être montrée pour la transformation inverse (2.73) pourvu qu'on restreigne convenablement la classe des fonctions complexes  $\tilde{f} : U \subseteq \mathbb{C} \rightarrow \mathbb{C}$  qui sont susceptibles de représenter une transformée de Laplace. Nous exigeons notamment que  $\tilde{f}$  soit analytique en  $\{z \in \mathbb{C} : \operatorname{Re}(z) > \gamma_0\}$  pour un  $\gamma_0 \in \mathbb{R}$ , ce qui reflète la condition qu'on puisse majorer sa transformée inverse  $f(t)$  par toute fonction exponentielle  $\propto e^{\gamma t}$  avec  $\gamma > \gamma_0$ . De plus, en vu de la manière dont la transformation de Laplace (2.68) est définie, nous exigeons aussi que  $\tilde{f}(z) \rightarrow 0$  pour  $|z| \rightarrow \infty$  avec  $\operatorname{Re}(z) > \gamma_0$ .<sup>2</sup> Ceci implique

$$\lim_{R \rightarrow \infty} \int_{\Gamma'} \tilde{f}(z) e^{zt} dz = 0 \quad (2.75)$$

pour tout  $t < 0$  où nous introduisons le chemin

$$\Gamma' : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow \mathbb{C}, \varphi \mapsto \Gamma(\varphi) = \gamma + Re^{i\varphi} \quad (2.76)$$

pour un  $\gamma > \gamma_0$  arbitraire, comme illustré dans la figure 2.3.

Définissons maintenant la “transformée inverse” de  $\tilde{f}$  suivant la prescription (2.73), c-à-d,

$$f : \mathbb{R} \rightarrow \mathbb{C}, t \mapsto f(t) = \lim_{R \rightarrow \infty} \frac{1}{2\pi i} \int_{\Gamma} \tilde{f}(z) e^{zt} dz \quad (2.77)$$

où  $\Gamma$  représente le chemin (2.72). Par le théorème de Cauchy (1.86) nous obtenons  $\int_{\Gamma} \tilde{f}(z) e^{zt} dz = \int_{\Gamma'} \tilde{f}(z) e^{zt} dz$  à cause de l'analyticité de  $\tilde{f}$  en tout  $z \in \mathbb{C}$  avec  $\operatorname{Re}(z) > \gamma_0$ , où nous choisissons le même paramètre  $\gamma$  dans les définitions (2.72) et (2.76) de  $\Gamma$  et  $\Gamma'$ . La relation (2.75) nous amène alors à déduire que  $f(t) = 0$  pour  $t < 0$ . Grace à cette dernière propriété nous pouvons évaluer la transformation de Laplace (2.68) de  $f$  en  $z = \gamma + i\omega$  suivant

$$\begin{aligned} \int_0^\infty f(t) e^{-zt} dt &= \int_{-\infty}^\infty f(t) e^{-zt} dt = \int_{-\infty}^\infty dt \lim_{R \rightarrow \infty} \frac{1}{2\pi i} \int_{\Gamma} dz' \tilde{f}(z') e^{z't} e^{-zt} \\ &= \lim_{R \rightarrow \infty} \int_{-\infty}^\infty i d\omega' \tilde{f}(\gamma + i\omega') \frac{1}{2\pi i} \int_{-R}^R dt e^{i(\omega' - \omega)t} \\ &= \tilde{f}(\gamma + i\omega) = \tilde{f}(z) \end{aligned} \quad (2.78)$$

pour tout  $\omega \in \mathbb{R}$ , où nous avons encore une fois échangé les deux intégrales et utilisé la relation (2.42). Par le principe de la continuation analytique nous obtenons finalement  $\int_0^\infty f(t) e^{-zt} dt = \tilde{f}(z)$  partout dans la région de définition  $U$  de  $\tilde{f}$ .

---

<sup>2</sup>Nous excluons ainsi la possibilité que  $f$  puisse être une distribution de Dirac, ce qui menerait à une transformée de Laplace constante :  $\tilde{f}(z) = c$  pour tout  $z \in \mathbb{C}$ .

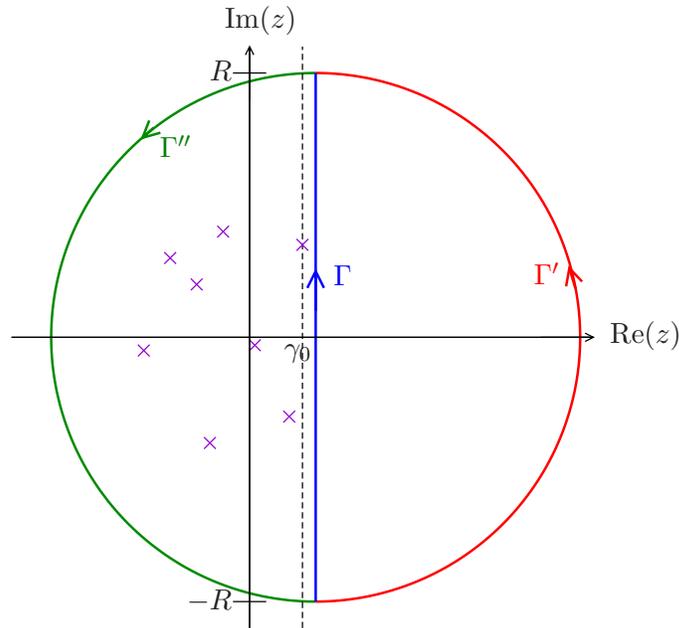


Figure 2.3: L'inversion de la transformée de Laplace est calculée par l'intégrale de  $\tilde{f}(z)e^{zt}$  le long du chemin  $\Gamma$  (2.72) qui suit une droite verticale avec la partie réelle  $\gamma > \gamma_0$  et qui est donc situé à droite de toutes les singularités de  $\tilde{f}$  (ces dernières étant marquées par des croix violettes). Par le théorème de Cauchy (1.86) cette intégrale de chemin peut être remplacée par une intégrale de  $\tilde{f}(z)e^{zt}$  le long du chemin  $\Gamma'$  (2.76) ce qui tend vers zéro pour  $R \rightarrow \infty$  si  $t < 0$ . Dans le cas contraire  $t > 0$  c'est l'intégrale le long du demi-cercle opposé  $\Gamma''$  qui tend vers zéro. Nous pouvons donc utiliser le théorème des résidus (1.113) pour évaluer la transformation inverse par la somme sur les résidus de la fonction  $\tilde{f}(z)e^{zt}$  associés à ces différentes singularités isolées si  $\tilde{f}$  est méromorphe.

Si la fonction  $f : \mathbb{R}_+ \rightarrow \mathbb{C}$  est dérivable, nous pouvons évaluer la transformée de Laplace de sa dérivée  $\dot{f} : \mathbb{R}_+ \rightarrow \mathbb{C}, t \mapsto \dot{f}(t) = \frac{d}{dt}f(t)$  suivant

$$\begin{aligned}\tilde{\dot{f}}(z) &= \int_0^\infty e^{-zt} \frac{d}{dt}f(t) dt = [f(t)e^{-zt}]_0^\infty + z \int_0^\infty f(t)e^{-zt} dt \\ &= z\tilde{f}(z) - f(0).\end{aligned}\tag{2.79}$$

Contrairement au cas de la transformation de Fourier, cette expression pour la transformée de la dérivée contient explicitement la valeur initiale de la fonction  $f$ . En appliquant la relation (2.79) également à  $\dot{\dot{f}}$ , nous pouvons évaluer la transformée de Laplace de la dérivée seconde  $\ddot{f}(t) = \frac{d^2}{dt^2}f(t)$  suivant

$$\tilde{\ddot{f}}(z) = z\tilde{\dot{f}}(z) - \dot{f}(0) = z^2\tilde{f}(z) - zf(0) - \dot{f}(0)\tag{2.80}$$

pourvu que la fonction  $f$  soit deux fois dérivable.

La transformation de Laplace peut donc être utilisée pour résoudre des équations différentielles linéaires définies avec des conditions initiales. Considérons p.ex. l'équation différentielle

$$\dot{f}(t) = \alpha f(t)\tag{2.81}$$

pour un  $\alpha \in \mathbb{C}$  avec la condition initiale  $f(0) = f_0$ . La transformation de Laplace de cette équation différentielle (2.81) nous fournit l'équation

$$\tilde{\dot{f}}(z) = z\tilde{f}(z) - f(0) = \alpha\tilde{f}(z)\tag{2.82}$$

que nous pouvons facilement résoudre : sa solution

$$\tilde{f}(z) = \frac{f_0}{z - \alpha}\tag{2.83}$$

possède un seul pôle à  $z = \alpha$  et peut être définie sur  $U = \mathbb{C} \setminus \{\alpha\}$ . La transformation inverse (2.73) peut donc être calculée par le théorème des résidus (1.113) où pour  $t > 0$  nous clôturons le chemin d'intégration par un demi-cercle qui passe par le côté gauche du plan complexe, comme il est illustré dans la figure 2.3. Nous obtenons ainsi

$$f(t) = \theta(t)f_0e^{\alpha t} = \begin{cases} 0 & : t < 0 \\ f_0e^{\alpha t} & : t > 0 \end{cases}\tag{2.84}$$

ce qui correspond bien à la solution recherchée de l'équation différentielle (2.81).

# Chapitre 3

## Les équations différentielles ordinaires

### 3.1 Classification

Les équations différentielles sont omniprésentes en physique. Citons en quelques exemples, pour commencer ce chapitre.

1. La désintégration d'un matériau radioactif est décrite par une équation tout à fait identique à l'équation (2.81) que nous venons de discuter dans le chapitre précédent, à savoir

$$\dot{n}(t) = -\gamma n(t). \quad (3.1)$$

Ici,  $n(t)$  représente le nombre des atomes radioactifs dans le spécimen en question et  $\gamma$  est le taux de désintégration, c-à-d, la densité de probabilité qu'un atome radioactif se désintègre à l'instant  $t$ .

2. La mécanique analytique regorge d'exemples d'équations différentielles grâce à la deuxième loi de Newton. Considérons p.ex. un pendule avec la longueur  $l$  qui oscille en présence de la gravité, comme illustré dans la figure 3.1(a). L'angle  $\varphi$  que le pendule forme avec la verticale évolue suivant l'équation

$$\ddot{\varphi}(t) = -\frac{g}{l} \sin \varphi(t) \quad (3.2)$$

où  $g$  est l'accélération de la pesanteur.

3. Toujours en mécanique analytique, un oscillateur harmonique unidimensionnel qui est amorti et forcé est de manière générale décrit par l'équation

$$\ddot{x}(t) + \gamma \dot{x}(t) + \omega_0 x(t) = f_0 \cos \omega t. \quad (3.3)$$

Ici,  $x$  représente la coordonnée de l'oscillateur,  $\omega_0$  est sa fréquence propre et  $\gamma$  son taux d'amortissement. Le forçage extérieur est caractérisé par l'amplitude  $f_0$  et la fréquence  $\omega_0$ .

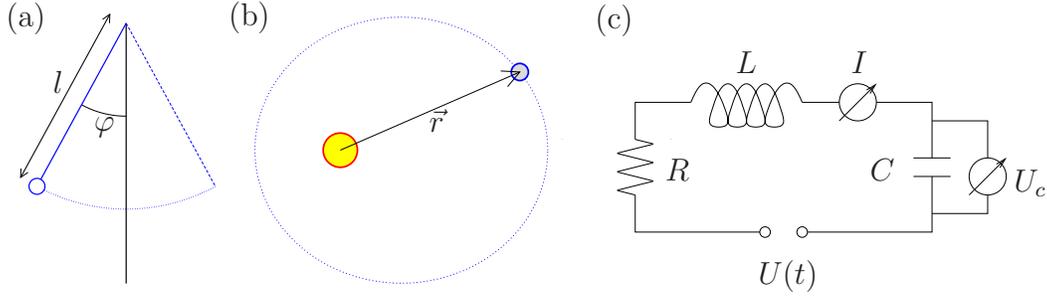


Figure 3.1: Trois exemples de systèmes physiques dont l'évolution temporelle est gouvernée par une équation différentielle : (a) un pendule qui oscille en la présence de la gravité ; (b) une planète qui tourne autour d'une étoile fixe ; (c) un circuit électrique connecté à une source de tension alternative.

4. Le mouvement d'une planète autour d'une étoile fixe est décrit par une équation newtonienne définie en trois dimensions. Soit  $\vec{r}$  le vecteur position de la planète, comme dessiné sur la figure 3.1(b). En supposant que l'étoile se trouve à l'origine du système des coordonnées, ce vecteur évolue selon l'équation

$$\ddot{\vec{r}}(t) = -\frac{\gamma M}{r^3(t)}\vec{r}(t) \quad (3.4)$$

avec  $r(t) \equiv |\vec{r}(t)|$ , où  $M$  est la masse de l'étoile et  $\gamma \simeq 6.674 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$  est la constante universelle de gravitation.

5. Des exemples d'équations différentielles peuvent aussi être trouvés dans le domaine de l'électricité. Considérons p.ex. un circuit électrique contenant une résistance  $R$ , une capacité  $C$ , une inductance  $L$  et une source de tension alternative  $U(t) = U_0 \cos \omega t$ , toutes montées en série comme il est montré dans la figure 3.1(c). Selon la loi des mailles de Kirchhoff, l'intensité  $I$  du courant traversant ce circuit satisfait à l'équation

$$U(t) = RI(t) + L\dot{I}(t) + U_c(t) \quad (3.5)$$

où  $U_c$  signifie la tension mesurée à la capacité. Cette dernière est liée au courant par l'équation

$$I(t) = C\dot{U}_c(t). \quad (3.6)$$

6. En électrostatique, l'équation de Poisson

$$\Delta\phi(\vec{r}) = -\frac{1}{\epsilon_0}\rho(\vec{r}) \quad (3.7)$$

permet de déterminer le potentiel électrique  $\phi(\vec{r})$  à la position  $\vec{r}$  en fonction de la densité des charges électriques  $\rho$  dans l'espace. Cette équation contient

l'opérateur de Laplace

$$\Delta = \frac{\partial^2}{\partial \mathbf{r}^2} \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (3.8)$$

ainsi que la permittivité du vide  $\epsilon_0$ .

7. Enfin, en mécanique quantique, une particule quantique est décrite par une fonction d'onde  $\psi$  qui évolue suivant l'équation de Schrödinger

$$i\hbar \frac{\partial}{\partial t} \psi(\vec{r}, t) = -\frac{\hbar^2}{2m} \Delta \psi(\vec{r}, t) + V(\vec{r}) \psi(\vec{r}, t) \quad (3.9)$$

contenant également le laplacien  $\Delta$ .  $\hbar = h/(2\pi)$  est la constante de Planck réduite,  $m$  est la masse de la particule et  $V(\vec{r})$  représente son énergie potentielle à la position  $\vec{r}$ , cette dernière étant supposée être indépendante du temps  $t$ .

Les exemples listés ci-dessus nous donnent quelques indications sur comment classifier et catégoriser les différents types d'équations différentielles. Tout d'abord, nous pouvons distinguer les *équations différentielles ordinaires*, où la fonction recherchée dépend d'une seule variable d'évolution (le temps généralement), des *équations aux dérivées partielles*, où elle présente des dérivées par rapport à plusieurs variables (spatio-temporelles typiquement). Très clairement, les deux derniers exemples (3.8) et (3.9) représentent des équations aux dérivées partielles tandis que tous les autres exemples listés ci-dessus représentent des équations différentielles ordinaires décrivant une évolution purement temporelle.

Du point de vue de la résolution d'une équation, il est très pertinent de distinguer les équations différentielles *linéaires* des équations différentielles *non linéaires*, ces dernières contenant des termes non linéaires en la fonction recherchée. Les équations (3.2) et (3.4) sont non linéaires tandis que tous les autres exemples listés ci-dessus affichent des équations différentielles linéaires. On est généralement capable de déterminer la solution exacte d'une équation différentielle linéaire alors que c'est rarement le cas pour une équation différentielle non linéaire.

Dans la catégorie des équations différentielles linéaires, il peut être intéressant de distinguer si l'équation en question est *homogène* ou *inhomogène*. Une équation différentielle linéaire est homogène par définition si tous les termes qu'elle contient sont linéaires dans la fonction recherchée, tandis qu'une équation différentielle linéaire inhomogène contient un ou plusieurs termes qui ne dépendent pas de la fonction recherchée. Ceci est le cas pour l'équation de Poisson (3.8), pour l'équation (3.3) qui décrit l'oscillateur harmonique amorti et forcé ainsi que pour le système d'équations (3.5) et (3.6) décrivant le circuit électrique, tandis que les équations linéaires (3.1) et (3.9) sont homogènes. Comme nous allons le voir plus loin dans ce chapitre, les solutions d'une équation différentielle linéaire forment

un espace vectoriel dans le cas d'une équation homogène et un espace affín dans le cas d'une équation inhomogène.

Il est aussi possible de catégoriser les équations différentielles par le nombre de composantes par lesquelles elles sont constituées. L'équation (3.4) est de caractère vectoriel et contient donc de fait plusieurs composantes (3 pour être précis), tout comme le système d'équations (3.5) et (3.6) décrivant le circuit électrique, tandis que tous les autres équations listées ci-dessus sont constituées d'une seule composante. Dans ce contexte, on peut argumenter que l'équation de Schrödinger (3.9) est en réalité une équation à deux composantes correspondant aux parties réelles et imaginaires de la fonction d'onde. Pour éviter un tel débat peu constructif et pour faciliter la discussion dans la suite de ce chapitre, nous allons supposer que la fonction recherchée est de manière générale une fonction à valeurs complexes, ce qui fait que l'équation de Schrödinger (3.9) appartient à la classe des équations différentielles à une composante.

Finalement, on peut aussi classifier les équations différentielles selon l'ordre maximal des dérivées qui y figurent. L'équation (3.1) et le système d'équations (3.5) et (3.6) représentent des équations différentielles du premier ordre tandis que toutes les autres équations listées ci-dessus sont des équations différentielles du deuxième ordre vu que des deuxièmes dérivées par rapport aux variables d'évolution y apparaissent. Cette catégorisation est un peu trop simpliste pour des équations aux dérivées partielles comme l'équation de Schrödinger (3.9) où l'ordre peut être différent pour les différentes variables d'évolution (spatiales et temporelle).

Dans la suite, nous allons restreindre la discussion à des équations différentielles ordinaires du premier ordre mais à plusieurs composantes. De manière générale, ces équations sont écrites comme

$$\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t), t) \quad (3.10)$$

où le vecteur  $\mathbf{y} \equiv (y_1, \dots, y_n) \in \mathbb{C}^n$  contient les  $n \in \mathbb{N}$  composantes de la fonction recherchée et le champ vectoriel  $\mathbf{f} : \mathbb{C}^n \times \mathbb{R} \rightarrow \mathbb{C}^n, (\mathbf{y}, t) \mapsto \mathbf{f}(\mathbf{y}, t)$  définit le système d'équations en question. Notons qu'il s'agit ici d'une notion très abstraite et basique de l'espace vectoriel  $\mathbb{C}^n$  qui n'est muni d'aucune définition de norme. Le système des équations (3.5) et (3.6) peut donc également être écrit sous la forme (3.10), avec  $y_1(t) \equiv I(t)$  et  $y_2(t) \equiv U_c(t)$ , même si le courant  $I$  et la tension  $U_c$  sont des grandeurs physiques ayant des unités différentes.

A première vue, il semble que la restriction à des équations de la forme (3.10) exclue toute équation différentielle de l'ordre  $n > 1$ , donc en particulier toutes les équations newtoniennes de la mécanique analytique. Ceci n'est pas le cas. Considérons l'équation différentielle à une composante

$$x^{(n)}(t) \equiv \frac{d^n x}{dt^n}(t) = g(x(t), \dot{x}(t), \dots, x^{(n-1)}(t), t) \quad (3.11)$$

qui est définie par la fonction  $g : \mathbb{C} \times \dots \times \mathbb{C} \times \mathbb{R} \rightarrow \mathbb{C}$ . Afin de décrire cette équation de l'ordre  $n \in \mathbb{N}$  dans le cadre de l'équation (3.10), nous introduisons le vecteur  $\mathbf{y}(t) \equiv (y_1(t), \dots, y_n(t))$  avec  $y_1(t) = x(t)$  et  $y_l(t) = x^{(l-1)}(t)$  pour  $l \in \{2, \dots, n\}$ . Il est ainsi aisé de montrer que l'équation (3.11) est parfaitement équivalente à l'équation (3.10) si nous définissons le champ vectoriel  $\mathbf{f} \equiv (f_1, \dots, f_n) : \mathbb{C}^n \times \mathbb{R} \rightarrow \mathbb{C}^n$  par  $f_l(\mathbf{y}, t) = y_{l+1}$  pour  $l \in \{1, \dots, n-1\}$  et  $f_n(\mathbf{y}, t) = g(y_1, \dots, y_{n-1}, t)$  pour tout  $(\mathbf{y}, t) \in \mathbb{C}^n \times \mathbb{R}$ .

## 3.2 L'existence et l'unicité de la solution

Avant d'aborder les différentes techniques de résolution pour les équations différentielles ordinaires, il est nécessaire de discuter sous quelles conditions la solution d'une équation différentielle ordinaire existe et est unique pour les conditions initiales données. Le théorème de *Picard-Lindelöf*, aussi connu sous le nom *Cauchy-Lipschitz*, fournit une réponse à cette question.

Pour cela, nous avons besoin d'une notion de continuité qui porte le nom du mathématicien LIPSCHITZ. Selon cette notion, une fonction  $\mathbf{f} : U \subseteq (\mathbb{C}^n \times \mathbb{R}) \rightarrow \mathbb{C}^n$ ,  $(\mathbf{y}, t) \mapsto \mathbf{f}(\mathbf{y}, t)$  définie sur l'ouvert  $U \subseteq (\mathbb{C}^n \times \mathbb{R})$  est appelée *localement lipschitzienne* par rapport à son premier argument (vectoriel) s'il existe pour tout  $(\mathbf{y}_0, t_0) \in U$  un voisinage ouvert  $V \subseteq U$  avec  $(\mathbf{y}_0, t_0) \in V$  et une constante réelle  $L > 0$  tel que

$$\|\mathbf{f}(\mathbf{y}_1, t) - \mathbf{f}(\mathbf{y}_2, t)\| \leq L\|\mathbf{y}_1 - \mathbf{y}_2\| \quad (3.12)$$

pour tous  $(\mathbf{y}_1, t), (\mathbf{y}_2, t) \in V$ , où  $\|\dots\|$  représente une norme définie sur  $\mathbb{C}^n$ . Très clairement, toute fonction qui est continûment dérivable est aussi localement lipschitzienne, la borne  $L$  pouvant être déterminée par la dérivée locale de la fonction, et toute fonction qui est localement lipschitzienne est aussi continue.

Le théorème de Picard-Lindelöf stipule que si la fonction  $\mathbf{f}$  satisfait à cette propriété et est donc localement lipschitzienne par rapport à son premier argument, l'équation différentielle

$$\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t), t) \quad (3.13)$$

possède une solution unique pour la condition initiale  $\mathbf{y}(t_0) = \mathbf{y}_0$ . Cette solution peut être définie sur un intervalle ouvert autour de  $t_0$ , c-à-d, sur  $I = ]t_1, t_2[ \subseteq \mathbb{R}$  avec  $t_1 < t_0 < t_2$ . Le théorème nous fournit aussi la relation

$$\mathbf{y}(t) = \mathbf{y}_0 + \int_{t_0}^t \mathbf{f}(\mathbf{y}(t'), t') dt' \quad (3.14)$$

à laquelle la solution satisfait pour tout  $t \in I$ , qui évidemment correspond à l'intégration de l'équation différentielle (3.13) de  $t_0$  à  $t$ . Cette dernière relation peut être utilisée pour prouver le théorème de Picard-Lindelöf, notamment en injectant itérativement des approximations successives  $\mathbf{y}^{(\nu)}$  ( $\nu \in \mathbb{N}_0$ ) pour la

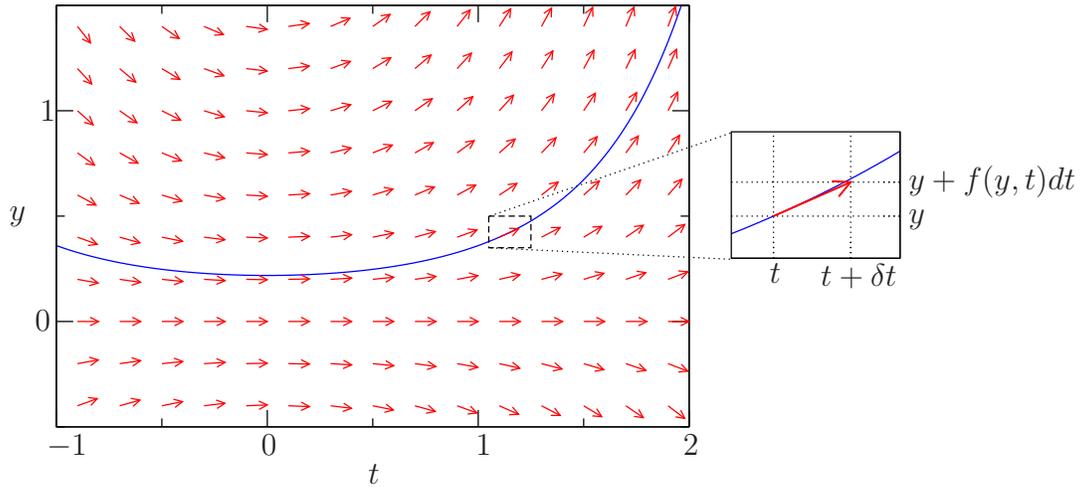


Figure 3.2: Illustration de la méthode de la résolution graphique, exemplifiée ici pour l'équation différentielle  $\dot{y}(t) = ty(t)$  dont la solution générale réelle est  $y(t) = \alpha \exp(t^2/2)$  pour une constante  $\alpha \in \mathbb{R}$ . La solution (ligne bleue, pour  $\alpha = 0,21843$ ) peut être tracée dans le graphique en partant de la condition initiale et en suivant les flèches qui indiquent la pente locale donnée par  $f(y, t) = ty$ , comme il est montré dans l'agrandissement.

fonction recherchée dans l'argument de  $\mathbf{f}$  dans l'équation (3.14), à commencer par  $\mathbf{y}^{(0)}(t) \equiv \mathbf{y}_0$  pour tout  $t \in I$ , et en montrant que la suite  $(\mathbf{y}^{(\nu)})_{\nu \in \mathbb{N}_0}$  des fonctions ainsi obtenues est convergente et converge vers une fonction qui satisfait à l'équation différentielle (3.13).

La relation (3.14) est le point de départ pour la conception de nombreux algorithmes permettant le calcul numérique de la solution de l'équation (3.13). Malgré le fait qu'elle ne fournisse pas directement sa solution analytique, elle peut toutefois être utilisée pour obtenir une vision qualitative du comportement de la fonction recherchée autour de ses valeurs initiales. Cela peut être assez facilement visualisé pour  $n = 1$ , c-à-d, pour le cas particulier d'une équation différentielle à une seule composante

$$\dot{y}(t) = f(y(t), t). \quad (3.15)$$

La solution de cette équation satisfait à la relation

$$y(t + \delta t) = y(t) + \int_t^{t+\delta t} f(y(t'), t') dt' \quad (3.16)$$

pour tout couple  $(y(t), t)$  où la fonction  $f : \mathbb{C} \times \mathbb{R} \rightarrow \mathbb{C}$  est localement lipschitzienne par rapport à son premier argument, pourvu que  $|\delta t|$  soit suffisamment petit tel que  $(y(t + \delta t), t + \delta t) \in V$ . Dans la limite  $\delta t \rightarrow 0$ , nous pouvons approximer l'intégrale apparaissant au côté gauche de l'équation (3.16) par l'aire d'un simple

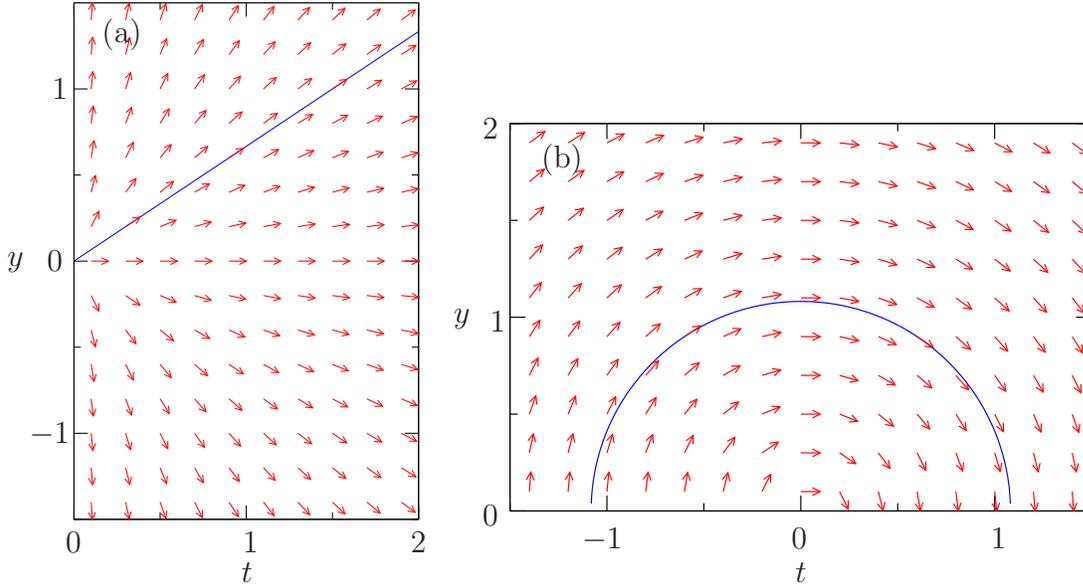


Figure 3.3: Illustration de la résolution graphique des équations différentielles (a)  $\dot{y}(t) = y(t)/t$  et (b)  $\dot{y}(t) = -t/y(t)$ .

rectangle riemannien (voir Eq. (1.78) pour la somme de Riemann), ce qui permet d'approcher la solution par

$$y(t + \delta t) \simeq y(t) + f(y(t), t)\delta t \quad (3.17)$$

ou, plus précisément encore,

$$y(t + \delta t) \simeq y(t) + f(y(t + \delta t/2), t + \delta t/2)\delta t, \quad (3.18)$$

cette dernière expression donnant naissance à la méthode *leapfrog* pour calculer numériquement la solution de l'équation (3.15).

