

Université Paris-Sud 11
L3 de Physique et Chimie

Méthodes Mathématiques
pour la Licence de Physique et Chimie

Jean-Luc Raimbault
Laboratoire de Physique des Plasmas
Ecole Polytechnique,
jean-luc.raimbault@lpp.polytechnique.fr

2010 - 2011

Les pages qui suivent présentent quelques méthodes mathématiques que vous aurez à utiliser dans vos cours de Physique et Chimie. Cet enseignement de Mathématique est structuré en 5 grandes parties :

- Variables complexes
- Equations différentielles
- Analyse dans \mathbb{R}^n
- Algèbre linéaire
- Analyse de Fourier.

Au sein de chacune de ses parties, plusieurs chapitres, allant du plus simple au plus compliqué, sont proposés. Les chapitres 1, 4, 5, 6, 9, 10, 12, 14 et 15 (cf. sommaire) seront traités en première intention et devraient être maîtrisés par tous les étudiants. Ils constituent la base du programme sur lequel vous serez interrogés. Les chapitres complémentaires, qui abordent des notions plus avancées, peut-être moins utiles pour certains d'entre vous, seront étudiés si le temps le permet et/ou proposés aux étudiants suffisamment à l'aise sur les chapitres de base.

Mettre en œuvre des méthodes mathématiques dans le contexte d'un problème de Physique ou Chimie suppose une connaissance des concepts mathématiques associés (ce que ce cours vous rappellera ou vous fera découvrir), et surtout une mise en pratique qui passe par la résolution de nombreux exercices. Cela suppose une présence assidue et active aux cours et aux travaux dirigés mais également un travail personnel important. Cette implication personnelle est déterminante et sera encouragée. L'objectif est d'arriver progressivement à identifier vos lacunes, puis à travailler - avec notre aide - à les combler, enfin à estimer *par vous-même* le niveau de compréhension que vous avez atteint. Pour vous y aider, des devoirs et tests vous seront régulièrement proposés et des livres d'exercices seront à votre disposition.

Enfin, il est bon de rappeler qu'un bagage mathématique s'entretient. Il vous faut donc prévoir de revenir périodiquement, tout au long de vos études (et même après!) sur des concepts et des méthodes que vous maîtriserez d'autant moins que vous les utiliserez de façon occasionnelle. Les livres sont faits pour ça. Les quelques indications suivantes pourront éventuellement vous guider dans la jungle des ouvrages disponibles.

Commençons par des ouvrages écrits généralement par des physiciens qui suivent une approche assez pragmatique.

1. *Mathematical Methods for Scientists and Engineers*, Donald McQuarrie, University Science Books, 2003.
Livre d'un célèbre physico-chimiste, excellent pédagogue. Le contenu est très progressif et contient beaucoup d'illustrations. A recommander pour débiter sur beaucoup de sujets de mathématiques.
2. *Mathematical Methods for Physicists*, G. B. Arfken and H. J. Weber, Harcourt/Academic Press, 2001.
Un livre de référence pour les utilisateurs de mathématiques en sciences appliquées. Style très direct, nombreux exercices et exemples d'applications en Physique.
3. Distributions et Transformation de Fourier, Ediscience (1971, 1978), McGraw Hill (1984, 1988, 1993).
Présente la théorie des distributions et la transformation de Fourier sous une forme très accessible au physicien. Applications à l'Optique.
4. *Série Schaum* chez Ediscience ou Mac Graw Hill.
Série dont les différents volumes sont spécialisés dans certains domaines des mathématiques. En particulier, on pourra consulter : Variables complexes, Algèbre linéaire, Equations différentielles, Calcul différentiel et intégral. L'approche est très très progressive, s'appuyant sur un minimum de cours, et un grand nombre d'exercices de difficulté croissante. A recommander pour faire le point et pour le travail personnel.
5. *Mathématiques pour l'ingénieur*, Nino Boccara, Ellipses, 1996.
4 petits volumes traitant chacun d'un sujet : fonctions analytiques, distributions, intégration, et probabilités. Les sujets sont souvent introduits par une démarche historique instructive. Exercices corrigés.
6. *Mathématiques pour la Physique*, Walter Appel, H-K Editions, 2002.
Un bon livre récent et rigoureux, qui fait le tour d'horizon de différents domaines des mathématiques utiles au physicien.

Les ouvrages suivants, écrits par des mathématiciens dans un style rigoureux, permettent d'affermir les bases ou d'acquérir une vision plus large de certains sujets mathématiques.

1. *Cours de Mathématiques*, J. Bass, Masson, 1968.

Un bon livre à l'ancienne, en 2 tomes, complet sur toutes les notions élémentaires, comprend de nombreux exercices.

2. *Principe d'analyse mathématique*, Walter Rudin, EdiScience International, 1995.
Un très bon livre d'analyse écrit par un mathématicien professionnel très pédagogue. Utile pour revoir les notions de base d'analyse.
3. *A course in mathematics for students for in physics*, P. Bamberg and S. Sternberg, Cambridge University Press, 2001.
Ouvrage en 2 tomes qui présente nombre de sujets traditionnels d'une façon souvent originale et profonde. A consulter pour l'ouverture d'esprit.
4. *An Introduction to the Mathematical Theory of Waves*, Roger Knobel, AMS, 2000.
Petit ouvrage sur un sujet spécifique : les ondes. Très simple, progressif et clair. Les ondes non-linéaires sont abordées. Illustration et exercices en utilisant MatLab.
5. *Equations différentielles et systèmes dynamiques*, J. Hubbard and B. West, traduit par V. Gautheron, Cassini, 1999.
Un livre sur les équations différentielles, écrit dans un esprit d'introduction à la théorie des systèmes dynamiques, donc selon le point de vue géométrique. Nombreuses illustrations.
6. *Dictionnaire des Mathématiques*, Encyclopaedia Universalis, Albin Michel, 1997.
Cet ouvrage regroupe les articles de mathématiques de la célèbre Encyclopédie. Ecrits par d'excellents spécialistes, ces articles de niveaux variés permettent en général d'avoir une vue d'ensemble sur un sujet particulier et sur ses liens avec d'autres domaines des mathématiques. Pas vraiment pour les débutants. A consulter en particulier pour son caractère synthétique.

Signalons enfin pour finir, 2 excellents ouvrages abordables à votre niveau, écrits par deux anciens professeurs de l'Université Paris-Sud, tous deux membres de l'Académie des Sciences.

1. *Méthodes mathématiques pour les sciences physiques*, J.-M. Bony, Editions de l'Ecole Polytechnique, 2000.
Contient l'analyse de Fourier, les fonctions d'une variable complexe et l'analyse hilbertienne. Diverses remarques sur les équations de la physique mathématique ou l'usage des différentielles en physique par exemple sont très instructives.
2. *Mathématiques pour la Licence de Physique Fondamentale*, J.-P. Kahane, Editions de l'Université Paris-Sud, 1992.
Offre beaucoup de recul et d'élégance sur des sujets mathématiques traditionnels. Un grand nombre de domaines abordés, complétés par des exercices en partie corrigés. A méditer.

Première partie

Variables complexes

Chapitre 1

Rappels sur les nombres complexes

Motivations historiques

Le corps des nombres complexes

Exponentielle d'un nombre complexe

Représentation des nombres complexes

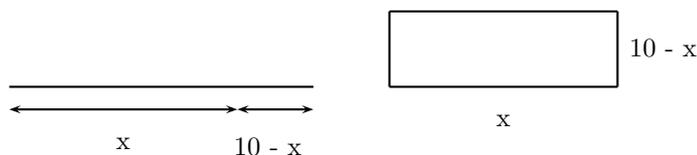
The imaginary numbers are a wonderful flight of God's spirit ; they are almost an amphibian between being and not being.

G. W. von Leibniz, 1702.

1.1 Motivations historiques

En 1545, le mathématicien italien Girolamo Cardan¹ proposa le problème suivant :

“ Comment diviser une droite de longueur 10 de telle sorte que le rectangle construit avec les 2 parties de la division ait une aire de 40? ”



Ce problème a pour solutions les racines de l'équation du second degré $x(10 - x) = 40$ dont on vérifiera aisément qu'elle admet les 2 racines $5 \pm \sqrt{-15}$. Il s'agissait d'un premier cas “concret” qui conduisit à s'interroger sur le sens à donner aux racines carrées de nombres négatifs. Un autre exemple plus simple est donné par l'équation $x^2 + 1 = 0$ qui n'a aucune solution dans \mathbb{R} puisque $x^2 + 1 \geq 1$.

Deux siècles plus tard, vers 1740, le mathématicien suisse Léonard Euler fut le premier à introduire la notation i pour désigner le “symbole”, $\sqrt{-1}$, soit $i \equiv \sqrt{-1}$, et donc, formellement, $i^2 = -1$. Ainsi, en utilisant formellement les règles de calcul établies pour les réels, les solutions du problèmes de Cardan peuvent-elles s'écrire

$$5 \pm \sqrt{-15} = 5 \pm \sqrt{15}i^2 = 5 \pm i\sqrt{15}$$

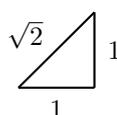
Cette écriture met donc en évidence de nouvelles grandeurs mathématiques qui s'expriment à l'aide d'un couple de réels $(5, \sqrt{15})$ et du symbole i . Cette première piste a été approfondie et formalisée au début du XIXème siècle par l'interprétation des nombres complexes (un terme dû au mathématicien allemand Carl Friedrich

1. Cardan (1501-1576) est un des plus fameux algébriste du moyen-âge puisqu'on lui doit en particulier la méthode de résolution des équations du 3ème degré. Rappelons pour mémoire que l'on sait résoudre explicitement toutes les équations algébriques jusqu'au 4ème degré (en les ramenant à des équations du 2ème ou du 3ème degré). Il faudra attendre Evariste Galois et ses successeurs pour prendre conscience que les équations de degré plus élevé ne peuvent pas en général s'exprimer à partir des fonctions élémentaires (ce qui ne veut pas dire qu'elles n'ont pas de solutions : des valeurs approchées peuvent facilement être calculées par ordinateur à l'aide d'algorithmes adéquats).

Gauss) comme des points d'un plan muni d'un repère cartésien, dont un axe est l'axe des nombres réels tandis que l'autre est celui des nombres "imaginaires".

1.2 Le corps des nombres complexes

Rappelons que l'ensemble des entiers naturels, \mathbb{N} , comprend 0 et tous les entiers positifs. A la différence de l'addition des entiers naturels qui ne fait pas sortir de \mathbb{N} , la soustraction et la division peuvent faire sortir de \mathbb{N} . L'introduction de l'ensemble des entiers relatifs, \mathbb{Z} , qui comprend 0 et tous les entiers positifs et négatifs, permet de rester dans \mathbb{Z} lorsqu'on soustrait mais pas lorsqu'on divise. Le problème est résolu par l'introduction de l'ensemble des rationnels, \mathbb{Q} , qui comprend 0 et tous les nombres de la forme m/n avec m et n entiers non nuls. Cependant, \mathbb{Q} ne contient pas tous les nombres puisque l'équation $r^2 = 2$ par exemple n'est satisfaite par aucun rationnel.²



On peut montrer que les rationnels complétés par les irrationnels constituent tous les points de la droite, ce qu'on appelle l'ensemble des nombres réels, \mathbb{R} .

De là, il est assez naturel de considérer tous les points du plan, que l'on peut obtenir comme l'ensemble des couples de points *ordonnés* de 2 nombres réels, que l'on notera (a, b) avec a et b éléments de \mathbb{R} . C'est le point de départ pour définir les nombres complexes.

Définition 1.2.1 *Un nombre complexe est un couple ordonné (a, b) de nombres réels.*

Pour pouvoir effectuer des calculs contrôlés avec les nombres complexes, il convient de définir les opérations suivantes.

Définition 1.2.2 *Soient $x = (a, b)$ et $y = (c, d)$ deux nombres complexes,*

1. *Egalité : $x = y \Leftrightarrow a = c$ et $b = d$,*
2. *Addition : $x + y = (a + c, b + d)$,*
3. *Multiplication : $xy = (ac - bd, ad + bc)$.*

Exercice 1.1 *Vérifier que ces règles sont compatibles avec une représentation des nombres complexes (a, b) par les matrices 2×2 :*

$$(a, b) \Leftrightarrow \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

Théorème 1.2.1 *L'ensemble des nombres complexes, noté \mathbb{C} , est un corps*

- dont l'élément neutre pour l'addition est le couple $(0, 0) \equiv 0 \in \mathbb{C}$,
- dont l'élément identité pour la multiplication est le couple $(1, 0) \equiv 1 \in \mathbb{C}$.

Exercice 1.2 *Montrer que pour $x \neq (0, 0)$,*

$$\frac{1}{x} = \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right)$$

Le nombre complexe $(0, 1)$ (à ne pas confondre avec $(1, 0)$: les couples sont ordonnés) mérite une attention particulière. On le baptise :

2. La démonstration peut s'effectuer par l'absurde. Supposons en effet que $r = m/n$ avec m, n entiers non nuls sans facteur commun. Alors, puisque $r^2 = 2$, $m^2 = 2n^2$, m^2 est donc pair. Donc m est pair car le carré d'un nombre impair est impair. Soit donc $m = 2p$ avec p entier, l'égalité $m^2 = 2n^2$ s'écrit donc $n^2 = 2p^2$. n^2 est donc pair ; on en déduit que n est également pair ; m et n sont donc tous deux pairs et ont donc 2 comme facteur commun, contrairement à l'hypothèse faite au début du raisonnement. On en déduit donc que l'équation $r = \sqrt{2} = m/n$ ne peut être satisfaite : $\sqrt{2}$ n'est donc pas rationnel.

Définition 1.2.3 On posera $i \equiv (0, 1)$, par définition.

Exercice 1.3 Vérifier que dans \mathbb{C} :

$$i^2 = -1$$

Les nombres complexes de la forme $(a, 0)$ forment un sous-corps de \mathbb{C} qui s'identifie avec \mathbb{R} . Remarquez que si $\lambda \in \mathbb{R}$, le produit de λ et du nombre complexe $z = (a, b)$ vérifie $\lambda z = (\lambda a, \lambda b)$.

1.3 Exponentielle complexe

L'exponentielle d'un nombre complexe peut être définie en généralisant la définition de l'exponentielle d'un nombre réel par une série.

Définition 1.3.1 Par définition, pour tout $z \in \mathbb{C}$,

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

Cette définition a un sens car on peut montrer que cette série est convergente $\forall z \in \mathbb{C}$. Cette définition constitue un premier exemple de fonction d'une variable complexe. L'étude des fonctions d'une (ou de plusieurs) variable(s) complexe(s) constitue un champ d'études mathématiques à part entière (c'est ce qu'on appelle la théorie des fonctions analytiques). La définition même de certaines fonctions d'une variable complexe, par prolongement de leurs définitions pour une variable réelle n'est pas toujours aussi simple que l'exemple de l'exponentielle pourrait le laisser penser. Des fonctions aussi usuelles que le logarithme ou la racine d'un nombre complexe ne peuvent en effet être définies que sur une partie du plan complexe.

Dans le cas particulier des imaginaires purs, $z \equiv iy$ où $y \in \mathbb{R}$, après avoir regroupés les termes pairs et impairs, on obtient :

$$e^{iy} = \left(1 - \frac{y^2}{2!} + \frac{y^4}{4!} - \dots \right) + i \left(y - \frac{y^3}{3!} + \dots \right) \equiv \cos y + i \sin y,$$

où on a utilisé la définition en termes de séries d'une sinus et du cosinus d'un nombre réel. Il s'agit de la formule dite d'Euler (1743), dite également représentation trigonométrique de l'exponentielle :

$$\forall x \in \mathbb{R}, \quad e^{ix} = \cos x + i \sin x$$

En remplaçant x par $-x$ dans la formule précédente, on a $e^{-ix} = \cos x - i \sin x$, et en combinant ces 2 expressions, on obtient les fonctions trigonométriques en termes d'exponentielles imaginaires :

$$\cos x = \frac{e^{ix} + e^{-ix}}{2} \quad \sin x = \frac{e^{ix} - e^{-ix}}{2i}$$

Exercice 1.4 Montrer que pour tout $x \in \mathbb{C}$

$$\cos(ix) = \cosh x \quad \text{et} \quad \sin(ix) = i \sinh x$$

Exercice 1.5 Montrer (en utilisant la formule du binôme) que pour a et b dans \mathbb{C}

$$e^{a+b} = e^a e^b$$

Une conséquence immédiate de cette propriété et de la formule d'Euler est la formule de De Moivre :

$$\forall x \in \mathbb{R}, \quad (\cos x + i \sin x)^n = \cos(nx) + i \sin(nx)$$

Exercice 1.6 Quelles relations trigonométriques peut-on dériver des identités suivantes ($t, t' \in \mathbb{R}$) ?

1. $|e^{it}| = 1$,
2. $e^{i(t+t')} = e^{it} e^{it'}$.

1.4 Représentation des nombres complexes

1.4.1 Représentation cartésienne

En utilisant les opérations définies précédemment, on peut écrire :

$$(x, y) = (x, 0) + (0, y) = (1, 0)(x, 0) + (0, 1)(y, 0) = 1x + iy \equiv x + iy$$

Ainsi,

$$z \equiv (x, y) \in \mathbb{C} \quad \Leftrightarrow \quad z = x + iy, \quad x \in \mathbb{R}, y \in \mathbb{R}.$$

C'est ce qu'on appelle *la représentation cartésienne* des nombres complexes. x et y étant respectivement *les parties réelles et imaginaires* de z .

Définition 1.4.1 Par définition, le conjugué du nombre complexe, z , noté \bar{z} , tel que :

$$\bar{z} \equiv x - iy$$

Exercice 1.7 Vérifier que

- $\overline{z + z'} = \bar{z} + \bar{z}'$,
- $\overline{zz'} = \bar{z}\bar{z}'$,
- $\overline{\bar{z}} = z$.

Calculons explicitement le produit d'un nombre complexe quelconque avec son conjugué :

$$z\bar{z} = (x + iy)(x - iy) = x^2 + y^2$$

qui est donc un nombre réel, positif ou nul.

Définition 1.4.2 Par définition, le module du nombre complexe, z est le nombre réel noté $|z|$ tel que :

$$|z| \equiv \sqrt{z\bar{z}} = \sqrt{x^2 + y^2} \in \mathbb{R}_+$$

Le module d'un nombre complexe jouit des mêmes propriétés que la valeur absolue pour les nombres réels (attention, le même symbole $|\cdot|$ est utilisé mais ne s'applique pas aux mêmes nombres). Bien qu'il n'y ait pas de relation d'ordre dans \mathbb{C} (on ne peut pas comparer 2 couples de nombres), la notion de module permet de définir une distance dans \mathbb{C} et, partant de là de développer l'analyse (notions de limite, continuité ...) dans le corps des nombres complexes.

1.4.2 Représentation géométrique

On peut passer de la représentation cartésienne à la représentation géométrique (module-argument) en utilisant les coordonnées polaires



$$z = x + iy = |z| \cos \theta + i |z| \sin \theta = |z| e^{i\theta}$$

θ est *l'argument* qui vérifie $\tan \theta = y/x$. L'argument, en tant qu'angle, est évidemment défini à 2π près ; si on impose à l'argument d'appartenir à l'intervalle $] -\pi, +\pi]$, il est déterminé de façon unique et s'appelle alors

“l’argument principal”. Par exemple, les représentations géométriques des nombres complexes de signes opposés $z_1 = 1 + i$ et $z_2 = -1 - i$ sont données respectivement par $z_1 = \sqrt{2}e^{+i\pi/4}$ et $z_2 = \sqrt{2}e^{-i3\pi/4}$. Comme on pouvait s’y attendre, on passe d’un nombre complexe à l’autre par une rotation de π autour de l’origine (symétrie de centre 0), les 2 valeurs des arguments (principaux) étant différenciées par les signes de x et y .

On notera en particulier qu’un nombre complexe est nul si et seulement si son module est nul. La représentation géométrique permet de donner une interprétation simple de plusieurs opérations sur les nombres complexes :

- $\bar{z} = |z|e^{-i\theta}$ (symétrie par rapport à l’axe Ox),
- $zz' = |z||z'|e^{-i(\theta+\theta')}$ (multiplication des modules et addition des arguments),
- $\frac{z}{z'} = \frac{|z|}{|z'|}e^{-i(\theta-\theta')}$ (division des modules et soustraction des arguments).

L’utilité de la représentation géométrique peut être illustré dans la recherche des racines de l’unité c’est-à-dire à déterminer les solutions dans \mathbb{C} de l’équation

$$z^n - 1 = 0$$

En utilisant la représentation géométrique, on obtient

$$z^n = 1 \quad \Leftrightarrow \quad |z|^n e^{in\theta} = 1 = 1e^{i2k\pi}, \quad k \in \mathbb{Z}$$

On doit donc avoir à la fois $|z|^n = 1$ (i.e. $|z| = 1$) et $n\theta = 2k\pi$. On en déduit donc les n solutions *distinctes* qui se répartissent uniformément sur le cercle unité :

$$z^n - 1 = 0 \quad \Leftrightarrow \quad z_k = e^{i2k\pi/n}, \quad k = 0, 1, 2, \dots, n-1.$$

On a mentionné au début de ce chapitre que les équations polynomiales n’avaient pas toujours de solutions dans \mathbb{R} . Dans le corps des nombres complexes, les choses sont beaucoup plus simples puisque l’on dispose du théorème suivant, appelé *théorème fondamental de l’algèbre*, dû à D’Alembert (1746), qui stipule :

Théorème 1.4.1 *Un polynôme de degré n à coefficients complexes possède exactement n racines (pas nécessairement distinctes) dans \mathbb{C} .*

Ainsi la théorie des équations algébriques est-elle plus harmonieuse dans l’ensemble des nombres complexes que dans l’ensemble des nombres réels. Il en va de même dans de nombreux domaines des mathématiques mettant en jeu les nombres complexes. Par exemple, si une fonction d’une variable complexe est dérivable une fois, elle est dérivable une infinité de fois (!), ce qui n’est certes pas le cas pour les fonctions d’une variable réelle.

Chapitre 2

Fonctions d'une variable complexe

De la même façon qu'une fonction d'une variable réelle à valeurs dans \mathbb{R} est définie par une prescription qui associe un nombre réel $x \in \mathbb{R}$ à un autre réel $f(x) \in \mathbb{R}$, il est possible de définir des fonctions d'une variable complexe à valeur dans \mathbb{C} .

En introduisant les notions de limite, continuité, dérivabilité, intégration ... on peut alors développer une analyse pour ces fonctions de variables complexes, que nous présentons succinctement dans ce qui suit.

2.1 Fonctions d'une variable complexe

Définition 2.1.1 On appelle fonction d'une variable complexe, une application f de \mathbb{C} dans \mathbb{C} :

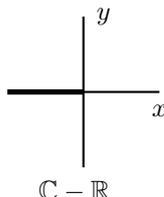
$$f : z = x + iy \mapsto Z = f(z) = X(x, y) + iY(x, y).$$

Si une seule valeur de $f(z)$ correspond à chaque valeur de z , f est dite *uniforme*; si plusieurs valeurs de $f(z)$ correspondent à chaque valeur de z , f est dite *multiforme*. Aucun calcul n'étant possible avec des fonctions multiformes, on peut toujours considérer une fonction multiforme comme un ensemble de fonction uniforme, et calculer avec l'une d'entre elle. Chacune des fonctions uniformes définies est *une branche (ou une détermination)* de la fonction, et l'élément choisi s'appelle *la branche (ou la détermination) principale*.