L'approximation (3.17) nous permet de résoudre l'équation différentielle (3.15) graphiquement, notamment en dessinant la pente locale de la solution, donnée par la fonction  $f(y, t)$ , par des flèches dans un diagramme  $y - t$ , comme il est illustré sur la figure 3.2. La solution peut être construite en suivant simplement la suite des flèches à partir du point  $(y_0, t_0)$  correspondant à la condition initiale  $y(t_0) = y_0$ . Cette méthode graphique permet même de déterminer les solutions analytiques de quelques équations différentielles sans effectuer un calcul à proprement parler. Considérons p.ex. l'équation différentielle

$$\dot{y}(t) = \frac{y(t)}{t} \quad (3.19)$$

pour  $t > 0$ . La pente d'une solution à cette équation est donnée par le rapport entre  $y$  et  $t$ , ce qui donne naissance à des flèches orientées en direction radiale

par rapport à l'origine dans le système des coordonnées  $y - t$ , comme montré sur la figure 3.3(a). Par conséquent, toute solution à cette équation différentielle suit une droite traversant l'origine et peut donc être exprimée par  $y(t) = \alpha t$  où la constante  $\alpha$  est déterminée par la condition initiale. Très clairement, nous avons  $\dot{y}(t) = y(t)/t = \alpha$  pour tout  $t$ , ce qui confirme que nous avons affaire à une solution à l'équation (3.19).

En revanche, l'équation

$$\dot{y}(t) = -\frac{t}{y(t)} \quad (3.20)$$

définie pour  $y > 0$  donne naissance à des flèches orientées perpendiculaires au rayon par rapport à l'origine dans le système des coordonnées  $y - t$ , comme illustré sur la figure 3.3(b). Par conséquent, les solutions de cette équation différentielle suivent des arcs de cercles concentriques autour de l'origine. Elles peuvent donc être exprimées sous forme analytique par

$$y(t) = \sqrt{\tau^2 - t^2} \quad (3.21)$$

où la valeur du paramètre  $\tau \in \mathbb{R}$  dépend de la condition initiale. Comme pour l'exemple précédent, il est aisé de montrer que la fonction (3.21) satisfait à l'équation différentielle (3.20).

### 3.3 La méthode de la séparation

Nous considérons à présent le cas particulier d'une équation différentielle réelle à une composante,

$$\dot{y}(t) = f(y(t), t), \quad (3.22)$$

dont le membre de droite peut s'écrire sous la forme

$$f(y, t) = g(y)h(t). \quad (3.23)$$

Nous supposons que  $h : I \rightarrow \mathbb{R}, t \mapsto h(t)$  est une fonction continue, définie sur l'intervalle ouvert  $I \subseteq \mathbb{R}$ , et  $g : U \rightarrow \mathbb{R}_+, y \mapsto g(y)$  est une fonction qui est localement lipschitzienne et ne rend que des valeurs positives sur son domaine de définition  $U \subseteq \mathbb{R}$ , ce dernier étant aussi un intervalle ouvert.

Cette équation différentielle peut être résolue par la *méthode de la séparation*. En pratique, cette méthode consiste à mettre tous les termes dépendant de la fonction recherchée  $y$  dans le membre de gauche de l'équation et tous les termes dépendant de la variable d'évolution  $t$  dans le membre de droite. Quant à la dérivée  $\dot{y}$  de la fonction recherchée, cette dernière peut être exprimée comme le rapport  $dy/dt$  entre la variation  $dy$  de la fonction recherchée et la variation  $dt$  de la variable d'évolution, tout à fait dans l'esprit de la définition (1.50) de

la dérivée. En traitant donc  $dy/dt$  comme un rapport ordinaire, nous pouvons achever la séparation des variables  $y$  et  $t$  dans l'équation différentielle

$$\dot{y} = \frac{dy}{dt} = g(y)h(t) \quad (3.24)$$

en divisant le membre de gauche et de droite de cette équation par  $g(y)$  (qui ne s'annule pas par définition) et en multipliant les deux membres de cette équation par  $dt$ . Nous obtenons ainsi

$$\frac{1}{g(y)}dy = h(t)dt. \quad (3.25)$$

Une simple intégration des deux membres de cette dernière équation,

$$\int \frac{1}{g(y)}dy = \int h(t)dt, \quad (3.26)$$

nous donne une relation implicite à laquelle la solution  $y : I \rightarrow U, t \mapsto y(t)$  pour la condition initiale  $y(t_0) = y_0$  avec  $t_0 \in I$  et  $y_0 \in U$  doit satisfaire, notamment

$$\int_{y_0}^{y(t)} \frac{1}{g(y')}dy' = \int_{t_0}^t h(t')dt'. \quad (3.27)$$

Le raisonnement développé ci-dessus ne peut pas être considéré comme une démonstration mathématique à proprement parler. Il convient donc de montrer qu'une fonction  $y$  qui satisfait à la relation (3.27) satisfait aussi à l'équation différentielle (3.24) ainsi qu'à la condition initiale  $y(t_0) = y_0$ . Par le théorème de Picard-Lindelöf, il s'agit donc de la solution unique de l'équation (3.24) pour la condition initiale en question.

Examinons d'abord si la fonction  $y$  satisfaisant à la relation (3.27) doit nécessairement satisfaire à la condition initiale  $y(t_0) = y_0$ . Pour ce faire, nous constatons d'abord que la fonction  $g$  est continue sur son domaine de définition  $U$ , vu qu'elle est localement lipschitzienne, et n'y atteint que des valeurs positives. La fonction  $1/g$  est donc également continue et strictement positive sur  $U$ . Sa primitive

$$\phi : U \rightarrow \mathbb{R}, y \mapsto \phi(y) = \int_{y_0}^y \frac{1}{g(y')}dy' \quad (3.28)$$

est donc strictement monotone sur  $U$ , avec  $\phi'(y) > 0$  pour tout  $y \in U$ . Par conséquent,  $\phi$  ne peut avoir qu'un seul zéro en  $U$ , ce dernier se trouvant en  $y = y_0$ . Comme le membre de droite de l'équation (3.27) s'annule trivialement pour  $t = t_0$ , nous pouvons conclure qu'une fonction  $y$  satisfaisant à la relation (3.27) doit nécessairement satisfaire à la condition initiale  $y(t_0) = y_0$ .

Examinons ensuite si une telle fonction satisfait aussi à l'équation différentielle (3.24). Pour ce faire, nous évaluons la relation (3.27) en  $t + \delta t$  et nous y soustrayons son évaluation en  $t$ . Ceci donne dans la limite  $\delta t \rightarrow 0$

$$\int_{t_0}^{t+\delta t} h(t') dt' - \int_{t_0}^t h(t') dt' = \int_t^{t+\delta t} h(t') dt' \simeq h(t) \delta t \quad (3.29)$$

pour le membre de droite de l'équation (3.27) et

$$\int_{y_0}^{y(t+\delta t)} \frac{1}{g(y')} dy' - \int_{y_0}^{y(t)} \frac{1}{g(y')} dy' = \int_{y(t)}^{y(t+\delta t)} \frac{1}{g(y')} dy' \simeq \frac{y(t+\delta t) - y(t)}{g(y(t))} \quad (3.30)$$

pour le membre de gauche. Comme, vu la définition de la dérivée, nous avons  $y(t + \delta t) - y(t) \simeq \dot{y}(t) \delta t$  dans la limite  $\delta t \rightarrow 0$ , nous obtenons en égalant les équations (3.29) et (3.30)

$$\frac{\dot{y}(t)}{g(y(t))} \delta t = h(t) \delta t, \quad (3.31)$$

ce qui prouve, après multiplication de cette équation par  $g(y(t))$ , que la fonction  $y$  satisfait à l'équation différentielle (3.24).

Grâce à cette méthode, nous pouvons facilement calculer la solution générale de l'équation différentielle

$$\dot{y}(t) = h(t) y^\alpha \quad (3.32)$$

pour tout  $\alpha \in \mathbb{R}$ . En évaluant la primitive de  $y^{-\alpha}$  par

$$\int y^{-\alpha} dy = \begin{cases} \frac{1}{1-\alpha} y^{1-\alpha} & : \alpha \neq 1 \\ \ln |y| & : \alpha = 1 \end{cases}, \quad (3.33)$$

nous obtenons par la relation (3.27)

$$\int_{t_0}^t h(t') dt' = \int_{y_0}^{y(t)} \frac{1}{y^\alpha} dy = \begin{cases} \frac{1}{1-\alpha} [(y(t))^{1-\alpha} - y_0^{1-\alpha}] & : \alpha \neq 1 \\ \ln |y(t)| - \ln |y_0| & : \alpha = 1 \end{cases}. \quad (3.34)$$

Pour  $\alpha = 1$ , cette équation est résolue par

$$y(t) = \operatorname{sgn}(y_0) \exp \left( \ln |y_0| + \int_{t_0}^t h(t') dt' \right) = y_0 \exp \left( \int_{t_0}^t h(t') dt' \right) \quad (3.35)$$

où  $\operatorname{sgn}(y_0) \equiv y_0/|y_0|$  est le signe de la valeur initiale  $y_0$  de la fonction recherchée. Pour  $\alpha \neq 1$ , nous obtenons la solution

$$y(t) = \left( y_0^{1-\alpha} + (1-\alpha) \int_{t_0}^t h(t') dt' \right)^{1/(1-\alpha)}. \quad (3.36)$$

Il est assez instructif de discuter le cas particulier d'une fonction  $h$  constante, c-à-d,  $h(t) = h_0$  pour tout  $t$ , et d'un exposant  $\alpha \in ]0, 1[$ , p.ex.  $\alpha = 2/3$ . La solution de l'équation différentielle (3.32) est donc donnée par

$$y(t) = \left( y_0^{1/3} + \frac{h_0}{3}(t - t_0) \right)^3 \quad (3.37)$$

dans ce cas. Si nous imposons la condition initiale  $y_0 = 0$  au temps initial  $t_0 = 0$ , nous obtenons ainsi la solution

$$y(t) = (h_0 t / 3)^3 \quad (3.38)$$

pour tout  $t \in \mathbb{R}$ . Ceci peut étonner un peu, vu que la fonction "nulle", c-à-d  $y(t) \equiv 0$  pour tout  $t$ , est aussi une solution à l'équation (3.37) satisfaisant à la même condition initiale  $y(0) = 0$ , ce qui semble contredire l'unicité de la solution.

La résolution de ce paradoxe apparent est que la fonction  $g(y) = y^{2/3}$  ne satisfait pas à la condition de Lipschitz (3.12), à savoir, pour ce cas particulier,

$$|y_1^{2/3} - y_2^{2/3}| < L|y_1 - y_2| \quad (3.39)$$

pour tout  $(y_1, t), (y_2, t) \in V$  où  $V$  est un voisinage de  $y_0$ . Vu que la dérivée  $g'(y) = 2y^{-1/3}/3$  diverge à l'origine, nous ne pouvons pas spécifier une constante  $L > 0$  telle que la condition (3.39) soit satisfaite pour  $y_0 = 0$ . Il n'y a donc aucune garantie mathématique qu'une solution à l'équation (3.37) soit unique pour une telle condition initiale.

## Exercices

3.1 Calculer la solution  $y : I \subseteq \mathbb{R} \rightarrow \mathbb{R}, t \mapsto y(t)$  de l'équation différentielle

- (a)  $\dot{y} = y^2 t$
- (b)  $\dot{y} = -(1 + y^2)t$
- (c)  $\dot{y} = (y^2 - 4)t$
- (d)  $\dot{y} = t\sqrt{4 - y^2}$
- (e)  $\dot{y} = t \exp(t - y)$
- (f)  $(4 + t^2)\dot{y} = 4 - y^2$
- (g)  $3\dot{y} = (2 + y - y^2)(1 - t)$
- (h)  $y\dot{y} = \cosh(t) \exp(y^2)$
- (i)  $y\dot{y} = \exp(t + y^2)$

pour la condition initiale  $y(0) = 1$ . Quel est l'intervalle maximal  $I$  dans lequel on peut définir la solution ?

3.2 Calculer la solution  $y : I \subseteq \mathbb{R} \rightarrow \mathbb{R}, t \mapsto y(t)$  de l'équation différentielle

(a)  $\dot{y} = 1 - y^2$

(b)  $\dot{y} = \sqrt{(1 - y^2)t}$

(c)  $(1 + y)\dot{y} = \exp(y^2 + 2y - t + 2)$

pour la condition initiale  $y(0) = 0$ . Quel est l'intervalle maximal  $I$  dans lequel on peut définir la solution ?

### 3.4 Les équations différentielles linéaires

Au vu de leur importance en physique et de la facilité de leur résolution, les équations différentielles linéaires méritent une discussion à part. Commençons par le cas des équations différentielles linéaires homogènes, c-à-d, pour lesquelles le membre de droite de l'équation  $\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t), t)$  est une fonction qui est strictement linéaire en la fonction recherchée. De manière générale, nous avons donc affaire à une équation différentielle à  $n \in \mathbb{N}$  composantes qui peut s'écrire sous la forme

$$\dot{\mathbf{y}}(t) = H(t)\mathbf{y}(t) \quad (3.40)$$

où

$$H : I \rightarrow \mathbb{C}^{n \times n}, t \mapsto H(t) = \begin{pmatrix} H_{11}(t) & \dots & H_{1n}(t) \\ \vdots & & \vdots \\ H_{n1}(t) & \dots & H_{nn}(t) \end{pmatrix} \quad (3.41)$$

est une fonction matricielle qui est continue sur son domaine de définition  $I \subseteq \mathbb{R}$ , ce dernier étant un intervalle ouvert. Si nous écrivons le vecteur de la fonction recherchée sous la forme

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{C}^n, \quad (3.42)$$

le membre de droite de l'équation (3.40) s'écrit explicitement

$$H(t)\mathbf{y} = \begin{pmatrix} H_{11}(t)y_1 + \dots + H_{1n}(t)y_n \\ \vdots \\ H_{n1}(t)y_1 + \dots + H_{nn}(t)y_n \end{pmatrix}. \quad (3.43)$$

Nous pouvons d'abord constater que la solution à l'équation différentielle (3.40) existe et est unique pour toutes les conditions initiales  $\mathbf{y}_0 \in \mathbb{C}^n$ . Il est, en effet, aisé de déterminer la constante  $L$  apparaissant dans la condition (3.12)

qui permet de vérifier la continuité de Lipschitz pour l'équation en question, notamment par une majoration de la norme matricielle

$$\|H(t)\|_m \equiv \max_{\mathbf{y} \in \mathbb{C}} \frac{\|H(t)\mathbf{y}\|}{\|\mathbf{y}\|} \quad (3.44)$$

au voisinage de  $t = t_0$ , où  $\|\mathbf{y}\|$  et  $\|H(t)\mathbf{y}\|$  représentent les normes vectorielles respectivement des vecteurs  $\mathbf{y}$  et  $H(t)\mathbf{y}$ . Tout jeu de conditions initiales  $\mathbf{y}_0 \in \mathbb{C}^n$  possède donc une solution unique  $\mathbf{y} : I \rightarrow \mathbb{C}^n, t \mapsto \mathbf{y}(t)$  de l'équation différentielle (3.40) avec  $\mathbf{y}(t_0) = \mathbf{y}_0$ .

La linéarité de l'équation (3.40) nous permet ensuite de démontrer avec aisance que si les fonctions  $\mathbf{y}_j : I \rightarrow \mathbb{C}^n, t \mapsto \mathbf{y}_j(t)$  avec  $j = 1, \dots, N$  sont toutes des solutions à l'équation différentielle (3.40), leurs conditions initiales étant arbitrairement choisies, alors la fonction

$$\mathbf{y} : I \rightarrow \mathbb{C}^n, t \mapsto \mathbf{y}(t) = \sum_{j=1}^N \alpha_j \mathbf{y}_j(t) \quad (3.45)$$

définie pour des constantes arbitraires  $\alpha_1, \dots, \alpha_N \in \mathbb{C}$  est aussi une solution à l'équation (3.40). L'espace des solutions à l'équation différentielle (3.40) est donc linéaire, c-à-d, il s'agit d'un espace vectoriel. Sa dimension est identique à la dimension  $n$  de l'espace vectoriel  $\mathbb{C}^n$ , vu que chaque solution  $\mathbf{y}$  peut être associée de manière unique à son jeu de conditions initiales  $\mathbf{y}_0 \in \mathbb{C}^n$ , ce dernier étant un élément d'un espace vectoriel à  $n$  dimensions. Par conséquent, il suffit de déterminer  $n$  solutions à l'équation différentielle (3.40) qui sont *linéairement indépendantes*, afin de pouvoir représenter toute autre solution à cette équation.

Concrètement, soit  $\{\varphi_1, \dots, \varphi_n\}$  avec

$$\varphi_j : I \rightarrow \mathbb{C}^n, t \mapsto \varphi_j(t) = \begin{pmatrix} \varphi_{j1}(t) \\ \vdots \\ \varphi_{jn}(t) \end{pmatrix} \quad (3.46)$$

pour tout  $j = 1, \dots, n$  un tel jeu de solutions à l'équation (3.40). L'indépendance linéaire de ces solutions s'exprime par le fait que si la combinaison linéaire

$$\sum_{j=1}^n c_j \varphi_j(t) = 0 \quad (3.47)$$

s'annule pour un  $t \in I$ , alors nous devons avoir  $c_j = 0$  pour tout  $j = 1, \dots, n$ . Supposons donc que ce jeu de solutions  $\varphi_1, \dots, \varphi_n$  satisfait à cette propriété d'indépendance linéaire. Nous pouvons alors montrer qu'il existe pour toute solution  $\mathbf{y} : I \rightarrow \mathbb{C}^n, t \mapsto \mathbf{y}(t)$  à l'équation différentielle (3.40) un jeu unique de coefficients  $\alpha_1, \dots, \alpha_n \in \mathbb{C}$  tel que nous avons

$$\mathbf{y}(t) = \sum_{j=1}^n \alpha_j \varphi_j(t) \quad (3.48)$$

pour tout  $t \in I$ . Cette solution générale de l'équation différentielle (3.40) peut être reformulée comme

$$\mathbf{y}(t) = \Phi(t)\boldsymbol{\alpha} \quad (3.49)$$

où nous définissons la fonction matricielle

$$\Phi : I \rightarrow \mathbb{C}^{n \times n}, t \mapsto \Phi(t) = (\varphi_1(t), \dots, \varphi_n(t)) = \begin{pmatrix} \varphi_{11}(t) & \dots & \varphi_{n1}(t) \\ \vdots & & \vdots \\ \varphi_{1n}(t) & \dots & \varphi_{nn}(t) \end{pmatrix}, \quad (3.50)$$

qui contient les solutions  $\varphi_j$  dans ses colonnes, et le vecteur

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{C}^n. \quad (3.51)$$

En présence des conditions initiales  $\mathbf{y}(t_0) = \mathbf{y}_0$ , ce dernier vaut concrètement  $\boldsymbol{\alpha} = (\Phi(t_0))^{-1} \mathbf{y}_0$ , c-à-d, nous avons

$$\mathbf{y}(t) = \Phi(t) (\Phi(t_0))^{-1} \mathbf{y}_0, \quad (3.52)$$

où  $(\Phi(t_0))^{-1}$  est l'inverse de la matrice  $\Phi(t_0)$ . On appelle un tel jeu  $\{\varphi_1, \dots, \varphi_n\}$  de solutions linéairement indépendantes un *système fondamental* de l'équation (3.40).

### 3.5 L'exponentielle de matrice

La solution (3.52) peut être explicitement déterminée dans le cas particulier d'une équation différentielle linéaire dont les coefficients  $H_{ij}$  caractérisant le membre de droite sont constants, c-à-d, ne dépendent pas de la variable d'évolution  $t$ . Nous avons donc affaire à l'équation

$$\dot{\mathbf{y}}(t) = H\mathbf{y}(t) \quad (3.53)$$

où  $H \in \mathbb{C}^{n \times n}$  est une matrice constante. De manière tout à fait analogue au cas unidimensionnel (cf. équation (3.35)), la solution générale de cette équation différentielle est donnée par

$$\mathbf{y}(t) = \exp[(t - t_0)H]\mathbf{y}_0 \quad (3.54)$$

avec  $\mathbf{y}_0 \in \mathbb{C}^n$ , où

$$\exp(\tau H) \equiv e^{\tau H} = \sum_{k=0}^{\infty} \frac{\tau^k}{k!} H^k \quad (3.55)$$

est la fonction exponentielle définie dans l'anneau des matrices complexes  $\mathbb{C}^{n \times n}$ . En effet, la fonction (3.54) satisfait à l'équation (3.53) vu que nous avons

$$\frac{d}{d\tau} \exp(\tau H) = \sum_{k=0}^{\infty} \frac{k\tau^{k-1}}{k!} H^k = H \sum_{k'=0}^{\infty} \frac{\tau^{k'}}{k'!} H^{k'} = H \exp(\tau H). \quad (3.56)$$

Par le théorème de Picard-Lindelöf, elle est donc la solution unique à l'équation différentielle (3.53) pour les conditions initiales  $\mathbf{y}(t_0) = \mathbf{y}_0$ .

Il nous reste à déterminer l'exponentielle de matrice  $e^{\tau H}$  afin d'achever le calcul de cette solution. Commençons par le cas où la matrice  $H \in \mathbb{C}^{n \times n}$  est diagonalisable. Dans ce cas, nous pouvons écrire  $H \in \mathbb{C}^{n \times n}$  sous la forme

$$H = BDB^{-1} \quad (3.57)$$

où

$$D = \text{diag}(E_1, \dots, E_n) \equiv \begin{pmatrix} E_1 & & (0) \\ & \ddots & \\ (0) & & E_n \end{pmatrix} \quad (3.58)$$

est une matrice diagonale contenant les valeurs propres  $E_j$  de  $H$  et

$$B = (\mathbf{v}_1, \dots, \mathbf{v}_n) \quad (3.59)$$

est une matrice inversible contenant dans ses colonnes les vecteurs propres  $\mathbf{v}_j$  de  $H$ , ces derniers satisfaisant à l'équation  $H\mathbf{v}_j = E_j\mathbf{v}_j$  pour tout  $j = 1, \dots, n$ . Nous pouvons donc évaluer l'exponentielle de matrice comme

$$\exp(\tau H) = \sum_{k=0}^{\infty} \frac{\tau^k}{k!} (BDB^{-1})^k = B \sum_{k=0}^{\infty} \frac{\tau^k}{k!} D^k B^{-1} = B \exp(\tau D) B^{-1}, \quad (3.60)$$

où nous avons fait usage de la relation

$$(BDB^{-1})^k = BD(B^{-1}B)D(B^{-1}B) \dots (B^{-1}B)DB^{-1} = BD^k B^{-1} \quad (3.61)$$

sachant que  $B^{-1}B = \mathbb{I} \equiv \text{diag}(1, \dots, 1)$  est la matrice unité. La fonction exponentielle de la matrice diagonale  $D$  se calcule facilement par

$$\exp(\tau D) = \text{diag}(e^{\tau E_1}, \dots, e^{\tau E_n}). \quad (3.62)$$

Dans le cas particulier d'une matrice  $H$  normale où, par définition,  $H$  commute avec sa conjuguée hermitienne  $H^\dagger$ ,

$$HH^\dagger = H^\dagger H, \quad (3.63)$$

nous pouvons diagonaliser  $H$  et choisir ses vecteurs propres  $\mathbf{v}_j$  tels que la matrice  $B$ , définie par l'équation (3.59), soit unitaire. Elle satisfait donc à la relation

$B^{-1} = B^\dagger$ , ce qui évidemment simplifie le calcul de l'exponentielle de matrice  $e^{\tau H}$  suivant l'équation (3.60). Ce cas d'une matrice normale est particulièrement important en physique vu que toutes les matrices hermitiennes (satisfaisant à  $H^\dagger = H$ ) et unitaires (satisfaisant à  $H^\dagger = H^{-1}$ ) sont des matrices normales.

Le cas que la matrice  $H$  soit diagonalisable est générique dans le sens qu'il se réalise avec une certitude numérique si l'on choisit les éléments de cette matrice par des nombres complexes aléatoires. Néanmoins, dans des cas exceptionnels, il est possible que l'on ne soit pas capable de déterminer une base de  $\mathbb{C}^n$  constituée de  $n$  vecteurs propres de  $H$  qui sont linéairement indépendants. Cela arrive généralement si la matrice  $H$  possède des valeurs propres dégénérées, c-à-d, qui correspondent à des zéros multiples de son polynôme caractéristique  $p_H(\lambda) = \det(\lambda \mathbb{I} - H)$ .

Dans cette situation, il convient de déterminer la *réduction de Jordan* de  $H$  afin de pouvoir calculer son exponentielle de matrice. Pour ce faire, supposons que la matrice  $H$  possède un maximum de  $k \leq n$  valeurs propres  $E_1, \dots, E_k$  tel que pour chacune des ces valeurs propres  $E_j$  on peut déterminer un vecteur propre associé  $\mathbf{v}_{j,1}$ , satisfaisant donc à la relation

$$H\mathbf{v}_{j,1} = E_j\mathbf{v}_{j,1}, \quad (3.64)$$

qui est linéairement indépendant de tous les autres vecteurs propres. Si  $k < n$ , nous pouvons, pour chaque valeur propre  $E_j$ , compléter la base de  $\mathbb{C}^n$  par des vecteurs  $\mathbf{v}_{j,2}, \dots, \mathbf{v}_{j,\nu_j}$  qui sont définis de manière récursive par

$$H\mathbf{v}_{j,i} = E_j\mathbf{v}_{j,i} + \mathbf{v}_{j,i-1}, \quad (3.65)$$

pour  $1 < i \leq \nu_j$ . Le nombre maximal  $\nu_j$  des vecteurs  $\mathbf{v}_{j,i}$  qui peuvent ainsi être construits est inférieur ou égal à la *multiplicité algébrique* de la valeur propre  $E_j$ , c-à-d, l'ordre avec lequel le zéro  $E_j$  apparaît dans le polynôme caractéristique  $p_H$  (sachant que nous pouvons avoir  $E_j = E_{j'}$  pour  $j \neq j'$ , comme c'est le cas pour la matrice unité  $\mathbb{I}$  où  $E_1 = \dots = E_n = 1$ ). Nous avons  $\nu_1 + \dots + \nu_k = n$ .

La réduction de Jordan est obtenue en représentant  $H$  dans la base des vecteurs  $\mathbf{v}_{j,i}$  ainsi construits. Cette dernière étant définie par la matrice

$$B = (\mathbf{v}_{1,1}, \dots, \mathbf{v}_{1,\nu_1}, \mathbf{v}_{2,1}, \dots, \mathbf{v}_{2,\nu_2}, \dots, \mathbf{v}_{k,1}, \dots, \mathbf{v}_{k,\nu_k}), \quad (3.66)$$

nous obtenons par les relations (3.64) et (3.65)

$$H = B\tilde{D}B^{-1} \quad (3.67)$$

avec la matrice diagonale par blocs

$$\tilde{D} = \begin{pmatrix} J_1 & & & (0) \\ & J_2 & & \\ & & \ddots & \\ (0) & & & J_k \end{pmatrix} \quad (3.68)$$

qui est définie en termes des *blocs de Jordan*

$$J_j = \begin{pmatrix} E_j & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & E_j \end{pmatrix} \quad (3.69)$$

de la taille  $\nu_j \times \nu_j$ . L'utilité de cette réduction de Jordan réside dans le fait que la matrice

$$\mathcal{I} = \begin{pmatrix} 0 & 1 & & (0) \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ (0) & & & 0 \end{pmatrix} \in \mathbb{R}^{\nu} \quad (3.70)$$

est *nilpotente* et satisfait à la propriété  $\mathcal{I}^l = \mathbb{O}$  pour tout  $l \geq \nu$ , où  $\mathbb{O} \in \mathbb{R}^{\nu}$  est la matrice nulle. Evidemment, cela simplifie le calcul de la série exponentielle (3.55) de la matrice  $\tau J_j$ . Concrètement, nous obtenons ainsi

$$\exp(\tau H) = B \exp(\tau \tilde{D}) B^{-1} \quad (3.71)$$

avec

$$\exp(\tau \tilde{D}) = \begin{pmatrix} e^{\tau J_1} & & & (0) \\ & e^{\tau J_2} & & \\ & & \ddots & \\ (0) & & & e^{\tau J_k} \end{pmatrix} \quad (3.72)$$

et

$$e^{\tau J_j} = e^{\tau E_j} \begin{pmatrix} 1 & \tau & \frac{\tau^2}{2} & \cdots & \frac{\tau^{\nu_j-1}}{(\nu_j-1)!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{\tau^2}{2} \\ & & & \ddots & \tau \\ (0) & & & & 1 \end{pmatrix}. \quad (3.73)$$

### Exercices

3.3 Calculer pour  $\mathbf{y} \in \mathbb{R}^2$  la solution générale réelle de l'équation différentielle

$\frac{d}{dt} \mathbf{y}(t) = A \mathbf{y}(t)$  avec

(a)  $A = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}$     (b)  $A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$     (c)  $A = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$

(d)  $A = \begin{pmatrix} 1 & -1 \\ 3 & 5 \end{pmatrix}$     (e)  $A = \begin{pmatrix} 2 & -1 \\ 1 & 4 \end{pmatrix}$     (f)  $A = \begin{pmatrix} 1 & 5 \\ -2 & -1 \end{pmatrix}$

(g)  $A = \begin{pmatrix} -1 & -3 \\ 2 & 4 \end{pmatrix}$     (h)  $A = \begin{pmatrix} 1 & 1 \\ 4 & 1 \end{pmatrix}$     (i)  $A = \begin{pmatrix} 8 & 6 \\ -9 & -7 \end{pmatrix}$

### 3.6 La fonction de Green

La solution générale d'une équation différentielle linéaire homogène étant déterminée, il s'avère plutôt facile de faire face à l'ajout d'un terme inhomogène dans une telle équation. Considérons donc à présent l'équation différentielle linéaire inhomogène

$$\dot{\mathbf{y}}(t) = H(t)\mathbf{y}(t) + \mathbf{b}(t) \quad (3.74)$$

où

$$H : I \rightarrow \mathbb{C}^{n \times n}, t \mapsto H(t) \quad (3.75)$$

est une fonction matricielle et

$$\mathbf{b} : I \rightarrow \mathbb{C}^n, t \mapsto \mathbf{b}(t) \quad (3.76)$$

une fonction vectorielle, toutes les deux étant continues sur leur domaine de définition  $I \subseteq \mathbb{R}$  qui est un intervalle ouvert. Nous constatons d'abord que la solution à cette équation différentielle (3.74) existe et est unique pour toutes les conditions initiales  $\mathbf{y}_0 \in \mathbb{C}^n$ , et ceci pour la même raison que dans le cas d'une équation différentielle linéaire homogène (3.40). En effet, la constante  $L$  permettant de vérifier la continuité de Lipschitz peut être définie d'exactly la même manière que dans ce dernier cas. Contrairement au cas homogène, une superposition quelconque  $\mathbf{y}(t) = \alpha_1 \mathbf{y}_1(t) + \alpha_2 \mathbf{y}_2(t)$  de deux solutions  $\mathbf{y}_j : I \rightarrow \mathbb{C}^n$  ( $j = 1, 2$ ) à l'équation (3.74) ne satisfait généralement pas à cette équation, évidemment à l'exception de quelques cas triviaux (comme  $\alpha_1 = 1$  et  $\alpha_2 = 0$ ).