Exemple Fonction racine carrée d'un nombre complexe :

$$f(z) = z^{1/2}$$

En utilisant la représentation géométrique de z , $z = \rho e^{i\theta} = \rho e^{i\theta} e^{i2k\pi}$, avec $k \in \mathbb{Z}$, on trouve que $f(z) = \rho^{1/2} e^{i\theta/2} e^{ik\pi} = \pm \rho^{1/2} e^{i\theta/2}$. Ainsi, z a-t-il 2 images : la fonction est bivaluée.



Les deux branches sont obtenues en empêchant z de faire un tour complet autour de l'origine. Dans ce cas particulier, on dit que l'origine est *un point de branchement*. On effectue ce qu'on appelle *une coupure* dans le plan complexe. Par exemple, on peut retirer du domaine de définition, l'ensemble des valeurs négatives, \mathbb{R}_- , avec pour choix "naturel" (dans le sens d'une définition qui prolonge celle de la racine d'un nombre réel) de détermination principale :

$$\text{si } z = \rho e^{i\theta} \text{ avec } \rho \in \mathbb{R}_+ \text{ et } \theta \in]-\pi, +\pi[, \quad w = \rho^{1/2} e^{i\theta/2}$$

Notez que le choix de \mathbb{R}_- n'est pas unique, tout autre demi-droite ferait également l'affaire.

2.2 Dérivation des fonctions d'une variable complexe

Définition 2.2.1 On dit qu'une fonction f est dérivable au sens complexe au point z_0 , si le quotient :

$$\frac{f(z) - f(z_0)}{z - z_0}$$

tend vers une limite, indépendamment de la façon dont z tend vers z_0 .

Cette limite unique est la dérivée de f en z_0 :

$$f'(z_0) = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

Soit U un disque ouvert de \mathbb{C} . On dit que f est holomorphe ou analytique dans U si f est dérivable en tout point de U .

La définition de la dérivabilité est donc formellement identique au cas réel. Le point important à souligner est que la valeur de la dérivée doit être unique, quelle que soit la façon dont on tend vers le point. Il s'agit d'une contrainte très forte.

Pour savoir si une fonction est dérivable on peut appliquer la définition ou utiliser les résultats du théorème suivant :

Théorème 2.2.1 Pour que la fonction $Z = f(z) = X(x, y) + iY(x, y)$ soit dérivable au point z_0 , il faut et il suffit que :

- X et Y , fonctions de (x, y) , soient dérivables en (x_0, y_0) ,
- et que, en ce point : $\partial_x X = \partial_y Y$ et $\partial_y X = -\partial_x Y$.

Si f est dérivable, sa dérivée $Z' = f'$ est telle que :

$$Z' = \partial_x X + i\partial_x Y = \partial_y Y - i\partial_y X.$$

Ce résultat est obtenu en appliquant la définition de la dérivabilité et en choisissant 2 façons possibles de faire tendre (x, y) vers (x_0, y_0) , par exemple, soit en suivant l'axe Ox ou soit en suivant l'axe Oy . La fonction n'est dérivable que si les 2 résultats sont identiques, ce qui conduit aux conditions du théorème. Les conditions sur les dérivées partielles s'appellent les conditions de Cauchy-Riemann. Il existe une version analogue lorsque la représentation géométrique des nombres complexes est utilisée. Par exemple, les conditions de Cauchy-Riemann montrent que la fonction définie par $f(z) = z^2 = (x^2 - y^2) + i2xy$ est dérivable en tout point de \mathbb{C} , de dérivée $f'(z) = 2z$, mais que la fonction $f(z) = \bar{z} = x - iy$ ne l'est en aucun point puisque $\partial_x X = -\partial_y Y$.

Les propriétés concernant la dérivabilité des sommes, produits, .. de fonctions, sont identiques à celles connues dans \mathbb{R} :

Théorème 2.2.2

Si f et g sont holomorphes dans un disque ouvert \mathcal{D} , il en est de même pour αf ($\alpha \in \mathbb{R}$), $f + g$, fg , f/g (pour $g(z) \neq 0$ dans \mathcal{D}), et pour $f \circ g$.

Plus généralement, on pourra utiliser les mêmes formules élémentaires pour le calcul des dérivées que celles utilisées dans \mathbb{R} .

En combinant les conditions de Cauchy-Riemann pour les fonctions dont les dérivées croisées sont égales, on trouve aussitôt que X ou Y vérifient l'équation $\Delta X = \Delta Y = 0$. X ou Y sont dites harmoniques et jouent un rôle important dans les problèmes physiques qui mettent en jeu l'équation de Laplace (cf. électrostatique et mécanique des fluides).

Théorème 2.2.3 Si la fonction $Z = f(z) = X(x, y) + iY(x, y)$ est holomorphe, et si X et Y ont des dérivées secondes continues, alors :

$$\Delta X \equiv \partial_{xx}^2 X + \partial_{yy}^2 X = 0, \quad \text{et} \quad \Delta Y \equiv \partial_{xx}^2 Y + \partial_{yy}^2 Y = 0.$$

On dit que X et Y sont harmoniques.

Enfin, la condition de dérivabilité est tellement forte dans \mathbb{C} que l'on obtient le résultat spectaculaire suivant :

Théorème 2.2.4 *Une fonction holomorphe est indéfiniment dérivable.*

2.3 Intégration des fonctions d'une variable complexe

L'intégrale curviligne d'une fonction d'une variable complexe est définie comme la somme de 2 intégrales curvilignes de fonctions de variables réelles :

Définition 2.3.1 *Soit f une fonction de module borné, définie dans un disque ouvert \mathcal{D} de \mathbb{C} . Soit \mathcal{C} un arc de courbe régulier contenu dans \mathcal{D} .*

L'intégrale curviligne de $f(z) = X + iY$ le long de \mathcal{C} est définie par :

$$\int_{\mathcal{C}} f(z)dz = \int_{\mathcal{C}} (X + iY)(dx + idy) = \int_{\mathcal{C}} (Xdx - Ydy) + i \int_{\mathcal{C}} (Ydx + Xdy)$$

Il n'est pas toujours nécessaire d'opérer cette décomposition. Dans certains cas, on peut également calculer directement l'intégrale à partir des variables complexes.

Exemple Soit à intégrer la fonction définie par $f(z) = 1/(z - a)$ avec $a \in \mathbb{C}$ le long d'un contour circulaire C_a qui entoure a .

La fonction est holomorphe dans $\mathbb{C} - \{a\}$, et dans ce cas la paramétrisation du contour est simple. Posons $z = a + Re^{i\theta}$ où R est le rayon du cercle C . Alors $dz = iRe^{i\theta} d\theta$ de sorte que

$$\int_{C_a} f(z)dz = \int_0^{2\pi} \frac{1}{Re^{i\theta}} iRe^{i\theta} d\theta = 2i\pi$$

Un certain nombre de résultats spécifiques importants concernent les intégrales curvilignes le long de contours fermés. Le théorème suivant, dit *théorème de Cauchy*, est fondamental.

Théorème 2.3.1 *Soit \mathcal{D} un disque ouvert de \mathbb{C} (ou plus généralement un domaine ouvert simplement connexe¹).*

Si f est holomorphe dans \mathcal{D} , et si la courbe \mathcal{C} fermée est contenue dans \mathcal{D} , alors :

$$\oint_{\mathcal{C}} f(z)dz = 0.$$

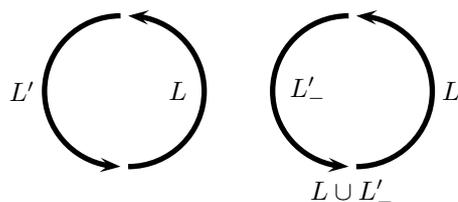
Une application directe de ce théorème montre que l'intégrale de l'exemple précédent $\oint_C dz/(z - a)$ s'annulerait si C est un cercle n'entourant pas a . Le théorème de Cauchy est une conséquence directe de la formule de Green-Riemann (ou formule du rotationnel) pour un champ de vecteurs à 2 composantes $\mathbf{V} = (V_x, V_y)$:

$$\iint_S (\partial_x V_y - \partial_y V_x) dx dy = \oint_C (V_x dx + V_y dy)$$

où S est la surface enclose par C . Le résultat est obtenu en appliquant cette formule aux champs $(X, -Y)$ et (Y, X) et en utilisant les conditions de Cauchy-Riemann.

En appliquant le théorème de Cauchy au contour $C = L \cup L'_-$ constitué de l'union du chemin L et du chemin L'_- parcouru en sens inverse, on montre que l'intégrale d'une fonction holomorphe dans un disque ouvert de \mathbb{C} ne dépend pas du chemin suivi :

1. Rappelons qu'un *domaine simplement connexe* est une partie de \mathbb{C} d'un seul morceau, sans trous.



Théorème 2.3.2 Si f une fonction holomorphe dans \mathcal{D} , domaine simplement connexe, alors l'intégrale joignant 2 points de \mathcal{D} ne dépend pas du chemin suivi. Plus explicitement, si \mathcal{L} et \mathcal{L}' sont 2 chemins dans \mathcal{D} , alors :

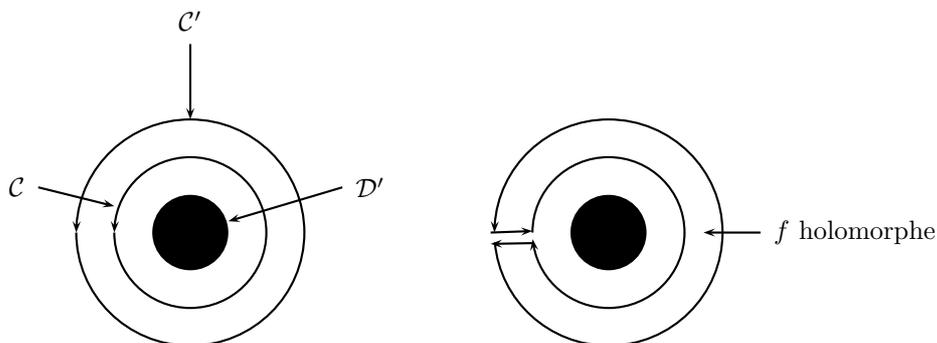
$$\int_{z_0}^z f(t)dt = \int_{\mathcal{L}} f(z)dz = \int_{\mathcal{L}'} f(z)dz, \quad \text{avec } z, z_0 \in \mathcal{D}.$$

Le théorème de Cauchy ne s'applique qu'aux fonctions holomorphes dans un domaine. L'intégrale d'une fonction sur un contour fermé entourant un domaine où la fonction n'est pas holomorphe n'est pas en général nul (l'exemple traité plus haut $\oint_{C_a} dz/(z-a) = 2i\pi$ en est un exemple). On dispose cependant du résultat suivant, très utile dans les calculs :

Théorème 2.3.3 Soit f une fonction holomorphe dans \mathcal{D} à l'exclusion d'une partie \mathcal{D}' où la fonction n'est pas holomorphe. Si les courbes fermées \mathcal{C} et \mathcal{C}' font chacune une fois le tour du domaine \mathcal{D}' , alors :

$$\oint_{\mathcal{C}} f(z)dz = \oint_{\mathcal{C}'} f(z)dz.$$

Ce résultat est obtenu par application du théorème de Cauchy en remarquant que f est holomorphe dans le domaine compris entre \mathcal{C} et \mathcal{C}' .



Il existe une autre classe de résultats importants, les *formules intégrales de Cauchy* qui montrent que si l'on connaît les valeurs d'une fonction analytique sur une courbe fermée \mathcal{C} , alors les valeurs de la fonction peuvent être calculées en tout point du domaine encerclé par la courbe \mathcal{C} .

Théorème 2.3.4 Soit f une fonction holomorphe dans un domaine simplement connexe \mathcal{D} ; soit \mathcal{C} une courbe fermée entièrement contenue à l'intérieur de \mathcal{D} , et entourant un domaine Δ . Soit $a \in \Delta$ mais n'appartenant pas à \mathcal{C} , alors :

$$f(a) = \frac{1}{2i\pi} \oint_{\mathcal{C}} \frac{f(z)}{(z-a)} dz,$$

$$f^{(n)}(a) = \frac{n!}{2i\pi} \oint_{\mathcal{C}} \frac{f(z)}{(z-a)^{(n+1)}} dz, \quad \text{pour } n \geq 1.$$

Par application du théorème 2.3.3, on peut reporter le calcul de l'intégrale de C sur un cercle γ_ϵ de centre a et de rayon ϵ que l'on choisira aussi petit que désiré. Alors,

$$\oint_C \frac{f(z)}{(z-a)} dz = \oint_{\gamma_\epsilon} \frac{f(z)}{(z-a)} dz = \oint_{\gamma_\epsilon} \frac{f(z) - f(a)}{(z-a)} dz + f(a) \oint_{\gamma_\epsilon} \frac{1}{(z-a)} dz$$

Le résultat est obtenu en considérant la limite $\epsilon \rightarrow 0$. La première intégrale est vaut $f'(a) \oint_{\gamma_\epsilon} dz \rightarrow 0$ et la deuxième $2i\pi f(a)$. Le résultat sur les dérivées est obtenue par dérivées successives sous le signe intégral.

2.4 Développement en série

Une autre conséquence importante de la dérivabilité d'une fonction en un point (caractère très fort, répétons-le), est la possibilité de développement de la fonction en série entière dans le voisinage du point.

Théorème 2.4.1 *Si f est holomorphe au point $z = a$, il existe un disque de centre a à l'intérieur duquel f est développable en série entière :*

$$f(z) = \sum_{n=0}^{\infty} A_n (z-a)^n,$$

$$A_n = \frac{f^{(n)}(a)}{n!}$$

Il s'agit donc d'un développement *en série de Taylor*. Un exemple de développement en série au voisinage de l'origine est donné par le résultat :

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n, \quad |z| < 1$$

que l'on peut également voir comme une généralisation au plan complexe d'un résultat connu sur les séries géométriques. On notera que le disque de convergence est limité par la rencontre de la singularité en $z = 1$, où la fonction n'est plus définie (et donc plus holomorphe).

Qu'en serait-il du développement de la série :

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

En utilisant le résultat du développement en série de $1/(1-z)$, on a aussitôt :

$$f(z) = \frac{1}{z} (1 + z + z^2 + \dots) = \frac{1}{z} + 1 + z + z^2 + \dots$$

Ainsi, cette fonction, non holomorphe en $z = 0$, admet-elle un développement en série de puissances positives et négatives (le terme en $1/z$) de la variable z . La généralisation de ce point de vue correspond au développement *en série de Laurent* :

Théorème 2.4.2 *Si f est holomorphe dans la couronne \mathcal{D} comprise entre 2 cercles concentriques de centre a , de rayons r et R , on peut y développer f suivant les puissances positives et négatives de $(z-a)$. On a le développement en série de Laurent :*

$$f(z) = \sum_{n=-\infty}^{\infty} A_n (z-a)^n,$$

$$A_n = \frac{1}{2i\pi} \oint_C \frac{f(s)}{(s-a)^{n+1}} ds,$$

C étant n'importe quelle courbe fermée régulière située à l'intérieur de \mathcal{D} et faisant une fois le tour de a dans le sens positif.

La démonstration de ces théorèmes est une conséquence des formules intégrales de Cauchy appliquées à des contours encerclant des domaines où la fonction $s \mapsto f(s)/(s-z)$ est holomorphe.

2.5 Méthode des résidus

Une fonction d'une variable réelle étant un cas particulier de fonction d'une variable complexe, il apparaît que certaines intégrales de fonctions réelles sont plus facilement calculables (voire seulement calculables) en passant dans le plan complexe. La méthode dite des *résidus* est une méthode de calcul d'intégrales curvilignes de fonction d'une variable complexe particulièrement efficace. Elle est utile, en particulier, pour le calcul des transformées de Laplace et de Fourier.

Commençons par introduire un peu de vocabulaire concernant les points singuliers des fonctions d'une variable complexe.

Définition 2.5.1 *Considérons le développement en série de Laurent d'une fonction f dont le développant en puissances négatives s'arrête au Nème terme :*

$$f(z) = \sum_{n=0}^{\infty} a_n(z-a)^n + \sum_{n=1}^N \frac{b_n}{(z-a)^n}$$

- La singularité de f en $z = a$ est appelée un pôle d'ordre N .
- Le pôle est dit simple si $N = 1$, multiple dans les autres cas.
- La singularité est dite essentielle si $N = \infty$.
- Une singularité (ou point singulier) est dite isolée si elle possède un voisinage dans lequel on ne trouve aucun autre point singulier pour la fonction considérée.

Par exemple, a est un pôle simple de la fonction définie par $f(z) = 1/(z-a)$, $f(z) = e^z/z^2$ est un pôle d'ordre 2 en $z = 0$, tandis que 0 est un point essentiel de la fonction définie par $f(z) = e^{1/z}$.

Théorème 2.5.1 *Soit \mathcal{D} un domaine simplement connexe, à l'intérieur duquel f possède une seule singularité a (pôle ou point essentiel isolé). On a :*

$$\oint_{\partial\mathcal{D}} f(z)dz = 2i\pi A_{-1}.$$

où A_{-1} est le coefficient du développement en série de Laurent de f au voisinage de a , et $\partial\mathcal{D}$ est la courbe limitant le domaine \mathcal{D} .

A_{-1} s'appelle le résidu de f en a .

Ce résultat est une conséquence directe du développement en série de Laurent (définition du coefficient A_n pour $n = -1$).

Théorème 2.5.2 *Soit \mathcal{D} un domaine non nécessairement connexe, dont la frontière $\partial\mathcal{D}$ est formée d'une ou plusieurs courbes fermées simples.*

Soit f une fonction holomorphe dans \mathcal{D} y compris sur sa frontière $\partial\mathcal{D}$, sauf en un nombre fini de points non situé sur $\partial\mathcal{D}$, et qui sont des pôles ou points essentiels isolés. On a

$$\oint_{\partial\mathcal{D}} f(z)dz = 2i\pi \sum_k A_{-1}^{(k)},$$

les $A_{-1}^{(k)}$ étant les résidus de f associés aux points singuliers $a_1, a_2, \dots, a_k, \dots \in \mathcal{D}$.

Ce résultat constitue à proprement parler le *théorème des résidus*. Un domaine non connexe est un domaine en plusieurs morceaux. La démonstration consiste à modifier les contours de façon à se ramener aux conditions d'applications du théorème précédent.

Le dernier point à évoquer concerne le calcul pratique des résidus. Si le pôle est simple, il existe une couronne de centre a où :

$$f(z) = \frac{A_{-1}}{z-a} + A_0 + A_1(z-a) + \dots \Rightarrow (z-a)f(z) = A_{-1} + A_0(z-a) + A_1(z-a)^2 + \dots$$

On peut donc obtenir A_{-1} en prenant la limite quand $z \rightarrow a$. Si le pôle est d'ordre plus élevé, 2 par exemple, on a de même :

$$(z-a)^2 f(z) = A_{-2} + A_{-1}(z-a) + A_0(z-a)^2 + \dots \Rightarrow \frac{d(z-a)^2 f(z)}{dz} = A_{-1} + 2A_0(z-a) + \dots$$

Le résidu étant obtenu en prenant encore la limite $z \rightarrow a$. D'une façon générale, on a donc :

Théorème 2.5.3 1. Soit a un pôle simple isolé d'une fonction f . Le résidu associé à a est donné par :

$$A_{-1} = \lim_{z \rightarrow a} [(z-a)f(z)].$$

2. Soit a un pôle multiple isolé d'ordre k d'une fonction f . Le résidu associé à a est donné par :

$$A_{-1} = \lim_{z \rightarrow a} \frac{1}{(k-1)!} \frac{d^{k-1}}{dz^{k-1}} [(z-a)^k f(z)].$$

Chapitre 3

Transformation de Laplace au sens des fonctions

Introduction

Définition de la transformée de Laplace

Caractérisation et holomorphie

Propriétés de la transformée de Laplace

Comportements asymptotiques

Inversion de la transformée de Laplace

3.1 Introduction

La transformée de Laplace appartient à la famille très vaste des transformées intégrales, qui établissent une relation entre une fonction f et sa transformée F sous la forme :

$$F(\omega) = \int_I K(\omega, t) f(t) dt$$

Une transformée particulière nécessite donc la définition du noyau $K(\omega, t)$ et de l'intervalle d'intégration I . Les transformations les plus utilisées sont celles de Fourier, pour laquelle on a :

$$I = \mathbb{R} \quad \text{et} \quad K(\omega, t) = e^{-i\omega t}, \quad \omega \in \mathbb{R} \quad (\text{Fourier}),$$

et celles de Laplace, pour laquelle on a :

$$I = \mathbb{R}^+ \quad \text{et} \quad K(\omega, t) = e^{-\omega t}, \quad \omega = \omega_r + i\omega_i \in \mathbb{C} \quad (\text{Laplace}).$$

Puisque ω est complexe, la transformation de Laplace peut être vue comme une généralisation de la transformation de Fourier, restreinte aux fonctions définies sur \mathbb{R}^+ . La restriction à \mathbb{R}^+ n'est guère contraignante dans les applications réalistes où $f(t)$ représente un signal physique à l'instant t qui ne peut exister de toute éternité. Il est en effet toujours possible de choisir l'instant où on démarre les mesures comme l'origine des temps. De ce point de vue, l'analyse de Fourier est plus adaptée à l'étude des régimes forcés, tandis que l'analyse de Laplace convient davantage pour l'étude des régimes transitoires.

En revanche, il est extrêmement bénéfique de passer de la variable réelle à la variable complexe qui rajoute le facteur de convergence $e^{-\omega_r t}$ dans l'intégrale, au moins dans une partie du plan complexe. Il en résulte qu'un grand nombre de fonctions admettent une transformée de Laplace, ce qui n'est pas le cas des transformées de Fourier.

Pour peu qu'ils soient linéaires, la transformée de Laplace est un outil très simple d'emploi pour résoudre les problèmes d'évolution (équations différentielles ou aux dérivées partielles, équations aux différences ou intégrales ...). Le principe général d'action de la transformée de Laplace sur les opérateurs d'évolution consiste en une réduction de l'ordre des opérateurs. Par transformée de Laplace, les équations différentielles deviennent

des équations algébriques, tandis que les équations aux dérivées partielles se transforment en des équations différentielles. Il en résulte une simplification efficace des problèmes qui permet souvent leur résolution analytique.

3.2 Définition de la transformée de Laplace

Définition 3.2.1 Soit f une fonction de la variable réelle, la transformée de Laplace de f , lorsqu'elle existe, est la fonction F de la variable complexe z définie par l'intégrale :

$$F(z) \equiv \int_{\mathbb{R}^+} f(t) e^{-zt} dt$$

Remarques

1. On appelle f , l'originale et sa transformée F , l'image.
2. Les notations utilisées pour les transformées de Laplace sont très variées et dépendent du domaine d'application. Afin de souligner sa nature d'élément de \mathbb{C} , nous avons noté la variable indépendante par z . Les lettres p et s sont également utilisées.
3. On remarquera que les valeurs de f pour $t < 0$ n'interviennent pas dans la définition. Une fonction f est dite *causale* si $f(t) = 0$ pour $t < 0$. On peut toujours rendre une fonction causale en la multipliant par la fonction de Heaviside H , ce que nous ferons couramment dans la suite.

La transformée de Laplace d'une fonction n'existe en général que dans une partie du plan complexe. Posons $z = x + iy$, l'existence de $F(z)$ impose que :

$$t \mapsto |f(t) e^{-zt}| = |f(t)| e^{-xt} \in L^1(\mathbb{R}^+)$$

Introduisons d'abord la notion d'abscisse de sommabilité.

Définition 3.2.2 Le nombre réel :

$$x_0 \equiv \inf\{x \in \mathbb{R} \mid t \mapsto |f(t)| e^{-xt} \in L^1(\mathbb{R}^+)\}$$

est appelé l'abscisse de sommabilité de f .

On peut maintenant donner le théorème d'existence suivant :

Théorème 3.2.1 Si f est une fonction d'abscisse de sommabilité x_0 , alors, la transformée de Laplace F existe dans le demi-plan ouvert $\Re z > x_0$.

En effet, posons $z = x + iy$,

$$|F(z)| \leq \int_{\mathbb{R}^+} |f(t)| e^{-xt} dt < \int_{\mathbb{R}^+} |f(t)| e^{-x_0 t} dt < +\infty,$$

pour $\Re z = x > x_0$.

En conséquence, F est bornée pour $\Re z > x_0$.

◇ Exemples

1. Fonction de Heaviside H .

L'abscisse de sommabilité est $x_0 = 0$, puisque $t \mapsto H(t)e^{-xt} \in L^1(\mathbb{R}^+)$ pour $x > 0$. Le calcul de la transformée est immédiat :

$$f(t) = H(t) \xrightarrow{F} F(z) = \int_{\mathbb{R}^+} e^{-zt} dt = \frac{1}{z} \quad \text{pour } \Re z > 0.$$

2. *Fonction puissance* $t \mapsto t^n$.

Commençons par le cas linéaire : $t \mapsto te^{-xt} \in L^1(\mathbb{R}^+)$ pour $x > 0$. On effectue le calcul par parties :

$$f(t) = t \xrightarrow{F} F(z) = \int_{\mathbb{R}^+} t e^{-zt} dt = \frac{1}{z} \int_{\mathbb{R}^+} 1 e^{-zt} dt = \frac{1}{z^2} \quad \text{pour } \Re z > 0.$$

Puis, par récurrence, pour $n \in \mathbb{N}$,

$$f(t) = t^n \xrightarrow{F} F(z) = \int_{\mathbb{R}^+} t^n e^{-zt} dt = \frac{n}{z} \int_{\mathbb{R}^+} t^{n-1} e^{-zt} dt = \frac{n!}{z^{n+1}} \quad \text{pour } \Re z > 0.$$

3. *Fonction exponentielle* $t \mapsto e^{at}$.

$$f(t) = e^{at} \xrightarrow{F} F(z) = \int_{\mathbb{R}^+} e^{at} e^{-zt} dt = \frac{1}{z-a} \quad \text{pour } \Re z > \Re a.$$

4. *Fonction* $t \mapsto 1/\sqrt{t}$.

$t^{-1/2} e^{-xt} \sim t^{-1/2}$ quand $t \rightarrow 0$ et $t^{-1/2} e^{-xt} \sim e^{-xt}$ quand $t \rightarrow +\infty$, donc $F(z)$ converge pour $x = \Re z > 0$.

$$f(t) = \frac{1}{\sqrt{t}} \xrightarrow{F} F(z) = \int_{\mathbb{R}^+} \frac{1}{\sqrt{t}} e^{-zt} dt = \int_{\mathbb{R}} e^{-zu^2} du = \sqrt{\frac{\pi}{z}} \quad \text{pour } \Re z > 0,$$

(la dernière égalité est obtenue en calculant le carré de l'intégrale). Pour le calcul de $z^{1/2}$, on choisira la détermination principale du logarithme de telle sorte qu'on obtienne le résultat usuel si z est réel.

3.3 Holomorphie

Commençons par préciser la relation entre la transformée de Fourier et la transformée de Laplace. Si x_0 est l'abscisse de sommabilité de f , la fonction $t \mapsto H(t)f(t)e^{-xt}$ est sommable sur \mathbb{R} pour $x > x_0$. La transformée de Laplace de f peut alors s'écrire comme une transformée de Fourier. En effet, posons $z = x + i2\pi y$:

$$F(x + i2\pi y) = \int_{\mathbb{R}^+} f(t) e^{-xt} e^{-i2\pi yt} dt = \int_{\mathbb{R}} [H(t)f(t)e^{-xt}] e^{-i2\pi yt} dt$$

soit encore :

$$F(x + i2\pi y) = \mathcal{F} [H(t)f(t)e^{-xt}] (y), \quad \text{pour } x > x_0$$

où \mathcal{F} désigne la transformée de Fourier.

Cette remarque facilitera certaines démonstrations.

Ainsi une transposition directe du résultat connu sur les transformées de Fourier conduit au résultat suivant, important dans la pratique, l'égalité des images par TL implique l'égalité des originaux :

$$F(z) = G(z) \quad \text{pour } \Re z > x_0, \quad \Rightarrow \quad f(t) = g(t) \quad (p.p.),$$

où x_0 est la plus grande des 2 abscisses de sommabilité des fonctions f et g .

Concernant les propriétés d'holomorphie, on a le résultat suivant :

Théorème 3.3.1 *Soit f une fonction d'abscisse de sommabilité x_0 :*

- L'abscisse de sommabilité de la fonction $t \in \mathbb{R}^+ \mapsto (-t)^m f(t)$ est x_0 .
- F est holomorphe dans le demi-plan $\Re z > x_0$, et

$$\frac{dF^m(z)}{dz^m} = \int_{\mathbb{R}^+} (-t)^m f(t) e^{-zt} dt$$

En effet, les 2 fonctions $t \mapsto f(t) e^{-xt}$ et $t \mapsto (-t)^m f(t) e^{-xt}$ ont le même comportement à l'infini, donc la même abscisse de sommabilité. Il faut justifier la dérivation sous le signe somme, ce qui résulte de l'inégalité :

$$|(-t)^m f(t) e^{-xt}| \leq t^m |f(t)| e^{-x_0 t}$$

pour tout $z = x + iy$ tel que $x > x_0$. La fonction majorante étant intégrable, on peut permuter la dérivation et le signe intégral, d'où le résultat. $F'(z)$ est fini en tant que TL : on en déduit donc l'holomorphie de F .

◇ Exemples

- $\frac{1}{(z+a)^2} = -\frac{d}{dz} \left(\frac{1}{z+a} \right) \xrightarrow{F} f(t) = t e^{-at}$.
- $\sin t \xrightarrow{F} \frac{1}{z^2+1}$ donc, $t \sin t \xrightarrow{F} -\frac{d}{dz} \left(\frac{1}{z^2+1} \right) = \frac{2z}{(z^2+1)^2}$.

3.4 Propriétés de la transformée de Laplace

Outre la propriété de linéarité qui découle de la définition intégrale de la transformée de Laplace, les propriétés de translation, conjugaison et dilatation qui suivent sont obtenues par de simples changements de variables (le vérifier).

□ **Linéarité**

$$\lambda f(t) + \mu g(t) \xrightarrow{F} \lambda F + \mu Lg, \quad \text{avec } \lambda, \mu \in \mathbb{C}.$$

□ **Translation**¹

$$\begin{aligned} H(t-t_0)f(t-t_0) &\xrightarrow{F} e^{-zt_0} F(z), \quad t_0 \in \mathbb{R}^+. \\ e^{-at} f(t) &\xrightarrow{F} F(z+a). \end{aligned}$$

□ **Conjugaison**

$$\bar{f}(t) \xrightarrow{F} \overline{F(\bar{z})}.$$

□ **Dilatation**

$$\lambda > 0, \quad f(\lambda t) \xrightarrow{F} \frac{1}{\lambda} F\left(\frac{z}{\lambda}\right).$$

◇ Exemples

- Fonction $t \mapsto t^n e^{at}$.
 $\frac{t^n}{n!} \xrightarrow{F} \frac{1}{z^{n+1}}$, donc $\frac{t^n e^{at}}{n!} \xrightarrow{F} \frac{1}{(z-a)^{n+1}}$, et pour finir $t^n e^{at} \xrightarrow{F} \frac{n!}{(z-a)^{n+1}}$.
- Original de $\frac{1}{z^2-2z+5}$.
 $\frac{1}{z^2-2z+5} = \frac{1}{(z-1)^2+4} \xrightarrow{F} \frac{e^t \sin 2t}{2}$.
- Fonction $t \mapsto e^{-at} \sin at$.
 $e^{-t} \sin t \xrightarrow{F} \frac{1}{(z+1)^2+1}$, donc $e^{-at} \sin at \xrightarrow{F} \frac{1}{a} \frac{1}{(z/a+1)^2+1} = \frac{a}{(z+a)^2+a^2}$

□ **Dérivation**

Une des applications importantes de la transformation de Laplace étant la résolution des équations différentielles, le théorème suivant est capital.

Théorème 3.4.1 Soit f une fonction continue sur \mathbb{R}^+ , sauf éventuellement en $t = 0$ où $\lim_{t \rightarrow 0^+} f(t) \equiv f(0^+)$ existe. On suppose en outre que f' est une fonction continue par morceaux qui admet une transformée de Laplace, alors :

$$f'(t) \xrightarrow{F} zF(z) - f(0^+),$$

La démonstration se fait par parties.

1. Les propriétés associées à la translation des variables sont parfois appelées "théorèmes du retard".

Ce résultat se généralise aisément (par récurrence) pour les dérivées d'ordres supérieurs :

$$f^{(n)}(t) \xrightarrow{F} z^n F(z) - z^{n-1} f(0^+) - z^{n-2} f'(0^+) - \dots - f^{(n-1)}(0^+).$$

L'apparente complication de la formule vient des sauts possibles à l'origine et de ses dérivées. On verra que ces termes sont pris automatiquement en compte dans le cadre des distributions.

Notons enfin que l'hypothèse de continuité pour les $(n-1)$ premières dérivées pour $t \neq 0$ est obligatoire pour une utilisation correcte de cette formule (cf. exercices).

◇ **Exemple**

Soit à résoudre l'équation différentielle

$$y''(t) + y(t) = \cos t, \quad y(0) = 1, \quad y'(0) = 0.$$

La transformée de Laplace $Y(z)$ s'écrit

$$Y(z) = \frac{z}{z^2 + 1} + \frac{z}{(z^2 + 1)^2} \Rightarrow y(t) = H(t) \left(\cos t + \frac{t}{2} \sin t \right).$$

□ **Intégration**

Ainsi, prendre la TL d'une dérivée revient essentiellement à multiplier par z . On ne sera pas surpris du résultat réciproque : une division par z correspond à une intégration de la fonction.

Théorème 3.4.2 Soit $\int_0^t f(t') dt'$ la primitive de f qui s'annule en 0, alors

$$\int_0^t f(t') dt' \xrightarrow{F} \frac{F(z)}{z},$$

Posons $g(t) \equiv \int_0^t f(t') dt'$. On a manifestement $g'(t) = f(t)$ et $g(0) = 0$. On a donc à la fois $f(t) \xrightarrow{F} F(z)$ et $g'(t) \xrightarrow{F} zG(z)$ par application du théorème précédent. L'identification de ces 2 résultats conduit au théorème $g(t) \xrightarrow{F} G(z) = F(z)/z$.

◇ **Exemple**

Original de $\frac{1}{z\sqrt{z}}$

$$\frac{1}{z\sqrt{z}} = \frac{1/\sqrt{z}}{z} \xleftarrow{F} \int_0^t \frac{1}{\sqrt{\pi t'}} dt' = 2\sqrt{\frac{t}{\pi}}$$

□ **Convolution**

Venons en maintenant aux propriétés liées au produit de convolution.

On rappelle que le produit de convolution $f \star g$ de 2 fonctions intégrables f et g est défini par la relation :

$$(f \star g)(t) \equiv \int_{\mathbb{R}} f(t') g(t-t') dt'$$

Supposons maintenant que f et g soient des fonctions causales. On a donc $f(t') = 0$ pour $t' < 0$ et $g(t-t') = 0$ pour $t' > t$. Le domaine d'intégration est donc restreint à l'intervalle $[0, t]$ dans le cas de fonctions causales :

$$(f \star g)(t) = \int_0^t f(t') g(t-t') dt' \quad (f \text{ et } g \text{ causales}).$$

On remarquera que $f \star g$ est elle-même causale, puisque $f(t') = 0$ pour $t' < 0$.

Le théorème central est le suivant.

Théorème 3.4.3 Soient f et g 2 fonctions causales qui admettent des TL, alors

$$(f \star g)(t) \xrightarrow{F} F(z).G(z) \text{ pour } \Re z > x_0,$$

où x_0 est la plus grande des 2 abscisses de sommabilités de f et g .

C'est une conséquence du théorème de Fubini.

◇ Exemples

1. Calcul de $H(t)t \star H(t)t^2$.

$$t \star t^2 \xrightarrow{F} \frac{1}{z^2} \cdot \frac{2}{z^3} = \frac{2}{z^5} \xleftarrow{F} 2 \frac{t^4}{4!} = \frac{t^4}{12}$$

2. Original de $\frac{1}{(z-1)(z-2)}$.

$$\frac{1}{(z-1)(z-2)} = \frac{1}{z-1} \cdot \frac{1}{z-2} \xleftarrow{F} H(t)e^t \star H(t)e^{2t} = \int_0^t e^{2(t-u)} e^u du = e^{2t} - e^t.$$

3.5 Comportements asymptotiques

□ Comportement à l'infini

On sait déjà que les transformées de Laplace sont bornées et holomorphes pour $\Re z > x_0$. Montrons en outre que la transformée de Laplace tend vers 0 à l'infini.

Théorème 3.5.1 Soit f une fonction d'abscisse de sommabilité x_0 , alors

$$\lim_{|z| \rightarrow +\infty} F(z) = 0, \text{ pour } \Re z > x_0.$$

En effet, posons $z = x_0 + Re^{i\theta}$. Prenons d'abord $|\theta| < \pi/2$, alors $\lim_{|z| \rightarrow +\infty} |f(t)e^{-zt}| = \lim_{R \rightarrow +\infty} |f(t)|e^{-x_0 t} e^{-R \cos \theta t} = 0$, puisque $\cos \theta > 0$. On obtient le résultat par application du théorème de convergence dominée. Lorsque $\theta = \pi/2$, on exprime la TL comme une TF et on utilise le lemme de Riemann-Lebesgue.

□ Théorème de la valeur finale

Comme pour la transformée de Fourier, il existe une correspondance entre le comportement d'une fonction f en $t = +\infty$ (ou en $t = 0$), et le comportement de sa transformée de Laplace F en $z = 0$ (ou en $z = +\infty$). On le voit empiriquement à partir de la définition de la transformée de Laplace, où l'on constate que l'intégrand $f(t)e^{-zt} \rightarrow 0$ lorsque $t \rightarrow +\infty$, sauf pour les petites valeurs de z , typiquement $t|z| \leq 1$. Le premier résultat précis, connu sous le nom de *théorème de la valeur finale* s'énonce :

Théorème 3.5.2 Soit $f \in L^1(\mathbb{R}^+)$. Si f a une limite $f(+\infty)$ lorsque $t \mapsto +\infty$, alors F vérifie :

$$\lim_{|z| \rightarrow 0^+} z F(z) = f(+\infty)$$

Lorsque $\Re z > 0$, $|e^{-zt} f'(t)| \leq |f'(t)|$. Donc par application du théorème de convergence dominée, on a $\lim_{|z| \rightarrow 0} (\int_{\mathbb{R}^+} f'(t) e^{-zt} dt) = \int_{\mathbb{R}^+} f'(t) dt = f(+\infty) - f(0^+)$. Par ailleurs le théorème sur la TL de $f'(t)$ donne $z F(z) - f(0^+)$. On obtient donc également $\lim_{|z| \rightarrow 0} (\int_{\mathbb{R}^+} f'(t) e^{-zt} dt) = \lim_{|z| \rightarrow 0} z F(z) - f(0^+)$, d'où le résultat.

□ Théorème de la valeur initiale

Théorème 3.5.3 Soit $f \in L^1(\mathbb{R}^+)$. Si f a une limite $f(0^+)$ lorsque $t \mapsto 0$, alors F vérifie :

$$\lim_{|z| \rightarrow +\infty} z F(z) = f(0^+)$$

$f'(t)$ a pour TL $zF(z) - f(0^+)$. Or toute TL doit tendre vers 0 lorsque $|z| \rightarrow +\infty$. Donc $\lim_{|z| \rightarrow +\infty} (zF(z) - f(0^+)) = 0$, d'où le résultat.

◇ **Exemples**

1. $f(t) = H(t)$,
 $F(z) = 1/z$, $f(0^+) = f(+\infty) = 1$,
on a bien $\lim_{|z| \rightarrow \infty} zF(z) = 1$, et $\lim_{|z| \rightarrow 0^+} zF(z) = 1$.
2. $f(t) = \cos t$,
 $F(z) = z/(z^2 + 1)$, $f(0^+) = 1$ et on a bien $\lim_{|z| \rightarrow \infty} zF(z) = 1$.
Le théorème de la valeur finale ne peut pas être utilisé car $f(+\infty)$ n'existe pas. Un calcul direct montre que $\lim_{|z| \rightarrow 0^+} zF(z) = 0$.

3.6 Inversion de la transformée de Laplace

On se pose maintenant le problème de déterminer l'original f lorsque la transformée de Laplace F est connue.

□ **Formule de Bromwich-Wagner**

Soit G une fonction holomorphe donnée. Le théorème suivant donne des conditions suffisantes sur G pour que celle-ci soit la transformée de Laplace d'une fonction.

Théorème 3.6.1 Soit G une fonction de la variable complexe telle que

- G soit holomorphe dans le demi-plan ouvert $\Re z > x_0$,
- $\lim_{|z| \rightarrow +\infty} |G(z)| = 0$, pour $\Re z > x_0$,
- Pour tout $x > x_0$, la fonction $y \in \mathbb{R} \mapsto G(x + iy)$ est sommable sur \mathbb{R} .

Soit \mathcal{B} une droite parallèle à l'axe imaginaire d'abscisse $x > x_0$. Cette droite est appelée droite de Bromwich. L'originale de la fonction G est donnée par l'intégrale :

$$\frac{1}{2i\pi} \int_{\mathcal{B}} G(z) e^{zt} dz.$$

Cette formule s'appelle la formule de Bromwich-Wagner.

Soit g l'original associé à la fonction G . Pour x donné tel que $x > x_0$, $G(z)$ peut s'exprimer comme une transformée de Fourier $G(x + i2\pi y) = \mathcal{F}[H(t)g(t)e^{-xt}](y)$. En utilisant la formule d'inversion de Fourier, on obtient :

$$H(t)g(t) = e^{+xt} \int_{\mathbb{R}} G(x + i2\pi y) e^{+i2\pi yt} dy = \frac{1}{2i\pi} \lim_{y \rightarrow +\infty} \int_{x-i2\pi y}^{x+i2\pi y} G(z) e^{zt} dz \equiv \frac{1}{2i\pi} \int_{\mathcal{B}} G(z) e^{zt} dz.$$

Remarques

1. G étant holomorphe pour $x > x_0$, toutes les singularités de G sont à gauche de \mathcal{B} .
2. Par application du théorème de Cauchy, la droite de Bromwich peut être déformée continument en n'importe quelle courbe pour peu qu'aucune des singularités de L ne soient franchies. En particulier le résultat ne doit pas dépendre de l'abscisse x de la droite de Bromwich.
3. Le fait que $G(z) \rightarrow 0$ lorsque $|z| \rightarrow +\infty$ est essentiel. L'intégrale définie dans le théorème peut exister sans correspondre pour autant à l'originale d'une transformée de Laplace. Par exemple, l'intégrale

$$\frac{1}{2i\pi} \int_{\mathcal{B}} e^{+z^2} e^{zt} dz = e^{+xt} \int_{\mathbb{R}} e^{(x+i2\pi y)^2} e^{+i2\pi yt} dy,$$

existe comme transformée de Fourier d'une gaussienne, mais ne peut pas être l'originale d'une transformée de Laplace, puisque $\lim_{|z| \rightarrow +\infty} e^{z^2} \neq 0$.

◇ **Exemples**

1. $G(z) = 1/z^n$ ($n \in \mathbb{N}^*$).

G est holomorphe dans \mathbb{C}^* , tend vers 0 lorsque z tend vers l'infini, et on a bien $y \mapsto 1/|z^n| = 1/(x^2 + 4\pi^2 y^2)^{n/2} \in L^1(\mathbb{R})$ pour $n > 1$.

Pour évaluer l'intégrale à calculer, soit :

$$\frac{1}{2i\pi} \int_{\mathcal{B}} \frac{e^{+zt}}{z^n} dz,$$

on utilise les contours dits de Bromwich représentés sur la figure 3.1. Pour $t > 0$, on utilise le théorème

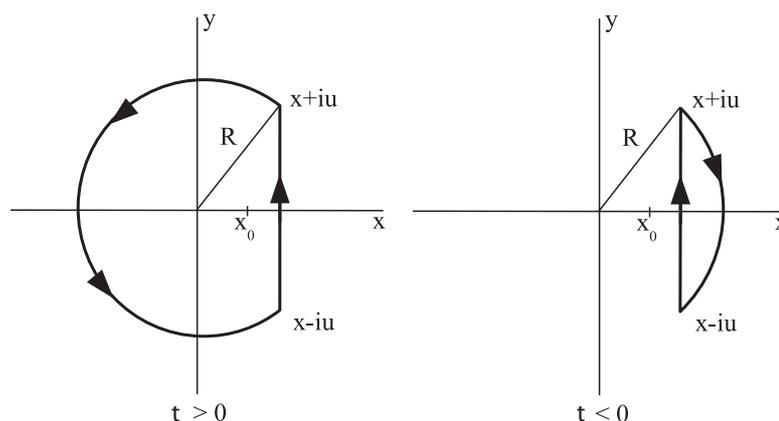


FIGURE 3.1 – Contours de Bromwich.

des résidus. Montrons que l'intégrale sur l'arc de cercle est nulle. On note d'abord que $u = \sqrt{(R + ix)}$, de sorte que $u \rightarrow +\infty$ quand $R \rightarrow +\infty$. Le domaine angulaire relatif à l'arc de cercle est donc l'intervalle $[\pi/2, 3\pi/2]$. Sur le demi-cercle où $z = R e^{i\theta}$ et $\cos \theta \leq 0$, on a $dz = iz d\theta$ de sorte que :

$$\left| \frac{e^{+zt}}{z^n} dz \right| = \left| \frac{e^{+zt}}{z^{n-1}} \right| = \frac{e^{tR \cos \theta}}{R^{n-1}} \leq \frac{1}{R^{n-1}},$$

qui tend vers 0 quand $R \rightarrow \infty$. Le seul résidu de la fonction étant le pôle $z = 0$, on peut donc écrire, pour $t > 0$

$$f(t) = \text{Res} \left(\frac{e^{+zt}}{z^n}, z = 0 \right)$$

Avec $e^{zt} = 1 + \dots + z^{n-1} \frac{t^{n-1}}{(n-1)!} + \dots$, on obtient le résultat attendu :

$$\frac{1}{2i\pi} \int_{\mathcal{B}} \frac{e^{+zt}}{z^n} dz = \frac{t^{n-1}}{(n-1)!}, \quad t > 0.$$

Comme toutes les singularités de la fonction sont à gauche de la droite de Bromwich, l'application du théorème de Cauchy au contour utilisé lorsque $t < 0$ conduit au deuxième résultat :

$$\frac{1}{2i\pi} \int_{\mathcal{B}} \frac{e^{+zt}}{z^n} dz = 0, \quad \text{pour } t < 0,$$

(l'intégrale sur l'arc de cercle s'annulant en conséquence de la même majoration que ci-dessus, avec maintenant, $t < 0$ et $\cos \theta > 0$).

La formule de Bromwich nous a donc permis de retrouver le résultat déjà connu :

$$G(z) = 1/z^n \quad \xleftarrow{F} \quad H(t)t^{n-1}/(n-1)!$$

Il est intéressant de noter que le théorème ne peut pas être appliqué à la fonction $G(z) = 1/z$ car elle n'est pas intégrable sur \mathbb{R} en tant que fonction de la variable y . La formule de Bromwich est cependant

bien définie dans ce cas et donne le bon résultat (1) (ce résultat peut être justifié à l'aide d'un autre théorème aux hypothèses plus faibles). Il est également important de remarquer que la fonction à intégrer $z \mapsto G(z) e^{zt}$ est holomorphe sur \mathbb{C}^* , et non pas seulement pour $\Re z > 0$, ce qui justifie les excursions dans le demi-plan $x < 0$.