Il est instructif de former la simple différence entre ces deux solutions, ce qui revient à définir la fonction

$$\mathbf{y} : I \rightarrow \mathbb{C}^n, t \mapsto \mathbf{y}(t) = \mathbf{y}_1(t) - \mathbf{y}_2(t). \quad (3.77)$$

Comme on peut facilement montrer en soustrayant les deux équations différentielles que l'on obtient si on injecte  $\mathbf{y}_1$  et  $\mathbf{y}_2$  dans l'équation (3.74), cette différence satisfait à l'équation différentielle homogène associée

$$\dot{\mathbf{y}}(t) = H(t)\mathbf{y}(t) \quad (3.78)$$

pour laquelle les stratégies et techniques de résolution ont été amplement discutées dans les deux sections précédentes. En particulier, nous avons déjà montré dans la section 3.4 que la solution générale de cette équation différentielle homogène (3.78) peut être représentée par une combinaison linéaire (3.48) de  $n$  solutions fondamentales  $\varphi_j : I \rightarrow \mathbb{C}^n$  ( $j = 1, \dots, n$ ) qui sont linéairement indépendantes. Soit  $\{\varphi_1, \dots, \varphi_n\}$  un tel système fondamental de l'équation (3.78). Nous pouvons alors affirmer que toute solution de l'équation différentielle inhomogène (3.74) peut être représentée par

$$\mathbf{y}(t) = \mathbf{y}_1(t) + \sum_{j=1}^n \alpha_j \varphi_j(t) \quad (3.79)$$

avec des coefficients  $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ , où  $\mathbf{y}_1 : I \rightarrow \mathbb{C}^n, t \mapsto \mathbf{y}_1(t)$  est une solution quelconque de cette équation (3.74). Autrement dit, il suffit de déterminer  $n$  solutions linéairement indépendantes de l'équation homogène (3.78) et une solution particulière de l'équation inhomogène (3.74) afin de pouvoir représenter toute autre solution de cette dernière équation. L'espace des solutions est donc *affin*.

Une stratégie générale pour calculer la solution d'une équation différentielle linéaire inhomogène consiste à déterminer sa *fonction de Green* associée. Pour expliquer cela, nous observons d'abord que si les fonctions  $\mathbf{y}_1, \mathbf{y}_2 : I \rightarrow \mathbb{C}^n$  sont respectivement des solutions aux équations différentielles inhomogènes

$$\dot{\mathbf{y}}_j(t) = H(t)\mathbf{y}_j(t) + \mathbf{b}_j(t) \quad (3.80)$$

avec des termes inhomogènes  $\mathbf{b}_j(t) : I \rightarrow \mathbb{C}^n$  pour  $j = 1, 2$ , alors la combinaison linéaire  $\mathbf{y} : I \rightarrow \mathbb{C}^n, t \mapsto \mathbf{y}(t) = \alpha\mathbf{y}_1(t) + \beta\mathbf{y}_2(t)$  avec  $\alpha, \beta \in \mathbb{C}$  est une solution de l'équation différentielle (3.74) pour  $\mathbf{b}(t) = \alpha\mathbf{b}_1(t) + \beta\mathbf{b}_2(t)$ . Cette propriété de linéarité, qui implique également les termes inhomogènes, peut être exploitée afin de déterminer de manière générale les solutions de l'équation différentielle (3.74) pour toute fonction continue  $\mathbf{b} : I \rightarrow \mathbb{C}^n$  jouant le rôle du terme inhomogène. Nous utilisons pour cela l'identité

$$\mathbf{b}(t) = \int_{t_1}^{t_2} \delta(t - \tau)\mathbf{b}(\tau)d\tau \quad (3.81)$$

où  $\delta$  représente la distribution de Dirac et où nous avons introduit  $t_1$  et  $t_2$  comme bornes de l'intervalle  $I = ]t_1, t_2[$ . Par le principe de superposition que nous venons de développer, il suffit donc de calculer les solutions de l'équation

$$\dot{\mathbf{y}}(t) = H(t)\mathbf{y}(t) + \delta(t - \tau)\mathbf{e}_j \quad (3.82)$$

pour tout  $\tau \in I$  et tout  $j = 1, \dots, n$ , où  $\mathbf{e}_j$  signifie le vecteur unitaire associé à la composante  $j$  de l'espace vectoriel  $\mathbb{C}^n$ . En effet, au vu des propriétés particulières de la fonction delta que nous avons discutées dans la section 2.1, cette tâche peut s'avérer relativement facile comparé au calcul direct d'une solution de l'équation (3.74).

Pour formaliser cette idée, nous introduisons une fonction matricielle

$$G : I \times I \rightarrow \mathbb{C}^{n \times n}, (t, \tau) \mapsto G(t, \tau) \quad (3.83)$$

qui satisfait à l'équation

$$\frac{\partial}{\partial t}G(t, \tau) = H(t)G(t, \tau) + \delta(t - \tau)\mathbb{I}, \quad (3.84)$$

où  $\mathbb{I}$  est la matrice unité en  $\mathbb{C}^{n \times n}$ . C'est la *fonction de Green* associée à l'équation différentielle (3.74). Une fois cette fonction  $G$  déterminée pour tout  $t \in I$  et tout

$\tau \in I$ , nous pourrons construire une solution à l'équation différentielle (3.74) par

$$\mathbf{y} : I \rightarrow \mathbb{C}^n, t \mapsto \mathbf{y}(t) = \int_{t_1}^{t_2} G(t, \tau) \mathbf{b}(\tau) d\tau. \quad (3.85)$$

Dans le cas particulier où les conditions initiales s'annulent, c-à-d,  $\mathbf{y}(t_0) = \mathbf{y}_0 = \mathbf{0}$ , il est utile de calculer cette fonction de Green avec ces mêmes conditions initiales, c-à-d,  $G(t_0, \tau) = \mathbb{O}$  pour tout  $\tau \in I$ , où  $\mathbb{O}$  est la matrice nulle en  $\mathbb{C}^{n \times n}$ . Nous obtenons ainsi par la construction (3.85) directement la solution recherchée.

### 3.7 La variation des constantes

La fonction de Green représente l'outil adéquat pour résoudre des équations différentielles linéaires inhomogènes dans un contexte très général, comprenant aussi les équations aux dérivées partielles. Néanmoins, dans le cas des équations différentielles ordinaires avec une seule ou très peu de composantes, il existe une autre approche qui est plus facile à mettre en œuvre. Cette approche s'inspire du constat que nous avons fait dans la section précédente, notamment que la solution générale d'une équation différentielle linéaire inhomogène est constituée, selon l'équation (3.79), de la solution générale de l'équation homogène associée et d'une solution particulière de l'équation inhomogène. Cette dernière solution particulière peut être obtenue en faisant varier les constantes  $\alpha_1, \dots, \alpha_n$  qui caractérisent la solution (3.48) de l'équation différentielle homogène.

Concrètement, la résolution de l'équation différentielle linéaire inhomogène

$$\dot{\mathbf{y}}(t) = H(t)\mathbf{y}(t) + \mathbf{b}(t), \quad (3.86)$$

avec  $\mathbf{y} \in \mathbb{C}^n$  et les fonctions continues  $H : I \rightarrow \mathbb{C}^{n \times n}$  et  $\mathbf{b} : I \rightarrow \mathbb{C}^n$  définies sur l'intervalle ouvert  $I \subseteq \mathbb{R}$ , se fait en deux étapes. La première étape consiste à déterminer la solution générale de l'équation différentielle homogène associée

$$\dot{\mathbf{y}}(t) = H(t)\mathbf{y}(t). \quad (3.87)$$

Comme nous l'avons discuté dans la section 3.4, cette solution générale peut être obtenue par un système fondamental  $\{\varphi_1, \dots, \varphi_n\}$  constitué de  $n$  solutions  $\varphi_j : I \rightarrow \mathbb{C}^n$  à l'équation homogène (3.87) qui sont linéairement indépendantes. La solution générale de l'équation (3.87) peut donc s'écrire sous la forme

$$\mathbf{y}(t) = \sum_{j=1}^n \alpha_j \varphi_j(t) = \Phi(t) \boldsymbol{\alpha} \quad (3.88)$$

avec

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{C}^n \quad (3.89)$$

le vecteur des constantes  $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ , où nous avons introduit la fonction matricielle

$$\Phi : I \rightarrow \mathbb{C}^{n \times n}, t \mapsto \Phi(t) = (\varphi_1(t), \dots, \varphi_n(t)) = \begin{pmatrix} \varphi_{11}(t) & \dots & \varphi_{n1}(t) \\ \vdots & & \vdots \\ \varphi_{1n}(t) & \dots & \varphi_{nn}(t) \end{pmatrix}. \quad (3.90)$$

Après avoir déterminé ce système fondamental, nous pouvons poser

$$\mathbf{y}_1(t) = \Phi(t)\boldsymbol{\alpha}(t) \quad (3.91)$$

comme solution particulière à l'équation différentielle inhomogène (3.86), où  $\boldsymbol{\alpha} : I \rightarrow \mathbb{C}^n$  est maintenant une fonction vectorielle à  $n$  composantes définie sur l'intervalle  $I$ . La dérivée première de cette candidate à une solution particulière par rapport à  $t$  est évaluée comme

$$\dot{\mathbf{y}}_1(t) = \dot{\Phi}(t)\boldsymbol{\alpha}(t) + \Phi(t)\dot{\boldsymbol{\alpha}}(t) = H(t)\Phi(t)\boldsymbol{\alpha}(t) + \Phi(t)\dot{\boldsymbol{\alpha}}(t), \quad (3.92)$$

où nous avons utilisé que chacune des colonnes de la fonction matricielle  $\Phi$  satisfait à l'équation différentielle (3.87). En injectant cette expression (3.92) dans le membre de gauche de l'équation (3.86), nous pouvons déduire que la fonction  $\mathbf{y}_1$  satisfait à cette équation différentielle (3.86) si et seulement si nous avons

$$\Phi(t)\dot{\boldsymbol{\alpha}}(t) = \mathbf{b}(t) \quad (3.93)$$

ou, de manière parfaitement équivalente vu que la matrice  $\Phi(t)$  est inversible,

$$\dot{\boldsymbol{\alpha}}(t) = (\Phi(t))^{-1}\mathbf{b}(t) \quad (3.94)$$

pour tout  $t \in I$ . Une intégration de cette dernière équation (3.94) sur  $t$  nous fournit

$$\boldsymbol{\alpha}(t) = \boldsymbol{\alpha}_0 + \int_{t_0}^t (\Phi(t'))^{-1}\mathbf{b}(t')dt' \quad (3.95)$$

avec le vecteur  $\boldsymbol{\alpha}_0 \in \mathbb{C}^n$  contenant les constantes d'intégration. En injectant cette dernière expression (3.95) dans l'équation (3.91), nous obtenons

$$\mathbf{y}_1(t) = \Phi(t)\boldsymbol{\alpha}_0 + \Phi(t) \int_{t_0}^t (\Phi(t'))^{-1}\mathbf{b}(t')dt' \quad (3.96)$$

comme solution générale de l'équation différentielle inhomogène (3.86).

Cette méthode de résolution est particulièrement facile à mettre en application en une dimension, c-à-d, dans le cas particulier d'une équation différentielle à une seule composante,

$$\dot{y}(t) = h(t)y(t) + b(t), \quad (3.97)$$

où  $h, b : I \rightarrow \mathbb{C}$  sont des fonctions continues qui sont définies sur l'intervalle ouvert  $I \subseteq \mathbb{R}$ . Comme nous avons déjà constaté dans la section 3.3 (cf. équation (3.35)), la solution générale de l'équation différentielle homogène associée  $\dot{y}(t) = h(t)y(t)$  est donnée par

$$y(t) = \alpha \exp\left(\int_{t_0}^t h(t') dt'\right) \quad (3.98)$$

avec une constante  $\alpha \in \mathbb{C}$ . En effectuant une variation de cette constante, ce qui revient à remplacer  $\alpha$  par  $\alpha(t)$  dans l'expression (3.98), nous obtenons

$$y_1(t) = \alpha(t) \exp\left(\int_{t_0}^t h(t') dt'\right) \quad (3.99)$$

comme candidate à une solution particulière à l'équation (3.97). En injectant dans l'équation (3.97) la dérivée première de cette fonction  $y_1$  par rapport à  $t$ , nous obtenons la relation

$$\dot{\alpha}(t) = \exp\left(-\int_{t_0}^t h(t') dt'\right) b(t) \quad (3.100)$$

à laquelle la fonction  $\alpha : I \rightarrow \mathbb{C}$  doit satisfaire afin que  $y_1$  soit une solution à l'équation (3.97). L'intégration sur  $t$  de cette dernière équation (3.100) nous donne

$$\alpha(t) = \alpha_0 + \int_{t_0}^t \exp\left(-\int_{t_0}^{t'} h(t'') dt''\right) b(t') dt' \quad (3.101)$$

avec la constante d'intégration  $\alpha_0 \in \mathbb{C}$ . Nous obtenons ainsi

$$y_1(t) = \alpha_0 \exp\left(\int_{t_0}^t h(t') dt'\right) + \int_{t_0}^t \exp\left(\int_{t'}^t h(t'') dt''\right) b(t') dt' \quad (3.102)$$

comme solution générale de l'équation différentielle (3.97).

## Exercices

3.4 Calculer la solution  $y : I \subseteq \mathbb{R} \rightarrow \mathbb{R}, t \mapsto y(t)$  de l'équation différentielle

- (a)  $(1 + t^2)\dot{y} = ty + t + t^3$
- (b)  $(1 - t^2)\dot{y} = 2ty + 1 + t$
- (c)  $(1 - t^2)\dot{y} = y + \sqrt{1 + t^3}(1 - t)$

pour la condition initiale  $y(0) = 1$ . Quel est l'intervalle maximal  $I$  dans lequel on peut définir la solution ?

3.5 Calculer la solution du système d'équations

$$\begin{array}{lll}
 \text{(a)} \quad \begin{cases} \dot{y}_1 = 3y_1 + 5y_2 - 2 \\ \dot{y}_2 = -2y_1 - 3y_2 - 1 \end{cases} & \text{(b)} \quad \begin{cases} \dot{y}_1 = -4y_2 + 2 \\ \dot{y}_2 = y_1 - 4y_2 \end{cases} & \text{(c)} \quad \begin{cases} \dot{y}_1 = 3y_1 + 4y_2 - 5 \\ \dot{y}_2 = -2y_1 - y_2 \end{cases} \\
 \text{(d)} \quad \begin{cases} \dot{y}_1 = 2y_1 + 5y_2 + 2 \\ \dot{y}_2 = -y_1 - 2y_2 + 2 \end{cases} & \text{(e)} \quad \begin{cases} \dot{y}_1 = -y_1 - y_2 + 2 \\ \dot{y}_2 = 5y_1 + y_2 + 2 \end{cases} & \text{(f)} \quad \begin{cases} \dot{y}_1 = 2y_1 + y_2 + 9 \\ \dot{y}_2 = -y_1 + 4y_2 - 9 \end{cases}
 \end{array}$$

pour les conditions initiales  $y_1(0) = y_2(0) = 1$ .

### 3.8 Les équations différentielles de l'ordre $n$

Comme nous l'avons déjà montré à la fin de la section 3.1, une équation différentielle de l'ordre  $n \in \mathbb{N}$  peut être considérée comme un cas particulier d'un système de  $n$  équations différentielles de l'ordre 1. D'un point de vue formel, ce cas a donc déjà été abordé dans les sections précédentes. Néanmoins, il convient toutefois de consacrer aux équations différentielles linéaires de l'ordre  $n > 1$  une section à part, au vu de l'importance dont jouissent surtout les équations différentielles linéaires du deuxième ordre en physique. Considérons donc l'équation différentielle à une composante<sup>1</sup>

$$x^{(n)}(t) + a_{n-1}(t)x^{(n-1)}(t) + \dots + a_1(t)\dot{x}(t) + a_0(t)x(t) = 0 \quad (3.103)$$

avec  $x : I \rightarrow \mathbb{C}$  la fonction recherchée et  $a_j : I \rightarrow \mathbb{C}$  ( $j = 0, \dots, n-1$ ) des fonctions continues qui sont définies sur l'intervalle ouvert  $I \subseteq \mathbb{R}$ . Comme nous l'avons déjà défini dans l'équation (3.11),  $x^{(n)}$  signifie la  $n$ -ième dérivée de la fonction  $x$  par rapport à  $t$ .

L'équation (3.103) peut être écrite sous la forme (3.11) si nous définissons

$$g(x, \dot{x}, \dots, x^{(n-1)}, t) = -a_0(t)x - a_1(t)\dot{x} - \dots - a_{n-1}(t)x^{(n-1)}. \quad (3.104)$$

Suivant la discussion à la fin de la section 3.1, elle est donc parfaitement équivalente à l'équation différentielle du premier ordre à  $n$  composantes

$$\dot{\mathbf{y}}(t) = H(t)\mathbf{y}(t) \quad (3.105)$$

avec le vecteur

$$\mathbf{y}(t) = \begin{pmatrix} x(t) \\ \dot{x}(t) \\ \vdots \\ x^{(n-1)}(t) \end{pmatrix} \in \mathbb{C}^n \quad (3.106)$$

<sup>1</sup>Le cas d'un système de plusieurs équations différentielles linéaires de l'ordre  $n$  peut être traité d'une manière tout à fait analogue. Pour ne pas inutilement compliquer la discussion, nous nous restreignons ici au cas d'une équation à une seule composante.

et la matrice

$$H(t) = \begin{pmatrix} 0 & 1 & & & & & (0) \\ & 0 & 1 & & & & \\ & & \ddots & \ddots & & & \\ (0) & & & \ddots & \ddots & & \\ & & & & 0 & 1 & \\ -a_0(t) & -a_1(t) & \dots & \dots & \dots & -a_{n-1}(t) & \end{pmatrix} \in \mathbb{C}^{n \times n}. \quad (3.107)$$

Suivant la discussion au début de la section 3.4, la solution à l'équation différentielle (3.105) existe et est unique pour tout jeu de conditions initiales  $\mathbf{y}(t_0) = \mathbf{y}_0 \in \mathbb{C}^n$ . Il en est donc de même pour l'équation différentielle (3.103) dont la solution existe et est unique pour tout jeu de valeurs initiales  $x(t_0), \dot{x}(t_0), \ddot{x}(t_0), \dots, x^{(n-1)}(t_0)$  de la fonction recherchée  $x$  ainsi que de ses  $n - 1$  premières dérivées.

L'analyse effectuée dans la section 3.4 nous enseigne aussi que la solution générale de l'équation différentielle (3.103) peut être représentée par la combinaison linéaire

$$x(t) = \sum_{j=1}^n \alpha_j x_j(t) \quad (3.108)$$

avec des coefficients  $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ , où  $\{x_1, \dots, x_n\}$  est un système fondamental de  $n$  solutions à l'équation (3.103) qui sont linéairement indépendantes. Plus précisément, suivant l'utilisation de ce concept dans la section 3.4, l'indépendance linéaire se traduit par la propriété que la matrice

$$\Phi(t) = \begin{pmatrix} x_1(t) & \dots & \dots & x_n(t) \\ \dot{x}_1(t) & \dots & \dots & \dot{x}_n(t) \\ \vdots & & & \vdots \\ x_1^{(n-1)}(t) & \dots & \dots & x_n^{(n-1)}(t) \end{pmatrix} \quad (3.109)$$

est inversible pour tout  $t \in I$  ou, de manière tout à fait équivalente, que son déterminant ne s'annule pas :  $\det \Phi(t) \neq 0$ , pour tout  $t \in I$ . On appelle la fonction  $W : I \rightarrow \mathbb{C}, t \mapsto W(t) = \det \Phi(t)$  le *wronskien* de l'équation différentielle (3.103), nommé d'après le mathématicien franco-polonais WRONSKI.

Dans le cas particulier d'une équation différentielle (3.103) avec des coefficients constants, c-à-d,  $a_j(t) \equiv a_j$  indépendant de  $t$  pour tout  $j = 0, \dots, n - 1$ , nous pouvons explicitement spécifier ce système fondamental et ainsi déterminer la solution générale de cette équation. La section 3.5 nous enseigne que cette solution générale peut s'écrire sous la forme d'une combinaison linéaire des fonctions exponentielles  $e^{\lambda t}$ , munies éventuellement des préfacteurs algébriques  $t^j$  avec  $j \in \mathbb{N}$ , où les exposants sont les valeurs propres de la matrice  $H$  (3.107). Pour calculer ces dernières, nous injectons la fonction

$$x(t) = \alpha e^{\lambda t} \quad (3.110)$$

dans l'équation différentielle (3.103), à savoir

$$x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + \dots + a_1\dot{x}(t) + a_0x(t) = 0 \quad (3.111)$$

dans le cas de coefficients constants. En calculant  $x^{(j)}(t) = \alpha\lambda^j e^{\lambda t}$  pour  $j = 1, \dots, n$ , nous obtenons ainsi l'équation polynômiale

$$\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 = 0 \quad (3.112)$$

à laquelle l'exposant  $\lambda$  doit satisfaire pour que la fonction (3.110) satisfasse à l'équation différentielle (3.111) et dont le membre de gauche correspond au polynôme caractéristique de la matrice  $H$ . Cette équation (3.112) peut être réécrite sous la forme

$$\prod_{j=1}^k [(\lambda - \lambda_j)^{\nu_j}] = 0, \quad (3.113)$$

où  $\lambda_1, \dots, \lambda_k \in \mathbb{C}$  sont les zéros de ce polynôme caractéristique avec leurs multiplicités respectives  $\nu_1, \dots, \nu_k \in \mathbb{N}$  satisfaisant à la relation  $\nu_1 + \dots + \nu_k = n$ . Dans le cas particulier (qui est générique en réalité comme nous l'avons souligné dans la section 3.5) que tous les zéros de ce polynôme sont des zéros simples, c-à-d,  $\nu_j = 1$  pour tout  $j = 1, \dots, k$  et  $k = n$ , nous pouvons choisir un système fondamental pour l'équation (3.111) qui est constitué des fonctions

$$x_j : I \rightarrow \mathbb{C}, t \mapsto x_j(t) = e^{\lambda_j t} \quad (3.114)$$

pour  $j = 1, \dots, n$ .

Dans le cas général (qui est plutôt exceptionnel en pratique) que nous avons affaire à des zéros non simples, on peut constituer un système fondamental par les fonctions

$$x_{j,l} : I \rightarrow \mathbb{C}, t \mapsto x_{j,l}(t) = \frac{t^l}{l!} e^{\lambda_j t} \quad (3.115)$$

pour  $j = 1, \dots, k$  et  $l = 1, \dots, \nu_j$ . Il est bien évident que ces fonctions sont linéairement indépendantes tel que nous avons  $\det \Phi(t) \neq 0$  pour la matrice (3.109). Pour prouver que la fonction  $x_{j,l}$ , définie par l'équation (3.115), satisfait à l'équation différentielle (3.111), il est utile de réécrire la dernière sous la forme

$$\prod_{j=1}^k \left[ \left( \frac{d}{dt} - \lambda_j \right)^{\nu_j} \right] x(t) = 0 \quad (3.116)$$

en parfaite analogie avec l'équation (3.113). Sachant que l'ordre dans lequel les facteurs apparaissent dans ce produit n'a pas d'importance, nous pouvons nous permettre de faire appliquer d'abord le terme correspondant à l'indice  $j$  à la fonction  $x_{j,l}$ . Avec

$$\dot{x}_{j,l}(t) = \frac{lt^{l-1}}{l!} e^{\lambda_j t} + \frac{\lambda_j t^l}{l!} e^{\lambda_j t} \quad (3.117)$$

nous évaluons

$$\left(\frac{d}{dt} - \lambda_j\right) x_{j,l}(t) = \begin{cases} x_{j,l-1}(t) & : l > 0 \\ 0 & : l = 0 \end{cases} \quad (3.118)$$

et donc, par récursion,

$$\left(\frac{d}{dt} - \lambda_j\right)^{\nu_j} x_{j,l}(t) = 0 \quad (3.119)$$

pourvu que  $l < \nu_j$ .

La résolution d'une équation différentielle linéaire de l'ordre  $n$  inhomogène peut être effectuée suivant la stratégie générique que nous avons discutée dans les deux sections précédentes 3.6 et 3.7. La solution générale de l'équation

$$x^{(n)}(t) + a_{n-1}(t)x^{(n-1)}(t) + \dots + a_1(t)\dot{x}(t) + a_0(t)x(t) = f(t) \quad (3.120)$$

pour une fonction continue  $f : I \rightarrow \mathbb{C}, t \mapsto f(t)$  est donnée par la somme de la solution générale de l'équation différentielle homogène associée (3.103) et une solution particulière à l'équation inhomogène (3.120). Dans le cas particulier des coefficients constants, c-à-d,  $a_j(t) \equiv a_j$  pour tout  $j = 0, \dots, n-1$ , et d'une fonction exponentielle

$$f(t) = f_0 e^{\lambda t} \quad (3.121)$$

comme terme inhomogène, avec  $f_0, \lambda \in \mathbb{C}$ , nous pouvons poser cette solution particulière comme

$$x_1(t) = x_0 e^{\lambda t}. \quad (3.122)$$

Son injection dans l'équation différentielle (3.120) donne l'équation

$$(\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0) x_0 e^{\lambda t} = f_0 e^{\lambda t} \quad (3.123)$$

à laquelle  $x_0 \in \mathbb{C}$  doit satisfaire afin que la fonction  $x_1$  représente une solution à l'équation (3.120). Cette équation est facilement résolue pour  $x_0$  par

$$x_0 = \frac{f_0}{\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0}, \quad (3.124)$$

pourvu que  $\lambda$  ne soit pas un zéro du polynôme constituant le dénominateur du membre de droite de cette expression. Dans le cas (plutôt exceptionnel) où  $\lambda$  est un zéro d'ordre  $\nu \in \mathbb{N}$  de ce polynôme, il convient de poser

$$x_1(t) = x_0 t^\nu e^{\lambda t} \quad (3.125)$$

pour obtenir une solution particulière à l'équation (3.120).

**Exercices**

3.6 Déterminer la solution générale réelle  $y(t) \in \mathbb{R}$  de l'équation différentielle

(a)  $y^{(2)}(t) - y^{(1)}(t) - 2y(t) = 1 - e^{-t}$

(b)  $y^{(3)}(t) - 3y^{(2)}(t) + 3y^{(1)}(t) - y(t) = 2e^{2t}$

(c)  $y^{(3)}(t) - 3y^{(1)}(t) - 2y(t) = \cos(t)$

(d)  $y^{(3)}(t) - y^{(2)}(t) + 4y^{(1)}(t) - 4y(t) = e^t$

(e)  $y^{(4)}(t) + 3y^{(2)}(t) - 4y(t) = 3 \sin t$

(f)  $y^{(3)}(t) - 4y^{(2)}(t) + 5y^{(1)}(t) - 2y(t) = e^{-t}$

(g)  $y^{(3)}(t) - 2y^{(2)}(t) + y^{(1)}(t) - 2y(t) = t + e^{-t}$

(h)  $y^{(5)}(t) - 2y^{(4)}(t) + 2y^{(3)}(t) - 2y^{(2)}(t) + y^{(1)}(t) = \cos 2t$

(i)  $y^{(3)}(t) - 2y^{(2)}(t) - 5y^{(1)}(t) + 6y(t) = \cos(t)$

(j)  $y^{(3)}(t) - 3y^{(2)}(t) + 4y^{(1)}(t) - 12y(t) = 1 + e^{-t}$

(k)  $y^{(3)}(t) - 2y^{(2)}(t) - 4y^{(1)}(t) + 8y(t) = \cos(2t)$



# Chapitre 4

## L'algèbre linéaire avec des fonctions

### 4.1 L'espace vectoriel $\mathbb{C}^n$

Avant de discuter les particularités d'un espace vectoriel qui est constitué des fonctions, il convient de révoir les propriétés de base de l'espace  $\mathbb{C}^n$ . Ce dernier est constitué des vecteurs colonnes à  $n$  composantes,

$$\mathbb{C}^n = \left\{ \mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} : v_1, \dots, v_n \in \mathbb{C} \right\}. \quad (4.1)$$

En tant qu'espace vectoriel, il s'agit d'un groupe commutatif muni d'une opération d'addition

$$+ : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}^n, (\mathbf{u}, \mathbf{v}) \mapsto \mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{pmatrix} = \mathbf{v} + \mathbf{u} \quad (4.2)$$

qui admet une multiplication par un scalaire provenant du corps des nombres complexes  $\mathbb{C}$ , notamment via

$$\cdot : \mathbb{C} \times \mathbb{C}^n \rightarrow \mathbb{C}^n, (\alpha, \mathbf{v}) \mapsto \alpha \mathbf{v} = \begin{pmatrix} \alpha v_1 \\ \vdots \\ \alpha v_n \end{pmatrix}, \quad (4.3)$$

de sorte que nous avons  $\alpha \mathbf{u} + \beta \mathbf{v} \in \mathbb{C}^n$  pour tous les vecteurs  $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$  et tous les nombres complexes  $\alpha, \beta \in \mathbb{C}$ .  $\mathbb{C}^n$  est un espace vectoriel à  $n$  dimensions. Une

base  $(\mathbf{e}_1, \dots, \mathbf{e}_n)$  de  $\mathbb{C}^n$  peut être formée par les vecteurs unitaires

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{e}_n = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \quad (4.4)$$

Sur l'espace vectoriel  $\mathbb{C}^n$  nous pouvons définir des transformations linéaires ou *endomorphismes*, c-à-d, des transformations

$$\mathcal{A} : \mathbb{C}^n \rightarrow \mathbb{C}^n, \mathbf{v} \mapsto \mathcal{A}(\mathbf{v}) \quad (4.5)$$

qui satisfont à la propriété

$$\mathcal{A}(\alpha \mathbf{u} + \beta \mathbf{v}) = \alpha \mathcal{A}(\mathbf{u}) + \beta \mathcal{A}(\mathbf{v}) \quad (4.6)$$

pour tous les vecteurs  $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$  et tous les scalaires  $\alpha, \beta \in \mathbb{C}$ . Une telle transformation  $\mathcal{A}$  est appelée *opérateur* dans la suite. Elle peut être représentée par une matrice carrée

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \in \mathbb{C}^{n \times n} \quad (4.7)$$

tel que nous avons

$$\mathcal{A}(\mathbf{v}) = A\mathbf{v} \equiv \begin{pmatrix} a_{11}v_1 + \dots + a_{1n}v_n \\ \vdots \\ a_{n1}v_1 + \dots + a_{nn}v_n \end{pmatrix} \quad (4.8)$$

pour tout  $\mathbf{v} \in \mathbb{C}^n$ .