2. $G(z) = \frac{e^{-az^{1/2}}}{z}$.

Cet exemple sera traité en TD. Indiquons seulement le choix pertinent du contour pour ce genre de fonction présentant des points de branchement. Ici, G présente un point de branchement en $z = 0$, et on doit choisir un contour qui évite 0 et qui présente une coupure; le choix standard est celui de la détermination principale du logarithme (coupure sur \mathbb{R}^-) qui permet de prolonger naturellement les résultats obtenus lorsque z est réel. Le contour est reporté sur la figure 3.2.

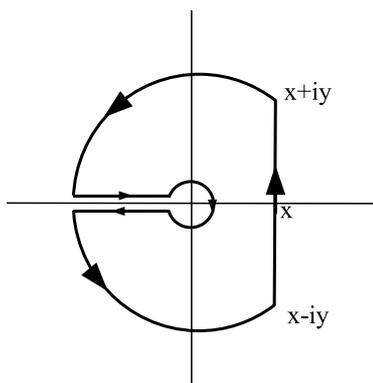


FIGURE 3.2 – Contour de Bromwich pour une fonction présentant un point de branchement à l'origine.

□ Décomposition en éléments simples

Lorsque l'image est une fraction rationnelle, il est plus simple d'effectuer une décomposition en éléments simples de la fonction plutôt que d'utiliser la formule de Bromwich.

Rappelons le principe de la décomposition d'une fraction rationnelle de la forme

$$G(z) = \frac{N(z)}{D(z)},$$

où N et D sont des polynômes tels que le degré de N soit inférieur à celui de D .

– A chaque facteur de la forme $(az + b)^n$ dans $D(z)$ correspond une décomposition de la forme :

$$\sum_{m=1}^n \frac{\alpha_m}{(az + b)^m}$$

– A chaque facteur de la forme $(az^2 + bz + c)^n$ dans $D(z)$ correspond une décomposition de la forme :

$$\sum_{m=1}^n \frac{\alpha_m z + \beta_m}{(az^2 + bz + c)^m}$$

– Les α_m et β_m sont ensuite déterminés par comparaison avec la fraction initiale $G(z)$.

◇ Exemples

1. $G(z) = \frac{1}{z(z+1)^2} = \frac{(z+1)-z}{z(z+1)^2} = \frac{1}{z(z+1)} - \frac{1}{(z+1)^2} = \frac{1}{z} - \frac{1}{z+1} - \frac{1}{(z+1)^2}$
 $g(t) = H(t) (1 - e^{-t} - t e^{-t})$.

2. $G(z) = \frac{1}{1+z+z^2} = \frac{1}{(z+1/2)^2 + 3/4}$
 $g(t) = \frac{2}{\sqrt{3}} H(t) e^{-t/2} \sin \frac{\sqrt{3}t}{2}$.

3. $G(z) = \frac{c+z}{(z-a)(z-b)} = \frac{(c+a)+(z-a)}{(z-a)(z-b)} = \frac{c+a}{a-b} \left[\frac{1}{z-a} - \frac{1}{z-b} \right] + \frac{1}{z-b} = \frac{c+a}{a-b} \frac{1}{z-a} - \frac{b+c}{a-b} \frac{1}{z-b}$
 $g(t) = H(t) \left[\frac{c+a}{a-b} e^{at} - \frac{b+c}{a-b} e^{bt} \right]$.

3.7 Remarque

Considérons la fonction constante $z \mapsto 1$. Puisqu'elle ne tend pas vers 0 à l'infini, elle ne peut être la transformée de Laplace d'une fonction au sens des fonctions. Comme pour les TF, il est possible, en passant des fonctions aux distributions, de définir des TL qui pourront croître à l'infini (pas plus vite qu'un polynôme tout de même).

3.8 Exercices

Exercice 3.1 Justifier l'existence ou la non-existence des transformées de Laplace des fonctions f suivantes :

$$\begin{aligned} a) \quad f(t) &= \sin t, \\ b) \quad f(t) &= \frac{\sin t}{t}, \\ c) \quad f(t) &= \sqrt{t}, \\ d) \quad f(t) &= \frac{\cos t}{t}, \\ e) \quad f(t) &= \int_0^t \frac{\cos \tau}{\sqrt{\tau}} d\tau, \\ f) \quad f(t) &= t^\alpha, \quad \alpha \in \mathbb{C}. \end{aligned}$$

Exercice 3.2 On rappelle que la transformée de Laplace de la fonction $t \mapsto e^{at}$ où a est un nombre complexe, est donnée par $F(z) = (z - a)^{-1}$ pour $\Re z > \Re a$.

En déduire les transformées suivantes :

$$\begin{aligned} f(t) = \sin \omega t &\Rightarrow F(z) = \frac{\omega}{z^2 + \omega^2}, \\ f(t) = \cos \omega t &\Rightarrow F(z) = \frac{z}{z^2 + \omega^2}, \\ f(t) = \sinh \omega t &\Rightarrow F(z) = \frac{\omega}{z^2 - \omega^2}, \\ f(t) = \cosh \omega t &\Rightarrow F(z) = \frac{z}{z^2 - \omega^2}, \\ f(t) = e^{-\gamma t} \sin \omega t &\Rightarrow F(z) = \frac{\omega}{(z + \gamma)^2 + \omega^2}, \\ f(t) = e^{-\gamma t} \cos \omega t &\Rightarrow F(z) = \frac{z + \gamma}{(z + \gamma)^2 + \omega^2}, \end{aligned}$$

où ω et γ sont des réels positifs.

Exercice 3.3 Soit f une fonction causale, a et t_0 des nombres réels strictement positifs. On cherche la solution de l'équation aux différences,

$$f(t) = a + f(t - t_0),$$

en utilisant la transformée de Laplace.

1. Montrer que la TL de f peut s'écrire, pour $\Re z > 0$ sous la forme d'une série :

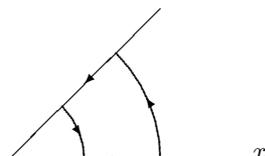
$$F(z) = \frac{a}{z} \sum_{n=0}^{+\infty} e^{-nzt_0}$$

2. En déduire la solution de l'équation aux différences par inversion et la représenter.

Exercice 3.4 On rappelle la définition de la fonction Gamma d'Euler :

$$\Gamma(z) \equiv \int_{\mathbb{R}^+} x^{z-1} e^{-x} dx,$$

1. Montrer que $\Gamma(z)$ existe dans le demi-plan $\Re z > 0$.
2. Calculer les 2 valeurs remarquables $\Gamma(1)$ et $\Gamma(1/2)$.
3. Montrer qu'on a $\Gamma(z+1) = z\Gamma(z)$ et en déduire que la fonction Gamma généralise la fonction factorielle définie pour $z \in \mathbb{N}$.
4. Utilisez le contour suivant pour montrer que :



$$f(t) = H(t) t^\alpha \Rightarrow F(z) = \frac{\Gamma(\alpha+1)}{z^{\alpha+1}}, \quad \Re \alpha > -1,$$

5. En déduire

$$f(t) = \frac{H(t-t_0)(t-t_0)^{\alpha-1}}{\Gamma(\alpha)} \Rightarrow F(z) = \frac{e^{-zt_0}}{z^\alpha}, \quad \Re \alpha > 0.$$

Exercice 3.5 1. Montrer que la fonction $f : t \mapsto f(t) = H(t) \ln t$ admet une transformée de Laplace mais pas une transformée de Fourier.

2. Soit $\lambda \in \mathbb{R}_*^+$. Montrer que F vérifie l'équation :

$$F(z/\lambda) = \lambda F(z) + \frac{\lambda \ln \lambda}{z}$$

3. Dériver l'expression précédente par rapport à λ et en déduire que $F'(z)$ obéit à l'équation différentielle :

$$F'(z) = -\frac{F(z)}{z} - \frac{1}{z^2}.$$

4. Intégrer l'équation précédente par la méthode de la variation de la constante et en déduire le résultat :

$$f(t) = H(t) \ln t \Rightarrow F(z) = -\frac{C + \log z}{z}, \quad \Re z > 0,$$

où C est la constante d'Euler-Mascheroni définie par l'intégrale : $C \equiv -\int_{\mathbb{R}^+} e^{-u} \ln u du \approx -0.577216$.

5. Déterminer sans aucun calcul la TL de la fonction $t \mapsto H(t) \ln t/\tau$ où τ est un nombre réel strictement positif.

Exercice 3.6 1. Soit f la fonction causale telle que $f(t) = \sin t$ pour $t > 0$.

Représentez les fonctions f , f' et f'' . Utilisez la relation entre $f''(t)$ et $f(t)$ pour calculer la transformée de Laplace $F(z)$.

2. Pour quelle raison ne peut-on pas utiliser la même méthode pour la fonction $g(t) = H(t) \sin|t - \pi|$?
3. Calculer $G(z)$ par une autre méthode.

Exercice 3.7 On rappelle que $(\pi/z)^{1/2}$ est la transformée de Laplace de la fonction $t \mapsto H(t)/\sqrt{t}$ pour $\Re z > 0$.

1. Déterminer la TL de la fonction $t \mapsto H(t)e^{\pm it}/\sqrt{t}$.
2. En utilisant la propriété concernant l'intégration de l'original, déterminer la TL de la fonction $t \mapsto H(t) \int_0^t e^{\pm i\tau}/\sqrt{\tau} d\tau$.
3. En déduire les transformées de Laplace des 2 intégrales de Fresnel :

$$C(t) \equiv H(t) \sqrt{\frac{2}{\pi}} \int_0^t \frac{\cos \tau}{\sqrt{\tau}} d\tau \quad \text{et} \quad S(t) \equiv H(t) \sqrt{\frac{2}{\pi}} \int_0^t \frac{\sin \tau}{\sqrt{\tau}} d\tau.$$

Exercice 3.8 On rappelle la définition des fonctions de Bessel de 1ère espèce d'ordre n ($n \in \mathbb{N}$) :

$$J_n(t) \equiv \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{i(n\theta - t \sin \theta)} d\theta$$

1. Etablir le résultat par une intégration directe :

$$f(t) = H(t) J_n(t) \quad \Rightarrow \quad F(z) = \frac{(\sqrt{1+z^2} - z)^n}{\sqrt{1+z^2}}.$$

2. Etablir la formule :

$$\int_0^t J_0(u) J_0(t-u) du = \sin t, \quad \text{pour } t > 0.$$

Exercice 3.9 On se propose de montrer que la fonction Beta d'Euler définie par la relation

$$B(p, q) \equiv \int_0^1 t^{p-1} (1-t)^{q-1} dt, \quad \Re p > 0, \quad \Re q > 0.$$

est associée à la fonction Gamma par la relation :

$$\Gamma(p+q) B(p, q) = \Gamma(p) \Gamma(q).$$

1. En effectuant le changement de variables $x = t^2$ dans la définition de la fonction Gamma, montrer que l'on peut écrire :

$$\Gamma(p) \equiv 2 \int_{\mathbb{R}^+} t^{2p-1} e^{-t^2} dt, \quad \Re p > 0.$$

2. Effectuez le produit $\Gamma(p)\Gamma(q)$ pour p et q tels que $\Re p > 0$ et $\Re q > 0$, et en déduire la relation entre la fonction Γ et la fonction B .

3. Montrer que $B(p, q)$ s'écrit comme un produit de convolution, et utilisez le théorème sur la transformée de Laplace du produit de convolution pour retrouver la relation entre B et Γ .

Exercice 3.10 On veut calculer la TL de la fonction $t \mapsto H(t) |\sin t|$.

1. Justifier l'identité

$$H(t) |\sin t| = \sum_{n=0}^{\infty} \Pi_{n\pi}(t) \sin(t - n\pi),$$

où $\Pi_{n\pi}$ est la fonction indicatrice de l'intervalle $[n\pi, (n+1)\pi]$.

2. En déduire le résultat :

$$f(t) = H(t) |\sin t| \quad \Rightarrow \quad F(z) = \frac{1}{1+z^2} \frac{1+e^{-\pi z}}{1-e^{-\pi z}}.$$

3. Utilisez la formule d'inversion pour retrouver le développement en série de Fourier de $t \mapsto |\sin t|$:

$$|\sin t| = \frac{2}{\pi} \left(1 + 2 \sum_{n=1}^{\infty} \frac{\cos 2nt}{1-4n^2} \right).$$

Exercice 3.11 1. Montrer que le théorème d'inversion peut être appliqué à la fonction :

$$F_1(z) = \frac{1}{1+z^n},$$

avec n entier supérieur ou égal à 2.

2. Calculer explicitement l'original pour $n = 2$.

Exercice 3.12 Calculer l'original des fonctions

$$F_2(z) = \frac{z}{(1+z)^3(z-1)^2}, \quad \text{et} \quad F_3(z) = \frac{e^{-az^{1/2}}}{z},$$

à l'aide de la formule d'inversion complexe en utilisant un contour de Bromwich adapté.

Exercice 3.13 1. Effectuer une décomposition en éléments simples de la fraction rationnelle

$$F(z) = \frac{2a^2z}{z^4 - a^4},$$

2. On suppose que F est la transformée de Laplace d'une fonction f . Utilisez le résultat de la question précédente pour déterminer f .

Exercice 3.14 Soit F la transformée de Laplace définie par la relation :

$$F(z) = \frac{1}{(z+a)(z+b)} \quad \text{avec} \quad a \neq b.$$

Déterminer l'original par les 3 méthodes suivantes :

1. décomposition en éléments simples,
2. théorème de convolution,
3. intégrale de Bromwich.

Deuxième partie

Equations différentielles

Chapitre 4

Introduction

Une équation différentielle est une équation mettant en jeu une fonction ainsi qu'un certain nombre de ses fonctions dérivées. La forme générale d'une équation différentielle d'ordre n s'écrit :

$$f(y, \dot{y}, \dots, y^{(n)}, t) = 0,$$

où y représente une fonction de la variable t , et $\dot{y}, \dots, y^{(n)}$ ses dérivées successives. D'un point de vue formel, le problème se pose donc de la même manière que pour les équations algébriques, mais avec la différence essentielle que l'inconnue n'est plus un nombre (réel ou complexe), mais une fonction, c'est à dire un être mathématique beaucoup plus compliqué.

L'usage des équations différentielles pour décrire le comportement des systèmes évoluant dans le temps est d'un usage universel dans toutes les sciences qui utilisent la modélisation mathématique. Cet outil commun à plusieurs disciplines ou sous-disciplines suggère bien souvent d'intéressantes analogies entre des domaines a priori sans relations. Dans ce chapitre, on commence par donner quelques exemples d'équations différentielles issues de différentes disciplines.

4.1 Mécanique.

La relation fondamentale de la mécanique, écrite à 1 dimension d'espace pour une particule ponctuelle, fournit une source intarissable d'équations différentielles. Dans un système d'unités adaptées, elle s'écrit

$$\ddot{x} = f(x, \dot{x}, t),$$

où x désigne la position de la particule, \dot{x} sa dérivée par rapport au temps (la vitesse), et où f représente les forces appliquées sur la particule. Cette équation, du second ordre en x , est généralement complétée par des *conditions initiales* qui spécifient la position et la vitesse à un instant origine : $x(0) = x_0, \dot{x}(0) = v_0$.

Il est utile de remarquer que cette équation du second ordre est équivalente à un système différentiel de 2 équations du 1er ordre. En effet, introduisons la vitesse $v \equiv \dot{x}$, l'équation précédente s'écrit aussi :

$$\begin{aligned}\dot{x} &= v, \\ \dot{v} &= f(x, v, t).\end{aligned}$$

Le plan (x, v) est appelé, aussi bien en physique qu'en mathématique, plan ou plus généralement *espace des phases*.

Dans le cas particulier où f ne dépend pas de x : $f = f(v, t)$, par exemple, dans le cas des mouvement dominés par les frottements, l'équation d'évolution de la vitesse : $\dot{v} = f(v, t)$ peut être résolue indépendamment de x . On obtient ensuite x par intégration de l'équation $\dot{x} = v$.

Si, a contrario, f ne dépend que de x : $f = f(x)$, l'équation obtenue en divisant les 2 équations différentielles

s'écrit

$$\begin{aligned}\frac{dv}{dx} &= \frac{f(x)}{v}, \\ v(0) &= v_0.\end{aligned}$$

On obtient donc encore une équation différentielle du 1er ordre, l'inconnue étant la fonction $v(x)$. Cette équation qui est séparable dans les variables v et x conduit directement à l'existence d'un *invariant* (l'énergie)

$$\frac{d}{dx} \left(\frac{1}{2} v^2(x) + W(x) \right) = 0,$$

où on a posé $W(x) \equiv -\int^x f(x')dx'$.

4.2 Dynamique des Populations.

De nombreuses modélisations de dynamique des populations (espèces animales, diffusion des virus, substances radioactives ou chimiques) ont été proposées. Parmi les plus simples, on peut citer celle attribuée à Malthus (1798) qui traduit la conservation du nombre d'individus N d'une espèce sous l'effet des naissances b et des décès d :

$$\begin{aligned}\dot{N} &= bN - dN, \\ N(0) &= N_0.\end{aligned}$$

Lorsque $b = 0$, on reconnaît dans cette équation la loi de décroissance exponentielle des substances radioactives si d est interprétée comme une constante de désintégration. Dans le cas où $b > d$, rien ne vient limiter la croissance de la population, ce qui n'est pas très réaliste. Verhulst (1836) a proposé un modèle phénoménologique non linéaire (modèle logistique) qui s'écrit

$$\begin{aligned}\dot{N} &= \alpha N \left(1 - \frac{N}{N_1} \right), \\ N(0) &= N_0,\end{aligned}$$

où α et N_1 sont des constantes positives. Ce modèle a un comportement très différent du modèle linéaire de Malthus. On montrera qu'il n'existe plus de solutions qui conduisent à l'extinction de l'espèce (la solution $N = 0$ est instable), le terme non linéaire conduisant à une stabilisation de la population vers la valeur limite $N = N_1$.

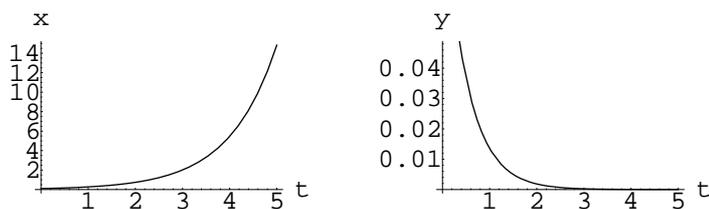
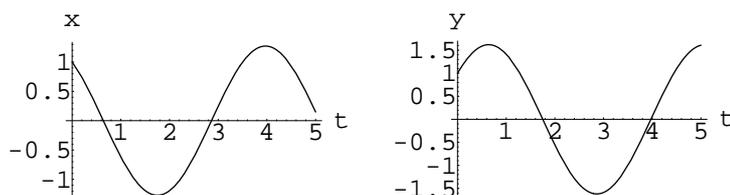
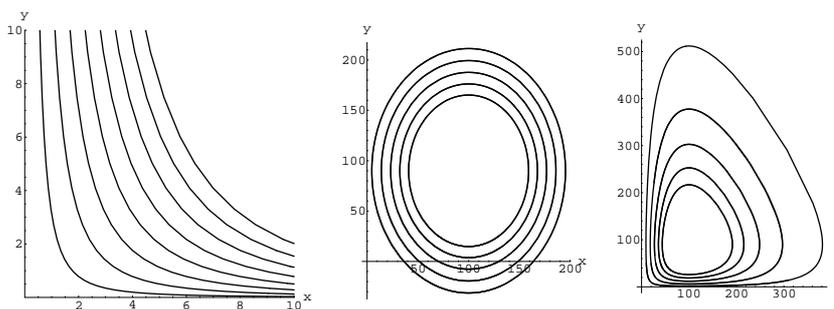
Une classe de modèles plus sophistiqués met en jeu 2 populations : une de proies (ou d'exploités) et une de prédateurs (ou d'exploiteurs). Les hypothèses suivantes sont vraisemblables :

1. en l'absence de prédateurs, les proies se multiplient proportionnellement à leur effectif.
2. en l'absence de proies, les prédateurs meurent proportionnellement à leur effectif.
3. le nombre de rencontres entre les 2 populations est proportionnel au produit des 2 populations. Chaque rencontre augmente le nombre de prédateurs et diminue le nombre de proies.

Ces hypothèses sont modélisées par le système non linéaire dit de Lokta-Volterra, où x désigne le nombre de proies et y le nombre de prédateurs

$$\begin{aligned}\dot{x} &= +a x - b xy = +ax \left(1 - \frac{b}{a} y \right), \\ \dot{y} &= -c y + d xy = -cy \left(1 - \frac{d}{c} x \right),\end{aligned}$$

et où a, b, c et d sont des constantes positives. Les solutions constantes $(0, 0)$ et $(c/d, a/b)$ sont manifestement des solutions du système. Les équations *linéarisées* autour de ces points particuliers ont un comportement très différent ainsi qu'il apparaît sur les figures (exponentiellement croissante ou décroissante autour de $(0, 0)$ et périodique autour de l'autre point). La solution au voisinage de $(0, 0)$ est manifestement instable, et nous montrerons que la solution périodique demeure stable pour le système non linéaire. Ce comportement est illustré sur les portraits de phase (représentation paramétrique des trajectoires) reportés sur la figure 1.3.

FIGURE 4.1 – Solutions du Modèle Linéarisé de Lotka Volterra autour de $(0, 0)$.FIGURE 4.2 – Solutions du Modèle Linéarisé de Lotka Volterra autour de $(c/d, a/b)$.FIGURE 4.3 – Portraits de Phase du Modèle de Lotka Volterra : cas linéarisé autour de $(0, 0)$ (à gauche), autour de $(c/d, a/b)$ (au centre) et cas non linéaire (à droite), pour différentes conditions initiales.

4.3 Equations aux Dérivées Partielles.

Mis à part les problèmes stationnaires à 1 dimension d'espace, la plupart des équations d'évolution ne sont pas des équations différentielles mais des *équations aux dérivées partielles*, c'est-à-dire des équations différentielles pour des fonctions de plusieurs variables. Il apparaît cependant que ces équations se ramènent à des équations différentielles lorsqu'on se limite à chercher des solutions sous une forme *séparable*.

Donnons deux exemples simples empruntés à l'électromagnétisme et à la mécanique quantique.

Les solutions de l'équation d'Helmholtz¹

$$\Delta\psi(x, y, z) + k^2\psi(x, y, z) = 0$$

cherchées sous la forme (séparable) : $\psi(x, y, z) = X(x)Y(y)Z(z)$, conduisent au système d'équations différentielles

1. L'opérateur Laplacien est défini par $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$

(le vérifier) :

$$\begin{aligned}\frac{d^2 X(x)}{dx^2} + l^2 X(x) &= 0, \\ \frac{d^2 Y(y)}{dy^2} + m^2 Y(y) &= 0, \\ \frac{d^2 Z(z)}{dz^2} + n^2 Z(z) &= 0,\end{aligned}$$

où l, m, n sont des constantes telles que $l^2 + m^2 + n^2 = k^2$.

D'une façon comparable, cherchons les solutions de l'équation de Schrödinger à 1 dimension d'espace dépendant du temps

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x) \psi(x, t) = i\hbar \frac{\partial \psi(x, t)}{\partial t},$$

sous la forme (séparable) $\psi(x, t) = \phi(x)f(t)$. On obtient aussitôt les 2 équations différentielles (le vérifier)

$$\begin{aligned}-\frac{\hbar^2}{2m} \frac{d^2 \phi(x)}{dx^2} + V(x)\phi(x) &= E\phi(x), \\ i\hbar f'(t) &= Ef(t).\end{aligned}$$

où E est une constante. La solution prend donc la forme bien connue $\psi(x, t) = \phi(x)e^{-iEt/\hbar}$.