Une telle matrice carrée  $A$  possède au moins un *vecteur propre*, c-à-d, un vecteur non nul  $\mathbf{v} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$  satisfaisant à la propriété qu'il existe un nombre complexe  $\lambda \in \mathbb{C}$  tel que  $A\mathbf{v} = \lambda\mathbf{v}$ , ce nombre complexe  $\lambda$  étant appelée *valeur propre* de la matrice  $A$ . Cette matrice  $A \in \mathbb{C}^{n \times n}$  est *diagonalisable* si elle possède  $n$  vecteurs propres qui sont linéairement indépendants. Comme nous l'avons vu dans la section 3.5 (cf. équations (3.57)–(3.59)), on peut, dans ce cas, la représenter comme

$$A = BDB^{-1} \quad (4.9)$$

où  $B \in \mathbb{C}^{n \times n}$  est une matrice inversible, contenant ces vecteurs propres dans ses colonnes, et

$$D = \begin{pmatrix} \lambda_1 & & (0) \\ & \ddots & \\ (0) & & \lambda_n \end{pmatrix} \in \mathbb{C}^{n \times n} \quad (4.10)$$

est une matrice diagonale contenant les valeurs propres associées. La matrice  $A$  est forcément diagonalisable si elle est *hermitique*, c-à-d,  $A = A^\dagger$ , où

$$A^\dagger = \begin{pmatrix} a_{11}^* & \dots & a_{n1}^* \\ \vdots & & \vdots \\ a_{1n}^* & \dots & a_{nn}^* \end{pmatrix} \in \mathbb{C}^{n \times n} \quad (4.11)$$

est la *conjuguée hermitienne* de  $A$ . Dans ce cas, ses valeurs propres  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$  sont réelles et ses vecteurs propres associés  $\mathbf{v}_1, \dots, \mathbf{v}_n$  peuvent être choisis tels que la matrice  $B = (\mathbf{v}_1, \dots, \mathbf{v}_n)$  qu'ils forment soit *unitaire*,  $B^{-1} = B^\dagger$ , ce qui permet la représentation  $A = BDB^\dagger$ .

Le concept d'un espace vectoriel ainsi défini est assez abstrait et basique, de la sorte qu'il peut p.ex. être constitué des vecteurs dont les composantes correspondent à des grandeurs physiques différentes (tel que le courant  $I$  et la tension  $U_c$  dans le cas du système des équations différentielles (3.5) et (3.6)) sans que cela soit en contradiction avec les propriétés constitutives de ce concept. En particulier, un tel espace vectoriel n'admet pas forcément des notions géométriques de distance ou d'angle entre deux vecteurs. Pour introduire de telles notions géométriques, il est nécessaire de définir une *norme* sur cet espace vectoriel. De manière générale, une telle norme est introduite par la définition d'un *produit scalaire*. Ce dernier, dans le cas de l'espace vectoriel  $\mathbb{C}^n$ , correspond à une transformation

$$\varphi : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}, (\mathbf{u}, \mathbf{v}) \mapsto \varphi(\mathbf{u}, \mathbf{v}) \quad (4.12)$$

- qui est linéaire par rapport au second argument, c-à-d,

$$\varphi(\mathbf{u}, \alpha\mathbf{v} + \beta\mathbf{w}) = \alpha\varphi(\mathbf{u}, \mathbf{v}) + \beta\varphi(\mathbf{u}, \mathbf{w}) \quad (4.13)$$

pour tous les vecteurs  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$  et tous les scalaires  $\alpha, \beta \in \mathbb{C}$ ,

- qui est *symétrique hermitienne*, c-à-d,

$$\varphi(\mathbf{u}, \mathbf{v}) = [\varphi(\mathbf{v}, \mathbf{u})]^* \quad (4.14)$$

pour tous  $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ , et

- qui est *positive*, c-à-d,

$$\varphi(\mathbf{v}, \mathbf{v}) > 0 \quad (4.15)$$

pour tout  $\mathbf{v} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ , sachant que nous obtenons trivialement  $\varphi(\mathbf{0}, \mathbf{0}) = 0$  par la propriété (4.13) car  $\mathbf{0} = 0 \cdot \mathbf{v}$  pour tout  $\mathbf{v} \in \mathbb{C}^n$ .

La norme du vecteur  $\mathbf{v} \in \mathbb{C}^n$  peut ainsi être définie par

$$\|\mathbf{v}\| = \sqrt{\varphi(\mathbf{v}, \mathbf{v})}. \quad (4.16)$$

Le choix naturel pour un produit scalaire donnant naissance à une norme via cette définition (4.16) est le produit scalaire *canonique*

$$\varphi(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^n u_j^* v_j \equiv \langle \mathbf{u} | \mathbf{v} \rangle. \quad (4.17)$$

Avec ce choix, la norme d'un vecteur  $\mathbf{v}$  correspond à la distance euclidienne du point  $(\operatorname{Re}(v_1), \operatorname{Im}(v_1), \dots, \operatorname{Re}(v_n), \operatorname{Im}(v_n)) \in \mathbb{R}^{2n \times 2n}$  par rapport à l'origine de l'espace  $\mathbb{R}^{2n \times 2n}$ . Ce choix peut s'avérer inapproprié si les composantes  $v_1, \dots, v_n$  du vecteur  $\mathbf{v}$  ne sont pas censées représenter le même poids relatif dans la définition d'une telle notion de distance. Dans ce cas, il convient, d'introduire la norme de l'espace vectoriel  $\mathbb{C}^n$  via le produit scalaire

$$\varphi(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^n w_j u_j^* v_j \quad (4.18)$$

où les coefficients de pondération positifs  $w_1, \dots, w_n \in \mathbb{R}_+$  permettent de tenir compte du poids relatif entre les différentes composantes du vecteur dans la définition de la notion d'une distance. D'une manière encore plus générale, un produit scalaire peut être défini par

$$\varphi(\mathbf{u}, \mathbf{v}) = \sum_{i,j=1}^n a_{ij} u_i^* v_j \equiv \langle \mathbf{u} | A | \mathbf{v} \rangle. \quad (4.19)$$

où

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = A^\dagger \in \mathbb{C}^{n \times n} \quad (4.20)$$

est une matrice hermitique dont les valeurs propres  $\lambda_1, \dots, \lambda_n \in \mathbb{R}_+$  sont strictement positives.

Muni ainsi d'une norme qui est définie par l'intermédiaire d'un tel produit scalaire via la relation (4.16), l'espace vectoriel  $\mathbb{C}^n$  est, par définition, un *espace hermitien*, ce qui est l'équivalent complexe d'un espace euclidien.  $\mathbb{C}^n$  est aussi un *espace de Banach* qui, par définition, signifie un espace vectoriel muni d'une norme par rapport à laquelle toute suite de Cauchy converge au sein de cet espace. Rappelons que la suite des vecteurs  $\{\mathbf{v}^{(l)}\}_{l \in \mathbb{N}}$  est une suite de Cauchy par définition si pour tout  $\epsilon > 0$  il existe un  $N \in \mathbb{N}$  tel que  $\|\mathbf{v}^{(l)} - \mathbf{v}^{(l')}\| < \epsilon$  pour tout  $l, l' > N$ .  $\mathbb{C}^n$  est donc un espace vectoriel *complet* dans le sens que la limite de toute suite de Cauchy dont tous les membres sont issus de  $\mathbb{C}^n$  fait également partie de  $\mathbb{C}^n$ .  $\mathbb{C}^n$  est aussi un *espace de Hilbert* qui, par définition, est un espace de Banach dont la norme est définie par l'intermédiaire d'un produit scalaire.

En pratique, ces notions de l'espace de Banach et de l'espace de Hilbert sont rarement mises en exergue dans le cadre des espaces vectoriels à une dimension

finie vu que dans ce cas la convergence des suites de Cauchy au sein de l'espace en question est une trivialité. Cela est toujours le cas pour l'espace

$$\mathbb{C}^\infty = \left\{ \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \end{pmatrix} \equiv (v_j)_{j \in \mathbb{N}} : v_1, v_2, \dots \in \mathbb{C} \text{ et } \sum_{j=1}^{\infty} |v_j|^2 < \infty \right\} \quad (4.21)$$

pour lequel on peut montrer qu'il est effectivement un espace de Hilbert si on y introduit une norme par l'intermédiaire du produit scalaire

$$\langle \mathbf{u} | \mathbf{v} \rangle = \sum_{j=1}^{\infty} u_j^* v_j. \quad (4.22)$$

Mais la situation est bien moins évidente pour d'autres espaces vectoriels à une dimension infinie, en particulier pour des espaces vectoriels constitués des *fonctions*.

## 4.2 L'espace vectoriel des fonctions

Soit  $I = ]a, b[ \subseteq \mathbb{R}$  un intervalle ouvert dans  $\mathbb{R}$ . Nous définissons par

$$C^\nu = \{ f : I \rightarrow \mathbb{C}, x \mapsto f(x) \text{ } \nu \text{ fois continûment dérivable} \} \quad (4.23)$$

avec  $\nu \in \{0, 1, \dots, \infty\}$  l'espace des fonctions définies sur  $I$  qui sont  $\nu$  fois continûment dérivables.  $C^0$  est l'espace des fonctions continues qui sont définies sur  $I$ , tandis que  $C^\infty$  contient toutes les fonctions analytiques définies sur cet intervalle. Bien évidemment, nous avons  $C^\nu \subset C^{\nu'}$  pour tout  $\nu > \nu'$ .

Il est aisé de montrer que  $C^\nu$  est un espace vectoriel défini sur le corps des nombres complexes. Cet espace est muni de l'opération d'addition

$$+ : C^\nu \times C^\nu \rightarrow C^\nu, (f, g) \mapsto f + g \quad (4.24)$$

définie par  $(f + g)(x) = f(x) + g(x)$  pour tout  $x \in I$  et de la multiplication scalaire

$$\cdot : \mathbb{C} \times C^\nu \rightarrow C^\nu, (\alpha, f) \mapsto \alpha f \quad (4.25)$$

définie par  $(\alpha f)(x) = \alpha f(x)$  pour tout  $x \in I$ . En effet, si les fonctions  $f, g \in C^\nu$  sont  $\nu$  fois continûment dérivables, alors trivialement  $\alpha f + \beta g$  est aussi  $\nu$  fois continûment dérivable pour tous les nombres complexes  $\alpha, \beta \in \mathbb{C}$ . Il est également aisé de montrer que  $C^\nu$  est un espace vectoriel à une dimension infinie. En effet, un ensemble infini  $(\varphi_0, \varphi_1, \varphi_2, \dots) \equiv (\varphi_j)_{j \in \mathbb{N}_0}$  des fonctions définies sur  $I$  qui sont linéairement indépendantes et  $\nu$  fois continûment dérivables pour tout  $\nu \in \{0, 1, \dots, \infty\}$  peut être constitué par les monômes  $\varphi_j : I \rightarrow \mathbb{C}, x \mapsto \varphi_j(x) = x^j$  pour tout  $j \in \mathbb{N}_0$ .

Une norme peut être introduite sur  $C^\nu$  par l'intermédiaire du produit scalaire défini par

$$\langle f|g \rangle = \int_a^b f^*(x)g(x)dx \quad (4.26)$$

pour toutes les fonctions  $f, g \in C^\nu$ . Ce produit scalaire est l'équivalent "continu" du produit scalaire canonique (4.17) que nous avons discuté dans le cadre de l'espace vectoriel  $\mathbb{C}^n$ , et peut effectivement être ramené à celui-là si on approche l'intégrale constituant le membre de droite de l'équation (4.26) par une somme de Riemann. Dans la suite, nous allons donc l'appeler *produit scalaire canonique* pour l'espace vectoriel  $C^\nu$ . La norme d'une fonction  $f \in C^\nu$  est donc définie par

$$\|f\| = \sqrt{\langle f|f \rangle}. \quad (4.27)$$

Cette définition peut être modifiée en

$$\|f\| = \sqrt{\varphi(f, f)} \quad (4.28)$$

où

$$\varphi : C^\nu \times C^\nu \rightarrow \mathbb{C}, (f, g) \mapsto \varphi(f, g) = \int_a^b f^*(x)g(x)w(x)dx \equiv \langle f|w|g \rangle \quad (4.29)$$

représente une généralisation du produit scalaire défini sur  $C^\nu$ . Ce dernier produit scalaire (4.29) admet la possibilité de pondérer différentes régions dans l'intervalle  $I$  de manière différente, notamment au travers d'une fonction réelle et strictement positive

$$w : I \rightarrow \mathbb{R}_+, x \mapsto w(x). \quad (4.30)$$

Il est l'équivalent "continu" du produit scalaire (4.18) que nous avons introduit dans le cadre de l'espace  $\mathbb{C}^n$ .

Un problème qui se pose avec cette définition de la norme est que celle-là ne peut pas être définie pour toutes les fonctions issues de  $C^\nu$ . Considérons p.ex. la fonction

$$f : I = ]a, b[ \rightarrow \mathbb{R}, x \mapsto f(x) = \frac{1}{\sqrt{w(x)(x-a)}} \quad (4.31)$$

qui fait bien partie de  $C^\nu$  pourvu que la fonction de pondération  $w$  soit  $\nu$  fois continûment dérivable. Le produit scalaire de  $f$  avec lui-même est évalué par

$$\varphi(f, f) = \int_a^b w(x)|f(x)|^2 dx = \int_a^b \frac{1}{(x-a)^2} dx = \infty \quad (4.32)$$

suivant la définition (4.29), ce qui signifie que la norme de  $f$  diverge et ne peut donc pas être définie. Un autre exemple est fourni par la fonction

$$f : I = ]-\infty, \infty[ \rightarrow \mathbb{R}, x \mapsto f(x) = \frac{1}{\sqrt{w(x)}} \quad (4.33)$$

dont la norme diverge également.

Ce problème peut être résolu en restreignant artificiellement l'espace  $C^\nu$  à des fonctions qui ont une norme finie par rapport à la définition (4.29) du produit scalaire  $\varphi$ . Introduisons donc l'espace

$$\mathcal{C}^\nu = \{f \in C^\nu : \|f\| = \sqrt{\varphi(f, f)} < \infty\}. \quad (4.34)$$

Comme la norme de toute combinaison linéaire  $\alpha f + \beta g \in C^\nu$  pour  $\alpha, \beta \in \mathbb{C}$  est finie si les normes des fonctions  $f, g \in C^\nu$  sont finies, il est aisé de montrer que  $\mathcal{C}^\nu$  est un sous-espace vectoriel de  $C^\nu$ . On l'appelle le sous-espace des fonctions *de carré intégrables*, aussi nommées *fonctions  $L^2$* , dans le cas particulier du produit scalaire canonique où nous avons  $w(x) = 1$  pour tout  $x \in I$ .

Un autre problème qui est moins facile à résoudre est que par rapport à la définition (4.28) de la norme, ni  $C^\nu$  ni son sous-espace des fonctions de carré intégrables  $\mathcal{C}^\nu$  ne sont complets, c-à-d, ne représentent des espaces de Banach. En effet, il est facile de construire des suites de fonctions issues de  $C^\nu$  qui sont des suites de Cauchy selon la norme (4.28) mais ne convergent pas au sein de  $C^\nu$ , dans le sens que leurs limites ne font pas partie de l'espace  $C^\nu$ . Considérons p.ex. la suite  $(g_n)_{n \in \mathbb{N}}$  des fonctions

$$g_n : I = \mathbb{R} \rightarrow \mathbb{R}, x \mapsto g_n(x) = \frac{1}{1 + x^{2n}} \quad (4.35)$$

dont chacune est analytique, impliquant ainsi  $g_n \in C^\infty \subset C^\nu$  pour tout  $\nu \in \mathbb{N}_0$  et tout  $n \in \mathbb{N}$ , et dont chacune est aussi de carré intégrable par rapport à la norme (4.27) qui est définie par l'intermédiaire du produit scalaire canonique. En majorant le module de la différence  $g_n(x) - g_{n'}(x)$  pour tout  $x \in \mathbb{R}$  et tous  $n, n' > N$  au delà d'un entier  $N \in \mathbb{N}$  suffisamment grand, ce qui peut être fait par

$$|g_n(x) - g_{n'}(x)| < \begin{cases} 1 - g_N(x) & : |x| < 1 \\ g_N(x) & : |x| > 1 \end{cases}, \quad (4.36)$$

il est aisé de montrer que  $(g_n)_{n \in \mathbb{N}}$  est une suite de Cauchy par rapport à cette norme. Cependant, la limite

$$\lim_{n \rightarrow \infty} g_n : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \lim_{n \rightarrow \infty} g_n(x) = \begin{cases} 1 & : |x| < 1 \\ 0 & : |x| > 1 \end{cases} \quad (4.37)$$

ne fait partie d'aucun espace  $C^\nu$  pour  $\nu \geq 0$  car cette fonction n'est pas continue en  $x = \pm 1$ .

Un autre exemple qui est encore plus frappant est fourni par la suite  $(h_n)_{n \in \mathbb{N}}$  des fonctions

$$h_n : I = \mathbb{R} \rightarrow \mathbb{R}, x \mapsto h_n(x) = n^{1/8} e^{-nx^2} \quad (4.38)$$

qui sont toutes parfaitement analytiques et de carré intégrables par rapport à la norme (4.27). Comme nous avons avec cette norme  $\|h_n\| = (\pi/2)^{1/4} n^{-1/8}$  et donc

$\lim_{n \rightarrow \infty} \|h_n\| = 0$ ,  $(h_n)_{n \in \mathbb{N}}$  est une suite de Cauchy. En revanche, nous évaluons  $\lim_{n \rightarrow \infty} h_n(0) = \lim_{n \rightarrow \infty} n^{1/8} = \infty$ , ce qui signifie que la limite de cette suite de Cauchy diverge à l'origine et ne peut donc pas du tout y être définie.

Une solution pour pallier ce problème d'incomplétude consiste à tout simplement “boucher” de tels trous que l'espace  $\mathcal{C}^\nu$  présente. En pratique, cette approche revient à définir la *complétion*  $\overline{\mathcal{C}^\nu}$  de l'espace  $\mathcal{C}^\nu$ . Par définition, cette complétion contient toutes les fonctions faisant partie de  $\mathcal{C}^\nu$  ainsi que toute autre fonction qui peut être formée par la limite d'une suite de Cauchy qui est constituée des fonctions issues de  $\mathcal{C}^\nu$ . Notons que de telles fonctions “limites” qui à la base ne font pas partie de  $\mathcal{C}^\nu$  sont largement majoritaires dans  $\overline{\mathcal{C}^\nu}$  par rapport aux éléments de  $\mathcal{C}^\nu$ , de la même manière que les nombres irrationnels sont largement majoritaires dans l'ensemble  $\mathbb{R}$  par rapport aux nombres rationnels. Notons aussi que la notion de dérivabilité dont est caractérisé l'espace  $\mathcal{C}^\nu$  par sa définition (4.1) perd toute sa signification lorsque l'on forme la complétion de cet espace, comme nous avons vu dans l'exemple (4.35) de la suite des fonctions  $g_n$ .

En abandonnant donc toute ambition de maintenir une condition de continuité ou dérivabilité dans un espace vectoriel des fonctions, nous sommes donc finalement en mesure de définir celui-ci en tant qu'espace de Hilbert, notamment par

$$\mathcal{H} = \{f : I \rightarrow \mathbb{C} : \|f\| = \sqrt{\varphi(f, f)} < \infty\}. \quad (4.39)$$

Cet espace contient toute fonction définie sur  $I$  dont la norme, définie par l'intermédiaire du produit scalaire (4.29), est finie. Cette définition d'un espace de Hilbert fait émerger un autre problème qui n'existe pas lorsque les membres de l'espace vectoriel des fonctions sont contraints de satisfaire à une condition de continuité, notamment concernant la discernabilité de deux fonctions  $f, g \in \mathcal{H}$  non identiques qui atteignent des mêmes valeurs presque partout en  $I$  : p.ex.  $f(x) = g(x)$  pour tout  $x \in I \setminus \{x_0\}$  avec  $x_0 \in I$ . Comme nous calculons

$$\|f - g\|^2 = \int_a^b |f(x) - g(x)|^2 w(x) dx = 0 \quad (4.40)$$

dans ce cas, la différence  $f - g$  de ces deux fonctions a la norme zéro, ce qui contredit le principe de positivité (4.15) auquel le produit scalaire doit satisfaire pour tout vecteur non nul.

Ce dernier problème peut être pallié en définissant l'espace de Hilbert (4.39) non pas en termes de fonctions à proprement parler, mais en termes de *classes d'équivalence* de fonctions. Les éléments de cet espace vectoriel sont donc des classes d'équivalence de fonctions où deux fonctions  $f, g : I \rightarrow \mathbb{C}$  qui sont de carré intégrables sont considérées comme étant équivalentes,  $f \sim g$ , si l'ensemble  $\{x \in I : f(x) \neq g(x)\}$  est de mesure nulle tel que nous avons la propriété (4.40). Le vecteur nul de cet espace de Hilbert est donc constitué par la classe d'équivalence des fonctions  $f$  satisfaisant à la propriété  $\|f\| = 0$ . Cette classe contient aussi la limite de la suite de Cauchy des fonctions  $h_n$  définies dans

l'équation (4.38), même si cette fonction limite ne peut pas du tout être définie en  $x = 0$ .

A ce stade, on peut se poser la question de savoir comment ça marche pour la fonction delta discutée dans la section 2.1 qui, par l'équation (2.4), satisfait aussi à la propriété  $\delta(x) = 0$  pour tout  $x \in I \setminus \{0\}$  (supposant que  $0 \in I$ ) sans que nous ayons  $\|\delta\| = 0$ . La réponse à cette question est que  $\delta$  ne fait pas partie de l'espace de Hilbert  $\mathcal{H}$ . En effet, on peut approcher cette fonction  $\delta$  par une suite  $(\delta_n)_{n \in \mathbb{N}}$  des fonctions

$$\delta_n : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \delta_n(x) = n\varphi(nx) \quad (4.41)$$

définies en termes d'une fonction générique  $\varphi : \mathbb{R} \rightarrow \mathbb{R}, \xi \mapsto \varphi(\xi)$  qui est de carré intégrable, tel que nous avons

$$\delta(x) = \lim_{n \rightarrow \infty} \delta_n(x) \quad (4.42)$$

en parfaite analogie avec l'équation (2.10). Mais une telle suite ne sera jamais une suite de Cauchy par rapport à la norme (4.28), car nous avons

$$\lim_{n \rightarrow \infty} \|\delta_n - \delta_{n'}\|^2 = \lim_{n \rightarrow \infty} \int_a^b (\delta_n(x) - \delta_{n'}(x))^2 dx = \infty \quad (4.43)$$

pour tout  $n' \in \mathbb{N}$ .

Notons finalement que le problème d'incomplétude de l'espace  $C^\nu$  peut aussi être résolu différemment, notamment en changeant la définition de la norme en

$$\|f\| = \sup_{x \in I} |f(x)| \quad (4.44)$$

pour tout  $f \in C^\nu$ . L'espace  $C^\nu = \{f \in C^\nu : \|f\| < \infty\}$  devient ainsi complet, et les suites des fonctions  $(g_n)_{n \in \mathbb{N}}$  et  $(h_n)_{n \in \mathbb{N}}$  que nous avons respectivement définies dans les équations (4.35) et (4.38) ne sont plus des suites de Cauchy par rapport à cette nouvelle norme. Cependant, bien que cet espace  $C^\nu$  représente désormais un espace de Banach, il ne s'agit pas d'espace de Hilbert car pour ce dernier la norme doit être définie par l'intermédiaire d'un produit scalaire.

### 4.3 Les opérateurs dans l'espace des fonctions

Comme nous avons introduit ce concept dans la section 4.1, un *opérateur* dans un espace vectoriel est un endomorphisme dans cet espace, c-à-d, une transformation linéaire qui est définie sur cet espace et a son image dans cet espace. Dans le cas de l'espace  $C^\nu$ , contenant, selon la définition (4.23), des fonctions définies sur l'intervalle  $I = ]a, b[ \subseteq \mathbb{R}$  qui sont  $\nu$  fois continûment dérivables, toute transformation  $\hat{A} : C^\nu \rightarrow C^\nu, f \mapsto \hat{A}f$  qui satisfait à la propriété de linéarité, à savoir

$$\hat{A}(\alpha f + \beta g) = \alpha \hat{A}f + \beta \hat{A}g \quad (4.45)$$

pour toutes les fonctions  $f, g \in C^\nu$  et tous les nombres complexes  $\alpha, \beta \in \mathbb{C}$ , est un opérateur.

Pour simplifier la discussion, considérons d'abord le cas  $\nu = \infty$  des fonctions analytiques définies sur l'intervalle  $I$ . En s'inspirant de la mécanique quantique, nous pouvons introduire sur  $C^\infty$  les opérateurs de *position* et *impulsion*

$$\hat{x} : C^\infty \rightarrow C^\infty, f \mapsto \hat{x}f, \quad (4.46)$$

$$\hat{p} : C^\infty \rightarrow C^\infty, f \mapsto \hat{p}f \quad (4.47)$$

qui sont définis tels que nous avons

$$(\hat{x}f)(x) = xf(x), \quad (4.48)$$

$$(\hat{p}f)(x) = -if'(x) \quad (4.49)$$

pour tout  $x \in I$ . Nous pouvons y introduire aussi tout autre opérateur formé par une combinaison additive ou multiplicative de  $\hat{x}$  ou  $\hat{p}$ , sachant que l'ensemble des opérateurs définis sur  $C^\nu$  forme un anneau, comme dans le cas des matrices carrées issues de  $\mathbb{C}^{n \times n}$ . Les opérateurs ainsi formés sont *locaux* dans le sens que l'évaluation en  $x$  de l'image d'une fonction sous un tel opérateur ne dépend que des valeurs de la fonction d'origine en  $x$  et dans son voisinage immédiat.

Des opérateurs non locaux peuvent également être introduits sur  $C^\infty$ . Un exemple est l'opérateur  $\hat{P} : C^\infty \rightarrow C^\infty, f \mapsto \hat{P}f$  qui effectue une opération de miroir par rapport au centre de l'intervalle  $I$ , définie tel que nous avons

$$(\hat{P}f)(x) = f(a + b - x) \quad (4.50)$$

pour tout  $x \in I$ . La convolution  $\hat{F}_g : f \mapsto \hat{F}_g f$  avec la fonction analytique  $g : \mathbb{R} \rightarrow \mathbb{C}$ , définie tel que nous avons

$$(\hat{F}_g f)(x) = \int_a^b f(x')g(x - x')dx' \quad (4.51)$$

pour tout  $x \in I$ , représente un autre exemple d'opérateur non local, mais ce dernier ne peut généralement pas être défini pour toute fonction analytique  $f \in C^\infty$  car l'intégrale constituant le membre de droite de l'équation (4.51) n'existe pas forcément pour toute fonction  $f \in C^\infty$ . Il en est de même pour la transformation de Fourier  $f \mapsto \hat{f}$ , définie dans l'équation (2.40), qui correspond aussi à un opérateur non local.

Le problème qu'un opérateur ne puisse pas être défini pour toute fonction issue de l'espace vectoriel en question est encore accentué lorsque l'on considère non pas l'espace  $C^\infty$  mais l'espace de Hilbert  $\mathcal{H}$  contenant, selon sa définition (4.39), toutes les fonctions définies sur l'intervalle  $I$  qui sont de carré intégrables<sup>1</sup>. Dans

---

<sup>1</sup>En réalité, comme nous l'avons souligné dans la section précédente, il ne s'agit pas des fonctions à proprement parler, mais des classes d'équivalence des fonctions qui atteignent des valeurs identiques partout en  $I$  à l'exception d'un ensemble de mesure nulle. Dans la suite, nous n'allons pas insister sur cette subtilité.

l'espace de Hilbert  $\mathcal{H}$ , l'opérateur d'impulsion  $\hat{p}$ , défini par l'équation (4.49), ne peut pas être appliqué à toute fonction  $f \in \mathcal{H}$  car la dérivée première n'existe pas pour des fonctions  $f \notin C^1$  qui ne sont pas continûment dérivables. Un problème d'un autre genre se pose pour l'opérateur de position  $\hat{x}$ , défini par l'équation (4.48). Celui-ci peut effectivement être appliqué à toute fonction  $f \in \mathcal{H}$ , mais l'image  $\hat{x}f$  n'est pas forcément contenue dans  $\mathcal{H}$ . Considérons p.ex. l'espace de Hilbert défini sur l'intervalle  $I = \mathbb{R}$  où la norme est définie par l'intermédiaire du produit scalaire canonique (4.26). La fonction

$$f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) = \frac{1}{\sqrt{1+x^2}} \quad (4.52)$$

est bien de carré intégrable et fait donc partie de l'espace  $\mathcal{H}$ . Mais cela n'est plus le cas pour  $\hat{x}f$  défini par

$$(\hat{x}f)(x) = \frac{x}{\sqrt{1+x^2}} \quad (4.53)$$

pour tout  $x \in \mathbb{R}$ .  $\hat{x}$  est donc un opérateur *non borné* dans cet espace de Hilbert, dans le sens que le rapport  $\|\hat{x}f\|/\|f\|$  ne peut pas être limité par une borne supérieure dans  $\mathcal{H}$ .

Pour pallier à ces complications qui surgissent dans le cas de l'espace de Hilbert  $\mathcal{H}$ , nous pouvons décider de restreindre la région de définition d'un opérateur  $\hat{A}$  à un sous-ensemble  $\mathcal{D} \subseteq \mathcal{H}$  qui est constitué de toutes les fonctions  $f$  auxquelles on peut effectivement appliquer  $\hat{A}$  et dont l'image  $\hat{A}f$  fait toujours partie de l'espace de Hilbert  $\mathcal{H}$ . Comme toute combinaison linéaire  $\alpha f + \beta g$  avec  $\alpha, \beta \in \mathbb{C}$  fait aussi partie de ce sous-ensemble  $\mathcal{D}$  si  $f$  et  $g$  en font partie,  $\mathcal{D}$  est effectivement un sous-espace vectoriel de  $\mathcal{H}$ . Pour des opérateurs qui ont une importance en physique et ne présentent pas des propriétés pathologiques, le sous-espace  $\mathcal{D}$  est normalement *dense* en  $\mathcal{H}$ , dans le sens que toute fonction  $f \in \mathcal{H}$  peut être approchée par une suite de Cauchy  $(f_n)_{n \in \mathbb{N}}$  constituée des fonctions  $f_n \in \mathcal{D}$  qui sont issues de ce sous-espace. Dans le cas de l'opérateur de position  $\hat{x}$ , la fonction définie dans l'équation (4.52) p.ex. peut être approchée par la suite des fonctions

$$f_n : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f_n(x) = \frac{\exp(-x^2/n)}{\sqrt{1+x^2}} \quad (4.54)$$

qui font toutes partie du sous-espace  $\mathcal{D}$  associé à  $\hat{x}$ . Comme l'espace  $\mathcal{C}^\infty$  des fonctions analytiques définies sur  $I$  qui sont de carré intégrables est, par la construction de cet espace de Hilbert, dense en  $\mathcal{H}$ , il en est de même pour le sous-espace  $\mathcal{D}$  associé à l'opérateur d'impulsion  $\hat{p}$ .

Supposons qu'on peut effectivement identifier comme région de définition pour l'opérateur  $\hat{A}$  un sous-espace  $\mathcal{D} \subseteq \mathcal{H}$  qui est dense en  $\mathcal{H}$  et pour lequel l'image de toute fonction issue de  $\mathcal{D}$  sous l'application de cet opérateur se trouve dans  $\mathcal{H}$ . Dans ce cas, on peut montrer que cet opérateur  $\hat{A} : \mathcal{D} \rightarrow \mathcal{H}$  possède un opérateur

adjoint  $\hat{A}^\dagger : \mathcal{D}^\dagger \rightarrow \mathcal{H}$  qui peut aussi être défini sur un sous-espace  $\mathcal{D}^\dagger \subseteq \mathcal{H}$  dense en  $\mathcal{H}$  tel que les images des fonctions issues de  $\mathcal{D}^\dagger$  sous l'application de  $\hat{A}^\dagger$  se trouvent dans  $\mathcal{H}$ . Par définition, cet opérateur adjoint satisfait à la propriété

$$\varphi(g, \hat{A}f) = \varphi(\hat{A}^\dagger g, f) \quad (4.55)$$

pour tout  $f \in \mathcal{D}$  et tout  $g \in \mathcal{D}^\dagger$ , où  $\varphi$  représente le produit scalaire (4.29) dont l'espace de Hilbert est muni. Dans le cas particulier où  $\hat{A}^\dagger$  est identique à  $\hat{A}$  (et donc trivialement  $\mathcal{D}^\dagger = \mathcal{D}$ ), on appelle cet opérateur *auto-adjoint*.

L'opérateur de position  $\hat{x}$  est effectivement un opérateur auto-adjoint car par la définition (4.48) de celui-ci nous avons

$$\int_a^b g^*(x)(\hat{x}f)(x)w(x)dx = \int_a^b (\hat{x}g)^*(x)f(x)w(x)dx \quad (4.56)$$

pour tout couple  $f, g \in \mathcal{D} \subset \mathcal{H}$ , indépendamment du choix de la fonction de pondération  $w$ . Quant à l'opérateur d'impulsion  $\hat{p}$ , celui-ci est auto-adjoint dans un espace de Hilbert qui est défini sur l'intervalle  $I = \mathbb{R}$  et muni du produit scalaire canonique caractérisé par  $w(x) = 1$  pour tout  $x \in \mathbb{R}$ . En effet, par la méthode de l'intégration par partie, nous calculons

$$\begin{aligned} \int_{-\infty}^{\infty} g^*(x)(\hat{p}f)(x)dx &= -i \int_{-\infty}^{\infty} g^*(x) \frac{d}{dx} f(x) dx \\ &= -i [g^*(x)f(x)]_{-\infty}^{\infty} + i \int_{-\infty}^{\infty} f(x) \frac{d}{dx} g^*(x) dx \\ &= \int_{-\infty}^{\infty} (\hat{p}g)^*(x)f(x)dx, \end{aligned} \quad (4.57)$$

sachant que les fonctions  $f$  et  $g$  sont contraintes de tendre vers zéro si leur argument tend vers  $\pm\infty$ , faute de quoi elles ne seraient pas de carré intégrables et ne feraient donc pas partie de l'espace de Hilbert  $\mathcal{H}$ .

## 4.4 Conditions de bord

Le raisonnement développé ci-dessus nous enseigne aussi que l'opérateur d'impulsion ne peut pas être auto-adjoint dans un espace de Hilbert  $\mathcal{H}$  qui est défini sur un intervalle  $I = ]a, b[ \subset \mathbb{R}$  dont au moins une borne ( $a$  ou  $b$ ) est finie, car dans ce cas-là  $[g^*(x)f(x)]_a^b$  ne s'annule pas forcément pour tout couple  $f, g \in \mathcal{H}$ . Il est effectivement possible de remédier à cette situation, notamment en fixant des *conditions de bord homogènes* aux bornes non infinies de l'intervalle  $I$ . Cela revient à introduire un sous-espace vectoriel de  $\mathcal{H}$  contenant toutes les fonctions définies sur l'intervalle  $I$  qui sont de carré intégrables et satisfont à ces conditions de bord.

Concrètement, dans le cas d'un intervalle fini,  $I = ]a, b[$  avec  $-\infty < a < b < \infty$ , et dans l'optique de rendre l'opérateur d'impulsion auto-adjoint, nous pouvons exiger que toute fonction faisant partie de l'espace de Hilbert s'annule aux bornes  $a$  et  $b$  de l'intervalle. Cet espace de Hilbert est ainsi défini comme

$$\mathcal{H} = \{f : I \rightarrow \mathbb{C} : \|f\| < \infty \text{ et } f(a) = f(b) = 0\}. \quad (4.58)$$

De telles conditions de bord sont nommées d'après le mathématicien DIRICHLET.

Une autre manière d'introduire des conditions de bord homogènes consiste à imposer que la dérivée première des fonctions faisant partie de l'espace de Hilbert s'annule aux bornes de l'intervalle. De telles conditions de bord sont nommées d'après le mathématicien NEUMANN. Nous obtenons ainsi l'espace de Hilbert

$$\mathcal{H} = \{f : I \rightarrow \mathbb{C} : \|f\| < \infty \text{ et } f'(a) = f'(b) = 0\}. \quad (4.59)$$

Plus précisément, comme toute fonction issue de l'espace de Hilbert n'est pas forcément dérivable, nous devons d'abord exiger ces conditions de bord de Neumann à tout membre du sous-espace  $\mathcal{C}^1$  des fonctions dérivables. Comme celui-ci est dense en  $\mathcal{H}$ , nous pouvons ensuite constituer l'espace de Hilbert (4.59) par l'union de toutes les limites des suites de Cauchy que l'on peut former dans  $\mathcal{C}^1$  en respectant ces conditions de bord.

Un autre exemple qui a une certaine importance en physique, notamment dans le contexte de la physique des solides, est celui des conditions de bord *périodiques*, nommées d'après les mathématiciens et physiciens BORN et VON KÁRMÁN. Dans cet exemple, nous exigeons que toute fonction faisant partie de l'espace de Hilbert ait la propriété que ses valeurs ainsi que les valeurs de ses dérivées supérieures soient identiques aux deux bornes de l'intervalle  $I$ . Nous avons ainsi affaire à l'espace de Hilbert

$$\mathcal{H} = \{f : I \rightarrow \mathbb{C} : \|f\| < \infty \text{ et } f^{(n)}(a) = f^{(n)}(b) = 0 \text{ pour tout } n \in \mathbb{N}_0\} \quad (4.60)$$

où  $f^{(n)}(x)$  signifie la  $n$ -ième dérivée de  $f$  par rapport à  $x$ , avec  $f^{(0)} \equiv f$ . De manière analogue à l'exemple précédent, cette condition doit d'abord être exigée pour toute fonction analytique  $f \in \mathcal{C}^\infty$  qui est de carré intégrable sur l'intervalle  $I$ , avant de constituer l'espace de Hilbert (4.60) par l'union des suites de Cauchy qui sont issues de ce sous-espace ainsi formé.

Comme on peut facilement imaginer, il existe une multitude d'autres conditions de bord que l'on peut imposer, p.ex. des conditions de bord mixtes du genre  $f(a) = f'(b) = 0$ . Il est cependant nécessaire que de telles conditions de bord respectent le principe de l'homogénéité, de sorte que l'on puisse les formuler par un jeu d'équations de la forme  $\sum_n (\alpha_n f^{(n)}(a) + \beta_n f^{(n)}(b)) = 0$  avec  $\alpha_n, \beta_n \in \mathbb{C}$  pour tout  $n \in \mathbb{N}_0$ . Considérons, comme contreexemple, l'espace

$$\mathcal{H} = \{f : I \rightarrow \mathbb{C} : \|f\| < \infty \text{ et } f(a) = 1, f(b) = 2\}. \quad (4.61)$$

Comme pour tout  $f \in \mathcal{H}$  nous avons  $2f(a) = 2$  et  $2f(b) = 4$ , la fonction  $2f$  ne fait pas partie de  $\mathcal{H}$  si  $f$  en fait partie. Cet espace (4.61) ne représente donc pas un espace vectoriel.

Notons finalement que l'espace de Hilbert

$$\mathcal{H} = \{f : \mathbb{R} \rightarrow \mathbb{C} : \|f\| = \sqrt{\langle f|f \rangle} < \infty\} \quad (4.62)$$

qui est défini sur l'intervalle  $I = ]-\infty, \infty[$  et muni du produit scalaire canonique présente, de manière implicite, aussi des conditions de bord, notamment celles de Dirichlet

$$\lim_{x \rightarrow -\infty} f(x) = \lim_{x \rightarrow \infty} f(x) = 0. \quad (4.63)$$

En effet, toute fonction  $f : \mathbb{R} \rightarrow \mathbb{C}$  qui est de carré intégrable sur  $\mathbb{R}$  doit satisfaire à ces conditions (4.63), faute de quoi nous avons  $\langle f|f \rangle = \int_{-\infty}^{\infty} f(x)dx = \infty$ .

## 4.5 Les polynômes orthogonaux

Nous avons déjà mentionné dans la section 4.2 que les monômes

$$\varphi_n : I \rightarrow \mathbb{C}, x \mapsto \varphi_n(x) = x^n \quad (4.64)$$

pour  $n \in \mathbb{N}_0$  forment un ensemble de fonctions analytiques qui sont linéairement indépendantes. Ils se prêtent donc particulièrement bien pour construire une base de l'espace de Hilbert

$$\mathcal{H} = \{f : I \rightarrow \mathbb{C} : \|f\| < \infty\}, \quad (4.65)$$

pourvu qu'ils en fassent partie, c-à-d, pourvu que leur norme

$$\|f\| = \sqrt{\varphi(f, f)}, \quad (4.66)$$

définie par l'intermédiaire du produit scalaire

$$\varphi : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}, (f, g) \mapsto \varphi(f, g) = \int_a^b f^*(x)g(x)w(x)dx, \quad (4.67)$$

soit finie. Cela est effectivement le cas si l'intervalle  $I = ]a, b[ \subseteq \mathbb{R}$  sur lequel l'espace de Hilbert est défini et la fonction de pondération  $w : I \rightarrow \mathbb{R}_+, x \mapsto w(x)$  qui caractérise le produit scalaire (4.67) sont tels que l'intégrale  $\int_a^b x^n w(x)dx$  existe pour tout  $n \in \mathbb{N}_0$ . Afin de construire avec ces monômes une base *orthonormée*  $(p_n)_{n \in \mathbb{N}_0} \equiv (p_0, p_1, \dots)$ , c-à-d, une base dont les membres satisfont à la condition

$$\varphi(p_n, p_{n'}) = \delta_{nn'} \quad (4.68)$$

pour tout  $n, n' \in \mathbb{N}_0$ , nous pouvons appliquer le *procédé de Gram-Schmidt* qui permet de successivement orthogonaliser les membres de la base. Ces derniers se transforment ainsi en *polynômes*

$$p_n : I \rightarrow \mathbb{R}, x \mapsto p_n(x) = \sum_{j=0}^n a_{nj}x^j \quad (4.69)$$

qui sont mutuellement orthogonaux par rapport au produit scalaire (4.67).

Concrètement, ce procédé de Gram-Schmidt peut être réalisé de manière inductive. Nous commençons avec le monôme de l'ordre zéro

$$\mathcal{P}_0 : I \rightarrow \mathbb{R}, x \mapsto \mathcal{P}_0(x) = \varphi_0(x) = 1 \quad (4.70)$$

qui, après une multiplication par un préfacteur afin de satisfaire à la condition (4.68) pour  $n = n' = 0$ , permet de constituer le premier membre de la base par la fonction constante

$$p_0(x) = \|\mathcal{P}_0\|^{-1} = \left( \int_a^b w(x)dx \right)^{-1/2}. \quad (4.71)$$

Supposons maintenant que nous avons construit  $n$  membres  $p_0, p_1, \dots, p_{n-1}$  de cette base en utilisant les monômes  $\varphi_0, \varphi_1, \dots, \varphi_{n-1}$ , où  $p_j$  est un polynôme de degré  $j$  comme indiqué dans l'équation (4.69). A un préfacteur global près, un polynôme réel de degré  $n$  peut être posé de manière générale comme

$$\mathcal{P}_n : I \rightarrow \mathbb{R}, x \mapsto \mathcal{P}_n(x) = x^n + \sum_{j=0}^{n-1} B_{n,j}p_j(x) \quad (4.72)$$

avec des coefficients  $B_{n,0}, \dots, B_{n,n-1} \in \mathbb{R}$ . Ce polynôme est rendu orthogonal à tous les membres de base déjà constitués si on choisit

$$B_{n,j} = -\varphi(p_j, \varphi_n) = - \int_a^b x^n p_j(x) w(x) dx \quad (4.73)$$

pour tout  $j = 0, \dots, n-1$ . En respectant la condition de normalisation  $\varphi(p_n, p_n) = 1$ , le membre suivant de la base est donc obtenu par

$$p_n(x) = \frac{1}{\|\mathcal{P}_n\|} \mathcal{P}_n(x). \quad (4.74)$$

La base orthonormée  $(p_0, p_1, \dots)$  étant ainsi construite, nous pouvons affirmer que tout membre  $f \in \mathcal{H}$  de l'espace de Hilbert peut être représenté par la combinaison linéaire

$$f = \sum_{n=0}^{\infty} f_n p_n \quad (4.75)$$

des polynômes orthogonaux  $p_n$  où les coefficients sont, de manière unique, donnés par

$$f_n = \int_a^b p_n(x) f(x) w(x) dx. \quad (4.76)$$

Cette affirmation est une variante du théorème de *Stone-Weierstrass* qui stipule que toute fonction continue  $f : [a, b] \rightarrow \mathbb{C}$  qui est définie sur un intervalle fermé  $[a, b]$  peut être approchée de manière uniforme par une suite  $(f^{(n)})_{n \in \mathbb{N}_0}$  de fonctions polynômiales étant de degré successivement croissant, de sorte que nous pouvons effectivement les écrire comme  $f^{(n)} = \sum_{j=0}^n f_j p_j$  avec des coefficients  $f_j \in \mathbb{C}$ .

Bien évidemment, la représentation (4.75) est aussi valable si  $f$  est une autre fonction polynomiale. Nous pouvons donc affirmer que tout polynôme de degré  $n$

$$q : I \rightarrow \mathbb{C}, x \mapsto q(x) = \sum_{j=0}^n \alpha_j x^j \quad (4.77)$$

avec  $\alpha_1, \dots, \alpha_n \in \mathbb{C}$  peut être représenté par la combinaison linéaire

$$q(x) = \sum_{j=0}^n q_j p_j(x) \quad (4.78)$$

dont les coefficients sont de manière unique donnés par

$$q_j = \int_a^b p_j(x) q(x) w(x) dx. \quad (4.79)$$

Cette dernière affirmation peut être facilement montrée en inversant les décompositions (4.69) des polynômes orthogonaux en monômes, permettant ainsi d'exprimer de manière unique le monôme  $\varphi_n$  en tant que combinaison linéaire des polynômes  $p_0, \dots, p_n$  pour tout  $n \in \mathbb{N}_0$ . Elle implique aussi, par la relation d'orthogonalité (4.68), que tout polynôme  $q$  de degré  $n$  est orthogonal, par rapport à la définition (4.67) du produit scalaire, à tout polynôme  $p_k$  avec  $k > n$  :

$$\varphi(p_k, q) = \int_a^b p_k(x) q(x) w(x) dx = 0. \quad (4.80)$$

Il s'ensuit que les polynômes orthogonaux non normalisés  $\mathcal{P}_n$ , définis par l'expression (4.72) avec les coefficients (4.73), satisfont à une relation de récursion de la forme

$$\mathcal{P}_{n+1}(x) = (x - a_n) \mathcal{P}_n(x) - b_{n-1} \mathcal{P}_{n-1}(x) \quad (4.81)$$

pour tout  $n \in \mathbb{N}_0$ , avec des coefficients  $a_n \in \mathbb{R}$  et  $b_n \in \mathbb{R}_+$  pour tout  $n \in \mathbb{N}_0$ , où  $b_{-1} = 0$ . Pour prouver cette relation de récursion, notons d'abord que les polynômes orthogonaux satisfont à la propriété

$$\int_a^b x p_n(x) p_k(x) w(x) dx = 0 \quad (4.82)$$

pour tout couple  $k, n \in \mathbb{N}_0$  avec  $|n-k| > 1$ . En effet, si nous supposons  $n < k$  sans perte de généralité, nous pouvons obtenir la relation (4.82) par l'orthogonalité (4.80) entre les polynômes  $p_k$  et

$$Q_n : I \rightarrow \mathbb{R}, x \mapsto Q_n(x) = x\mathcal{P}_n(x), \quad (4.83)$$

ce dernier étant de degré  $n+1 < k$ . La décomposition de ce polynôme  $Q_n$  dans la base  $(p_j)_{j \in \mathbb{N}_0}$  des polynômes orthogonaux s'écrit donc logiquement

$$Q_n(x) = \varphi(p_{n+1}, Q_n)p_{n+1}(x) + \varphi(p_n, Q_n)p_n(x) + \varphi(p_{n-1}, Q_n)p_{n-1}(x) \quad (4.84)$$

par la relation (4.82).

Notons ensuite que le polynôme  $Q_n - \mathcal{P}_{n+1}$ , défini par

$$\begin{aligned} Q_n(x) - \mathcal{P}_{n+1}(x) &= x \left( x^n + \sum_{j=0}^{n-1} B_{n,j} p_j(x) \right) - \left( x^{n+1} + \sum_{j=0}^n B_{n+1,j} p_j(x) \right) \\ &= \sum_{j=0}^{n-1} B_{n,j} x p_j(x) - \sum_{j=0}^n B_{n+1,j} p_j(x) \end{aligned} \quad (4.85)$$

pour tout  $x \in I$  selon les définitions (4.72) et (4.83), est effectivement un polynôme de degré  $n$  et, en tant que tel, orthogonal au polynôme  $p_{n+1}$  à cause de la relation (4.80). Nous avons donc  $\varphi(p_{n+1}, Q_n - \mathcal{P}_{n+1}) = 0$  et ainsi, par la linéarité du produit scalaire et la relation (4.74),

$$\varphi(p_{n+1}, Q_n) = \varphi(p_{n+1}, \mathcal{P}_{n+1}) = \|\mathcal{P}_{n+1}\|, \quad (4.86)$$

ce qui nous donne  $\varphi(p_{n+1}, Q_n)p_{n+1}(x) = \mathcal{P}_{n+1}(x)$  pour tout  $x$ . Comme  $Q_{n-1} - \mathcal{P}_n$  est un polynôme de degré  $n-1$  pour  $n > 0$ , nous avons

$$\varphi(\mathcal{P}_n, Q_{n-1} - \mathcal{P}_n) = \|\mathcal{P}_n\| \varphi(p_n, Q_{n-1} - \mathcal{P}_n) = 0 \quad (4.87)$$

et ainsi

$$\begin{aligned} \|\mathcal{P}_n\|^2 &= \varphi(\mathcal{P}_n, \mathcal{P}_n) = \varphi(\mathcal{P}_n, Q_{n-1}) = \int_a^b \mathcal{P}_n(x) x \mathcal{P}_{n-1}(x) w(x) dx \\ &= \int_a^b x \mathcal{P}_n(x) \mathcal{P}_{n-1}(x) w(x) dx = \varphi(Q_n, \mathcal{P}_{n-1}), \end{aligned} \quad (4.88)$$

ce qui prouve  $\varphi(p_{n-1}, Q_n) > 0$  pour tout  $n > 0$ . En résumé, nous pouvons donc réécrire la décomposition (4.84) de  $Q_n$  dans la base des polynômes orthogonaux comme

$$Q_n(x) = x\mathcal{P}_n(x) = \mathcal{P}_{n+1}(x) + a_n \mathcal{P}_n(x) + b_{n-1} \mathcal{P}_{n-1}(x) \quad (4.89)$$

avec  $a_n = \varphi(\mathcal{P}_n, Q_n) / \varphi(\mathcal{P}_n, \mathcal{P}_n)$  pour tout  $n \in \mathbb{N}_0$  et

$$b_{n-1} = \begin{cases} \varphi(\mathcal{P}_n, \mathcal{P}_n) / \varphi(\mathcal{P}_{n-1}, \mathcal{P}_{n-1}) > 0 & : n > 0 \\ 0 & : n = 0 \end{cases}, \quad (4.90)$$

ce qui prouve la relation (4.81).

Considérons, à présent, très concrètement le cas de l'espace de Hilbert (4.65) qui est défini sur l'intervalle  $I = ]-1, 1[$  et muni du produit scalaire canonique, c-à-d,  $w(x) = 1$  pour tout  $x \in I$ . Les premiers trois polynômes orthogonaux non normalisés (4.72) et normalisés (4.74) se construisent successivement via

$$\mathcal{P}_0(x) = 1, \quad (4.91)$$

$$p_0(x) = \left( \int_{-1}^1 dy \right)^{-1/2} = \frac{1}{\sqrt{2}}, \quad (4.92)$$

$$\mathcal{P}_1(x) = x - p_0 \int_{-1}^1 yp_0 dy = x, \quad (4.93)$$

$$p_1(x) = \mathcal{P}_1(x) \left( \int_{-1}^1 y^2 dy \right)^{-1/2} = \sqrt{\frac{3}{2}} x, \quad (4.94)$$

$$\mathcal{P}_2(x) = x^2 - p_1(x) \int_{-1}^1 y^2 p_1(y) dy - p_0 \int_{-1}^1 y^2 p_0 dy = x^2 - \frac{1}{3}, \quad (4.95)$$

$$p_2(x) = \mathcal{P}_2(x) \left( \int_{-1}^1 \mathcal{P}_2^2(y) dy \right)^{-1/2} = \sqrt{\frac{45}{8}} \left( x^2 - \frac{1}{3} \right), \quad (4.96)$$

où nous utilisons  $\int_{-1}^1 x^k p_n(x) dx = 0$  si  $k+n$  est un nombre impair. Nous obtenons ainsi les *polynômes de Legendre*  $P_n$  qui, par convention, sont définis par

$$P_n(x) = \frac{(2n)!}{2^n n!^2} \mathcal{P}_n(x), \quad (4.97)$$

notamment tel que nous avons  $P_n(1) = 1$ , pour tout  $n \in \mathbb{N}_0$ . Concrètement, les cinq premiers polynômes de Legendre sont évalués par

$$P_0(x) = 1, \quad (4.98)$$

$$P_1(x) = x, \quad (4.99)$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1), \quad (4.100)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x), \quad (4.101)$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3). \quad (4.102)$$

Ils satisfont à la relation de récursion

$$nP_n(x) - (2n-1)xP_{n-1}(x) + (n-1)P_{n-2}(x) = 0 \quad (4.103)$$

pour tout  $n \in \mathbb{N}$ .

Dans le cas de l'intervalle infini  $I = \mathbb{R}$  et de la fonction de pondération

$$w : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto w(x) = \exp(-x^2), \quad (4.104)$$

grâce à laquelle les polynômes ont une norme finie et font donc partie de l'espace de Hilbert (4.65), nous obtenons par ce procédé les polynômes orthogonaux et normalisés par

$$p_n(x) = \frac{1}{\sqrt{\sqrt{\pi}2^n n!}} H_n(x) \quad (4.105)$$

pour tout  $n \in \mathbb{N}_0$ , où  $H_0, H_1, \dots$  sont les *polynômes d'Hermite*. Les cinq premiers représentants de ces polynômes sont évalués par

$$H_0(x) = 1, \quad (4.106)$$

$$H_1(x) = 2x, \quad (4.107)$$

$$H_2(x) = 4x^2 - 2, \quad (4.108)$$

$$H_3(x) = 8x^3 - 12x, \quad (4.109)$$

$$H_4(x) = 16x^4 - 48x^2 + 12. \quad (4.110)$$

Les polynômes d'Hermite satisfont à la relation de récursion

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x) \quad (4.111)$$

pour tout  $n \in \mathbb{N}_0$ .

Dans le cas de l'intervalle semi-infini  $I = \mathbb{R}_+ = ]0, \infty[$ , nous pouvons choisir la fonction de pondération

$$w : \mathbb{R}_+ \rightarrow \mathbb{R}, x \mapsto w(x) = e^{-x} \quad (4.112)$$

afin de rendre les polynômes normalisables sur cet intervalle. Nous obtenons ainsi les *polynômes de Laguerre*  $L_0, L_1, \dots$  comme polynômes orthogonaux non normalisés qui satisfont à la relation  $L_n(0) = 1$  pour tout  $n \in \mathbb{N}_0$ . Leurs trois premiers représentants sont donnés par

$$L_0(x) = 1, \quad (4.113)$$

$$L_1(x) = 1 - x, \quad (4.114)$$

$$L_2(x) = 1 - 2x + \frac{1}{2}x^2. \quad (4.115)$$

Les polynômes de Laguerre satisfont à la relation de récursion

$$(n+1)L_{n+1}(x) = (2n+1-x)L_n(x) - nL_{n-1}(x) \quad (4.116)$$

pour tout  $n \in \mathbb{N}_0$ .

Une variante de ces polynômes est obtenue si nous modifions la définition de la fonction de pondération (4.112) en

$$w(x) = x^\alpha e^{-x} \quad (4.117)$$

avec  $\alpha > -1$ . Nous obtenons ainsi les *polynômes de Laguerre généralisés*

$$L_n^{(\alpha)}(x) = \sum_{k=0}^n \frac{1}{k!} \binom{n+\alpha}{n-k} (-x)^k, \quad (4.118)$$

où nous utilisons la définition

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \equiv \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)} \quad (4.119)$$

pour tout  $n, k \in \mathbb{R}_+$  avec  $n \geq k$  et  $\Gamma : \mathbb{R}_+ \rightarrow \mathbb{R}, x \mapsto \Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dx$ . Les polynômes de Laguerre généralisés satisfont à la relation de récursion

$$(n+1)L_{n+1}^{(\alpha)}(x) = (2n+1+\alpha-x)L_n^{(\alpha)}(x) - (n+\alpha)L_{n-1}^{(\alpha)}(x) \quad (4.120)$$

pour tout  $n \in \mathbb{N}_0$ .

Enfin, des polynômes orthogonaux peuvent aussi être construits dans un espace de Hilbert qui est défini sur un intervalle fini et muni d'un produit scalaire non canonique. Sur l'intervalle  $] -1, 1[$  nous pouvons ainsi obtenir les *polynômes de Tchebychev*, de première espèce,  $T_n$ , pour la fonction de pondération

$$w : I \rightarrow \mathbb{R}, x \mapsto w(x) = \frac{1}{\sqrt{1-x^2}} \quad (4.121)$$

et de seconde espèce,  $U_n$ , pour

$$w : I \rightarrow \mathbb{R}, x \mapsto w(x) = \sqrt{1-x^2}. \quad (4.122)$$

Ces polynômes de Tchebychev satisfont aux propriétés

$$\cos(n\theta) = T_n(\cos \theta), \quad (4.123)$$

$$\sin(n\theta) = U_n(\cos \theta) \sin \theta \quad (4.124)$$

pour tout  $n \in \mathbb{N}_0$  et tout  $\theta \in \mathbb{R}$ .

# Chapitre 5

## Eléments de la théorie spectrale

### 5.1 L'équation de Sturm-Liouville

Après avoir discuté les propriétés générales d'un espace vectoriel constitué de fonctions, nous sommes suffisamment armés pour pouvoir résoudre le problème qui est posé en mécanique quantique par l'équation de Schrödinger stationnaire à une dimension

$$-\frac{\hbar^2}{2m}\psi''(x) + V(x)\psi(x) = E\psi(x), \quad (5.1)$$

décrivant le profil spatial de la fonction d'onde  $\psi$  d'une particule de masse  $m$  qui se meut avec l'énergie totale  $E$  dans le paysage d'énergie potentielle  $V$ . Cette équation (5.1) représente un cas particulier d'une équation différentielle linéaire du deuxième ordre à une composante complexe, c-à-d, d'une équation de la forme (3.103) pour  $n = 2$ . Dans un souci, propre au mathématicien, de ne pas inutilement restreindre l'analyse à des cas trop particuliers, nous étendons la discussion dans ce chapitre à une équation générale du type (3.103) pour  $n = 2$  ayant des coefficients réels.

Considérons donc l'équation différentielle

$$\psi''(x) + a_1(x)\psi'(x) + a_0(x)\psi(x) = 0 \quad (5.2)$$

définie sur l'intervalle ouvert  $I = ]a, b[ \subseteq \mathbb{R}$ , où  $a_j : I \rightarrow \mathbb{R}, x \mapsto a_j(x)$  pour  $j = 0, 1$  sont des fonctions continues qui sont définies sur l'intervalle  $I$  et atteignent des valeurs réelles. Il s'avère utile, pour la suite de la discussion, de formuler cette équation (5.2) différemment. Nous définissons dans ce but la fonction

$$p : I \rightarrow \mathbb{R}_+, x \mapsto p(x) = p_0 \exp\left(\int_{x_1}^x a_1(x') dx'\right) \quad (5.3)$$

pour un  $p_0 > 0$  et un  $x_1 \in I$ , et la fonction

$$q : I \rightarrow \mathbb{R}, x \mapsto q(x) = -a_0(x)p(x). \quad (5.4)$$

On peut montrer que  $p$  est continûment dérivable et satisfait à l'équation différentielle  $p'(x) = a_1(x)p(x)$ , et que  $q$  est une fonction continue. Avec ces deux fonctions  $p$  et  $q$ , l'équation différentielle (5.2) peut être réécrite de manière tout à fait équivalente comme

$$\psi''(x) + \frac{p'(x)}{p(x)}\psi'(x) - \frac{q(x)}{p(x)}\psi(x) = 0 \quad (5.5)$$

ou

$$p(x)\psi''(x) + p'(x)\psi'(x) - q(x)\psi(x) = 0 \quad (5.6)$$

ou

$$\frac{d}{dx} \left[ p(x) \frac{d}{dx} \psi(x) \right] - q(x)\psi(x) = 0, \quad (5.7)$$

cette dernière formulation (5.7) de l'équation étant nommée après les mathématiciens STURM et LIOUVILLE. L'équation de Schrödinger (5.1) est récupérée dans le cas particulier  $p(x) = \hbar^2/(2m)$ , impliquant  $a_1(x) = 0$ , et  $q(x) = V(x) - E$ , impliquant  $a_0(x) = 2m(E - V(x))/\hbar^2$ , pour tout  $x \in I$ .