Chapitre 5

Rappels sur les équations différentielles

Dans ce chapitre, on rappelle quelques résultats concernant les équations différentielles (linéaires et non linéaires). On insiste sur le fait que *toutes* les équations différentielles linéaires du 1er ordre, à coefficients constants ou pas, sont exactement intégrables. Ce résultat n'est malheureusement pas généralisable au cas des systèmes différentiels linéaires où l'on doit se contenter de résultats explicites lorsque le système est à coefficients constants. On montre également dans ce chapitre qu'un système différentiel quelconque *et* sa condition initiale peuvent être reformulés en une seule équation intégrale. La preuve de la convergence des itérations successives issues de cette formulation intégrale du problème, conduit au théorème de Cauchy-Lipschitz, qui garantit *localement* l'existence et l'unicité des solutions des systèmes différentiels linéaires sous des hypothèses assez faibles.

5.1 Terminologie

Un système différentiel linéaire d'ordre n est un système d'équations différentielles linéaires de la forme

$$\begin{aligned}\dot{y}_1(t) &= a_{11}(t)y_1(t) + \cdots + a_{1n}(t)y_n(t) + b_1(t), \\ &\dots \\ \dot{y}_n(t) &= a_{n1}(t)y_1(t) + \cdots + a_{nn}(t)y_n(t) + b_n(t),\end{aligned}$$

où $y_1 \cdots y_n$ sont les fonctions inconnues à déterminer, et où les a_{ij} et b_i sont supposées données.

Ce système différentiel peut manifestement s'écrire comme une seule équation différentielle dans \mathbb{R}^n :

$$\dot{\mathbf{y}}(t) = A(t)\mathbf{y}(t) + \mathbf{b}(t),$$

où A est la matrice des coefficients a_{ij} , et où on a introduit les vecteurs de \mathbb{R}^n $\mathbf{y} = (y_1, \dots, y_n)$, $\dot{\mathbf{y}} \equiv (\dot{y}_1, \dots, \dot{y}_n)$ et $\mathbf{b} = (b_1, \dots, b_n)$. L'équation (ou le système) est dit *homogène* si $\mathbf{b} = \mathbf{0}$, et *non homogène* lorsque $\mathbf{b} \neq \mathbf{0}$.

Toutes les autres formes pour \mathbf{b} correspondent aux cas des *équations différentielles non linéaires*. La majorité des équations différentielles non triviales rencontrées dans les applications sont non linéaires. Nous expliquerons plus loin, comment la *linéarisation* de ces équations autour de certains points remarquables peut donner des informations, au moins locales, sur le comportement des solutions d'une équation non linéaire.

Lorsque \mathbf{b} ne dépend pas explicitement du temps, mais seulement de $\mathbf{y}(t)$, on dit que l'équation différentielle est *autonome*. Le système étudié est alors invariant par translation dans le temps : 2 particules partant d'un même point à des instants différents suivront la même trajectoire. Autrement dit, si $\mathbf{y}(t)$ est solution d'une équation différentielle autonome, la solution décalée dans le temps de t_0 , $\mathbf{y}(t - t_0)$, est également solution. Dans le cas général des équations *non autonomes*, la trajectoire suivie au cours du temps ne dépend pas seulement de la position initiale, mais également de l'instant de départ.

Il est possible d'associer une équation autonome dans \mathbb{R}^{n+1} à une équation non autonome dans \mathbb{R}^n

Théorème 5.1.1 *Toute équation non autonome d'ordre n , $\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t), t)$, peut être transformée en une équation autonome d'ordre $n + 1$ par adjonction d'une variable supplémentaire $\tau(t) \equiv t$.*

En effet, puisque $\tau = t$, l'équation $\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t), t)$ est équivalente à :

$$\begin{aligned}\dot{\mathbf{y}}(t) &= \mathbf{f}(\mathbf{y}(t), \tau(t)), \\ \dot{\tau}(t) &= 1.\end{aligned}$$

qui constitue bien une équation différentielle d'ordre $n + 1$ pour les variables (\mathbf{y}, τ) . □

Le prix à payer est donc une augmentation de la dimensionalité du système étudié, ce qui peut considérablement compliquer la représentation et la compréhension des solutions.

5.2 Quelques conséquences de la linéarité

Un premier résultat important, parfois appelé *principe de superposition* (mais c'est un théorème!), affirme que toute combinaison linéaire de solutions d'une équation différentielle linéaire homogène est aussi une solution. Plus précisément :

Théorème 5.2.1 *Soit $A(t)$ une matrice $n \times n$ fonction continue de t . Si $\mathbf{y}_1(t)$ et $\mathbf{y}_2(t)$ sont 2 solutions de l'équation différentielle linéaire homogène $\dot{\mathbf{y}}(t) = A(t)\mathbf{y}(t)$, alors toute combinaison linéaire des 2 solutions :*

$$C_1\mathbf{y}_1(t) + C_2\mathbf{y}_2(t),$$

est aussi une solution.

Ce résultat évident est une conséquence directe de la linéarité (le vérifier).

Un autre résultat bien connu affirme que la solution générale d'une équation différentielle linéaire non homogène est obtenue comme la somme d'une solution particulière de l'équation non homogène et de la solution de l'équation homogène associée. Soit

Théorème 5.2.2 *Soit $\mathbf{y}_p(t)$ une solution particulière de l'équation non homogène*

$$\dot{\mathbf{y}}(t) = A(t)\mathbf{y}(t) + \mathbf{b}(t)$$

$\mathbf{y}(t)$ est solution générale de l'équation différentielle précédente si et seulement si l'on a

$$\mathbf{y}(t) = \mathbf{y}_p(t) + \mathbf{y}_h(t),$$

où $\mathbf{y}_h(t)$ est solution de l'équation homogène $\dot{\mathbf{y}}(t) = A(t)\mathbf{y}(t)$.

On a d'abord que

$$\dot{\mathbf{y}}_p(t) + \dot{\mathbf{y}}_h(t) = A(t)\mathbf{y}_p(t) + \mathbf{b}(t) + A(t)\mathbf{y}_h(t) = A(t)(\dot{\mathbf{y}}_p(t) + \dot{\mathbf{y}}_h(t)) + \mathbf{b}(t),$$

et réciproquement, si $\mathbf{y}(t) = \mathbf{y}_p(t) + \mathbf{y}_h(t)$ est solution de l'équation non homogène, alors

$$\dot{\mathbf{y}}(t) - \dot{\mathbf{y}}_p(t) = A(t)\mathbf{y}(t) + \mathbf{b}(t) - A(t)\mathbf{y}_p(t) - \mathbf{b}(t) = A(t)(\mathbf{y}(t) - \mathbf{y}_p(t)),$$

est donc bien solution de l'équation homogène. □

Ces 2 résultats très généraux ne donnent pas de moyens opérationnels pour déterminer de façon explicite les solutions des équations différentielles homogènes ou inhomogènes. Nous verrons au Chapitre 3 qu'il est possible, à l'aide d'outils d'algèbre linéaire, d'obtenir des solutions explicites mais seulement dans le cas plus simple où la matrice A ne dépend pas de la variable t .

5.3 Deux solutions explicites importantes

Dans cette section, on rappelle 2 exemples de solutions explicites fréquemment rencontrées dans les applications.

5.3.1 Equation différentielle linéaire du 1er ordre

Un premier cas très simple où le calcul explicite est possible correspond au cas de la dimension $n = 1$, c'est-à-dire au cas d'une seule équation différentielle linéaire que l'on écrira sous la forme

$$\dot{y}(t) = a(t)y(t) + b(t),$$

dont on cherche la solution avec pour condition initiale $y(0) = y_0$.

L'équation homogène est à variables séparables : $\frac{dy}{y} = a(t)dt$, et on en déduit que la solution générale de l'équation homogène est donnée par :

$$y_h(t) = y_0 e^{\int_0^t a(\tau) d\tau}.$$

Une façon élémentaire pour obtenir la solution de l'équation inhomogène consiste à utiliser la méthode dite de la *variation de la constante*. Autrement dit, on cherche la solution sous la forme $y(t) = C(t)e^{\int_0^t a(\tau) d\tau}$. En dérivant, on obtient $\dot{y}(t) = \dot{C}(t)e^{\int_0^t a(\tau) d\tau} + C(t)a(t)e^{\int_0^t a(\tau) d\tau}$. Comme on doit avoir par ailleurs $\dot{y}(t) = a(t)y(t) + b(t)$, on en déduit que la fonction $t \mapsto C(t)$ est déterminée par l'équation différentielle $\dot{C}(t) = b(t)e^{-\int_0^t a(\tau) d\tau}$ avec la condition initiale $C(0) = y_0$.

On est donc arrivé au résultat suivant :

Théorème 5.3.1 Soit $t \mapsto a(t)$ et $t \mapsto b(t)$ des fonctions continues, la solution générale de l'équation différentielle linéaire

$$\dot{y}(t) = a(t)y(t) + b(t),$$

est donnée par :

$$\begin{aligned} y(t) &= g(t) \left[y_0 + \int_0^t \frac{b(\tau)}{g(\tau)} d\tau \right], \\ g(t) &\equiv e^{\int_0^t a(\tau) d\tau}. \end{aligned}$$

Ainsi toutes les équations différentielles linéaires du 1er ordre (homogènes ou inhomogènes) sont exactement intégrables.

Faisant le lien avec le résultat de la section précédente, on pourra remarquer que $g(t) \int_0^t b(\tau)/g(\tau) d\tau$ est une solution particulière explicite de l'équation non homogène (le vérifier), tandis que $y_0 g(t)$ est la solution générale de l'équation homogène. $g(t)$, solution de l'équation homogène pour la condition initiale particulière $g(0) = 1$, est appelée la *résolvante* de l'équation différentielle.

Exemple Considérons une particule soumise à la fois à un champ de forces sinusoïdales et à une force de frottement proportionnelle à la vitesse ; le principe fondamental de la dynamique s'écrit

$$\begin{aligned} m\dot{v} &= -kv + F_0 + F_1 \cos(\omega t), \\ v(0) &= v_0, \end{aligned}$$

k, F_0, F_1 étant des constantes.

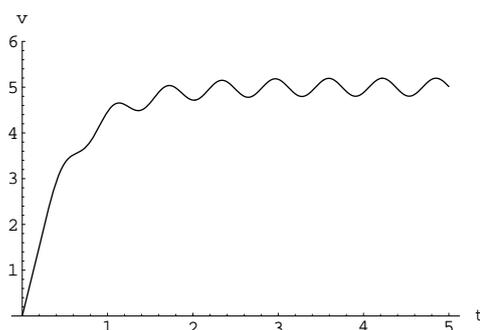
La fonction $g(t)$ étant $e^{-kt/m}$, la solution du problème s'écrit

$$v(t) = e^{-kt/m} \left[v_0 + \int_0^t e^{+k\tau/m} \left(\frac{F_0}{m} + \frac{F_1}{m} \cos(\omega\tau) \right) d\tau \right],$$

soit

$$v(t) = v_0 e^{-kt/m} + \frac{F_0}{k} (1 - e^{-kt/m}) + \frac{F_1}{k^2 + (m\omega)^2} (k \cos(\omega t) + m\omega \sin(\omega t) - k e^{-kt/m})$$

On voit donc que la solution oscille autour de la solution asymptotique $v = F_0/k$, tandis que les premières oscillations apparaissent pour un temps caractéristique de l'ordre de m/k . La position de la particule peut ensuite être obtenue par intégration de l'équation $\dot{x} = v$.

FIGURE 5.1 – Solution de $\dot{v} = -2v + 10 - 2 \cos(10t)$, $v(0) = 0$.

Exercice 5.1 N_i désignant la concentration d'un nucléide i et λ_i la constante de désintégration associée, le cas d'une filiation radioactive s'écrit

$$\begin{aligned}\dot{N}_1 &= -\lambda_1 N_1, \\ \dot{N}_2 &= +\lambda_1 N_1 - \lambda_2 N_2.\end{aligned}$$

Les conditions initiales sont telles que $N_1(0) = N_0$ et $N_2(0) = 0$. Résoudre l'équation différentielle associée à N_1 . En déduire l'équation différentielle vérifiée par N_2 et la résoudre. On distinguera les cas $\lambda_1 = \lambda_2$ et $\lambda_1 \neq \lambda_2$.

5.3.2 Equations différentielles du second ordre à coefficients constants

Un autre cas intéressant par le nombre de ses applications est celui des *équations différentielles du second ordre à coefficients constants*. La mécanique et l'électromagnétisme fournissent, parmi d'autres domaines, des exemples incontournables de telles équations. On peut citer en particulier l'équation qui décrit le mouvement d'une masse m rappelée par un ressort linéaire de constante de raideur k , en présence de frottement d'amortissement constant α (*oscillateur harmonique amorti*); si $x(t)$ désigne le déplacement par rapport à la position d'équilibre, on a :

$$m\ddot{x}(t) + \alpha\dot{x}(t) + kx(t) = 0.$$

De même, la charge $q(t)$ circulant dans un circuit RLC série alimenté par un générateur fournissant une tension $E(t)$ s'écrit

$$L\ddot{q}(t) + R\dot{q}(t) + \frac{1}{C}q(t) = E(t).$$

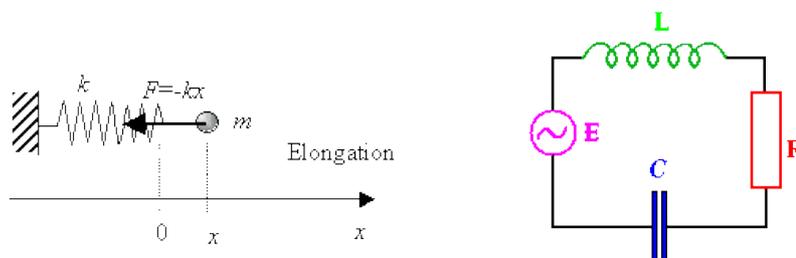


FIGURE 5.2 – Oscillateurs mécaniques et électriques.

Rappelons la démarche habituellement suivie pour déterminer les solutions de ce type d'équation différentielle :

Théorème 5.3.2 *Pour résoudre l'équation différentielle à coefficients constants :*

$$y''(x) + a y'(x) + b y(x) = 0,$$

on cherche des solutions sous la forme $y(x) = e^{rx}$ ce qui conduit à l'équation caractéristique

$$r^2 + ar + b = 0.$$

3 cas sont ensuite à considérer :

1. *L'équation caractéristique a 2 racines réelles r_1 et r_2 , et la solution générale s'écrit :*

$$y(x) = c_1 e^{r_1 x} + c_2 e^{r_2 x}$$

2. *L'équation caractéristique a 1 racine double r , et la solution générale s'écrit :*

$$y(x) = e^{rx} (c_1 + c_2 x)$$

3. *L'équation caractéristique a 2 racines complexes conjuguées $\alpha \pm i\beta$, et la solution générale s'écrit :*

$$y(x) = e^{\alpha x} (c_1 \cos \beta x + c_2 \sin \beta x).$$

Les constantes c_1 et c_2 dépendent du choix des conditions initiales.

Au vu de ces résultats, il est clair que la famille des exponentielles joue un rôle déterminant dans la solution des équations différentielles à coefficients constants. On peut déjà s'en rendre compte en étudiant les exemples très simples, mais fondamentaux, qui suivent.

Exemple Considérons pour commencer l'équation différentielle $\ddot{x}(t) = x(t)$. Quelles sont les fonctions qui ne sont pas modifiées lorsqu'on les dérive 2 fois ? Il s'agit précisément des fonctions exponentielles $e^{\pm t}$, et l'ensemble des solutions de cette équation différentielle peut s'écrire : $x(t) = A e^t + B e^{-t}$, où A et B sont des constantes arbitraires. Une solution équivalente est donnée par $x(t) = C \cosh t + D \sinh t$.

Quelles sont maintenant les fonctions qui changent de signe lorsqu'on les dérive 2 fois ? C'est-à-dire qui satisfont l'équation différentielle $\ddot{x}(t) = -x(t)$. Il s'agit encore de fonctions exponentielles, mais avec un argument imaginaire pur : $e^{\pm it}$. Les solutions générales s'écrivent donc : $x(t) = A e^{it} + B e^{-it}$, ou $x(t) = C \cos t + D \sin t$.

À la différence des 2 cas précédents, les fonctions qui s'annulent deux fois lorsqu'on les dérive : $\ddot{x}(t) = 0$, ne correspondent plus à des solutions de type exponentielle, mais à des fonctions linéaires : $x(t) = A + B t$.

Exercice 5.2 *On considère l'équation différentielle : $\ddot{x} = 2\gamma\dot{x} - x$, avec $\gamma \in \mathbb{R}$.*

- *Quelle est la solution de cette équation pour $\gamma = -1$?*
- *Discuter le comportement des racines de l'équation caractéristique au voisinage de $\gamma = -1$ (on pourra poser $\gamma = -1 + \epsilon$, avec $|\epsilon| \ll 1$).*

On propose de justifier la démarche du théorème 5.3.2 par une méthode qui n'utilise que la solution des équations différentielles homogènes ou non homogènes du 1er ordre :

Exercice 5.3 *L'équation différentielle à coefficients constants $y''(x) + a y'(x) + b y(x) = 0$, peut s'écrire formellement sous la forme :*

$$(D^2 + a D + b I) y = 0,$$

où on a introduit l'opérateur de dérivation $D = \frac{d}{dx}$ et l'identité I .

1. *Etablir les égalités*

$$y'' + a y' + b y = (D - r_1 I) (D - r_2 I) y = (D - r_2 I) (D - r_1 I) y,$$

où r_1 et r_2 sont des constantes, réelles ou complexes, que l'on déterminera.

2. Pour trouver les solutions générales y de l'équation différentielle, on pose $z = (D - r_2 I)y$. Les solutions sont donc obtenues par la résolution des équations différentielles du 1er ordre :

$$\begin{aligned}(D - r_1 I) z &= 0, \\ (D - r_2 I) y &= z.\end{aligned}$$

Intégrer successivement ces 2 équations différentielles et retrouver les solutions générales de l'équation différentielle annoncées plus haut.

La méthode proposée dans l'exercice précédent revient en fait à présenter l'équation différentielle initiale sous la forme d'un système différentiel dont la matrice est triangulaire. Elle peut être étendue sans difficulté à des équations d'ordre plus élevé avec ou sans second membre ainsi que le montrent les exemples suivants.

Exercice 5.4 Utilisez une démarche similaire à celle de l'exercice précédent pour montrer que l'équation différentielle :

$$\ddot{y}(t) - 3\dot{y}(t) + 2y(t) = 0,$$

a pour solution

$$y(t) = A e^{-2t} + B e^t + C t e^t,$$

où A, B et C sont des constantes arbitraires.

Exercice 5.5 Montrer que la solution générale de l'équation différentielle

$$\ddot{y}(t) + 3\dot{y}(t) + 4y(t) = \sin 2t,$$

est donnée par

$$y(t) = A e^{-t} + e^{-t} (B \cos t + C \sin t) - \frac{1}{10} \sin 2t.$$

On pourra commencer par résoudre l'équation sans second membre puis on cherchera une solution particulière de l'équation avec second membre.

Exercice 5.6 Montrer que la solution générale de l'équation différentielle décrivant un oscillateur harmonique forcé

$$\ddot{y}(t) + a^2 y(t) = \sin bt, \quad a, b \in \mathbb{R}$$

est donnée par

$$y(t) = A \cos at + B \sin at + \frac{1}{a^2 - b^2} \sin bt,$$

lorsque $b \neq a$, mais par

$$y(t) = A \cos at + B \sin at - \frac{t}{2a} \cos at,$$

à la résonance (i.e. lorsque $b = a$).

5.4 Système d'équations différentielles du 1er ordre

On a déjà souligné l'avantage qu'il pouvait y avoir à remplacer une équation différentielle d'ordre élevé par un système d'équations différentielles d'ordre plus bas. Cette procédure peut être rendue systématique, et simplifiée, à la fois la démonstration des théorèmes et la détermination des solutions.

Considérons par exemple le cas bien connu du pendule placé dans un champ de gravité. L'équation différentielle non linéaire du second ordre décrivant le mouvement s'écrit dans un système d'unités ad hoc

$$\ddot{x} + \sin x = 0,$$

où x représente une variable angulaire. En mécanique, il est naturel d'introduire la variable conjuguée \dot{x} qui représente la vitesse angulaire, ce qui permet d'écrire l'équation précédente du 2ième ordre sous la forme d'un système de 2 équations différentielles du 1er ordre :

$$\begin{aligned}\dot{x} &= y, \\ \dot{y} &= -\sin x.\end{aligned}$$

Plus généralement,

Théorème 5.4.1 *Toute équation différentielle (explicite) d'ordre n ($n \geq 1$), linéaire ou non linéaire, peut être transformée en un système de n équations différentielles du 1er ordre.*

Soit donc l'équation différentielle explicite d'ordre n :

$$y^{(n)} = f(y, \dot{y}, \dots, y^{(n-1)}, t).$$

Introduisons n nouvelles fonctions : $z_1 = y, z_2 = \dot{y}, \dots, z_n = y^{(n-1)}$; l'équation différentielle est alors équivalente au système différentiel (en général non linéaire) portant sur les nouvelles variables z_1, \dots, z_n .

$$\begin{aligned}\dot{z}_1 &= z_2, \\ &\dots, \\ \dot{z}_{n-1} &= z_n, \\ \dot{z}_n &= f(z_1, z_2, \dots, z_n, t).\end{aligned}$$

□

On notera que la réciproque n'est pas vérifiée : un système différentiel arbitraire de n équations différentielles du 1er ordre, ne peut pas, en général, être ramené à une seule équation différentielle d'ordre n .

Exemple L'équation différentielle linéaire :

$$y^{(n)}(t) = a_1(t)y(t) + a_2(t)\dot{y}(t) + \dots + a_n(t)y^{(n-1)},$$

peut s'écrire comme un système différentiel d'ordre n associé à la matrice

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 \\ a_1(t) & a_2(t) & \dots & a_{n-1}(t) & a_n(t) \end{pmatrix}$$

D'une façon plus systématique, on sera amené à étudier les systèmes de n équations différentielles du 1er ordre, que l'on définira par les équations

$$\begin{aligned}\dot{y}_1 &= f_1(y_1, \dots, y_n, t), \\ &\dots \\ \dot{y}_n &= f_n(y_1, \dots, y_n, t),\end{aligned}$$

complétées par les conditions initiales

$$y_1(0) = y_{01}, \dots, y_n(0) = y_{0n}.$$

Introduisons les vecteurs de \mathbb{R}^n $\mathbf{y} = (y_1, \dots, y_n)$, $\mathbf{y}_0 = (y_{01}, \dots, y_{0n})$, et la fonction vectorielle $\mathbf{f} = (f_1, \dots, f_n)$. Le système différentiel précédent peut être écrit sous la forme d'une seule équation différentielle dans \mathbb{R}^n

$$\begin{aligned}\dot{\mathbf{y}}(t) &= \mathbf{f}(\mathbf{y}(t), t), \\ \mathbf{y}(0) &= \mathbf{y}_0,\end{aligned}$$

où on retiendra que \mathbf{y} et \mathbf{y}_0 sont des éléments de \mathbb{R}^n .