Comme l'équation de Sturm-Liouville (5.7) est une équation différentielle linéaire du deuxième ordre, nous pouvons affirmer, suivant la discussion dans la section 3.8, que sa solution est déterminée de manière unique par un jeu de conditions initiales pour  $\psi$  et  $\psi'$  en un  $x = x_0 \in I$ . Il s'ensuit qu'une solution non triviale,  $\psi \not\equiv 0$ , ne peut pas posséder des zéros non simples. En effet, si la solution  $\psi$  et sa dérivée première  $\psi'$  s'annulent en un  $x_0 \in I$ , le jeu des conditions initiales en  $x_0$  qui en résulte,  $\psi(x_0) = 0$  et  $\psi'(x_0) = 0$ , sera satisfait par la solution triviale  $\psi \equiv 0$ , c-à-d,  $\psi(x) = 0$  pour tout  $x \in I$ , qui, par l'unicité de la solution, sera donc la solution unique dans ce cas.

Comme nous l'avons déjà constaté dans la section 3.8 (cf. équation (3.108)), toute solution à l'équation de Sturm-Liouville (5.7) peut être représentée par une combinaison linéaire unique de deux solutions  $\varphi_1, \varphi_2 : I \rightarrow \mathbb{C}$  à cette équation qui sont linéairement indépendantes, c-à-d, pour lesquelles le wronskien

$$W(x) = \det \begin{pmatrix} \varphi_1(x) & \varphi_2(x) \\ \varphi_1'(x) & \varphi_2'(x) \end{pmatrix} = \varphi_1(x)\varphi_2'(x) - \varphi_2(x)\varphi_1'(x) \quad (5.8)$$

ne s'annule pas en  $I$ . De manière tout à fait générale, ce wronskien est, dans le cas de l'équation de Sturm-Liouville, donné par

$$W(x) = \frac{c}{p(x)} \quad (5.9)$$

pour une constante  $c \in \mathbb{C}$ , et ne dépend donc pas du tout de  $x$  dans le cas de l'équation de Schrödinger. En effet, on peut facilement calculer

$$\begin{aligned} \frac{d}{dx} [p(x)W(x)] &= \frac{d}{dx} [p(x) (\varphi_1(x)\varphi_2'(x) - \varphi_2(x)\varphi_1'(x))] \\ &= \varphi_1(x) \frac{d}{dx} \left[ p(x) \frac{d}{dx} \varphi_2(x) \right] - \varphi_2(x) \frac{d}{dx} \left[ p(x) \frac{d}{dx} \varphi_1(x) \right] \\ &= \varphi_1(x)q(x)\varphi_2(x) - \varphi_2(x)q(x)\varphi_1(x) = 0 \end{aligned} \quad (5.10)$$

en utilisant le fait que  $\varphi_1$  et  $\varphi_2$  satisfont à l'équation (5.7).

Si une première solution de l'équation de Sturm-Liouville (5.7) a déjà été déterminée par la fonction  $\varphi_1 : I \rightarrow \mathbb{C}, x \mapsto \varphi_1(x)$ , la résolution de l'équation différentielle linéaire inhomogène du premier ordre

$$\varphi_2'(x) = \frac{\varphi_1'(x)}{\varphi_1(x)}\varphi_2(x) + \frac{c}{p(x)\varphi_1(x)} \quad (5.11)$$

qui résulte des équations (5.8) et (5.9) permettra d'en construire une seconde solution qui est linéairement indépendante de  $\varphi_1$ . Concrètement, nous obtenons par la méthode de la variation de la constante (cf. section 3.7) l'expression générale

$$\varphi_2(x) = \int_{x_0}^x \frac{\varphi_1(x)}{p(x')\varphi_1^2(x')} dx' \quad (5.12)$$

pour cette seconde solution, où la constante  $x_0 \in I$  doit, en pratique, être choisie telle que l'intégrale constituant le membre de droite de l'équation (5.12) existe. Cette seconde solution peut donc être utilisée pour compléter le système fondamental  $(\varphi_1, \varphi_2)$  associé à l'équation de Sturm-Liouville (5.7).

## 5.2 Conditions aux limites

Le défi majeur qui est à relever dans le cadre de l'équation de Schrödinger stationnaire (5.1) en mécanique quantique consiste à déterminer les solutions de cette équation non pas pour un jeu de conditions initiales en un  $x_0 \in I = ]a, b[$  donné mais pour un jeu de *conditions aux limites* spécifiées aux bornes  $a, b$  de l'intervalle sur lequel ces solutions sont définies. Dans le cas le plus courant d'un intervalle infini  $I = \mathbb{R}$ , on cherche les solutions  $\psi$  à l'équation de Schrödinger (5.1) qui sont des fonctions de carré intégrables sur  $\mathbb{R}$ . Suivant le raisonnement développé à la fin de la section 4.4 (cf. équation (4.63)), elles satisfont donc nécessairement aux conditions aux limites de Dirichlet

$$\lim_{x \rightarrow -\infty} \psi(x) = \lim_{x \rightarrow \infty} \psi(x) = 0, \quad (5.13)$$

faute de quoi l'intégrale  $\int_{-\infty}^{\infty} |\psi(x)|^2 dx$  n'existerait pas. De telles conditions aux limites de Dirichlet,  $\psi(a) = \psi(b) = 0$ , peuvent aussi être imposées en présence d'un intervalle fini  $I = ]a, b[$  avec  $-\infty < a < b < \infty$ , notamment dans le cas d'une particule quantique qui est enfermée dans une boîte à une dimension limitée par des murs durs de hauteur infinie. En présence d'une configuration d'énergie potentielle  $V$  qui est parfaitement symétrique par rapport à l'origine, une condition aux limites de Neumann, notamment  $\psi'(0) = 0$ , peut être imposée en la borne inférieure de l'intervalle  $I = ]0, \infty[$  (ou  $I = ]0, b[$  si on a affaire à un mur dur à  $x = \pm b$ ) afin de déterminer les solutions à l'équation de Schrödinger stationnaire qui sont symétriques par rapport à l'origine. Enfin, dans le cas d'un potentiel

parfaitement périodique, c-à-d,  $V(x + L) = V(x)$  pour tout  $x \in \mathbb{R}$  pour une période  $L > 0$ , on peut s'intéresser à calculer toutes les solutions de l'équation de Schrödinger stationnaire (5.1) qui sont aussi périodiques avec la même période. Cela peut se faire en imposant les conditions aux limites périodiques  $\psi(0) = \psi(L)$  et  $\psi'(0) = \psi'(L)$  sur l'intervalle fini  $I = ]0, L[$ .

De manière générale, nous nous intéressons donc, dans la suite de ce chapitre, à des solutions de l'équation de Sturm-Liouville (5.7) qui satisfont à un jeu de deux conditions aux limites homogènes, spécifiées aux bornes  $x = a$  et  $x = b$  de l'intervalle  $I = ]a, b[ \subseteq \mathbb{R}$  sur lequel les fonctions  $p$  et  $q$  sont définies. Comme détaillé ci-dessus, de telles conditions aux limites homogènes peuvent p.ex. être les conditions de Dirichlet  $\psi(a) = 0$  et  $\psi(b) = 0$ , les conditions de Neumann  $\psi'(a) = 0$  et  $\psi'(b) = 0$  ou alors les conditions périodiques  $\psi(a) = \psi(b)$  et  $\psi'(a) = \psi'(b)$ . Notons que l'imposition de ces dernières est judicieuse seulement si  $p$ ,  $p'$  et  $q$  satisfont aussi à ces conditions périodiques, c-à-d,  $p(a) = p(b)$ ,  $p'(a) = p'(b)$  et  $q(a) = q(b)$ , faute de quoi nous avons, en général,  $\psi''(a) \neq \psi''(b)$  pour la solution de l'équation de Sturm-Liouville (5.7), ce qui signifie que cette dernière solution n'est pas une fonction périodique à proprement parler.

Plus généralement, un jeu de deux conditions aux limites homogènes peut être généré par deux opérateurs  $\Lambda_j : \psi \mapsto \Lambda_j \psi$  pour  $j = 1, 2$  qui transforment  $\psi$  en une fonction à deux arguments, tel que nous avons pour tout couple  $(x, y) \in I \times I$

$$(\Lambda_j \psi)(x, y) \equiv (\Lambda_j \psi)_{x,y} = \alpha_j \psi(x) + \beta_j \psi(y) + \gamma_j \psi'(x) + \delta_j \psi'(y) \quad (5.14)$$

pour  $j = 1, 2$ , avec des coefficients  $\alpha_j, \beta_j, \gamma_j, \delta_j \in \mathbb{C}$ . Les deux conditions aux limites peuvent donc être formulées par  $(\Lambda_1 \psi)_{a,b} = 0$  et  $(\Lambda_2 \psi)_{a,b} = 0$ . Bien évidemment, nous devons exiger que les deux jeux de coefficients  $(\alpha_1, \beta_1, \gamma_1, \delta_1)$  et  $(\alpha_2, \beta_2, \gamma_2, \delta_2)$  soient linéairement indépendants, c-à-d, tels que  $c_1(\alpha_1, \beta_1, \gamma_1, \delta_1) \neq c_2(\alpha_2, \beta_2, \gamma_2, \delta_2)$  pour tout couple  $c_1, c_2 \in \mathbb{C} \setminus \{0\}$ , afin de ne pas obtenir deux fois la même condition.

Étudions d'abord ce qui se passe si on impose une seule de ces deux conditions aux limites. Nous constatons qu'en règle générale<sup>1</sup> les solutions à l'équation différentielle (5.7) qui satisfont à la condition aux limites  $(\Lambda_1 \psi)_{a,b} = 0$  forment un espace vectoriel unidimensionnel. Cet espace vectoriel peut être explicitement constitué par l'ensemble des fonctions  $\psi : I \rightarrow \mathbb{C}$  pour lesquelles il existe un  $\alpha \in \mathbb{C}$  tel que

$$\psi(x) = \alpha [(\Lambda_1 \varphi_2)_{a,b} \varphi_1(x) - (\Lambda_1 \varphi_1)_{a,b} \varphi_2(x)] \quad (5.15)$$

pour tout  $x \in I$ , où  $(\varphi_1, \varphi_2)$  est un système fondamental de l'équation (5.7). Une exception notable apparaît dans le cas d'une équation de Sturm-Liouville (5.7) avec  $p(x) = p_0 > 0$  et  $q(x) = q_0 < 0$  pour tout  $x \in I = ]a, b[$ , notamment lorsque l'on impose une condition périodique,  $\psi(a) = \psi(b)$  ou  $\psi'(a) = \psi'(b)$ , et lorsqu'il existe un  $l \in \mathbb{N}$  tel que  $k(b - a) = 2\pi l$  avec  $k = (-q_0/p_0)^{1/2}$ . Dans ce

<sup>1</sup> "règle générique" serait une terminologie mieux adaptée ici, du point de vue mathématique.

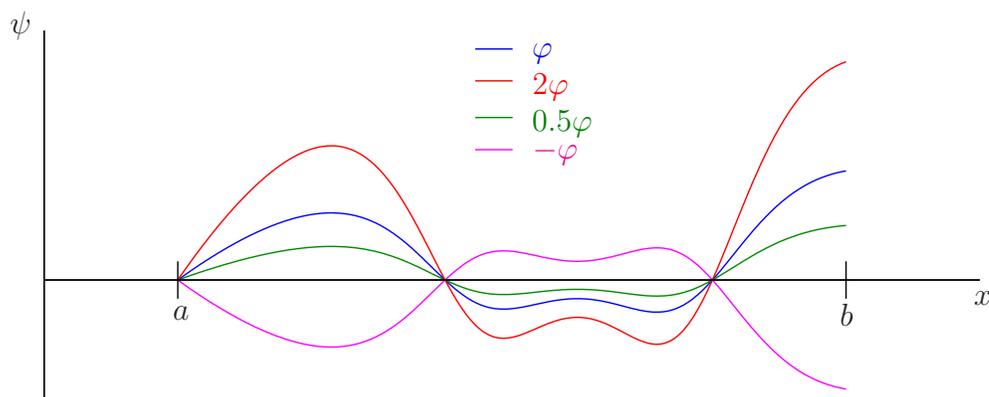


Figure 5.1: Schéma des graphiques d'une fonction  $\varphi : I = ]a, b[ \rightarrow \mathbb{R}$  ainsi que de ses multiples  $2\varphi$ ,  $0.5\varphi$  et  $-\varphi$ . Si  $\varphi$  satisfait à la condition de Dirichlet en  $x = a$ ,  $\varphi(a) = 0$ , mais non pas à celle en  $x = b$ , c-à-d,  $\varphi(b) \neq 0$ , il en est de même pour tout multiple non nul de  $\varphi$ . Comme l'ensemble des solutions satisfaisant à la condition de Dirichlet en  $x = a$  forme un espace vectoriel unidimensionnel, nous pouvons inférer que seule la solution triviale  $\psi \equiv 0$  satisfait aux deux conditions de Dirichlet dans ce cas. Le graphique illustre également que des solutions non triviales satisfaisant à ces deux conditions aux limites existent si et seulement si la borne supérieure  $b$  coïncide avec un des nœuds de la fonction  $\varphi$ , ce qui peut donc être considéré comme un cas exceptionnel.

cas particulier, la solution générale de cette équation de Sturm-Liouville,  $\psi(x) = A \cos(kx) + B \sin(kx)$  pour tout  $x \in I$ , satisfait à cette condition périodique pour tout couple  $A, B \in \mathbb{C}$ , ce qui signifie que l'espace vectoriel des solutions satisfaisant à la condition aux limites  $(\Lambda_1\psi)_{a,b} = 0$  est un espace à deux dimensions.

Supposons ensuite qu'il existe une solution non triviale,  $\psi \neq 0$ , à l'équation de Sturm-Liouville (5.7) qui satisfait aux deux conditions aux limites  $(\Lambda_1\psi)_{a,b} = 0$  et  $(\Lambda_2\psi)_{a,b} = 0$ . Au vu de l'homogénéité de l'expression (5.14), nous pouvons inférer que tout multiple  $\alpha\psi$  de  $\psi$  avec  $\alpha \in \mathbb{C}$  est également une solution de l'équation (5.7) satisfaisant au même jeu de conditions aux limites. Les solutions recherchées en présence de ces deux conditions aux limites forment donc, en règle générale, un espace vectoriel unidimensionnel. Encore une fois, le cas de l'équation de Sturm-Liouville avec  $p(x) = p_0 > 0$  et  $q(x) = q_0 < 0$  pour tout  $x \in I = ]a, b[$  en combinaison avec les deux conditions périodiques  $\psi(a) = \psi(b)$  et  $\psi'(a) = \psi'(b)$  et avec la propriété  $(-q_0/p_0)^{1/2}(b-a) = 2\pi l$  pour un  $l \in \mathbb{N}$  constitue une exception notable à cette règle. Dans ce dernier cas, les solutions recherchées forment un espace à deux dimensions car toute solution à l'équation de Sturm-Liouville satisfait à ces deux conditions aux limites.

Supposons maintenant qu'il existe solution  $\varphi : I \rightarrow \mathbb{C}$  à l'équation de Sturm-Liouville (5.7) qui satisfait à une des deux conditions aux limites,  $(\Lambda_1\varphi)_{a,b} = 0$ ,

mais non pas à l'autre,  $(\Lambda_2\varphi)_{a,b} \neq 0$ . Comme illustré sur la figure 5.1 pour le cas particulier des conditions de Dirichlet, nous pouvons, dans ce cas, inférer qu'en règle générale<sup>2</sup> la solution triviale  $\psi \equiv 0$  est la seule solution qui satisfait aux deux conditions aux limites  $(\Lambda_1\varphi)_{a,b} = 0$  et  $(\Lambda_2\varphi)_{a,b} = 0$ . Comme la propriété  $(\Lambda_2\varphi)_{a,b} \neq 0$  doit être considérée comme étant générique pour une solution  $\varphi$  satisfaisant à  $(\Lambda_1\varphi)_{a,b} = 0$ , dans le sens qu'elle apparaît forcément lorsque l'on choisit les bornes  $a, b$  de l'intervalle  $I$  au hasard, nous pouvons affirmer qu'une solution non triviale à l'équation de Sturm-Liouville qui satisfait aux deux conditions aux limites  $(\Lambda_1\varphi)_{a,b} = 0$  et  $(\Lambda_2\varphi)_{a,b} = 0$  n'existe généralement pas.

Notons enfin que si nous imposons un jeu de conditions aux limites dont au moins une est inhomogène, c-à-d,  $(\Lambda_1\varphi)_{a,b} = c_1$  et  $(\Lambda_2\varphi)_{a,b} = c_2$  avec  $(c_1, c_2) \neq (0, 0)$ , l'équation de Sturm-Liouville (5.7) possède une solution unique si et seulement si seule la solution triviale  $\psi \equiv 0$  satisfait à l'équation (5.7) en présence des conditions aux limites homogènes associées  $(\Lambda_1\varphi)_{a,b} = 0$  et  $(\Lambda_2\varphi)_{a,b} = 0$ . En effet, s'il existe une solution non triviale à l'équation (5.7) satisfaisant à  $(\Lambda_1\psi)_{a,b} = 0$  et  $(\Lambda_2\psi)_{a,b} = 0$ , nous pourrions ajouter à  $\varphi$  un multiple  $\alpha\psi$  avec  $\alpha \in \mathbb{C}$  pour ainsi obtenir une autre solution  $\tilde{\varphi} \equiv \varphi + \alpha\psi$  satisfaisant aux conditions aux limites  $(\Lambda_1\tilde{\varphi})_{a,b} = c_1$  et  $(\Lambda_2\tilde{\varphi})_{a,b} = c_2$ .

### 5.3 L'opérateur de Sturm-Liouville

Pour obtenir davantage d'informations sur les propriétés des solutions non triviales de l'équation de Sturm-Liouville en présence des conditions aux limites homogènes, il est utile de changer la perspective et de considérer ces solutions non triviales comme des *fonctions propres* d'un opérateur qui agit sur un espace de Hilbert constitué de fonctions. Concrètement, dans le cas des conditions aux limites

$$(\Lambda_1\psi)_{a,b} = 0 \quad \text{et} \quad (\Lambda_2\psi)_{a,b} = 0 \quad (5.16)$$

imposées aux bornes de l'intervalle  $I = ]a, b[ \subseteq \mathbb{R}$ , où les expressions  $(\Lambda_1\psi)_{a,b}$  et  $(\Lambda_2\psi)_{a,b}$  sont définies via l'équation (5.14) en termes des jeux de coefficients  $(\alpha_1, \beta_1, \gamma_1, \delta_1)$  et  $(\alpha_2, \beta_2, \gamma_2, \delta_2)$  qui sont linéairement indépendants, cet espace de Hilbert est défini comme

$$\mathcal{H} = \{\psi : I \rightarrow \mathbb{C} : \|\psi\| < \infty \text{ et } (\Lambda_1\psi)_{a,b} = (\Lambda_2\psi)_{a,b} = 0\} \quad (5.17)$$

où la norme

$$\|\psi\| = \sqrt{\langle \psi | \psi \rangle} \quad (5.18)$$

---

<sup>2</sup>Le cas  $p(x) = p_0 > 0$  et  $q(x) = q_0 < 0$  pour tout  $x \in I = ]a, b[$  avec les conditions aux limites périodiques et  $(-q_0/p_0)^{1/2}(b-a) = 2\pi l$  pour un  $l \in \mathbb{N}$  constitue toujours une exception à cette règle générale. Dans ce cas très spécial, les solutions satisfaisant à ces deux conditions aux limites forment un espace vectoriel unidimensionnel.

est définie par l'intermédiaire du produit scalaire canonique

$$\varphi : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}, (\phi, \psi) \mapsto \varphi(\phi, \psi) = \langle \phi | \psi \rangle = \int_a^b \phi^*(x)\psi(x)dx. \quad (5.19)$$

$\mathcal{H}$  contient donc toute fonction définie sur l'intervalle  $I$  qui est de carré intégrable et satisfait aux conditions aux limites (5.16).

Sur cet espace de Hilbert nous définissons l'opérateur de *Sturm-Liouville*

$$\hat{L} : \mathcal{D} \subset \mathcal{H} \rightarrow \mathcal{H}, \psi \mapsto \hat{L}\psi \quad (5.20)$$

par

$$(\hat{L}\psi)(x) = -\frac{d}{dx} [p(x)\psi'(x)] + q(x)\psi(x) \quad (5.21)$$

pour tout  $x \in I$ , où  $p$  et  $q$  sont les fonctions qui constituent la définition de l'équation de Sturm-Liouville (5.7). Cette dernière équation peut donc être réécrite comme  $(\hat{L}\psi)(x) = 0$ . Strictement parlant, comme nous l'avons fait remarquer dans la section 4.3, la région de définition de cet opérateur  $\hat{L}$  doit être restreinte à un sous-espace vectoriel  $\mathcal{D} \subset \mathcal{H}$  contenant toute fonction  $\psi$  qui est deux fois dérivable et pour laquelle  $\hat{L}\psi$  est de carré intégrable. Nous pouvons cependant supposer que les fonctions  $p$  et  $q$  sont suffisamment peu pathologiques de sorte que ce sous-espace  $\mathcal{D}$  est dense en  $\mathcal{H}$ .

Etudions à présent sous quelles conditions cet opérateur de Sturm-Liouville est auto-adjoint par rapport au produit scalaire canonique (5.19). Cette propriété est satisfaite si nous avons  $\langle \phi | \hat{L}\psi \rangle = \langle \psi | \hat{L}\phi \rangle^*$  pour tout couple  $\psi, \phi \in \mathcal{H}$  (ou, plus précisément, tout couple  $\psi, \phi \in \mathcal{D} \subset \mathcal{H}$ ), où nous avons introduit la notation  $\langle \phi | \hat{L}\psi \rangle \equiv \langle \phi | \hat{L}\psi \rangle$  pour ce qu'on appelle *élément de matrice* de l'opérateur  $\hat{L}$  par rapport aux fonctions  $\phi$  et  $\psi$ . En utilisant la méthode de l'intégration par partie, nous calculons

$$\begin{aligned} \langle \phi | \hat{L}\psi \rangle - \langle \psi | \hat{L}\phi \rangle^* &= \int_a^b \phi^*(x)(\hat{L}\psi)(x)dx - \int_a^b \psi(x)(\hat{L}\phi^*)(x)dx \\ &= \int_a^b \phi^*(x) \left( -\frac{d}{dx} \left[ p(x) \frac{d}{dx} \psi(x) \right] + q(x)\psi(x) \right) dx \\ &\quad - \int_a^b \psi(x) \left( -\frac{d}{dx} \left[ p(x) \frac{d}{dx} \phi^*(x) \right] + q(x)\phi^*(x) \right) dx \\ &= - \left[ \phi^*(x)p(x) \frac{d}{dx} \psi(x) \right]_a^b + \int_a^b p(x) \left( \frac{d\phi}{dx} \phi(x) \right)^* \frac{d\psi}{dx}(x) dx \\ &\quad + \left[ \psi(x)p(x) \frac{d}{dx} \phi^*(x) \right]_a^b - \int_a^b p(x) \frac{d\psi}{dx}(x) \left( \frac{d\phi}{dx}(x) \right)^* dx \\ &= \left[ p(x) \left( -\phi^*(x) \frac{d}{dx} \psi(x) + \psi(x) \frac{d}{dx} \phi^*(x) \right) \right]_a^b. \quad (5.22) \end{aligned}$$

Nous pouvons donc affirmer que l'opérateur de Sturm-Liouville (5.21) est auto-adjoint par rapport au produit scalaire canonique (5.19) si l'imposition des conditions aux limites (5.16) implique la relation

$$p(a) [(\phi'(a))^* \psi(a) - \phi^*(a) \psi'(a)] = p(b) [(\phi'(b))^* \psi(b) - \phi^*(b) \psi'(b)] \quad (5.23)$$

pour tout couple  $\psi, \phi \in \mathcal{H}$ . Ceci est effectivement le cas pour les conditions aux limites qui sont le plus couramment utilisées en physique, à savoir les conditions de Dirichlet  $\psi(a) = \psi(b) = 0$ , les conditions de Neumann  $\psi'(a) = \psi'(b) = 0$ , des conditions mixtes du genre  $\psi'(a) = \psi(b) = 0$ , ainsi que les conditions périodiques  $\psi(a) = \psi(b)$  et  $\psi'(a) = \psi'(b)$  pourvu que nous ayons également  $p(a) = p(b)$ . La relation (5.23) est aussi satisfaite dans le cas particulier où la fonction  $p$  s'annule aux bornes de l'intervalle, c-à-d,

$$\lim_{x \rightarrow a} p(x) = \lim_{x \rightarrow b} p(x) = 0, \quad (5.24)$$

et ceci, dans ce dernier cas, indépendamment des conditions aux limites imposées.

Pour la suite de la discussion dans ce chapitre, nous supposons que les conditions pour que l'opérateur de Sturm-Liouville (5.21) soit auto-adjoint sont effectivement réunies. Dans ce cas, nous pouvons inférer des propriétés intéressantes concernant ses *fonctions propres*, ce dernier concept signifiant l'équivalent des vecteurs propres pour un espace vectoriel constitué de fonctions. Soient donc  $\psi, \phi \in \mathcal{H} \setminus \{0\}$  deux fonctions propres de  $\hat{L}$ , c-à-d, deux fonctions non nulles,  $\psi \neq 0$  et  $\phi \neq 0$ , pour lesquelles il existe un couple de *valeurs propres* associées  $\lambda, \mu \in \mathbb{C}$  tel que

$$\hat{L}\psi(x) = \lambda\psi(x), \quad (5.25)$$

$$\hat{L}\phi(x) = \mu\phi(x) \quad (5.26)$$

pour tout  $x \in I$ . Nous obtenons ainsi, par le fait que  $\hat{L}$  est auto-adjoint, l'identité

$$\begin{aligned} \langle \phi | \hat{L} | \psi \rangle &= \int_a^b \phi^*(x) (\hat{L}\psi)(x) dx = \lambda \int_a^b \phi^*(x) \psi(x) dx = \lambda \langle \phi | \psi \rangle \quad (5.27) \\ &= \langle \psi | \hat{L} | \phi \rangle^* = \int_a^b \psi(x) (\hat{L}\phi)^*(x) dx = \mu^* \int_a^b \psi(x) \phi^*(x) dx = \mu^* \langle \phi | \psi \rangle \end{aligned}$$

qui peut se réécrire sous la forme

$$(\lambda - \mu^*) \langle \phi | \psi \rangle = 0. \quad (5.28)$$

Dans le cas particulier  $\phi \equiv \psi$ , où nous avons logiquement  $\lambda = \mu$ , la relation (5.28) implique  $\lambda = \lambda^*$  puisque, par la définition d'une fonction propre,  $\psi \neq 0$  et donc  $\langle \psi | \psi \rangle > 0$ . Les valeurs propres associées à l'opérateur de Sturm-Liouville sont donc réelles. Si, en revanche, nous avons affaire à un couple de fonctions

propres  $\psi, \phi$  dont les valeurs propres sont différentes, c-à-d,  $\lambda \neq \mu$ , et donc aussi  $\lambda \neq \mu^*$  par le raisonnement ci-dessus, nous pouvons inférer de l'équation (5.28) la relation  $\langle \phi | \psi \rangle = 0$ . Les fonctions propres de  $\hat{L}$  dont les valeurs propres associées sont différentes sont donc *orthogonales*. Comme  $\hat{L}$  n'a pas de partie imaginaire, nous pouvons les choisir telles qu'elles atteignent des valeurs purement réelles. En effet, si  $\psi : I \rightarrow \mathbb{C}$  satisfait à l'équation  $\hat{L}\psi = \lambda\psi$  pour une valeur propre donnée  $\lambda \in \mathbb{R}$ , il en est de même pour sa partie réelle  $\text{Re}(\psi)$  que nous pouvons donc choisir comme fonction propre associée à  $\lambda$ .

Il est à noter que cette propriété d'orthogonalité est spécifique au choix que nous avons fait pour le produit scalaire dont l'espace de Hilbert (5.17) est muni, à savoir le produit scalaire canonique (5.19). En effet, si nous changeons ce choix en imposant sur cet espace de Hilbert le produit scalaire

$$\varphi(\phi, \psi) = \int_a^b \phi^*(x)\psi(x)w(x)dx \equiv \langle \phi | w | \psi \rangle \quad (5.29)$$

défini au travers de la fonction de pondération positive

$$w : I \rightarrow \mathbb{R}_+, x \mapsto w(x), \quad (5.30)$$

l'opérateur  $\hat{L}$  n'est, en général, plus auto-adjoint. Cependant, il est, dans ce cas, facile d'introduire un autre opérateur

$$\tilde{L} \equiv \frac{1}{w(x)}\hat{L} = \frac{1}{w(x)} \left( -\frac{d}{dx} [p(x)\psi'(x)] + q(x)\psi(x) \right) \quad (5.31)$$

pour lequel on peut montrer qu'il est auto-adjoint par rapport à ce nouveau produit scalaire. En effet, nous avons

$$\varphi(\phi, \tilde{L}\psi) = \int_a^b \phi^*(x) \frac{1}{w(x)} (\hat{L}\psi)(x) w(x) dx = \langle \phi | \hat{L} | \psi \rangle, \quad (5.32)$$

$$\varphi(\tilde{L}\phi, \psi) = \int_a^b \frac{1}{w(x)} (\hat{L}\phi)^*(x) \psi(x) w(x) dx = \langle \psi | \hat{L} | \phi \rangle^* \quad (5.33)$$

ainsi que  $\langle \phi | \hat{L} | \psi \rangle = \langle \psi | \hat{L} | \phi \rangle^*$  car  $\hat{L}$  est auto-adjoint par rapport au produit scalaire canonique (5.19), ce qui implique  $\varphi(\phi, \tilde{L}\psi) = \varphi(\tilde{L}\phi, \psi)$  pour tout couple  $\psi, \phi \in \mathcal{H}$ . Par conséquent, en généralisant le raisonnement que nous avons effectué ci-dessus pour l'opérateur  $\hat{L}$ , les valeurs propres de  $\tilde{L}$  sont réelles et les vecteurs propres de  $\tilde{L}$  qui sont associés à des valeurs propres différentes sont orthogonaux par rapport à ce nouveau produit scalaire (5.29). En ce qui concerne l'opérateur de Sturm-Liouville  $\hat{L}$ , nous pouvons affirmer que ses *fonctions propres généralisées*  $\psi$ , satisfaisant à la relation

$$(\hat{L}\psi)(x) = \lambda w(x)\psi(x) \quad (5.34)$$

pour tout  $x \in I$  avec une valeur propre associée  $\lambda \in \mathbb{R}$ , sont orthogonales par rapport au produit scalaire (5.29) si leurs valeurs propres sont différentes.

## 5.4 Le théorème de Sturm

Davantage de propriétés caractérisant les valeurs propres et fonctions propres de l'opérateur de Sturm-Liouville peuvent être inférées grâce à un théorème remarquable qui porte le nom du mathématicien STURM. Pour ce théorème, nous introduisons dans l'espace de Hilbert défini sur l'intervalle ouvert  $I = ]a, b[ \subseteq \mathbb{R}$  deux opérateurs de Sturm-Liouville  $\hat{L}_1, \hat{L}_2$  avec

$$\hat{L}_j = -\frac{d}{dx}p(x)\frac{d}{dx} + q_j(x) \quad (5.35)$$

pour  $j = 1, 2$ , où  $p : I \rightarrow \mathbb{R}_+$  est continûment dérivable et  $q_1, q_2 : I \rightarrow \mathbb{R}$  sont deux fonctions continues qui satisfont à la propriété  $q_1(x) \geq q_2(x)$  pour tout  $x \in I$ , et ceci tel que l'ensemble des points  $x \in I$  où  $q_1(x) = q_2(x)$  est de mesure nulle. Supposons que chacun de ces deux opérateurs  $\hat{L}_j$  possède une fonction propre réelle  $\varphi_j : I \rightarrow \mathbb{R}$  à la valeur propre 0, c-à-d, une solution non triviale,  $\varphi_j \neq 0$ , de l'équation  $\hat{L}_j\varphi_j(x) = 0$  pour tout  $x \in I$ . Supposons aussi que  $\varphi_1$  possède deux zéros dans  $I$ , notamment en  $x = \alpha$  et  $x = \beta$ , avec  $a < \alpha < \beta < b$ . Dans ce cas, nous pouvons affirmer que  $\varphi_2$  possède un zéro dans l'intervalle  $]\alpha, \beta[$ , c-à-d, il existe un  $x_0 \in I$  avec  $\alpha < x_0 < \beta$  tel que  $\varphi_2(x_0) = 0$ .

Pour la démonstration de ce théorème de Sturm, nous supposons sans perte de généralité que les deux zéros  $\alpha, \beta$  de la fonction  $\varphi_1$  sont voisins, c-à-d, que  $\varphi_1$  ne possède aucun autre zéro dans l'intervalle ouvert  $]\alpha, \beta[$ . En effet, si ce théorème est démontré pour ce cas particulier de deux zéros voisins, il s'appliquera automatiquement au cas plus général de deux zéros non voisins car l'intervalle  $]\alpha, \beta[$  sera, dans ce cas, composé de plusieurs sous-intervalles qui sont bornés par des paires de zéros voisins de  $\varphi_1$  et dont chacun contiendra donc un zéro de  $\varphi_2$ . Comme la fonction  $\varphi_1$  est deux fois continûment dérivable en tant que solution à l'équation de Sturm-Liouville  $\hat{L}_1\varphi_1(x) = 0$ , elle ne peut pas changer de signe dans l'intervalle  $]\alpha, \beta[$ . Nous supposons sans perte de généralité qu'elle atteint des valeurs positives dans cet intervalle, c-à-d,  $\varphi_1(x) > 0$  pour tout  $x \in ]\alpha, \beta[$ . Comme par hypothèse nous avons  $\hat{L}_j\varphi_j(x) = 0$  pour tout  $x \in I$  et tout  $j = 1, 2$ , nous pouvons calculer par la méthode de l'intégration par partie

$$\begin{aligned} 0 &= \int_{\alpha}^{\beta} \left( \varphi_2(x)(\hat{L}_1\varphi_1)(x) - \varphi_1(x)(\hat{L}_2\varphi_2)(x) \right) dx \\ &= \int_{\alpha}^{\beta} \left( -\varphi_2(x)\frac{d}{dx} [p(x)\varphi_1'(x)] + \varphi_1(x)\frac{d}{dx} [p(x)\varphi_2'(x)] \right) dx \\ &\quad + \int_{\alpha}^{\beta} (q_1(x) - q_2(x)) \varphi_1(x)\varphi_2(x) dx \\ &= [p(x)(-\varphi_2(x)\varphi_1'(x) + \varphi_1(x)\varphi_2'(x))]_{\alpha}^{\beta} \\ &\quad + \int_{\alpha}^{\beta} (q_1(x) - q_2(x)) \varphi_1(x)\varphi_2(x) dx, \end{aligned} \quad (5.36)$$

ce qui nous donne, avec  $\varphi_1(\alpha) = \varphi_1(\beta) = 0$ , l'identité

$$p(\beta)\varphi_2(\beta)\varphi_1'(\beta) - p(\alpha)\varphi_2(\alpha)\varphi_1'(\alpha) = \int_{\alpha}^{\beta} (q_1(x) - q_2(x)) \varphi_1(x)\varphi_2(x)dx. \quad (5.37)$$

Supposons, par l'absurde, que  $\varphi_2$  ne contient aucun zéro dans l'intervalle  $]\alpha, \beta[$ . Comme cette fonction est deux fois continûment dérivable, elle ne peut pas changer de signe dans  $]\alpha, \beta[$  et doit donc être soit strictement positive soit strictement négative dans cet intervalle. Si elle est strictement positive, c-à-d,  $\varphi_2(x) > 0$  pour tout  $x \in ]\alpha, \beta[$ , le membre de droite de l'équation (5.37) est également positif, c-à-d,

$$\int_{\alpha}^{\beta} (q_1(x) - q_2(x)) \varphi_1(x)\varphi_2(x)dx > 0, \quad (5.38)$$

car  $\varphi_1(x) > 0$  pour tout  $x \in ]\alpha, \beta[$  et car, par hypothèse, les deux fonctions  $q_1$  et  $q_2$  satisfont à la propriété  $q_1(x) \geq q_2(x)$  pour tout  $x \in I$  avec l'ensemble des points  $x$  où  $q_1(x) = q_2(x)$  étant de mesure nulle dans  $I$ , ce qui implique l'existence d'au moins un intervalle ouvert de taille finie  $I' \subseteq ]\alpha, \beta[$  dans lequel nous avons  $q_1(x) > q_2(x)$  pour tout  $x \in I'$ . En revanche, le membre de gauche de l'équation (5.37) ne peut pas être positif dans ce cas, c-à-d,

$$p(\beta)\varphi_2(\beta)\varphi_1'(\beta) - p(\alpha)\varphi_2(\alpha)\varphi_1'(\alpha) \leq 0, \quad (5.39)$$

car  $p(\alpha)$  et  $p(\beta)$  sont positifs et nous devons avoir  $\varphi_1'(\alpha) > 0$  et  $\varphi_1'(\beta) < 0$  si  $\varphi_1(\alpha) = \varphi_1(\beta) = 0$  et  $\varphi_1(x) > 0$  pour tout  $x \in ]\alpha, \beta[$ , sachant que  $\varphi_1$  ne peut pas posséder des zéros non simples en tant que solution à une équation de Sturm-Liouville, comme nous l'avons déjà fait remarquer dans la section 5.1. A contrario, si  $\varphi_2(x) < 0$  pour tout  $x \in ]\alpha, \beta[$ , nous obtenons avec un raisonnement tout à fait analogue que le membre de droite de l'équation (5.37) est négatif tandis que le membre de gauche de l'équation (5.37) ne peut pas être négatif. La supposition que  $\varphi_2$  ne contienne aucun zéro dans l'intervalle  $]\alpha, \beta[$  n'est donc pas tenable, ce qui prouve le théorème de Sturm.

Un corollaire intéressant de ce théorème est obtenu pour le choix particulier d'une fonction  $q_1$  satisfaisant à  $q_1(x) \geq 0$  pour tout  $x \in I$ , de sorte que l'ensemble des zéros de  $q_1$  dans  $I$  est de mesure nulle, et de la fonction  $q_2 \equiv 0$ . En choisissant une fonction constante pour la solution non triviale de l'équation  $\hat{L}_2\varphi_2 = 0$ , nous pouvons inférer du théorème de Sturm qu'il n'existe pas de solution non triviale de l'équation  $\hat{L}_1\varphi_1 = 0$  possédant des zéros en  $x = \alpha$  et  $x = \beta$  avec  $a < \alpha < \beta < b$ . En effet, si  $\varphi_1 : I \rightarrow \mathbb{R}$  est une solution non triviale de l'équation  $\hat{L}_1\varphi_1 = 0$ , nous pouvons, en analogie avec l'équation (5.36), calculer

$$\begin{aligned} 0 &= \int_{\alpha}^{\beta} (\hat{L}_1\varphi_1)(x)dx = \int_{\alpha}^{\beta} \left( -\frac{d}{dx} [p(x)\varphi_1'(x)] + q_1(x)\varphi_1(x) \right) dx \\ &= p(\alpha)\varphi_1'(\alpha) - p(\beta)\varphi_1'(\beta) + \int_{\alpha}^{\beta} q_1(x)\varphi_1(x)dx, \end{aligned} \quad (5.40)$$

ce qui mène inévitablement à une contradiction si  $\alpha$  et  $\beta$  sont deux zéros voisins de  $\varphi_1$  car, dans ce cas, le membre de droite de l'équation (5.40) sera strictement positif si  $\varphi_1(x) > 0$  pour tout  $x \in ]\alpha, \beta[$  et strictement négatif sinon.

Une telle contradiction est aussi obtenue si la solution non triviale de l'équation  $\hat{L}_1\varphi_1 = 0$  satisfait aux conditions de Neumann,  $\varphi_1'(\alpha) = \varphi_2'(\beta) = 0$ , aux conditions mixtes,  $\varphi_1(\alpha) = \varphi_2'(\beta) = 0$  ou  $\varphi_1'(\alpha) = \varphi_2(\beta) = 0$ , ou aux conditions périodiques,  $\varphi_1(\alpha) = \varphi_2(\beta)$ ,  $\varphi_1'(\alpha) = \varphi_2'(\beta)$  et  $p(\alpha) = p(\beta)$ . La preuve de cette affirmation peut être fournie de manière tout à fait analogue au cas des conditions de Dirichlet  $\varphi_1(\alpha) = \varphi_2(\beta) = 0$  discuté ci-dessus si  $\varphi_1$  ne possède aucun zéro dans l'intervalle  $] \alpha, \beta[$ . Elle s'avère triviale si  $\varphi_1$  possède au moins deux zéros dans  $] \alpha, \beta[$  car il suffit, dans ce cas, de restreindre l'intégrale dans l'équation (5.40) à un intervalle borné par deux zéros voisins de  $\varphi_1$  afin de démontrer la contradiction. Si  $\varphi_1$  possède exactement un zéro dans l'intervalle en question, il convient de restreindre l'intégrale dans l'équation (5.40) à l'intervalle  $] \alpha, x_0[$  ou à l'intervalle  $] x_0, \beta[$  avec  $x_0 \in ] \alpha, \beta[$  étant le zéro de  $\varphi_1$ , et d'utiliser le raisonnement développé ci-dessus pour les conditions de Dirichlet (ou les conditions mixtes, dans le cas  $\varphi_1'(\alpha) = \varphi_2'(\beta) = 0$ ) avec  $\varphi_1$  étant strictement positif ou strictement négatif dans cet intervalle restreint. Notons qu'en présence des conditions périodiques ce dernier cas d'existence d'un seul zéro de  $\varphi_1$  dans  $] \alpha, \beta[$  implique forcément  $\varphi_1(\alpha) = \varphi_2(\beta) = 0$  grâce à la continuité de la fonction  $\varphi_1$  et grâce au fait que ce zéro doit être simple si  $\varphi_1$  est la solution d'une équation de Sturm-Liouville.

## 5.5 Le spectre de l'opérateur de Sturm-Liouville

Comme nous l'avons annoncé au début de la section précédente, le théorème de Sturm nous permet d'inférer davantage de propriétés caractérisant les valeurs propres et fonctions propres de l'opérateur de Sturm-Liouville. Dans l'optique de pouvoir effectuer des démonstrations sans devoir faire face à trop de complications, nous exemplifions cela, dans cette section, pour le cas particulier d'un intervalle fermé  $I = [a, b] \subset \mathbb{R}$ , avec  $-\infty < a < b < \infty$ , aux bornes duquel des conditions aux limites de Dirichlet sont imposées. Nous considérons donc l'espace de Hilbert

$$\mathcal{H} = \left\{ \psi : I \rightarrow \mathbb{C} : \int_a^b |\psi(x)|^2 dx < \infty \text{ et } \psi(a) = \psi(b) = 0 \right\} \quad (5.41)$$

sur lequel est défini l'opérateur de Sturm-Liouville

$$\hat{L} : \mathcal{D} \subset \mathcal{H} \rightarrow \mathcal{H}, \psi \mapsto \hat{L}\psi, \quad (5.42)$$

notamment par

$$(\hat{L}\psi)(x) = -\frac{d}{dx} [p(x)\psi'(x)] + q(x)\psi(x) \quad (5.43)$$

pour tout  $x \in I$ , où  $p : I \rightarrow \mathbb{R}_+$ ,  $x \mapsto p(x) > 0$  est continûment dérivable sur  $I$  et  $q : I \rightarrow \mathbb{R}$ ,  $x \mapsto q(x)$  est une fonction continue. Comme  $I = [a, b]$  est un intervalle fermé et donc de taille finie, la fonction  $q$  possède sur cet intervalle un minimum global fini, qui est nommé par

$$q_0 = \min_{x \in I} q(x). \quad (5.44)$$

Nous avons déjà constaté dans la section 5.3 que l'opérateur de Sturm-Liouville  $\hat{L}$  est auto-adjoint en présence de ces conditions aux limites de Dirichlet. Ses valeurs propres sont donc réelles et les fonctions propres de  $\hat{L}$  associées à des valeurs propres différentes sont orthogonales et peuvent être choisies telles qu'elles atteignent des valeurs purement réelles. Par le théorème de Sturm, nous pouvons aussi inférer que l'opérateur de Sturm-Liouville  $\hat{L}$  ne possède pas de valeur propre en-dessous de la borne inférieure  $q_0$  des valeurs de la fonction  $q$ . Le spectre de cet opérateur  $\hat{L}$  est donc *borné vers le bas*. En effet, si nous supposons, par l'absurde, l'existence d'une valeur propre  $\lambda$  de  $\hat{L}$  en dessous de cette borne inférieure, c-à-d,  $\lambda < q_0$ , nous pouvons obtenir une contradiction avec le corollaire que nous avons discuté à la fin de la section 5.4, notamment au travers de l'opérateur

$$\hat{L}_1 = \hat{L} - \lambda = -\frac{d}{dx}p(x)\frac{d}{dx} + q(x) - \lambda \quad (5.45)$$

pour lequel, par ce corollaire, une solution non triviale de l'équation  $\hat{L}_1\varphi_1 = 0$  satisfaisant aux conditions aux limites de Dirichlet en  $x = a$  et  $x = b$  ne peut pas exister.

Une deuxième conséquence du théorème de Sturm concerne les zéros des fonctions propres de l'opérateur de Sturm-Liouville. Soient  $\lambda_1, \lambda_2 \in \mathbb{R}$  avec  $\lambda_2 > \lambda_1$  deux valeurs propres de  $\hat{L}$  avec les fonctions propres associées  $\psi_1, \psi_2 : I \rightarrow \mathbb{R}$ , satisfaisant donc à l'équation  $\hat{L}_j\psi_j(x) = 0$  pour tout  $x \in I$  avec  $\hat{L}_j \equiv \hat{L} - \lambda_j$  pour  $j = 1, 2$ . Le théorème de Sturm, qui est appliqué ici avec le choix des fonctions  $q_j \equiv q - \lambda_j$  dans l'expression (5.35), nous permet d'inférer qu'entre chaque paire de zéros voisins de  $\psi_1$  se trouve un zéro de  $\psi_2$ . La fonction  $\psi_2$  possède donc forcément plus de zéros que  $\psi_1$ .

Afin de comprendre de manière plus détaillée combien de zéros possède quelle fonction propre de  $\hat{L}$ , nous définissons la fonction  $\psi_\lambda : I \rightarrow \mathbb{R}$ ,  $x \mapsto \psi_\lambda(x)$  telle qu'elle représente la solution unique de l'équation différentielle  $(\hat{L} - \lambda)\psi_\lambda = 0$  pour le jeu des conditions initiales  $\psi_\lambda(a) = 0$  et  $\psi'_\lambda(a) = 1$  définies en  $x = a$ . En tant que solution d'une équation de Sturm-Liouville,  $\psi_\lambda$  est une fonction continue qui ne peut pas posséder des zéros non simples. Elle est également continue par rapport au paramètre  $\lambda$  que l'on peut interpréter comme un deuxième argument de cette fonction. Nous pouvons donc conclure qu'une variation de ce paramètre  $\lambda$  ne peut mener à un changement du nombre des zéros de  $\psi_\lambda$  se trouvant dans  $I$  que si, par cette variation,  $\lambda$  passe au travers d'une valeur propre  $\lambda_n$  de  $\hat{L}$ , cette dernière étant caractérisée par la propriété  $\psi_{\lambda_n}(b) = 0$ . Autrement dit, les

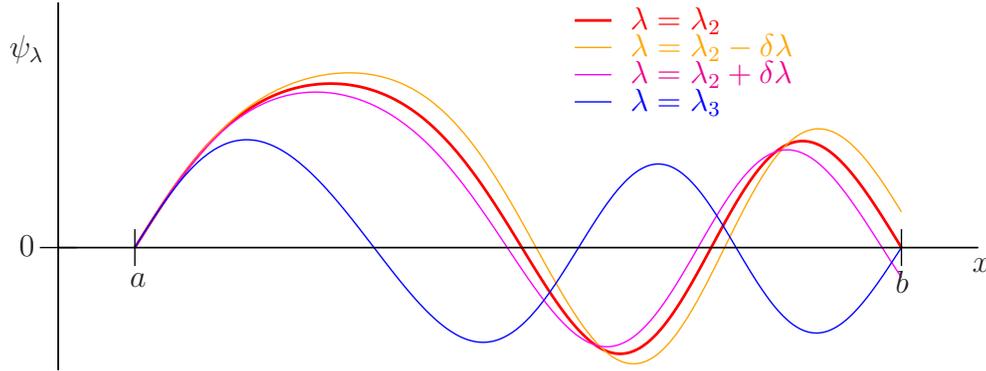


Figure 5.2: Graphiques des fonctions  $\psi_{\lambda_2}$ ,  $\psi_{\lambda_2 - \delta\lambda}$ ,  $\psi_{\lambda_2 + \delta\lambda}$  et  $\psi_{\lambda_3}$ , où  $\psi_\lambda$  avec  $\lambda \in \mathbb{R}$  est la solution unique de l'équation différentielle  $(\hat{L} - \lambda)\psi_\lambda = 0$  pour les conditions initiales  $\psi_\lambda(a) = 0$  et  $\psi'_\lambda(a) = 1$ .  $\lambda_2$  et  $\lambda_3$  sont deux valeurs propres consécutives du spectre de l'opérateur de Sturm-Liouville  $\hat{L}$ , ce dernier étant défini sur l'espace de Hilbert des fonctions de carré intégrable définies sur l'intervalle  $I = ]a, b[$  qui satisfont aux conditions aux limites de Dirichlet en  $x = a$  et  $x = b$ . Le calcul a été effectué pour le cas de l'équation de Schrödinger (5.1) en présence de l'énergie potentielle  $V(x) = 40 V_0 \sin(2\pi(x - a)/L)$  avec  $V_0 = \hbar^2/(mL^2)$  et  $L = b - a$ , où les valeurs propres associées aux deuxième et troisième état excité sont respectivement données par  $\lambda_2 \simeq 52,8739 V_0$  et  $\lambda_3 \simeq 81,5196 V_0$  et où nous avons choisi  $\delta\lambda = 2 V_0$ .

zéros de  $\psi_\lambda$  entrent un par un dans l'intervalle  $I$  par sa borne supérieure lorsque le paramètre  $\lambda$  est augmenté de manière continue, comme il est illustré dans la figure 5.2, et nous avons affaire à une valeur propre chaque fois qu'un nouveau zéro apparaît à la borne supérieure de l'intervalle.

En s'appuyant sur ce raisonnement, il est possible d'affirmer que l'opérateur de Sturm-Liouville  $\hat{L}$  possède un nombre infini de valeurs propres  $\lambda_n$  avec  $n \in \mathbb{N}_0$ . Etant liées à la présence d'un zéro de la fonction  $\psi_\lambda$  en  $x = b$  qui doit forcément être un zéro simple, ces valeurs propres sont *discrètes*, de sorte que nous pouvons les ordonner comme  $-\infty < \lambda_0 < \lambda_1 < \lambda_2 < \dots$ , c-à-d, tel que  $\lambda_i < \lambda_j$  pour tout couple  $i, j \in \mathbb{N}_0$  avec  $i < j$ . Nous pouvons aussi affirmer, au travers du comportement de la fonction  $\psi_\lambda$  par rapport à des variations du paramètre  $\lambda$ , que la fonction propre  $\psi_n$  associée à la valeur propre  $\lambda_n$  possède exactement  $n$  zéros à l'intérieur de l'intervalle  $I$ .

En considérant toujours un intervalle fermé de taille finie  $I = [a, b]$ , nous pouvons généraliser ces constats pour d'autres jeux de conditions aux limites pour lesquels l'opérateur de Sturm-Liouville  $\hat{L}$  est auto-adjoint, à savoir les conditions de Neumann, mixtes ou périodiques, ce dernier cas apportant toutefois quelques petits amendements. En présence de ces conditions aux limites, l'opérateur  $\hat{L}$  possède un nombre infini de valeurs propres  $\lambda_n \in \mathbb{R}$  avec  $n \in \mathbb{N}_0$ . Comme pour

les conditions aux limites de Dirichlet, le spectre de ces valeurs propres est discret et borné par dessous. Plus précisément, aucune valeur propre ne peut exister en-dessous de la valeur minimale  $q_0$  de la fonction  $q$  dans l'intervalle  $I$ , suivant le raisonnement que nous avons effectué à la fin de la section 5.4. Nous pouvons donc, de manière très générale, ordonner ces valeurs propres tel que nous avons  $-\infty < \lambda_0 < \lambda_1 < \lambda_2 < \dots$ , c-à-d, tel que  $\lambda_i < \lambda_j$  pour tout couple  $i, j \in \mathbb{N}_0$  avec  $i < j$ . La fonction propre  $\psi_n$  de  $\hat{L}$  qui est associée à la valeur propre  $\lambda_n$  possède  $n$  zéros à l'intérieur de l'intervalle  $I$ .

Une exception notable à cette règle générale apparaît dans le cas des conditions aux limites périodiques. Dans ce cas,  $\psi_{2l-1}$  possède généralement, tout comme  $\psi_{2l}$ ,  $2l$  zéros à l'intérieur de  $I$  pour tout  $l \in \mathbb{N}$ , sauf si un de ces zéros coïncide avec les bornes de cet intervalle. Dans le cas particulier d'un opérateur de Sturm-Liouville défini avec les fonctions constantes  $p(x) = p_0 > 0$  et  $q(x) = q_0$  pour tout  $x \in I$  en combinaison avec les conditions aux limites périodiques, une *dégénérescence* surgit entre des paires des valeurs propres de  $\hat{L}$ , de sorte que nous avons  $-\infty < \lambda_0 < \lambda_1 = \lambda_2 < \lambda_3 = \lambda_4 < \dots$ , ces paires de valeurs propres  $\lambda_{2l-1} = \lambda_{2l}$  étant associées aux fonctions propres  $\psi_{2l-1}(x) = \sin(k_l x)$  et  $\psi_{2l}(x) = \cos(k_l x)$  avec  $k_l = 2l\pi/(b-a)$  pour tout  $l \in \mathbb{N}$ .

Comme nous l'avons déjà mentionné ci-dessus, les fonctions propres  $\psi_n$  associées à des valeurs propres différentes sont orthogonales. Par une simple adaptation de leurs préfacteurs (combinée, dans le cas des dégénérescences, avec des procédures d'orthogonalisation appliquées aux sous-espaces propres de dimension supérieure à 1), elles peuvent donc être choisies telles qu'elles satisfont à

$$\langle \psi_n | \psi_{n'} \rangle = \int_a^b \psi_n^*(x) \psi_{n'}(x) dx = \delta_{nn'} \quad (5.46)$$

pour tout couple  $n, n' \in \mathbb{N}_0$ . Il est aussi possible, quoique techniquement plus compliqué, de montrer que l'ensemble de ces fonctions propres ainsi normalisées  $(\psi_n)_{n \in \mathbb{N}_0}$  forme un *système complet*, c-à-d, une base dans l'espace de Hilbert  $\mathcal{H}$  constitué des fonctions de carré intégrables qui sont définies sur l'intervalle  $I$  et satisfont aux deux conditions aux limites imposées. Cela implique que nous avons pour toute fonction  $f \in \mathcal{H}$

$$\lim_{N \rightarrow \infty} \left\| f(x) - \sum_{n=0}^N f_n \psi_n(x) \right\| = 0 \quad (5.47)$$

pour tout  $x \in I$  avec

$$f_n = \langle \psi_n | f \rangle = \int_a^b \psi_n^*(x) f(x) dx, \quad (5.48)$$

la convergence dans l'expression (5.47) étant uniforme sur l'intervalle  $I$ . Nous

pouvons ainsi représenter cette fonction dans la base  $(\psi_n)_{n \in \mathbb{N}_0}$  par

$$f(x) = \sum_{n=0}^N \langle \psi_n | f \rangle \psi_n(x) \quad (5.49)$$

pour tout  $x \in I$ .

La généralisation de ces considérations pour le cas d'un intervalle ouvert  $I = ]a, b[ \subseteq \mathbb{R}$  s'avère assez compliquée et n'est pas discutée en détail dans le cadre de ce cours. Une première complication est apportée par le fait que les valeurs de la fonction  $q$  ne possèdent pas forcément une borne inférieure sur un intervalle ouvert, même si cet intervalle est toujours de taille finie. L'existence d'une valeur propre minimale de l'opérateur de Sturm-Liouville n'est donc plus garantie par le théorème de Sturm dans ce cas. Notons que ce constat n'exclut pas strictement l'existence des opérateurs de Sturm-Liouville qui possèdent une valeur propre minimale malgré le fait que les valeurs de leur fonction  $q$  n'ont pas de borne inférieure, comme montre l'exemple de l'équation de Schrödinger (5.1) en présence du potentiel coulombien

$$V(r) = -\frac{e^2}{4\pi\epsilon_0 r} \quad (5.50)$$

défini sur  $I = \mathbb{R}_+$ , qui décrit l'interaction entre l'électron et le proton dans l'atome d'hydrogène. Mais il existe bien d'autres exemples pour des potentiels pour lesquels l'hamiltonien de l'équation de Schrödinger (5.1) ne possède pas d'état fondamental, comme p.ex.  $V(x) = -\alpha/x^2$  défini sur  $\mathbb{R}_+$  ou  $V(x) = -\alpha x^2$  défini sur  $\mathbb{R}$ , avec  $\alpha > 0$ .

Si la fonction  $q$  caractérisant l'opérateur de Sturm-Liouville  $\hat{L}$  possède effectivement un infimum dans l'intervalle ouvert  $I$ , ce qui garantit l'existence d'une borne inférieure pour les valeurs propres de cet opérateur, nous pouvons rencontrer une autre complication par le fait que l'ensemble  $(\psi_n)$  des fonctions propres associées à ces valeurs propres ne forme pas forcément un système complet dans l'espace de Hilbert constitué des fonctions qui sont de carré intégrables sur  $I$ . Toujours dans le cadre de l'équation de Schrödinger, nous pouvons fournir un exemple pour ce cas de figure par le potentiel de *Lennard-Jones*

$$V(r) = V_0 \left[ \left( \frac{r_0}{r} \right)^{12} - \left( \frac{r}{r_0} \right)^6 \right] \quad (5.51)$$

défini sur  $I = \mathbb{R}_+$ , qui représente un modèle heuristique pour l'énergie potentielle d'interaction entre deux atomes neutres de la même espèce en fonction de leur distance  $r$ , où l'échelle de longueur caractéristique  $r_0$  peut être interprétée comme le diamètre effectif de l'atome en question. Pour le choix spécifique  $V_0 = 3000 \hbar^2 / (mr_0^2)$ , l'hamiltonien qui correspond à ce potentiel (5.51) possède, comme indiqué dans la figure 5.3(a), au total 8 valeurs propres dont les fonctions propres associées sont de carré intégrables sur  $\mathbb{R}_+$ . L'ensemble de ces fonctions propres ne peut donc pas former une base de l'espace de Hilbert défini

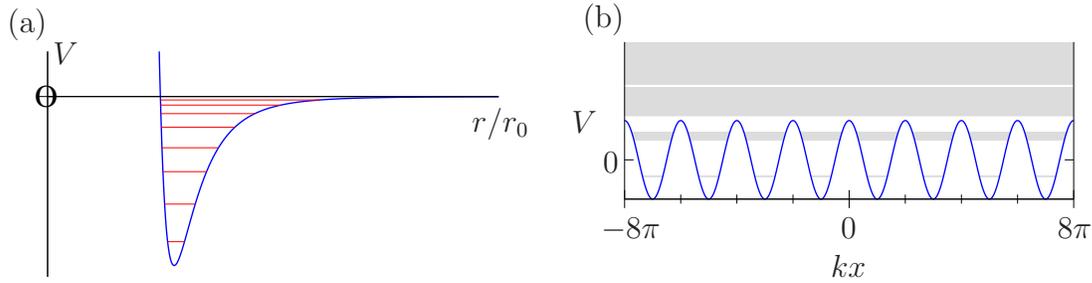


Figure 5.3: Deux exemples de potentiels dans le cadre de l'équation de Schrödinger qui n'admettent pas de spectre de valeurs propres discrètes dont les fonctions propres associées forment une base complète de l'espace de Hilbert des fonctions de carré intégrables. (a) Potentiel de *Lennard-Jones* (5.51), représentant un modèle pour le potentiel d'interaction entre deux atomes en fonction de leur distance  $r$ . L'hamiltonien correspondant à ce potentiel possède, pour le choix spécifique  $V_0 = 3000 \hbar^2 / (mr_0^2)$ , au total 8 valeurs propres (indiquées en rouge) dont les fonctions propres associées sont de carré intégrables sur l'intervalle  $]0, \infty[$ . (b) Potentiel périodique (5.53) représentant un modèle unidimensionnel pour un matériau cristallin. L'hamiltonien correspondant à ce potentiel ne possède aucune valeur propre dont la fonction propre associée est de carré intégrable sur  $I = \mathbb{R}$ . En revanche, il possède des bandes de valeurs propres appartenant à un spectre continu, dont les fonctions propres associées sont quasi périodiques et satisfont à la propriété  $\psi(x + 2\pi/k) = \psi(x) \exp(i\alpha)$  pour tout  $x \in I$  avec une constante  $\alpha \in \mathbb{R}$ . Ces bandes sont indiquées par les couches grises pour le choix spécifique  $V_0 = 0.625 \hbar^2 k^2 / m$ .

sur l'intervalle  $I = \mathbb{R}_+$ . Ce problème apparaît aussi dans le cadre de l'atome d'hydrogène où l'ensemble des fonctions propres de l'hamiltonien du potentiel coulombien (5.50) ne représente pas un système complet dans l'espace de Hilbert, malgré le fait que le nombre de ses valeurs propres est infini.