Exercice 5.7 On considère le système différentiel :

$$\begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} = \begin{pmatrix} +y(t) \\ -x(t) \end{pmatrix}$$

1. Montrer que x et y satisfont la même équation différentielle du second ordre, et donner sa solution générale en fonction de 2 constantes arbitraires.
2. Les conditions initiales du système différentiel sont $x(0) = x_0$ et $y(0) = y_0$. Avec quelles conditions initiales doivent être résolues les équations différentielles du second ordre en x et en y ?
3. Montrer que la solution du système différentiel peut s'écrire sous la forme

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \mathcal{R}(t) \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

où $\mathcal{R}(t)$ est une matrice que l'on précisera.

5.5 Equation intégrale

Un formalisme encore plus compact réunit en une seule *équation intégrale* le système différentiel et sa condition initiale :

Théorème 5.5.1 Soit $\mathbf{f} : U \times I \rightarrow \mathbb{R}^n$ une fonction continue, définie sur le produit d'un ouvert U de \mathbb{R}^n et d'un intervalle I de \mathbb{R} . L'équation différentielle $\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t), t)$ qui passe par \mathbf{y}_0 à $t = 0$, est équivalente à l'équation intégrale¹

$$\mathbf{y}(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(\mathbf{y}(\tau), \tau) d\tau.$$

La condition initiale est trivialement satisfaite et le système différentiel est retrouvé directement par dérivation (le vérifier). \square

Il est alors tentant de substituer l'expression de $\mathbf{y}(t)$ sous l'intégrale et de générer ainsi la suite des *approximations de Picard*. Partant d'une fonction d'essai $\mathbf{y}^{(0)}(t)$ (un choix courant consiste à prendre la condition initiale $\mathbf{y}^{(0)}(t) = \mathbf{y}_0$), l'itération d'ordre (p) est générée à partir de la précédente par la relation :

$$\mathbf{y}^{(p)}(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(\mathbf{y}^{(p-1)}(\tau), \tau) d\tau.$$

Le problème est évidemment de savoir si la suite ainsi générée converge vers une fonction solution du système différentiel étudié et si cette solution est unique.

Exemple Cherchons les premières approximations de Picard de l'équation différentielle

$$\begin{aligned} \dot{y}(t) &= 1 + y^2(t), \\ y(0) &= 0 \end{aligned}$$

Partant de $y^{(0)}(t) = y(0) = 0$, on trouve successivement

$$\begin{aligned} y^{(1)}(t) &= \int_0^t (1 + 0^2) d\tau = t, \\ y^{(2)}(t) &= \int_0^t (1 + \tau^2) d\tau = t + \frac{t^3}{3}, \\ y^{(3)}(t) &= \int_0^t \left(1 + \left(\tau + \frac{\tau^3}{3} \right)^2 \right) d\tau = t + \frac{t^3}{3} + \frac{2t^5}{15} + \frac{t^7}{63}, \\ &\dots \end{aligned}$$

1. Plus précisément, il s'agit d'une équation intégrale non linéaire de Volterra.

Cette équation différentielle peut être intégrée exactement (elle est à variables séparables). On reconnaît le début du développement de la solution $y(t) = \tan t$ dans les premières approximations reportées ci-dessus.

5.6 Théorème d'existence et d'unicité

Dans cette section, on donne le théorème fondamental de Cauchy-Lipschitz qui garantit *localement* l'existence et l'unicité des solutions des équations différentielles, éventuellement non linéaires, dans des conditions peu contraignantes. La démonstration de ce résultat passe par la preuve de la convergence des itérations de Picard introduites auparavant. Nous en donnons la formulation dans le cas d'une seule équation différentielle.

Théorème 5.6.1 *Soit $f : I \times U \rightarrow \mathbb{R}$ une fonction continue, où U et I sont des intervalles bornés de \mathbb{R} qui contiennent respectivement les points $y = y_0$ et $t = 0$. On suppose en outre que f vérifie la condition dite de Lipschitz :*

$$\forall t \in I, \forall x, y \in U, \quad |f(t, x) - f(t, y)| \leq \lambda |x - y|,$$

avec $\lambda > 0$.

Il existe un intervalle $J \subseteq I$ sur lequel est défini l'unique solution de l'équation différentielle $\dot{y}(t) = f(t, y(t))$ qui satisfait la condition initiale $y(0) = y_0$.

On admettra que ce théorème admet une généralisation au cas d'un système d'équations différentielles.

Hormis la continuité, la condition moins triviale imposée par le théorème est la condition de Lipschitz. Une condition plus parlante qui implique l'inégalité de Lipschitz, est que f et $\frac{\partial f}{\partial y}$ soient continues dans $I \times U$. En effet, les fonctions continues sur des intervalles bornés sont bornées. Puis par le théorème des accroissements finis, on a pour tout $(t, y_1), (t, y_2) \in I \times U$,

$$|f(t, y_1) - f(t, y_2)| = \left| \frac{\partial f}{\partial y}(c) \right| |y_2 - y_1|,$$

où $c \in [y_1, y_2]$. La condition de Lipschitz résulte alors du caractère borné de la dérivée partielle.

Il est intéressant de constater que si la dérivée partielle n'est pas bornée, il peut exister plusieurs solutions, comme le montre l'exemple suivant

Exemple Considérons l'équation différentielle

$$\dot{y} = y^{1/3},$$

avec pour condition initiale $y(0) = 0$. La fonction $y \mapsto y^{1/3}$ est continue. En revanche sa dérivée n'est pas définie en 0. Les conditions du théorème ne sont donc pas remplies et on vérifiera que $y = 0$ et $y = \pm \sqrt{8/27} t^{3/2}$ sont solutions de l'équation différentielle.

Commentaires Le théorème de Cauchy-Lipschitz garantit en particulier l'unicité des solutions des équations différentielles pour une condition initiale donnée. Autrement dit, à 2 conditions initiales différentes, correspondent 2 solutions différentes pour toutes les valeurs antérieures ou postérieures de t . En terme dynamique, 2 trajectoires partant de 2 points initiaux différents ne peuvent se couper ou même se toucher. Cette remarque nous permettra plus loin de représenter facilement sur un même schéma, différentes solutions des équations différentielles non linéaires correspondant à différentes conditions initiales.

L'équation différentielle étant définie pour $t \in I$, le théorème de Cauchy-Lipschitz ne garantit l'existence de solutions (bornées), que dans un intervalle J plus petit que I . Le fait que les solutions n'existent dans le cas général que sur un intervalle plus petit que l'intervalle de définition de l'équation différentielle est associé au phénomène dit *d'explosion* en temps fini de la solution. Donnons un exemple de ce comportement.

Exemple On considère l'équation différentielle $\dot{y} = y^2$ définie pour toute valeur $t \in \mathbb{R}$, avec la condition initiale $y(0) = 25$. Cette équation différentielle non linéaire est à variables séparables et il est facile de trouver l'unique solution :

$$y(t) = \frac{1}{1/25 - t}.$$

Il est clair que la solution n'est pas définie pour $t = 1/25$. Le domaine d'existence de la solution est $J =]-\infty, 1/25[\subset \mathbb{R}$.

Ce phénomène d'explosion en temps fini des solutions est la signature d'un comportement non linéaire. On peut en effet démontrer que l'équation différentielle linéaire du type $\dot{y}(t) = a(t)y(t) + b(t)$ avec $y(0) = y_0$ où a et b sont des fonctions continues pour tout $t \in I$, admet une unique solution définie dans I tout entier.

Chapitre 6

Systemes Différentiels Linéaires

Les équations différentielles linéaires - au moins celles à coefficients constants - sont les seules équations différentielles pour lesquelles les solutions peuvent être formulées d'une façon systématique et explicite. Cette formulation s'appuie essentiellement sur des résultats d'algèbre linéaire que nous aurons ainsi l'occasion de retravailler. Une conséquence de l'existence et de l'unicité des solutions des systèmes différentiels exposée dans le chapitre précédent, est la possibilité, pour les équations linéaires, de relier la solution à l'instant origine, avec la solution à un instant t ultérieur, au moyen d'un opérateur d'évolution, appelé *propagateur*. Ce point de vue, à fort contenu géométrique, éclaire le problème de la résolution des équations différentielles sous un jour nouveau qui s'avèrera également utile pour l'interprétation des équations différentielles non linéaires.

6.1 Exponentielle de matrice

Dans cette section on donne quelques définitions et propriétés qui généralisent aux matrices des résultats bien connus pour les éléments de \mathbb{R} ou \mathbb{C} .

Définition 6.1.1 Soit A une matrice $n \times n$. L'exponentielle de A est définie par

$$e^A \equiv \sum_{k=0}^{\infty} \frac{A^k}{k!} \equiv I + A + \frac{A^2}{2!} + \dots$$

Soit α un nombre positif tel que $|a_{ij}| \leq \alpha$ pour tout i, j variant de 1 à n . Manifestement les éléments de A^2 qui s'écrivent $a_{i1}a_{1j} + \dots + a_{in}a_{nj}$ sont eux mêmes majorés par $n\alpha^2$, et plus généralement les éléments de A^k sont majorés par $n^{k-1}\alpha^k$. Ainsi chacun des éléments de e^A est tel que :

$$|(e^A)_{ij}| \leq 1 + \alpha + \frac{n\alpha^2}{2!} + \dots + \frac{n^{k-1}\alpha^k}{k!} + \dots \leq e^{n\alpha},$$

ce qui montre que tous les éléments d' e^A existent, et donc que e^A est définie. □

Démontrons maintenant 2 propriétés qui auront une importance capitale pour la suite.

Théorème 6.1.1 Soient A une matrice $n \times n$ à coefficients constants.

$$\begin{aligned} e^{(s+t)A} &= e^{sA} e^{tA}, \quad s, t \in \mathbb{R} \\ \frac{d}{dt} (e^{tA}) &= A e^{tA}, \quad t \in \mathbb{R} \end{aligned}$$

La première propriété est démontrée à l'aide de la formule du binôme :

$$\begin{aligned} e^{(s+t)A} &= \sum_{n=0}^{\infty} \frac{(s+t)^n A^n}{n!} = \sum_{n=0}^{\infty} \frac{A^n}{n!} \left[\sum_{k=0}^n \frac{n!}{(n-k)!k!} s^{n-k} t^k \right] \\ &= \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \frac{A^n s^{n-k} t^k}{(n-k)!k!} = \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{A^{m+k} s^m t^k}{m!k!} = e^{sA} e^{tA}. \end{aligned}$$

Pour démontrer la deuxième propriété, on écrit la définition de la dérivée

$$\frac{d}{dt} (e^{tA}) = \lim_{s \rightarrow 0} \frac{e^{(t+s)A} - e^{tA}}{s} = e^{tA} \lim_{s \rightarrow 0} \frac{e^{sA} - 1}{s} = e^{tA} A$$

□

Le cas particulier $s = 1$ et $t = -1$ montre que l'inverse de e^A est la matrice e^{-A} . La première propriété ne doit pas être généralisée aux cas de matrices A et B qui ne commutent pas :

$$e^{A+B} \neq e^A e^B \neq e^B e^A, \text{ en général.}$$

Exercice 6.1 Montrer que si A est une matrice réelle antisymétrique (c'est à dire telle que $A^T = -A$), alors e^{tA} est une matrice orthogonale (c'est à dire que $e^{tA} (e^{tA})^T = I$).

Indication : calculer $\frac{d}{dt} [e^{tA} (e^{tA})^T]$.

6.2 Propagateur

Reprenons de la suite des approximations de Picard et tirons partie du fait que A est indépendant de t . On a donc, formellement :

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{y}_0 + \int_0^t A \mathbf{y}(\tau) d\tau \\ &= \left[I + A \left(\int_0^t dt_1 \right) + A^2 \int_0^t \left(\int_0^{t_1} dt_2 \right) dt_1 + \dots \right] \mathbf{y}_0 \\ &= \left[I + At + \frac{A^2 t^2}{2!} + \dots \right] \mathbf{y}_0 \\ &= \sum_{n=1}^{\infty} \frac{A^n}{n!} \mathbf{y}_0 = e^{tA} \mathbf{y}_0 \end{aligned}$$

Les propriétés mêmes de l'exponentielle d'un opérateur permettent de retrouver très aisément ce résultat.

Théorème 6.2.1 Soit A une matrice $n \times n$ à coefficients constants (indépendants de t), l'équation différentielle linéaire homogène

$$\begin{aligned} \dot{\mathbf{y}}(t) &= A \mathbf{y}(t), \\ \mathbf{y}(0) &= \mathbf{y}_0, \end{aligned}$$

admet pour unique solution

$$\mathbf{y}(t) = e^{tA} \mathbf{y}_0.$$

Résultat évident puisqu'on a manifestement $e^0 = I$ et $\dot{\mathbf{y}}(t) = A e^{tA} \mathbf{y}_0 = A \mathbf{y}(t)$. L'unicité est garantie par le théorème de Cauchy-Lipschitz. □

Ce résultat fondamental met donc en jeu un opérateur e^{tA} qui permet de passer de la position à l'instant initial \mathbf{y}_0 à la position à l'instant t ; pour cette raison e^{tA} est appelé *propagateur*.

La forme explicite de la solution dans le cas non homogène est donnée par le théorème suivant :

Théorème 6.2.2 Soient J un intervalle de \mathbb{R} qui contient le point $t = 0$, A une matrice $n \times n$ qui ne dépend pas de t et $t \mapsto \mathbf{b}(t)$ une fonction continue de $t \in J$, le système différentiel linéaire non homogène

$$\begin{aligned} \dot{\mathbf{y}}(t) &= A \mathbf{y}(t) + \mathbf{b}(t), \\ \mathbf{y}(0) &= \mathbf{y}_0, \end{aligned}$$

admet une solution unique pour tout $t \in J$.

Cette solution est donnée explicitement par

$$\mathbf{y}(t) = e^{tA} \mathbf{y}_0 + \int_0^t e^{(t-\tau)A} \mathbf{b}(\tau) d\tau.$$

Soit $G(t) = e^{tA}$ le propagateur associé au système homogène. La solution de l'équation non homogène est obtenue en cherchant une solution sous la forme $\mathbf{y}(t) = G(t)\mathbf{x}(t)$. On a d'une part

$$\dot{\mathbf{y}} = \dot{G}\mathbf{x} + G\dot{\mathbf{x}} = AG\mathbf{x} + G\dot{\mathbf{x}},$$

et d'autre part $\dot{\mathbf{y}} = A\mathbf{y} + \mathbf{b} = AG\mathbf{x} + \mathbf{b}$. Il s'ensuit que \mathbf{x} satisfait

$$\begin{aligned} \dot{\mathbf{x}}(t) &= G^{-1}(t)\mathbf{b}(t), \\ \mathbf{x}(0) &= \mathbf{y}_0, \end{aligned}$$

qui admet pour unique solution

$$\mathbf{x}(t) = \mathbf{y}_0 + \int_0^t G^{-1}(\tau)\mathbf{b}(\tau) d\tau.$$

□

Ainsi la seule connaissance du propagateur permet donc de résoudre les systèmes différentiels linéaires avec ou sans second membre. Pour cette raison, le propagateur est également appelé *résolvante*.

6.3 Calcul pratique du propagateur

Les résultats précédents, pour concis qu'ils soient, ne nous dispensent cependant pas de calculer explicitement l'exponentielle d'un opérateur, ce qui peut être tout de même délicat, surtout en dimension élevée.

Si la matrice A est diagonalisable, les choses sont simples.

Théorème 6.3.1 Soit A une matrice $n \times n$ diagonalisable. Le propagateur de l'équation différentielle homogène à coefficients constants : $\dot{\mathbf{y}}(t) = A\mathbf{y}(t)$ est donné par :

$$e^{tA} = Pe^{tD}P^{-1},$$

où D est la matrice diagonale dont les éléments diagonaux sont les valeurs propres de A , et où P est la matrice dont les colonnes sont les vecteurs propres de A .

En effet, si A est diagonalisable, on a $P^{-1}AP = D$. Comme $A^n = PD^nP^{-1}$, avec D^n diagonale, on en déduit que e^{tA} est elle-même diagonale, avec pour éléments $(e^{t\lambda_1}, \dots, e^{t\lambda_n})$, de sorte que $e^{tA} = Pe^{tD}P^{-1}$. □

Dans le cas où A n'est pas diagonalisable, on peut s'appuyer sur le théorème suivant que nous ne démontrerons pas¹ :

Théorème 6.3.2 Toute matrice A possède une décomposition unique $A = S + N$, où

- S est diagonalisable,
- N est nilpotente (i.e. pour k suffisamment grand $N^k = 0$),
- S et N commutent.

Comme S et N commutent, on en déduit que $e^{tA} = e^{tS}e^{tN}$. S étant diagonalisable, e^{tS} se calcule comme expliqué ci-dessus, tandis que puisque N est nilpotente, on a, en appliquant la définition de l'exponentielle d'une matrice : $e^{tN} = I + tN + \dots + \frac{t^{k-1}}{(k-1)!}N^{k-1}$.

1. ce théorème est une conséquence du théorème de réduction à la forme normale de Jordan.

6.4 Equations différentielles en dimension 2

6.4.1 Diagonalisation des matrices 2×2

Soit A une matrice 2×2 quelconque à coefficients réels. Lorsqu'il existe un vecteur $v \neq 0$ qui satisfait :

$$Av = \lambda v,$$

on dit que λ est la *valeur propre* de A correspondant au *vecteur propre* v . Géométriquement, l'opérateur linéaire associé à la matrice A contracte ou dilate ces vecteurs de l'espace dans leur propre direction.

Posons

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

Le système $Av = \lambda v$ est équivalent à la résolution du système d'équations :

$$(A - \lambda I) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0,$$

qui n'admet de solutions non nulles, que si le déterminant est nul, soit :

$$\det(A - \lambda I) = (a - \lambda)(d - \lambda) - bc = 0.$$

En introduisant la trace de la matrice (somme des éléments diagonaux) et le déterminant de A , on montre aisément que les valeurs propres sont les racines du *pôlynome caractéristique* $P(\lambda)$ suivant :

$$P(\lambda) = \lambda^2 - \lambda \operatorname{tr} A + \det A = 0.$$

Les solutions s'écrivent

$$\lambda_{1,2} = \frac{\operatorname{tr} A \pm \sqrt{(\operatorname{tr} A)^2 - 4 \det A}}{2}$$

On peut noter que

$$\begin{aligned} \operatorname{tr} A &= \lambda_1 + \lambda_2, \\ \det A &= \lambda_1 \lambda_2. \end{aligned}$$

Les vecteurs propres sont obtenus par résolution du système linéaire $(A - \lambda I)v = 0$.

3 cas sont donc possibles

1. $(\operatorname{tr} A)^2 - 4 \det A > 0$: 2 valeurs propres réelles distinctes.
2. $(\operatorname{tr} A)^2 - 4 \det A = 0$: 1 valeur propre réelle double.
3. $(\operatorname{tr} A)^2 - 4 \det A < 0$: 2 valeurs propres complexes conjuguées.

Exercice 6.2 On considère le système différentiel

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{avec} \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Montrer que x (ou y) satisfait l'équation différentielle du 2^{ème} ordre : $P(D)x = 0$, où P représente le pôlynome caractéristique et D l'opérateur $\frac{d}{dx}$.

En déduire que x peut être obtenu par résolution du système auxiliaire :

$$\begin{pmatrix} \dot{x} \\ \dot{z} \end{pmatrix} = U \begin{pmatrix} x \\ z \end{pmatrix} \quad \text{avec} \quad U = \begin{pmatrix} \lambda_2 & 1 \\ 0 & \lambda_1 \end{pmatrix},$$

où λ_1 et λ_2 sont les valeurs propres de A , et z une variable auxiliaire.

Résoudre le système précédent et en déduire les expressions de x (ou y) en fonction de λ_1 et λ_2 .

6.4.2 Forme explicite du propagateur en dimension 2

Dans cette section on utilise le fait que les puissances successives d'une matrice ne sont pas indépendantes pour donner une forme explicite compacte au propagateur e^{tA} . Le théorème de Cayley-Hamilton joue un rôle fondamental dans cette réduction. $P(X)$ désignant le polynôme caractéristique évoqué plus haut, ce théorème affirme que $P(A) = 0$ (le vérifier), soit plus explicitement :

$$(A - \lambda_1 I)(A - \lambda_2 I) = 0.$$

1. *Valeurs propres réelles distinctes* : $\lambda_1 \neq \lambda_2$. La matrice A peut s'écrire (le vérifier) :

$$A = \lambda_1 P_1 + \lambda_2 P_2,$$

où on a introduit les 2 opérateurs de projection

$$P_1 = \frac{A - \lambda_2 I}{\lambda_1 - \lambda_2}, \quad P_2 = \frac{A - \lambda_1 I}{\lambda_2 - \lambda_1}.$$

Exercice 6.3 Soient $P_1 = (A - \lambda_2 I)/(\lambda_1 - \lambda_2)$ et $P_2 = (A - \lambda_1 I)/(\lambda_2 - \lambda_1)$, montrer que

- P_1 (resp. P_2) projette tout vecteur \mathbf{v} du plan sur la direction du vecteur propre \mathbf{v}_1 (resp. \mathbf{v}_2) parallèlement à la direction de l'autre vecteur propre \mathbf{v}_2 (resp. \mathbf{v}_1).
- $P_1^n = P_1$ et $P_2^n = P_2$ pour tout $n \geq 1$

Ces 2 opérateurs vérifient en outre, l'égalité $P_1 + P_2 = I$ (par définition), ainsi que $P_1 P_2 = P_2 P_1 = 0$, par application du théorème de Cayley-Hamilton.

Comme P_1 et P_2 commutent, on a $e^{tA} = e^{t\lambda_1 P_1} e^{t\lambda_2 P_2}$, et en utilisant les propriétés des projecteurs énoncés plus haut, on trouve

$$e^{tA} = e^{\lambda_1 t} P_1 + e^{\lambda_2 t} P_2.$$

Exercice 6.4 Etablir la relation $e^{tA} = e^{\lambda_1 t} P_1 + e^{\lambda_2 t} P_2$.

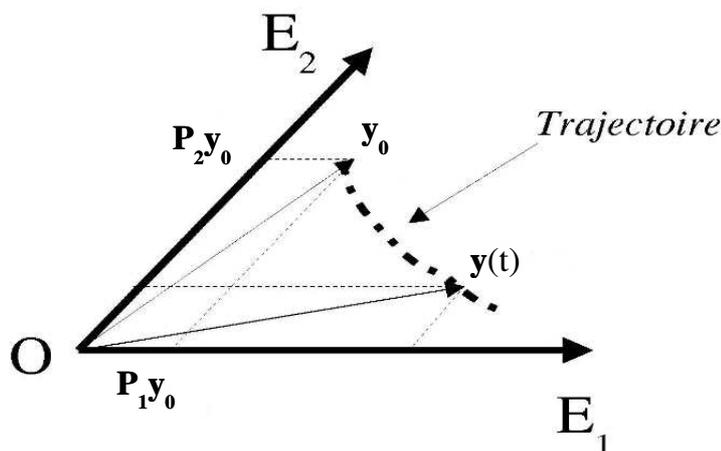


FIGURE 6.1 – Action du propagateur sur le vecteur initial \mathbf{y}_0 dans le cas de 2 valeurs propres réelles et distinctes. E_1 et E_2 désignent les sous-espaces propres engendrés par les vecteurs propres \mathbf{v}_1 et \mathbf{v}_2 . La figure est faite dans le cas $\lambda_1 > 0$ et $\lambda_2 < 0$.

On constate donc que l'action de l'opérateur sur le vecteur initial \mathbf{y}_0 , se ramène à une projection sur chacun des sous-espaces propres, accompagnée d'une dilation ou d'une contraction dans la direction de ces espaces propres.

2. *Valeurs propres complexes conjuguées* : $\lambda_{1,2} = \alpha \pm i\beta$.

Le théorème de Cayley-Hamilton s'écrit :

$$(A - \lambda_1 I)(A - \lambda_2 I) = 0 \Leftrightarrow (A - \alpha I)^2 = -\beta^2 I,$$

soit encore

$$A - \alpha I = \pm \beta J \quad \text{avec} \quad J = \begin{pmatrix} 0 & -1 \\ +1 & 0 \end{pmatrix},$$

puisque $J^2 = -I$. Comme par ailleurs,

$$e^{tA} = e^{t\alpha I} e^{\pm t\beta J},$$

le propagateur s'écrit sous la forme

$$e^{tA} = e^{\alpha t} \mathcal{R}(\pm \beta t),$$

où $\mathcal{R}(x)$ est la matrice de rotation

$$\mathcal{R}(x) = \begin{pmatrix} +\cos x & -\sin x \\ +\sin x & +\cos x \end{pmatrix}.$$

L'action géométrique du propagateur dans cette situation se ramène donc à une rotation qui rapproche ou éloigne de l'origine selon le signe de α .

Exercice 6.5 Soit

$$J = \begin{pmatrix} 0 & -1 \\ +1 & 0 \end{pmatrix},$$

montrer que

$$e^{tJ} = \begin{pmatrix} +\cos t & -\sin t \\ +\sin t & +\cos t \end{pmatrix}.$$

Montrer que ce propagateur permet de résoudre l'équation différentielle $\ddot{y}(t) + y(t) = 0$.

3. Valeur propre double : $\lambda_1 = \lambda_2 = \lambda$.

Le théorème de Cayley-Hamilton s'écrit dans ce cas particulier :

$$(A - \lambda I)^2 = 0.$$

En utilisant la décomposition $A = \lambda I + (A - \lambda I)$ et en remarquant que λI et $(A - \lambda I)$ commutent, on obtient aussitôt la forme réduite

$$e^{tA} = e^{\lambda t} [I + t(A - \lambda I)]$$

On a donc obtenu le résultat suivant

Théorème 6.4.1 Soit A une matrice 2×2 à coefficients constants, λ_1 et λ_2 les valeurs propres de A , l'équation différentielle linéaire homogène

$$\dot{\mathbf{y}}(t) = A\mathbf{y}(t),$$

$$\mathbf{y}(0) = \mathbf{y}_0,$$

admet pour solution : $\mathbf{y}(t) = e^{tA}\mathbf{y}_0$, avec

1. $e^{tA} = [e^{\lambda_1 t} P_1 + e^{\lambda_2 t} P_2]$ si $\lambda_1 \neq \lambda_2$ réelles,
2. $e^{tA} = e^{\lambda t} [I + t(A - \lambda I)]$ si $\lambda_1 = \lambda_2 = \lambda$,
3. $e^{tA} = e^{\alpha t} \mathcal{R}(\beta t)$ si $\lambda_{1,2} = \alpha \pm i\beta$,

où P_1 (resp. P_2) est l'opérateur de projection sur la direction du vecteur propre \mathbf{v}_1 (resp. \mathbf{v}_2) parallèlement à la direction de l'autre vecteur propre \mathbf{v}_2 (resp. \mathbf{v}_1), et où $\mathcal{R}(\beta t)$ est l'opérateur de rotation de l'angle βt dans le sens trigonométrique.

Ces résultats peuvent être utilisés pour résoudre les exercices suivants.

Exemple On considère le système différentiel

$$\dot{x} = +7x + 4y,$$

$$\dot{y} = -8x - 5y.$$

Comme $\text{tr}A = 2$ et $\det A = -3$, on a

$$\lambda_{1,2} = \frac{2 \pm \sqrt{4+12}}{2} = -1, +3.$$

L'application de la formule correspondant aux cas des valeurs propres réelles et différentes donne immédiatement

$$e^{tA} = \frac{e^{+3t}}{4} \begin{pmatrix} +8 & +4 \\ -8 & -4 \end{pmatrix} + \frac{e^{-t}}{-4} \begin{pmatrix} +4 & +4 \\ -8 & -8 \end{pmatrix}$$

soit,

$$e^{tA} = \begin{pmatrix} 2e^{3t} - e^{-t} & e^{3t} - e^{-t} \\ -2e^{3t} + 2e^{-t} & -e^{3t} + 2e^{-t} \end{pmatrix}$$

La solution pour des conditions initiales x_0, y_0 est donc :

$$\begin{aligned} x(t) &= (2e^{3t} - e^{-t})x_0 + (e^{3t} - e^{-t})y_0, \\ y(t) &= (-2e^{3t} + 2e^{-t})x_0 + (-e^{3t} + 2e^{-t})y_0 \end{aligned}$$

Exercice 6.6 Déterminer le propagateur et les solutions du système différentiel

$$\begin{aligned} \dot{x} &= +2x + y, \\ \dot{y} &= +2y. \end{aligned}$$

L'exercice suivant, montre qu'à contrario, si l'on sait résoudre un système différentiel linéaire, on obtient immédiatement l'exponentielle de l'opérateur associé.

Exercice 6.7 On considère le système différentiel du 1er ordre suivant :

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} 1 & a & 1 \\ 0 & 1 & b \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

où a, b et c sont des réels avec $c \neq 1$.

Résoudre ce système, équation par équation, en commençant par l'équation différentielle en z , et en déduire l'exponentielle de la matrice associée au système différentiel.

Dans le cas diagonalisable, on peut obtenir une généralisation facile en dimension quelconque.

Exercice 6.8 Soit A une matrice $n \times n$ diagonalisable. On note $\lambda_1, \dots, \lambda_n$ ses valeurs propres et $\mathbf{v}_1, \dots, \mathbf{v}_n$ les vecteurs propres associés. Vérifier que la solution de l'équation différentielle $\dot{\mathbf{y}}(t) = A\mathbf{y}(t)$ satisfaisant $\mathbf{y}(0) = \mathbf{y}_0$ est donnée par :

$$\mathbf{y}(t) = \sum_{i=1}^n c_i e^{t\lambda_i} \mathbf{v}_i,$$

où les coefficients c_i sont les composantes de \mathbf{y}_0 dans la base des vecteurs propres : $\mathbf{y}_0 = \sum_{i=1}^n c_i \mathbf{v}_i$.

Voici, enfin, un exercice relatif à une équation différentielle non homogène.

Exercice 6.9 La trajectoire d'une particule chargée en présence d'un champ magnétique statique et homogène et d'un champ électrique homogène mais dépendant du temps, peut être ramenée au système suivant :

$$\begin{aligned} \dot{v}_1(t) &= +\omega_c v_2(t) + \mathcal{E}_1(t), \\ \dot{v}_2(t) &= -\omega_c v_1(t) + \mathcal{E}_2(t), \end{aligned}$$

où v_1 et v_2 représentent les composantes cartésiennes de la vitesse dans le plan orthogonal au champ magnétique, ω_c une fréquence caractéristique (fréquence cyclotron), $\mathcal{E}_1(t)$ et $\mathcal{E}_2(t)$ les composantes convenablement normalisées du champ électrique orthogonal à la direction du champ magnétique.

Déterminer le propagateur et en déduire la solution du système différentiel.

Chapitre 7

Analyse qualitative des équations différentielles

On a donné au Chapitre 2 les théorèmes qui garantissent l'existence des solutions des équations différentielles sous des conditions pas trop contraignantes. En revanche, ces théorèmes ne donnent pas de moyen explicite pour exprimer ces solutions. En fait, dans le cas non linéaire, à l'exception de quelques cas accidentels, il n'existe pas de moyen systématique pour trouver les solutions.

On doit donc abandonner l'espoir d'une solution explicite qui permette une estimation quantitative, et se tourner vers des approches plus qualitatives. Ce point de vue qui peut apparaître plus limité de prime abord s'est en fait avéré extrêmement fécond, surtout lorsqu'il est appliqué à des systèmes d'ordre assez élevé. Au début du 20ème siècle, sous l'impulsion d'Henri Poincaré, cette démarche fut à l'origine du renouveau de l'étude des *systèmes dynamiques* et a donné naissance à *la théorie du chaos*.

7.1 Exemple

Limitons nous aux équations différentielles autonomes pour lesquelles la variable t n'intervient pas explicitement ; on considère donc

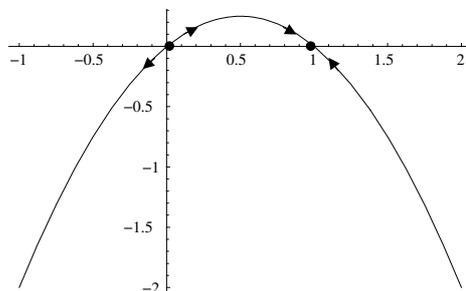
$$\begin{aligned}\dot{\mathbf{y}} &= \mathbf{f}(\mathbf{y}), \\ \mathbf{y}(0) &= \mathbf{y}_0.\end{aligned}$$

Les solutions $\bar{\mathbf{y}}$ de l'équation $\mathbf{f}(\mathbf{y}) = 0$ sont des points particuliers pour lesquels $\dot{\mathbf{y}} = \mathbf{0}$; les solutions qui passent par ces points n'évoluent pas avec t : on dit que ce sont les solutions d'équilibre. Les $\bar{\mathbf{y}}$ sont appelés *points critiques (ou fixes)*, (*ou d'équilibre*). L'importance de l'analyse locale des solutions au voisinage des points critiques tient au fait qu'elle permet une compréhension du comportement global des solutions dans les cas les plus simples.

Avant de développer ces idées d'un point de vue plus général, considérons un cas particulier en dimension 1, par exemple le modèle logistique évoqué dans l'introduction, pour des valeurs particulières des paramètres α et N_1 ($\alpha = N_1 = 1$) :

$$\begin{aligned}\dot{N} &= N(1 - N), \\ N(0) &= N_0.\end{aligned}$$

Les points fixes sont solutions de $N(1 - N) = 0$, soit $\bar{N} = 0$ et $\bar{N} = 1$. La représentation de $f(N) = N(1 - N)$ (cf. Fig 7.1) montre clairement que $f(N) > 0$ pour $0 < N < 1$; on en déduit que $\dot{y} > 0$ et donc que y est une fonction croissante dans ce même intervalle. On aboutit à la conclusion opposée dans les intervalles complémentaires. Les flèches reportées sur la figure indiquent le sens des variations de y . On dit que $\bar{N} = 0$ est un *point fixe instable*, puisque toute solution voisine de ce point fixe tend à s'en écarter ; a contrario, le point fixe $\bar{N} = 1$ est dit *stable*. Comme on sait que les trajectoires correspondant à différentes valeurs initiales N_0 ne peuvent se

FIGURE 7.1 – Application Logistique $f(N)=N(1-N)$

couper, il est facile de tracer l'évolution qualitative des solutions correspondant à des valeurs initiales placées de part et d'autre des points fixes (cf. Fig 7.2). Ce type de représentation présentant un ensemble de solutions

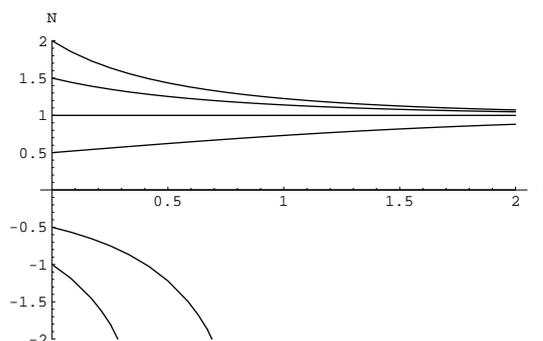


FIGURE 7.2 – Solutions du Modèle Logistique

caractéristiques sur un même schéma, permet de comprendre d'un coup d'oeil le comportement du système étudié. On voit que toutes les solutions telles que $N_0 < 0$ explosent en temps fini, tandis que toutes les solutions positives (les seules physiques puisque N représente une population) convergent vers la valeur limite $N_0 = 1$. On notera également que les solutions stationnaires constituent des barrières infranchissables (d'après le théorème de Cauchy-Lipschitz, les solutions ne peuvent se croiser). Cet exemple montre clairement que l'analyse locale des points fixes suffit à une compréhension du comportement global de ce modèle.

7.2 Classification des points fixes (1 dimension)

La classification des points fixes à 1 dimension est très simple. Comme on va le voir on peut en distinguer 3 types qualifiés *d'attractifs* (ou *stables*), *de répulsifs* (ou *instables*) ou *de points selles* (ou *semistables*). Soit \bar{y} un point fixe, c'est à dire une des solutions de l'équation $f(\bar{y}) = 0$. Le comportement local de l'équation différentielle est donné par un développement de Taylor au voisinage de \bar{y} qui conduit à l'équation

$$\dot{y} = f'(\bar{y})(y - \bar{y}) + \mathcal{O}((y - \bar{y})^2),$$

où $f'(\bar{y})$ désigne la dérivée de f par rapport à y calculée en $y = \bar{y}$. Le développement de Taylor à cet ordre permet de discuter les cas où $f'(\bar{y}) \neq 0$ qui correspond aux cas des points stables et instables; le sens de variation de y signalé sur la Figure 7.3 par des flèches est déterminé à partir du signe du produit $f'(\bar{y})(y - \bar{y})$.

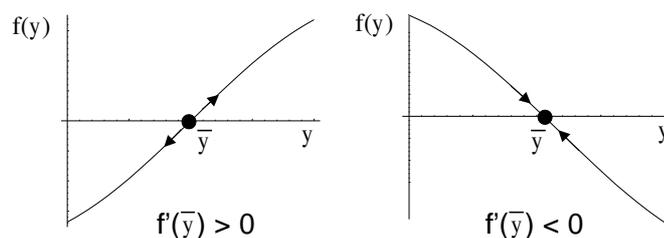


FIGURE 7.3 – Points fixes stables (à droite) et instables (à gauche).

On remarquera que la situation rencontrée dans le modèle logistique correspond à un cas de raccordement des 2 cas représentés ci-dessus.

Lorsque $f'(\bar{y}) = 0$, c'est le signe de $f''(\bar{y})$ qui permet de discuter les différents cas de points selles; les variations de y étant alors données par le signe de $f''(\bar{y})$ (développement de Taylor au 2ième ordre), voire par le signe de $f'''(\bar{y})(y - \bar{y})^3$ si $f''(\bar{y})$ est elle aussi nulle. Les différentes situations sont reportées sur la figure 7.4.

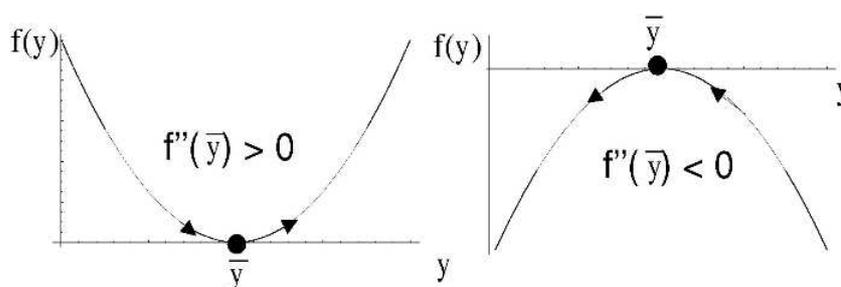


FIGURE 7.4 – Exemples de Points Selles

Les points selles présentent une autre particularité par rapport aux deux autres types de points fixes : ils sont *structurellement instables*. On entend par là que la moindre perturbation (toujours présente dans les systèmes physiques) peut faire disparaître le point selle au profit de l'apparition d'une paire de points fixes stable-instable. Lorsque cette transition est obtenue sous l'effet d'une variation d'un paramètre μ inhérent au modèle, on parle de *bifurcation*. La figure 7.5 illustre une telle bifurcation lorsque l'effet d'une variation du paramètre μ revient à translater le schéma comprenant le point selle vers le bas. Un tel effet ne peut se produire dans le cas des points fixes stables ou instables qui sont dits par conséquent *structurellement stables*.

Exercice 7.1 La décroissance d'une population suite à des collisions à 2 corps peut être modélisée par l'équation différentielle

$$\dot{N}(t) = -kN^2,$$

où $N(t)$ représente le nombre d'individus de la population à l'instant t , et où k est une constante positive. On notera N_0 le nombre d'individus à l'instant $t = 0$.

- Quel est le type de cette équation différentielle ?
- Après avoir déterminé les points critiques, tracer l'allure qualitative des solutions sans aucun calcul.
- Intégrer l'équation différentielle et donner la forme explicite des solutions.
- Mettre en évidence que la solution diverge en un temps fini.

Exercice 7.2 On considère l'équation

$$\begin{aligned} \dot{v} &= -g - k|v|, \\ v(0) &= v_0, \end{aligned}$$

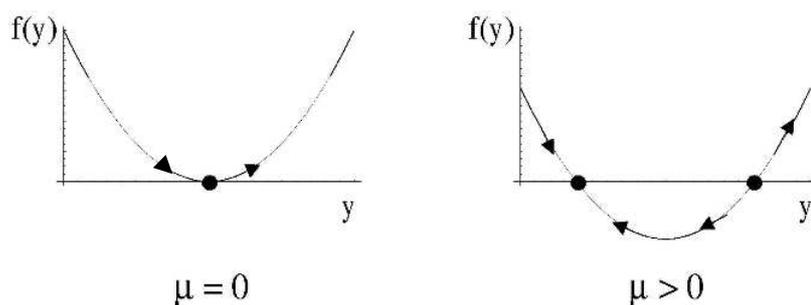


FIGURE 7.5 – Exemple de Bifurcation

où g représente l'accélération de la pesanteur, v la vitesse et k une constante positive.

- Donner une interprétation physique simple de cette équation.
- Montrer que ce système n'a qu'un seul point fixe stable, et utiliser cette information pour représenter les solutions dans les cas suivants : $v_0 > 0$, $v_0 = 0$, $-(g/k)^{1/2} < v_0 < 0$, $v_0 < -(g/k)^{1/2}$.

Exercice 7.3 On considère la réaction chimique $A + B \rightarrow C$. Les réactifs A et B sont en concentration initiale $A(0) = A_0$ et $B(0) = B_0$. La vitesse de production du produit obéit à l'équation différentielle

$$\dot{C}(t) = \alpha (C(t) - A_0) (C(t) - B_0),$$

où α est la constante (positive) de réaction.

On veut étudier comment la concentration du produit C au cours du temps dépend - qualitativement et quantitativement - de sa concentration initiale $C(0) = C_0$.

- Déterminer les points critiques et représenter qualitativement l'allure des solutions. On distinguera les 2 cas : $A_0 = B_0$ et $A_0 \neq B_0$.

Chapitre 8

Stabilité des systèmes différentiels

L'étude de la stabilité des systèmes est évidemment déterminante pour les applications (le système peut-il exister au voisinage d'un point de fonctionnement particulier en présence de perturbations?), mais également intéressante du point de vue des différents types de comportements globaux que le système peut présenter. Cette approche a été développée en dimension 1 dans le chapitre précédent, et s'appuie sur une classification des points critiques. La classification des points critiques en dimension quelconque est un problème mathématique ardu qui se rattache à *la théorie des singularités des fonctions différentiables*.

Bien que l'on puisse imaginer toute sorte de matrices définissant un système différentiel linéaire, on verra qu'il n'y a qu'un nombre limité de comportements asymptotiques (i.e aux temps longs) pour les systèmes différentiels linéaires, qui ne dépendent que des valeurs propres de la matrice associée. On montre également qu'une première approche de la stabilité (locale) des systèmes différentiels non linéaires peut être obtenue en *linéarisant* le système étudié au voisinage de ses points critiques.

8.1 Stabilité des systèmes différentiels linéaires

L'origine joue un rôle particulier pour les systèmes linéaires. Le point $\mathbf{y} = \mathbf{0}$ est en effet le seul point stationnaire ($\dot{\mathbf{y}} = \mathbf{0}$) de l'équation différentielle $\dot{\mathbf{y}} = A\mathbf{y}$.

Définition 8.1.1 *L'origine est dite asymptotiquement stable si toute solution y du système linéaire satisfait*

$$\lim_{t \rightarrow +\infty} y(t) = 0.$$

On dispose alors du résultat suivant (valable en dimension quelconque) connu sous le nom de *Critère de Routh* :

Théorème 8.1.1 *L'origine est asymptotiquement stable si et seulement si toutes les valeurs propres de A sont de partie réelle strictement négative.*

Cela résulte directement en dimension 2 du théorème 6.4.1. □

Dans le cas de la dimension 2, comme $\lambda_1 \lambda_2 = \det A$ et $\lambda_1 + \lambda_2 = \operatorname{tr} A$, on peut également dire que l'origine est asymptotiquement stable si $\det A > 0$ et $\operatorname{tr} A < 0$.

8.1.1 Portraits de phase en dimension 2

Introduisons un peu de terminologie et discutons qualitativement le comportement des solutions selon le signe des valeurs propres.

1. Si λ_1 et λ_2 sont toutes 2 négatives et différentes, on parle de *nœud* (asymptotiquement stable), ou de *puits*.
2. Si λ_1 et λ_2 sont toutes 2 positives et différentes, on parle aussi de *nœud* (instable), ou de *source*.

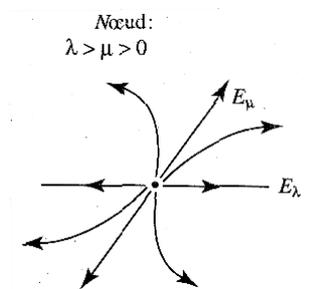


FIGURE 8.1 – Source.

3. Si une des valeurs propres est positive et l'autre négative, il existe au moins une direction instable et on parle de *point selle* ou de *col*.

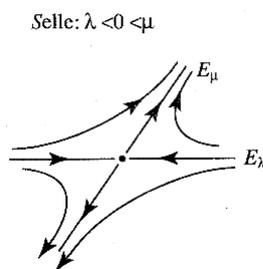


FIGURE 8.2 – Point Selle.

4. Dans le cas de valeurs propres identiques, il n'y a qu'un seul vecteur propre¹ et on parle de *nœud impropre* (stable ou instable).

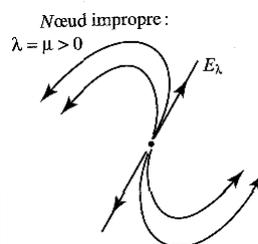


FIGURE 8.3 – Nœud Impropre Instable.

5. Dans le cas des valeurs propres conjuguées, les trajectoires sont des spirales logarithmiques lorsque la partie réelle des valeurs propres $\alpha \neq 0$ et on parle de *foyers* (stable ou instable). Dans le cas où $\alpha = 0$, les trajectoires sont périodiques, il s'agit de *centres*.

En dimension 2, l'ensemble des comportements peut être représenté dans le plan $\det A - \text{tr}A$, ce que montre la figure 8.5.

1. sauf si la matrice est de la forme λI , auquel cas, tout vecteur est vecteur propre.