Dans le cadre de la théorie spectrale, ce problème est techniquement résolu en introduisant la notion du *spectre continu*. À l'opposé du spectre discret rassemblant les valeurs propres de l'opérateur de Sturm-Liouville  $\hat{L}$  dont les fonctions propres sont de carré intégrables, ce spectre continu est constitué des valeurs  $\lambda \in \mathbb{R}$  pour lesquelles l'équation de Sturm-Liouville

$$(\hat{L} - \lambda)\psi = 0 \quad (5.52)$$

admet une solution non triviale sous forme d'une fonction  $\psi : I \rightarrow \mathbb{C}$  qui n'est pas de carré intégrable et ne fait donc pas partie de l'espace de Hilbert  $\mathcal{H}$ , mais qui peut être approchée par une suite des fonctions issues de  $\mathcal{H}$  qui converge de manière ponctuelle en tout  $x \in I$ . De telles fonctions  $\psi$  font partie de ce qu'on appelle le *bord* de l'espace de Hilbert  $\mathcal{H}$ . Elles peuvent être utilisées pour compléter la base des fonctions propres de l'opérateur de Sturm-Liouville dans le

cadre d'un espace de Hilbert élargi qui inclut le bord de  $\mathcal{H}$ .

En général, les valeurs propres  $\lambda$  associées à de telles solutions non triviales de l'équation de Sturm-Liouville (5.52) issues du bord de l'espace de Hilbert  $\mathcal{H}$  n'apparaissent pas de manière isolée comme c'est le cas pour un spectre discret. Concrètement, dans les cas du potentiel coulombien (5.50) ainsi que du potentiel de Lennard-Jones (5.51), toutes les valeurs positives  $\lambda \in \mathbb{R}_+$  représentent des valeurs propres dans ce sens, ce qui justifie effectivement l'appellation "continu" pour ce spectre. Un autre cas de figure apparaît pour le potentiel périodique

$$V(x) = V_0 \cos(kx) \quad (5.53)$$

défini sur l'intervalle  $I = \mathbb{R}$ , en présence duquel l'hamiltonien de l'équation de Schrödinger (5.1) ne possède aucune valeur propre dont la fonction propre associée est de carré intégrable sur  $\mathbb{R}$ . En revanche, il existe pour ce potentiel un nombre infini de *bandes* de valeurs propres appartenant à un spectre continu, dont les fonctions propres associées sont issues du bord de l'espace de Hilbert<sup>3</sup>. Ces bandes sont séparées par des interstices, qui sont aussi appelés *bandes interdites* (ou *band gaps* en anglais), dans lesquels aucune solution non triviale de l'équation de Sturm-Liouville (5.52) (avec  $\hat{L}$  étant l'hamiltonien du potentiel périodique) ne peut exister ni dans l'espace de Hilbert ni dans son bord.

Notons enfin, en guise de remarque, qu'il existe aussi la notion d'un *spectre fractal* qui se situe conceptuellement entre les notions de spectre discret et de spectre continu. Une discussion approfondie de toutes ces notions constitue un cours à part entière et dépasse le cadre de ces notes.

## 5.6 La fonction de Green

En mécanique quantique, la nécessité de connaître les propriétés spectrales de l'hamiltonien décrivant une particule quantique apparaît concrètement lorsqu'il s'agit de résoudre l'équation de Schrödinger dépendant du temps (3.9), cette dernière étant définie pour un potentiel qui ne dépend pas du temps. En présence d'une spécification du profil spatial initial de la fonction d'onde  $\psi$ , cette résolution peut être entamée par une transformation de Laplace de  $\psi$  par rapport au temps. Suivant la procédure que nous avons décrite dans la section 2.4 (cf. équation (2.79)), cette transformation de Laplace transforme la dérivée temporelle de  $\psi$ , constituant le membre de gauche de l'équation (3.9), en une simple multiplication de la fonction d'onde par la coordonnée réciproque, cette dernière étant effectivement donnée par l'énergie de la particule en question. De ce produit est soustrait le profil initial de  $\psi$ . Nous obtenons ainsi une équation de Schrödinger stationnaire qui est munie d'un terme inhomogène, ce dernier étant constitué du profil initial de  $\psi$ .

---

<sup>3</sup>Ces fonctions propres  $\psi$  sont *quasi périodiques* dans le sens qu'il existe pour chacune d'elles un  $\alpha \in \mathbb{R}$  tel que  $\psi(x + 2\pi/k) = \psi(x) \exp(i\alpha)$  pour tout  $x \in I$ .

D'un point de vue plus général, il convient donc de discuter la résolution de l'équation de Sturm-Liouville en présence d'un terme inhomogène. Afin d'éviter les complications que nous venons d'évoquer à la fin de la section précédente, nous restreignons cette discussion au cas d'un intervalle fermé  $I = [a, b]$  avec  $-\infty < a < b < \infty$ . Sur cet intervalle est défini l'espace de Hilbert

$$\mathcal{H} = \left\{ \psi : I \rightarrow \mathbb{C} : \int_a^b |\psi(x)|^2 dx < \infty \text{ et } (\Lambda_1 \psi)_{a,b} = (\Lambda_2 \psi)_{a,b} = 0 \right\} \quad (5.54)$$

constitué des fonctions qui sont de carré intégrables sur  $I$  et satisfont aux conditions aux limites homogènes  $(\Lambda_j \psi)_{a,b} = 0$  pour  $j = 1, 2$ . Ces conditions aux limites sont telles que l'opérateur de Sturm-Liouville

$$\hat{L} = -\frac{d}{dx} p(x) \frac{d}{dx} + q(x) \quad (5.55)$$

est auto-adjoint par rapport au produit scalaire canonique. Nous considérons, plus particulièrement, les conditions aux limites pour lesquelles nous avons discuté en détail les propriétés spectrales de l'opérateur  $\hat{L}$  dans la section précédente, à savoir les conditions de Dirichlet, de Neumann, mixtes ou périodiques.

Notre but dans cette section est de déterminer pour un  $\lambda \in \mathbb{C}$  la solution de l'équation différentielle inhomogène

$$(\lambda - \hat{L})\psi(x) = f(x) \quad (5.56)$$

satisfaisant aux conditions aux limites  $(\Lambda_1 \psi)_{a,b} = (\Lambda_2 \psi)_{a,b} = 0$ , où  $f : I \rightarrow \mathbb{C}$  est une fonction continue. Pour ce faire, nous pouvons nous inspirer de la procédure qui a été détaillée dans la deuxième partie de la section 3.6 et décomposer le terme inhomogène suivant l'équation (3.81) comme

$$f(x) = \int_a^b f(x') \delta(x - x') dx', \quad (5.57)$$

où  $\delta$  représente la distribution de Dirac. La solution recherchée de l'équation différentielle (5.56) est donc, suivant l'expression (3.85), donnée par

$$\psi(x) = \int_a^b f(x') G(x, x') dx' \quad (5.58)$$

où

$$G : I \times I \rightarrow \mathbb{C}, (x, x') \mapsto G(x, x') \quad (5.59)$$

est la fonction de Green associée à l'équation (5.56). Cette fonction de Green satisfait à l'équation différentielle inhomogène

$$(\lambda - \hat{L})G(x, x') = \left( \lambda + \frac{\partial}{\partial x} p(x) \frac{\partial}{\partial x} - q(x) \right) G(x, x') = \delta(x - x') \quad (5.60)$$

ainsi qu'aux conditions aux limites

$$(\Lambda_1 G)_{x=a, x=b} = (\Lambda_2 G)_{x=a, x=b} = 0 \quad (5.61)$$

en ce qui concerne son premier argument.

Nous pouvons maintenant faire valoir le travail que nous avons effectué dans la section 5.5 et utiliser le fait que l'ensemble des fonctions propres de l'opérateur de Sturm-Liouville défini sur un intervalle fermé en présence des conditions aux limites de Dirichlet, de Neumann, mixtes ou périodiques forme une base dans l'espace de Hilbert (5.54). Cette *base propre*  $(\psi_n)_{n \in \mathbb{N}_0}$  de l'opérateur  $\hat{L}$  est donc constituée de ses fonctions propres  $\psi_n : I \rightarrow \mathbb{C}, x \mapsto \psi_n(x)$  satisfaisant, pour tout  $n \in \mathbb{N}_0$ , à

$$\hat{L}\psi_n(x) = \lambda_n \psi_n(x) \quad (5.62)$$

pour tout  $x \in I$  avec les valeurs propres associées  $\lambda_n \in \mathbb{R}$ . Comme les fonctions propres associées à des valeurs propres différentes sont orthogonales, nous pouvons, par une adaptation des préfacteurs des fonctions  $\psi_n$ , choisir cette base telle qu'elle soit orthonormée, c-à-d, telle que nous ayons  $\int_a^b \psi_n^*(x) \psi_{n'}(x) dx = \delta_{nn'}$  pour tout couple  $n, n' \in \mathbb{N}_0$ .

La solution de l'équation (5.60) peut être déterminée en décomposant cette équation dans la base propre orthonormée de l'opérateur  $\hat{L}$ . En posant

$$G(x, x') = \sum_{n=0}^{\infty} g_n(x') \psi_n(x) \quad (5.63)$$

avec

$$g_n(x') = \int_a^b \psi_n^*(x) G(x, x') dx \quad (5.64)$$

pour tout  $n \in \mathbb{N}_0$  et tout  $x' \in I$ , et en utilisant la décomposition

$$\delta(x - x') = \sum_{n=0}^{\infty} \psi_n^*(x') \psi_n(x) \quad (5.65)$$

de la fonction delta dans la base propre de  $\hat{L}$ , qui résulte de l'identité

$$\int_a^b \psi_n^*(x) \delta(x - x') dx = \psi_n^*(x') \quad (5.66)$$

pour tout  $n \in \mathbb{N}_0$  et tout  $x' \in I$ , nous pouvons réécrire l'équation (5.60) comme

$$\begin{aligned} 0 &= (\lambda - \hat{L})G(x, x') - \delta(x - x') \\ &= \sum_{n=0}^{\infty} \left( g_n(x') (\lambda - \hat{L})\psi_n(x) - \psi_n^*(x') \psi_n(x) \right) \\ &= \sum_{n=0}^{\infty} [(\lambda - \lambda_n) g_n(x') - \psi_n^*(x')] \psi_n(x) \end{aligned} \quad (5.67)$$

en utilisant la relation (5.62). Comme  $(\psi_n)_{n \in \mathbb{N}_0}$  est un ensemble de fonctions qui sont linéairement indépendantes, cette relation (5.67) est satisfaite si et seulement si nous avons

$$g_n(x') = \frac{\psi_n^*(x')}{\lambda - \lambda_n} \quad (5.68)$$

pour tout  $x' \in I$  et tout  $n \in \mathbb{N}_0$ . Nous obtenons ainsi l'expression explicite

$$G(x, x') = \sum_{n=0}^{\infty} \frac{\psi_n^*(x')\psi_n(x)}{\lambda - \lambda_n} \quad (5.69)$$

pour la fonction de Green associée à l'équation de Sturm-Liouville inhomogène (5.56). La résolution de cette équation peut donc être réalisée en pratique par la *diagonalisation* de l'opérateur de Sturm-Liouville, qui consiste en la détermination de tous ses valeurs propres  $\lambda_n$  et les fonctions propres  $\psi_n$  associées.

L'approche décrite ci-dessus échoue généralement si  $\lambda = \lambda_n$  pour un  $n \in \mathbb{N}_0$ . Dans ce cas-là, l'équation différentielle (5.56) n'admet pas de solution respectant les conditions aux limites imposées si la décomposition de la fonction  $f$  dans la base propre de  $\hat{L}$  fournit une composante non nulle par rapport à  $\psi_n$ , c-à-d, si  $\int_a^b \psi_n^*(x)f(x)dx \neq 0$ . Il est à noter que ce cas de figure n'apparaît pas dans le cadre de la résolution de l'équation de Schrödinger dépendant du temps par une transformation de Laplace car la contrainte de convergence de l'intégrale (2.68) définissant cette transformation induit, en pratique, la présence d'une partie imaginaire positive dans le paramètre  $\lambda$ , qui ne coïncide donc pas avec une valeur propre de l'hamiltonien.

## 5.7 Les polynômes classiques

La théorie de Sturm-Liouville fournit non seulement le cadre adapté pour résoudre l'équation de Schrödinger, notamment au travers du calcul des valeurs propres et vecteurs propres de l'hamiltonien, comme nous venons de le discuter. Elle permet aussi de déterminer des solutions polynomiales à des équations différentielles linéaires du deuxième ordre dont les coefficients sont des polynômes. Considérons, concrètement, l'équation différentielle

$$Q(x)\psi''(x) + l(x)\psi'(x) + \lambda\psi(x) = 0 \quad (5.70)$$

avec

$$Q(x) = q_0 + q_1x + q_2x^2, \quad (5.71)$$

$$l(x) = l_0 + l_1x, \quad (5.72)$$

où  $q_0, q_1, q_2, l_0, l_1, \lambda \in \mathbb{R}$  sont des constantes réelles. Les degrés maximaux de ces polynômes (5.71) et (5.72) ont été judicieusement choisis de sorte qu'une solution

polynomiale à l'équation différentielle (5.70) puisse effectivement exister. En effet, si nous ajoutons p.ex. un terme cubique à l'expression (5.71) pour le polynôme  $Q$ , le degré du polynôme correspondant au terme  $Q(x)\psi''(x)$  serait supérieur au degré du polynôme  $\psi$  ainsi qu'à celui du polynôme correspondant au terme  $l(x)\psi'(x)$ , ce qui excluerait l'existence d'une solution polynomiale à l'équation (5.70).

Posons alors le polynôme du degré  $n \in \mathbb{N}_0$

$$\psi(x) = \sum_{j=0}^n C_j x^j, \quad (5.73)$$

avec  $C_n \neq 0$ , comme solution à l'équation différentielle (5.70). En évaluant sa première et deuxième dérivée comme

$$\psi'(x) = \sum_{j=1}^n j C_j x^{j-1} = \sum_{j=0}^{n-1} (j+1) C_{j+1} x^j, \quad (5.74)$$

$$\begin{aligned} \psi''(x) &= \sum_{j=2}^n j(j-1) C_j x^{j-2} = \sum_{j=1}^{n-1} (j+1)j C_{j+1} x^{j-1} \\ &= \sum_{j=0}^{n-2} (j+2)(j+1) C_{j+2} x^j \end{aligned} \quad (5.75)$$

et en injectant ces expressions (5.73–5.75) dans l'équation (5.70), nous obtenons l'identité

$$\begin{aligned} &\sum_{j=0}^{n-2} [(j+2)(j+1)q_0 C_{j+2} + (j+1)(jq_1 + l_0) C_{j+1} + (j(j-1)q_2 + jl_1 + \lambda) C_j] x^j \\ &+ [n((n-1)q_1 + l_0) C_n + ((n-1)(n-2)q_2 + (n-1)l_1 + \lambda) C_{n-1}] x^{n-1} \\ &+ [n(n-1)q_2 + nl_1 + \lambda] C_n x^n = 0 \end{aligned} \quad (5.76)$$

pour tout  $x \in \mathbb{R}$ . Comme le membre de gauche de l'équation (5.76) représente effectivement un polynôme de degré  $n$  qui est identique à la fonction nulle, tous les coefficients devant les monômes  $x^j$  pour  $j = 0, \dots, n$  doivent s'annuler dans l'équation (5.76). Cela concerne, en particulier, le coefficient  $[n(n-1)q_2 + nl_1 + \lambda] C_n$  devant le monôme  $x^n$ . Comme  $C_n \neq 0$  par hypothèse pour un polynôme de degré  $n$ , nous pouvons conclure qu'une solution non triviale de l'équation (5.70) sous forme d'un polynôme existe si et seulement si nous avons  $\lambda = \lambda_n$  pour un  $n \in \mathbb{N}_0$ , avec

$$\lambda_n = -n((n-1)q_2 + l_1). \quad (5.77)$$

Il est, en principe, possible de construire le polynôme qui, dans ce cas, résout l'équation différentielle (5.70) en exploitant l'annulation des autres coefficients devant les monômes  $x^j$  avec  $j < n$ . Si nous choisissons, sans perte de généralité,  $C_n = 1$  (sachant que de toute manière la solution polynomiale non triviale de l'équation (5.70) est unique à un préfacteur global près), nous obtenons par l'annulation du coefficient devant le monôme  $x^{n-1}$  dans l'équation (5.76) l'expression pour  $C_{n-1}$ . Nous pouvons ensuite procéder de manière récursive et déterminer l'expression pour  $C_j$  par les expressions déjà calculées pour  $C_{j+1}$  et  $C_{j+2}$ , en exploitant l'annulation du coefficient devant  $x^j$  pour tout  $j = n-2, \dots, 0$ . Cette méthode fonctionnant a priori sans faille, elle est assez fastidieuse à mettre en évidence et apporte peu d'éclaircissement en ce qui concerne les caractéristiques des polynômes résultant de ce protocole de construction.

Nous poursuivons donc, dans cette section, une autre approche pour tenter de déterminer le polynôme  $\psi_n$  qui résout l'équation (5.70) si  $\lambda = \lambda_n$  pour un  $n \in \mathbb{N}_0$ , c-à-d, l'équation

$$Q(x)\psi_n''(x) + l(x)\psi_n'(x) + \lambda_n\psi_n(x) = 0. \quad (5.78)$$

Cette approche s'inspire de la manière dont nous avons introduit, dans la section 5.1, l'équation de Sturm-Liouville (5.7) à partir d'une expression assez générale (5.2) pour une équation différentielle linéaire du deuxième ordre. En effet, nous pouvons réécrire l'équation (5.78) sous la forme

$$\begin{aligned} \lambda_n\psi_n(x) &= \left(-Q(x)\frac{d^2}{dx^2} - l(x)\frac{d}{dx}\right)\psi_n(x) = Q(x)\left(-\frac{d^2}{dx^2} - \frac{l(x)}{Q(x)}\frac{d}{dx}\right)\psi_n(x) \\ &= \frac{Q(x)}{p(x)}\left(-\frac{d}{dx}p(x)\frac{d}{dx}\right)\psi_n(x), \end{aligned} \quad (5.79)$$

où nous avons, en parfaite analogie avec la définition (5.3) dans la section 5.1, introduit la fonction

$$p(x) = p_0 \exp\left(\int_{x_1}^x \frac{l(x')}{Q(x')} dx'\right) \quad (5.80)$$

pour une constante  $p_0 > 0$  et un  $x_1 \in \mathbb{R}$  choisi tel que l'intégrale constituant l'argument de la fonction exponentielle existe. Définie selon l'équation (5.80), cette fonction  $p$  est par construction analytique et n'atteint que de valeurs positives. Elle est donc parfaitement qualifiée pour assurer que

$$\hat{L} = -\frac{d}{dx}p(x)\frac{d}{dx} \quad (5.81)$$

est un opérateur de Sturm-Liouville selon la définition de ce concept dans la section 5.3, ceci notamment pour le cas particulier que la fonction  $q$  constituant le second terme de cet opérateur suivant l'expression (5.21) est identique à la fonction nulle. La relation (5.79) peut donc être réécrite sous la forme

$$(\hat{L}\psi_n)(x) = \lambda_n w(x)\psi_n(x) \quad (5.82)$$

avec

$$w(x) = \frac{p(x)}{Q(x)}. \quad (5.83)$$

L'étape suivante consiste à identifier un intervalle  $I = ]a, b[ \subseteq \mathbb{R}$  qui est tel que  $Q(x) > 0$  pour tout  $x \in I$  et que nous avons

$$\lim_{x \rightarrow a} p(x) = \lim_{x \rightarrow b} p(x) = 0. \quad (5.84)$$

Cette dernière contrainte peut effectivement être satisfaite si

$$\lim_{x \rightarrow a} \int_{x_1}^x \frac{l(x')}{Q(x')} dx' = \lim_{x \rightarrow b} \int_{x_1}^x \frac{l(x')}{Q(x')} dx' = -\infty, \quad (5.85)$$

ce qui implique que chacune des deux bornes  $a$  et  $b$  de l'intervalle  $I$  doit soit représenter un zéro du polynôme  $Q$  (tant que ce zéro-là ne coïncide pas avec celui du polynôme  $l$ ) soit être identifiée à  $\pm\infty$ . Nous pouvons ainsi définir sur cet intervalle  $I$  l'espace de Hilbert

$$\mathcal{H} = \{f : I \rightarrow \mathbb{C} : \|f\| = \sqrt{\varphi(f, f)} < \infty\} \quad (5.86)$$

qui est muni du produit scalaire

$$\varphi(\phi, \psi) = \int_a^b \phi^*(x)\psi(x)w(x)dx, \quad (5.87)$$

ce dernier étant caractérisé par la fonction de pondération  $w : I \rightarrow \mathbb{R}_+, x \mapsto w(x)$  qui est définie par l'expression (5.83) et qui n'atteint que des valeurs positives sur  $I$  grâce à la manière dont nous avons construit cet intervalle.

Suivant le raisonnement que nous avons développé dans la section 5.3, nous pouvons affirmer que l'opérateur

$$\tilde{L} \equiv \frac{1}{w(x)} \hat{L} \quad (5.88)$$

est auto-adjoint sur cet espace de Hilbert, grâce aux relations (5.32) et (5.33) et grâce au fait que la condition (5.23) pour que  $\hat{L}$  soit auto-adjoint par rapport au produit scalaire canonique est satisfaite au travers des conditions de bord (5.84) de la fonction  $p$ . Les polynômes recherchés  $\psi_n$ , satisfaisant à l'équation (5.82), sont donc des fonctions propres de  $\tilde{L}$  aux valeurs propres associées  $\lambda_n$ . Comme nous l'avons souligné à la fin de la section 5.3, ils sont, par conséquent, orthogonaux par rapport au produit scalaire (5.87), c-à-d, nous avons

$$\varphi(\psi_n, \psi_{n'}) = \int_a^b \psi_n^*(x)\psi_{n'}(x)w(x)dx = 0 \quad (5.89)$$

pour tout couple  $n, n' \in \mathbb{N}_0$  avec  $n \neq n'$ . Par le théorème de Sturm, nous pouvons aussi inférer qu'entre deux zéros voisins du polynôme  $\psi_n$  se trouve un zéro du polynôme  $\psi_{n'}$  si  $\lambda_n < \lambda_{n'}$ .

La tâche de calculer les solutions polynomiales à l'équation différentielle (5.70) est donc ramenée à la détermination des *polynômes orthogonaux* qui sont associés au produit scalaire (5.87) caractérisé par la fonction de pondération (5.83) et la spécification des bornes  $a$  et  $b$  de l'intervalle  $I \subseteq \mathbb{R}$ . Le protocole général pour construire ces polynômes orthogonaux a été discuté en détail dans la section (4.5). Nous pouvons donc effectivement résoudre le problème auquel cette section est consacrée par l'application d'une simple relation de récursion à trois termes de la forme (4.81), pourvu que cette relation soit connue pour les polynômes orthogonaux en question.

Considérons, comme premier exemple, le cas d'une fonction  $Q$  qui est constante et vaut  $Q(x) = 1$  pour tout  $x \in \mathbb{R}$ . En anticipant que la dérivée de la fonction linéaire  $l$  doit être négative afin d'assurer l'existence d'un intervalle  $I$  aux bornes duquel la fonction  $p$ , définie par l'expression (5.80), s'annule, nous choisissons, sans perte de généralité,  $l(x) = -2x$  pour tout  $x \in \mathbb{R}$ , sachant que la multiplication de ce terme par un préfacteur positif et l'ajout d'une constante ne font que redéfinir l'échelle et l'origine de l'abscisse associée à la coordonnée  $x$ . L'équation différentielle (5.70) s'écrit ainsi

$$\psi''(x) - 2x\psi'(x) + \lambda\psi(x) = 0 \quad (5.90)$$

pour ce cas.

Comme nous avons  $q_2 = 0$  et  $l_1 = -2$  selon les définitions générales (5.71) et (5.72) des fonction  $Q$  et  $l$ , des solutions polynomiales non triviales de l'équation (5.90) existent, suivant l'équation (5.77), pour les valeurs propres  $\lambda = \lambda_n = 2n$  avec  $n \in \mathbb{N}_0$ . En choisissant  $p_0 = 1$  et  $x_1 = 0$ , les fonctions  $p$  et  $w$ , définies respectivement par les équations (5.80) et (5.83), sont obtenues comme

$$w(x) = p(x) = \exp(-x^2) \quad (5.91)$$

pour tout  $x \in \mathbb{R}$ . L'intervalle  $I$  est donc logiquement identifié à  $] -\infty, \infty[$ . Sans faire davantage de calcul, nous pouvons déjà à ce stade conclure que ce sont les *polynômes d'Hermite*  $H_n$ , c-à-d, les polynômes orthogonaux associé à l'intervalle  $I = \mathbb{R}$  et à la fonction de pondération  $w(x) = \exp(-x^2)$ , qui résolvent l'équation différentielle (5.90) pour  $\lambda = 2n$  avec  $n \in \mathbb{N}_0$ .

Dans le cas d'une fonction  $Q$  qui est linéaire en  $x$  et vaut  $Q(x) = x$  pour tout  $x \in \mathbb{R}$ , nous pouvons sans perte de généralité choisir  $l(x) = l_0 - x$  pour tout  $x \in \mathbb{R}$ , sachant que la dérivée de  $l$  doit être négative afin d'assurer l'existence d'un intervalle dans lequel la fonction de pondération  $w$  est positive et aux bornes duquel  $p$  s'annule. En substituant  $l_0 = \nu + 1$ , nous obtenons ainsi l'équation différentielle

$$x\psi''(x) + (\nu + 1 - x)\psi'(x) + \lambda\psi(x) = 0 \quad (5.92)$$

pour laquelle, suivant l'équation (5.77) en combinaison avec les valeurs  $q_2 = 0$  et  $l_1 = -1$  pour ce cas-ci, des solutions polynomiales non triviales existent si  $\lambda = \lambda_n = n$  avec  $n \in \mathbb{N}_0$ . Nous évaluons pour ce cas, suivant l'équation (5.80)

$$p(x) = p_0 \exp \left( \int_{x_1}^x \frac{\nu + 1 - x'}{x'} dx' \right) = x^{\nu+1} e^{-x}, \quad (5.93)$$

en choisissant  $x_1 = 1$  et  $p_0 = e^{-1}$ , et ainsi, suivant l'équation (5.83),

$$w(x) = x^\nu e^{-x}. \quad (5.94)$$

L'intervalle  $I$  dans lequel  $w$  est positif et aux bornes duquel  $p$  s'annule est donc logiquement identifié à  $\mathbb{R}_+$ , avec la contrainte  $\nu > -1$ . Nous obtenons ainsi les *polynômes de Laguerre généralisés*  $L_n^{(\nu)}$ , définis par l'expression (4.118), comme solutions polynomiales à l'équation (5.92) pour  $\lambda = n \in \mathbb{N}_0$ .

Un intervalle de taille fini peut résulter du choix d'un polynôme quadratique pour  $Q$ , notamment si ce dernier atteint des valeurs positives sur cet intervalle et s'annule à ses bornes. Considérons donc le cas du polynôme  $Q(x) = 1 - x^2$  pour tout  $x \in \mathbb{R}$ . En définissant  $\mu = (l_0 - l_1)/2 - 1$  et  $\nu = -(l_0 + l_1)/2 - 1$ , nous obtenons avec l'expression générale (5.72) pour le polynôme  $l$  l'équation différentielle

$$(1 - x^2)\psi''(x) + [\mu - \nu - (\mu + \nu + 2)x]\psi'(x) + \lambda\psi(x) = 0 \quad (5.95)$$

qui, selon la formule (5.77) évaluée pour  $q_2 = -1$  et  $l_1 = -\mu - \nu - 2$ , admet des solutions polynomiales non triviales si  $\lambda = \lambda_n = n(n + \mu + \nu + 1)$  pour un  $n \in \mathbb{N}_0$ . Nous calculons pour ce cas, suivant l'équation (5.80),

$$\begin{aligned} p(x) &= p_0 \exp \left( \int_{x_1}^x \frac{(\mu + 1)(1 - x') - (\nu + 1)(1 + x')}{(1 - x')(1 + x')} dx' \right) \\ &= (1 + x)^{\mu+1} (1 - x)^{\nu+1} \end{aligned} \quad (5.96)$$

avec le choix  $x_1 = 0$  et  $p_0 = 1$ , et ainsi, suivant l'équation (5.83),

$$w(x) = (1 + x)^\mu (1 - x)^\nu. \quad (5.97)$$

Pourvu que nous ayons  $\mu > -1$  et  $\nu > -1$ , nous pouvons identifier  $I = ] - 1, 1[$  comme intervalle dans lequel  $w$  est positif et aux bornes duquel  $p$  s'annule.

Les polynômes orthogonaux qui sont associés à cet intervalle  $] - 1, 1[$  et à la fonction de pondération (5.97) pour les paramètres  $\mu, \nu > -1$  sont appelés *polynômes de Jacobi*. Cette classe des polynômes de Jacobi contient la famille des *polynômes de Gegenbauer*, ces derniers étant définis pour le cas particulier  $\mu = \nu$ , c-à-d, pour la fonction de pondération  $w(x) = (1 - x^2)^\nu$  pour tout  $x \in ] - 1, 1[$ . Dans cette famille des polynômes de Gegenbauer se trouvent en particulier les *polynômes de Legendre* qui sont associés à la fonction de pondération  $w(x) = 1$

pour tout  $x \in ]-1, 1[$  et que l'on obtient donc pour le cas particulier  $\mu = \nu = 0$ . Ces polynômes de Legendre  $P_n$  satisfont donc à l'équation différentielle

$$(1 - x^2)P_n''(x) - 2xP_n'(x) + n(n + 1)P_n(x) = 0 \quad (5.98)$$

pour tout  $n \in \mathbb{N}_0$ . La famille des polynômes de Gegenbauer contient également les *polynômes de Tchebychev* de la première et la seconde espèce, qui sont respectivement obtenus pour les choix  $\mu = \nu = -1/2$  et  $\mu = \nu = 1/2$ .

Au vu de leurs propriétés particulières, les polynômes d'Hermite, de Laguerre et de Jacobi sont appelés *polynômes classiques*. Ils jouent un rôle important dans la résolution de l'équation de Schrödinger stationnaire, notamment pour les cas des hamiltoniens de l'oscillateur harmonique et de l'atome d'hydrogène.