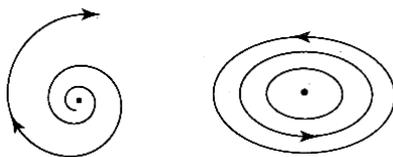
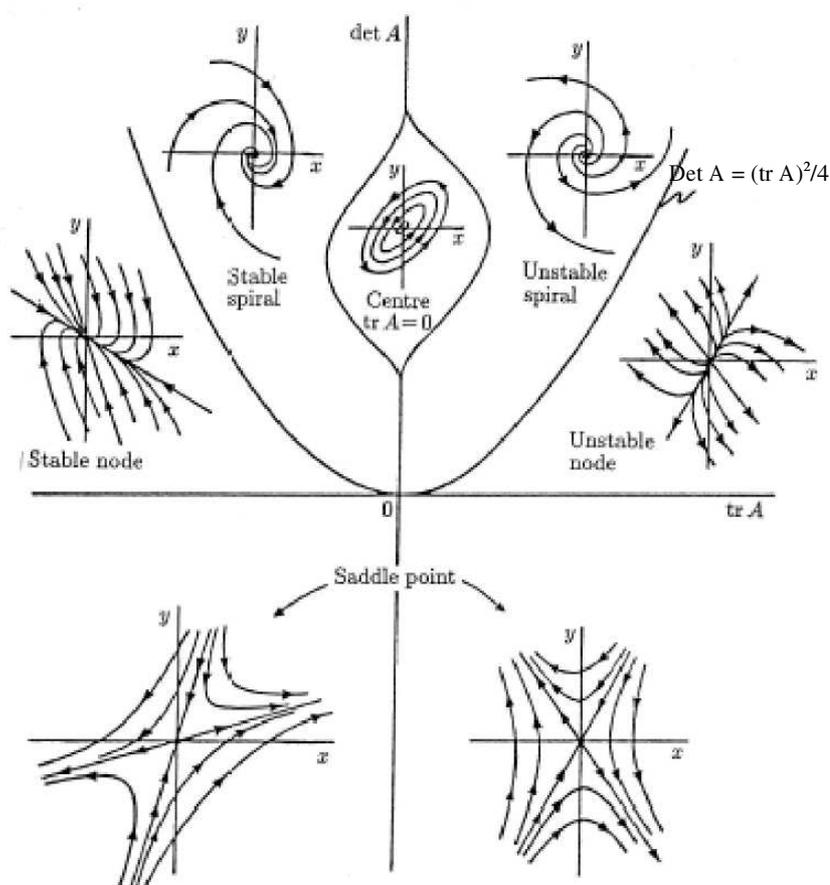


FIGURE 8.4 – Foyer instable (à gauche), et Centres (à droite).

FIGURE 8.5 – Portraits de Phases dans le plan $\det A - \text{tr} A$

8.2 Stabilité des systèmes différentiels non linéaires

Dans cette section, on se contentera d'aborder brièvement la stabilité locale des systèmes différentiels non linéaires en dimension 2 par la méthode de linéarisation.

8.2.1 Linéarisation

Le cas général des systèmes différentiels en dimension 2 peut s'écrire

$$\begin{aligned}\dot{x}_1 &= f_1(x_1, x_2), \\ \dot{x}_2 &= f_2(x_1, x_2)\end{aligned}$$

Soit (\bar{x}_1, \bar{x}_2) un point fixe du système. On peut procéder à une *linéarisation* du système autour de ce point fixe par le changement de variables $x_1 = \bar{x}_1 + \delta x_1$, $x_2 = \bar{x}_2 + \delta x_2$. δx_1 et δx_2 décrivent de petites fluctuations autour du point fixe. En utilisant le fait que $f_1(\bar{x}_1, \bar{x}_2) = f_2(\bar{x}_1, \bar{x}_2) = 0$, les fonctions f_1 et f_2 peuvent être décrites localement par leurs approximations linéaires :

$$\begin{aligned} f_1^{\text{lin}}(x_1, x_2) &= \delta x_1 \frac{\partial f_1}{\partial x_1}(\bar{x}_1, \bar{x}_2) + \delta x_2 \frac{\partial f_1}{\partial x_2}(\bar{x}_1, \bar{x}_2), \\ f_2^{\text{lin}}(x_1, x_2) &= \delta x_1 \frac{\partial f_2}{\partial x_1}(\bar{x}_1, \bar{x}_2) + \delta x_2 \frac{\partial f_2}{\partial x_2}(\bar{x}_1, \bar{x}_2). \end{aligned}$$

Comme on a bien sûr $\delta \dot{x}_1 = \dot{x}_1$ et $\delta \dot{x}_2 = \dot{x}_2$, on doit étudier le système différentiel à coefficients constants suivant

$$\begin{pmatrix} \delta \dot{x}_1 \\ \delta \dot{x}_2 \end{pmatrix} = \left(\begin{array}{cc} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{array} \right) \Big|_{(\bar{x}_1, \bar{x}_2)} \begin{pmatrix} \delta x_1 \\ \delta x_2 \end{pmatrix}$$

On est donc ramené à l'étude de la stabilité d'un système à coefficients constants ; cela permet d'apprécier si les fluctuations tendent à croître avec le temps, rendant ainsi le système instable autour du point critique étudié, ou si ces fluctuations régressent au cours du temps, ce qui correspond à la situation stable.

– **Remarque** – La question se pose bien entendu de savoir dans quelle mesure l'étude de la stabilité linéaire nous renseigne sur la stabilité du système lorsque les termes non-linéaires sont pris en compte. Lorsque les 2 valeurs propres ont une partie réelle strictement négative, la stabilité linéaire implique la stabilité non-linéaire. Dans le cas des systèmes instables et lorsque les 2 valeurs propres sont de parties réelles strictement positives, un système qui est instable par stabilité linéaire le demeure lorsque les contributions non-linéaires sont pris en compte. En revanche, lorsqu'au moins une des parties réelles des valeurs propres est nulle, c'est-à-dire dans le cas des centres, la prise en compte des termes non-linéaires peut conduire à des résultats différents de ceux obtenus par linéarisation.

– **Exemple** – On considère le système différentiel

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= -2y - \sin x. \end{aligned}$$

On veut étudier la stabilité du système au voisinage du point critique $(0, 0)$. La matrice des dérivées premières s'écrit

$$\begin{pmatrix} 0 & 1 \\ -\cos x & -2 \end{pmatrix},$$

qui vaut en $(0, 0)$:

$$A = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix}.$$

Comme $\text{tr}A = -2 < 0$ et $\det A = +1 > 0$, on en déduit que le point critique $(0, 0)$ est asymptotiquement stable. \square

Exercice 8.1 Le modèle de Lokta-Volterra s'écrit

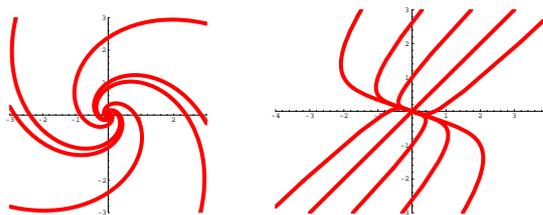
$$\begin{aligned} \dot{x} &= +a x - b xy, \\ \dot{y} &= -c y + d xy. \end{aligned}$$

où a, b, c, d sont des constantes positives.

1. Déterminer les points critiques du modèle.
2. Déterminer et résoudre les systèmes différentiels linéarisés autour de ces points fixes. En déduire la nature de ces points fixes.

Exercice 8.2 L'équation de Newton d'un pendule avec frottements s'écrit

$$\ddot{\theta}(t) = -\omega^2 \sin \theta(t) - \rho \dot{\theta}(t),$$

FIGURE 8.6 – $0 < \rho < 2\omega$ (à gauche), $\rho > 2\omega$ (à droite)

où ρ est une constante positive.

Les portraits de phase des systèmes linéarisés autour du point critique $(0,0)$ sont représentés sur la figure suivante pour différentes valeurs de ω et ρ .

Effectuer les calculs que vous jugez nécessaires pour une compréhension au moins qualitative de ces différents comportements.

8.2.2 Cycles limites

On a déjà discuté le comportement des systèmes différentiels linéaires aux temps longs en dimension 2. Parmi les solutions présentant une certaine forme de stabilité, rappelons que nous avons distingué en particulier les solutions asymptotiquement stables (puits et foyers stables) et les solutions périodiques (centres).

Les systèmes différentiels non linéaires peuvent également admettre des solutions asymptotiquement stables, mais possèdent un autre type de comportement générique connu sous le nom de *cycles limites*. A la différence des centres où à chaque condition initiale correspondait une orbite différente, les cycles limites se comportent comme des *attracteurs* pour toutes les conditions initiales dans un certain voisinage (on parle de *bassin d'attraction*) du cycle limite.

Un exemple caractéristique de système non linéaire présentant un cycle limite est fourni par *l'oscillateur non linéaire de Van der Pol* qui trouve de nombreuses applications en électricité et en mécanique. L'équation associée s'écrit :

$$\ddot{x} + (x^2 - 1)\dot{x} + x = 0.$$

Cette équation décrit une oscillation harmonique perturbée par un terme de frottement, positif si $|x| > 1$, et négatif pour $|x| < 1$. Si la condition initiale est telle que $|x| < 1$, le mouvement se trouve amplifié par le terme de frottement, cette amplification diminuant lorsque l'amplitude tend vers la valeur $|x| = 1$. A contrario, pour de fortes amplitudes, c'est-à-dire si la condition initiale est telle que $|x| > 1$, l'oscillateur est dissipatif et ramène donc la solution vers des amplitudes plus faibles. La figure suivante présente le plan de phase de l'oscillateur de Van der Pol pour plusieurs conditions initiales.

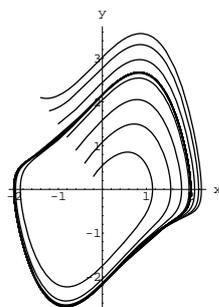


FIGURE 8.7 – Cycle Limite de l'Oscillateur de Van der Pol

Il n'est pas facile de prévoir le nombre et la position des cycles limites, même en dimension 2. L'exercice

suivant, où un calcul explicite est possible, permet de mieux comprendre ce comportement caractéristique des systèmes non linéaires.

Exercice 8.3 On considère le système dynamique suivant :

$$\begin{aligned}\dot{x}_1 &= +x_2 + \mu x_1 (1 - x_1^2 - x_2^2), \\ \dot{x}_2 &= -x_1 + \mu x_2 (1 - x_1^2 - x_2^2),\end{aligned}$$

où μ est un paramètre réel.

1. Montrer que le système linéarisé est associé à la matrice

$$\begin{pmatrix} \mu & +1 \\ -1 & \mu \end{pmatrix}$$

Diagonaliser cette matrice et discuter la stabilité du système en fonction des valeurs du paramètre μ .

2. Montrer que le changement de variable $x_1 = r \cos \theta$ et $x_2 = r \sin \theta$ permet de réécrire le système différentiel non linéaire sous la forme

$$\begin{aligned}\dot{r} &= \mu r(1 - r^2), \\ \dot{\theta} &= -1\end{aligned}$$

Intégrer ce système avec les conditions initiales $r = r_0, \theta = \theta_0$ et montrer qu'il admet la solution

$$\begin{aligned}r^2(t) &= \frac{r_0^2}{r_0^2 + (1 - r_0^2) e^{-2\mu t}}, \\ \theta(t) &= \theta_0 - t\end{aligned}$$

3. Discuter la limite de $r(t)$ lorsque $t \rightarrow \infty$ dans les 2 cas $\mu > 0$ et $\mu < 0$. On distinguera les cas $r_0 = 0$ et $r_0 \neq 0$. Commenter les résultats obtenus.

Le changement de comportement qualitatif observé lorsque le paramètre μ passe des valeurs négatives aux valeurs positives est un exemple de bifurcation dite de Hopf.

Troisième partie

Analyse dans \mathbb{R}^n

Chapitre 9

Différentier et Intégrer

Dérivées et différentielles

Circulation et Flux

Opérateurs différentiels

Formule de Stokes

Singularités

Différentielles exactes

Théorèmes de Helmholtz

Théorème de Leibnitz

La plupart des grandeurs physiques rencontrés en Physique sont des fonctions de plusieurs variables. Ainsi en va-t-il des *champs* : champs électriques, champs de vitesses, ..etc .. qui, en représentation eulérienne, sont des fonctions des variables d'espace $\vec{x} \in \mathbb{R}^3$ et du temps t .

On considèrera dans la suite, pour fixer les idées, et simplifier l'écriture, des fonctions f de 3 variables indépendantes que l'on notera x, y et z . La plupart des résultats donnés dans ce chapitre se généralise aisément aux cas d'un nombre quelconque (mais fini) de variables.

9.1 Dérivées et différentielles

1. Dérivées partielles

La dérivée partielle d'une fonction de plusieurs variables est définie d'une façon très comparable au cas des fonctions d'une seule variable ; la dérivée partielle $\partial_x f$ de la fonction f par rapport à la variable x est définie par la relation :

$$\partial_x f \equiv f_x \equiv \frac{\partial f}{\partial x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y, z) - f(x, y, z)}{\Delta x},$$

et de même pour les autres composantes. Cette définition montre que la dérivée partielle est évaluée lorsqu'une seule variable varie, les autres restant fixées pendant l'opération.

En général les dérivées partielles peuvent elles-même dépendre des variables x, y, z . Il est donc possible de définir par le même procédé des dérivées successives d'une variable : f_{xx} , et à la différence des fonctions d'une seule variable des dérivées croisées : f_{xy} . Une question très naturelle se pose de savoir si l'ordre dans lequel on prend les dérivées importe. Le résultat mathématique précis est le suivant

Théorème 9.1.1 *Si f_{xy} et f_{yx} sont continues en un point, alors ces dérivées sont égales en ce point.*

Faute de quoi, f_{xy} et f_{yx} ne sont pas forcément égales.

Par exemple, l'égalité des dérivées secondes croisées conduit en thermodynamique aux relations dites de Maxwell.

2. Différentielle. Différentiabilité

La *variation* Δf de la fonction f entre les points $\vec{r} = (x, y, z)$ et $\vec{r} + \Delta\vec{r} = (x + \Delta x, y + \Delta y, z + \Delta z)$ s'écrit

$$\Delta f \equiv f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y, z),$$

La *différentielle* est l'approximation linéaire de Δf définie par

$$df \equiv f_x \Delta x + f_y \Delta y + f_z \Delta z.$$

Ces 2 grandeurs ne sont égales que pour des fonctions linéaires.

La fonction f est *différentiable au point* \vec{r} si la différence entre la variation de la fonction et sa différentielle tend vers 0 lorsque $\Delta \vec{r} \rightarrow 0$, plus vite que $\Delta \vec{r}$ lui-même :

$$\Delta f - df = \|\Delta \vec{r}\| \epsilon(\vec{r}, \Delta \vec{r}) \quad \text{avec} \quad \lim_{\Delta \vec{r} \rightarrow 0} \epsilon(\vec{r}, \Delta \vec{r}) = 0.$$

Dans ces conditions, on voit que df est une excellente approximation de la variation Δf . Soit en particulier la fonction $f(x, y, z) = x$; il est clair que l'on a pour cette fonction $df = dx = \Delta x$, puisque $f_x = 1$ et $f_y = f_z = 0$. C'est la raison pour laquelle on note en général la différentielle d'une fonction quelconque :

$$df = f_x dx + f_y dy + f_z dz,$$

que l'on appelle la *différentielle totale* de la fonction f au point \vec{r} .

Comme pour les fonctions d'une seule variable, la différentiabilité en un point implique la continuité de la fonction au même point. En revanche, il ne suffit pas que les dérivées partielles existent en un point pour qu'une fonction y soit différentiable, il faut en outre que les dérivées partielles y soient continues (on dit dans ce cas que la fonction appartient à la classe des fonctions \mathcal{C}^1).

Exercice 9.1 Montrer que $c(x, t) = (4\pi Dt)^{-1/2} e^{-x^2/4Dt}$ satisfait l'équation de diffusion

$$c_t = D c_{xx}$$

3. Fonctions composées

Considérons la fonction $f(x, y, z)$ dans le cas où x, y et z sont elles-mêmes des fonctions d'une autre variable que nous noterons t . Cette situation se rencontre fréquemment en mécanique où $\vec{r}(t) \equiv (x(t), y(t), z(t))$ décrit la trajectoire d'une particule au cours du temps. On peut donc définir une fonction $g = f \circ \vec{r}$ dépendant de la variable t par la relation :

$$g(t) \equiv f(x(t), y(t), z(t)).$$

Il est facile d'établir le résultat suivant qui généralise la relation bien connue pour la composition de 2 fonctions u et v d'une seule variable : $(u \circ v)' = (u' \circ v) v'$

Théorème 9.1.2 Supposons que f_x, f_y et f_z soient continues et x, y, z des fonctions dérivables de la variable t , alors la fonction $g : t \mapsto f(x(t), y(t), z(t))$ est dérivable et sa dérivée vaut :

$$\frac{dg}{dt} = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_z \frac{dz}{dt}.$$

Un abus de notation, très fréquent en physique, consiste à confondre les fonctions f et g , ce qui conduit à écrire la relation précédente sous la forme :

$$\frac{df}{dt} = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_z \frac{dz}{dt}.$$

Cette pratique est justifiée dans le cadre de la physique, où les symboles mathématiques sont associés à des grandeurs physiques bien déterminées, mais il convient toutefois d'être conscient que les *fonctions* f et g sont différentes. Par opposition aux dérivées partielles, df/dt est appelée *dérivée totale*.

Une situation un peu plus générale peut être rencontrée lorsque les fonctions x, y et z dépendent elles-mêmes de plusieurs variables. Par exemple, à partir de la fonction $f(x, y, z)$, lorsque $x = x(s, t), y = y(s, t), z = z(s, t)$, on peut définir une fonction des 2 variables s et t : $g(s, t) \equiv f(x(s, t), y(s, t), z(s, t))$, qui admet donc 2 dérivées partielles données par les relations

$$\begin{aligned} g_s &= f_x x_s + f_y y_s + f_z z_s, \\ g_t &= f_x x_t + f_y y_t + f_z z_t. \end{aligned}$$

Les dérivées partielles g_s et g_t sont également couramment écrites comme les dérivées totales df/ds et df/dt . Ce genre de dérivée en chaînes, porte le nom de *chain rule* dans la littérature anglo-saxonne.

Exemple. Supposons que l'on mesure la température T à intervalles de temps réguliers le long d'une trajectoire définie par la fonction $\vec{r}(t)$. On connaît donc la fonction $T(\vec{r}(t), t)$; les variations infinitésimales de température sont données par la différentielle totale (avec l'abus de notation signalé plus haut) :

$$\frac{dT}{dt} = T_t \frac{dt}{dt} + T_x \frac{dx}{dt} + T_y \frac{dy}{dt} + T_z \frac{dz}{dt},$$

qu'il est plus courant d'écrire sous la forme¹ :

$$\frac{dT}{dt} = \frac{\partial T}{\partial t} + \frac{d\vec{r}}{dt} \cdot \vec{\nabla} T,$$

où on a introduit le gradient de T : $\vec{\nabla} T \equiv (\partial T/\partial x, \partial T/\partial y, \partial T/\partial z)$.

Cette forme met clairement en évidence le fait que les variations de température pendant le temps dt ont 2 origines distinctes, une qui provient des variations de T au point considéré, et une autre qui est une conséquence du déplacement effectué pendant le temps dt . En mécanique des fluides, cette dérivée totale est appelée *dérivée lagrangienne* ou encore dérivée particulaire ou dérivée convective. L'idée est toujours la même, il s'agit d'une information obtenue en suivant la trajectoire d'une particule mobile. L'autre point de vue, dit eulérien, consiste à effectuer les mesures en des points fixes de l'espace et à y mesurer les variations temporelles $\partial T/\partial t$, et de façon complémentaire à effectuer des mesures à un instant précis, dans différentes directions de l'espace : $\vec{\nabla} T$. Comme on le voit, la seule connaissance de la dérivée totale ne permet pas d'accéder à chacune des dérivées partielles, tandis que la connaissance de toutes les dérivées partielles et de la trajectoire permet de reconstruire la dérivée totale.

Exercice 9.2 1. Montrer que $g(x, t) \equiv f(x \pm vt)$ est solution de l'équation aux dérivées partielles

$$g_t \mp v g_x = 0$$

On posera $g = f \circ u$, avec $u(x, t) = x \pm vt$.

2. Montrer que $g(x, t) \equiv f(x/\sqrt{t})$ ramène l'équation aux dérivées partielles

$$g_t - D g_{xx} = 0,$$

à l'équation différentielle $D f_{uu} + \frac{u}{2} f_u = 0$, où $u(x, t) = x/\sqrt{t}$.

9.2 Circulation et flux

Le but de cette section est de rappeler le mode de calcul des intégrales d'une fonction de plusieurs variables, non pas sur un intervalle ou une partie du plan, ce qui devrait être acquis, mais sur une courbe ou une surface quelconque. Les notions physiques sous-jacentes sont les importantes notions de *circulation* et de *flux*. Bien qu'une définition rigoureuse et profonde de ce type d'intégrales met en jeu la notion de forme différentielle, on se limite dans la suite à l'approche intuitive utilisée en Physique qui suffit dans la plupart des cas.

1. *Intégrales sur des parties de \mathbb{R}^n*

Commençons par un bref rappel sur les intégrales dans des parties $U_n \subset \mathbb{R}^n$, c'est-à-dire sur le calcul des nombres notés

$$\int_{U_1} f(x) dx, \quad \int_{U_2} f(x, y) dS, \quad \int_{U_3} f(x, y, z) dV$$

D'un point de vue pratique, on se débrouille en général pour ramener le calcul des intégrales à 2 ou 3 dimensions au calcul d'intégrales à une seule dimension. Dans cette réduction, un point capital tient dans le fait que les éléments différentiels dS et dV peuvent s'interpréter comme des éléments de surface $dS = dx dy$ et de volume $dV = dx dy dz$.

L'autre résultat important que l'on se contente de rappeler est le théorème dit du changement de variables dont l'utilisation est souvent suggérée par les symétries particulières du problème étudié.

1. pour bien souligner la différence avec les dérivées partielles, on note quelquefois les dérivées totales par le symbole D/Dt

Théorème 9.2.1 Soient U et V 2 parties ouvertes de \mathbb{R}^3 , $J(u, v, w)$ le jacobien d'une bijection φ de U sur V définie par les relations

$$\varphi : (u, v, w) \in U \mapsto (x = x(u, v, w), y = y(u, v, w), z = z(u, v, w)) \in V$$

De façon explicite, on a donc :

$$J(u, v, w) \equiv \begin{vmatrix} x_u & y_u & z_u \\ x_v & y_v & z_v \\ x_w & y_w & z_w \end{vmatrix}$$

Soit f une fonction intégrable sur V , alors $(f \circ \varphi) \cdot |J|$ est intégrable sur U , et on a :

$$\int_V f(x, y, z) dx dy dz = \int_U f(x(u, v, w), y(u, v, w), z(u, v, w)) |J| du dv dw.$$

On a bien sûr le même résultat sur \mathbb{R} et \mathbb{R}^2 , le jacobien devenant respectivement $J(u) = dx/du$ et $J(u, v) = x_u y_v - x_v y_u$.

2. Circulation et intégrales curvilignes

Supposons que l'on s'intéresse à l'intégrale $\int_C \vec{A}(\vec{r}) \cdot d\vec{r}$ qui correspond à la circulation du vecteur \vec{A} le long de la trajectoire représentée par le chemin C (un travail si \vec{A} est une force). Il existe deux façons équivalentes de calculer cette intégrale selon l'information dont on dispose. Si la trajectoire (dans le plan Oxy pour simplifier) est connue sous la forme d'une fonction $y = f(x)$, il suffit d'exprimer le produit scalaire dans la base cartésienne (\vec{e}_x, \vec{e}_y) pour ramener l'intégrale à une somme d'intégrales d'une seule variable (x ou y) :

$$\int_C \vec{A}(\vec{r}) \cdot d\vec{r} = \int_C A_x(x, y) dx + \int_C A_y(x, y) dy$$

Comme les composantes² du vecteur \vec{A} dépendent en général des variables x et y , il convient dans ce calcul d'utiliser l'équation de la courbe $y = f(x)$ afin d'éliminer x ou y dans chacune des intégrales. Selon la géométrie du problème étudié, on peut également être amené à utiliser un autre système de coordonnées. La trajectoire associée au chemin C peut être également connue sous forme paramétrique (c'est souvent le cas en Mécanique). Dans ce cas la position le long de la trajectoire est donnée par le vecteur $\vec{r}(t) = x(t)\vec{e}_x + y(t)\vec{e}_y + z(t)\vec{e}_z$, où x , y et z sont des fonctions connues du paramètre t (on peut penser au temps) qui varie entre les deux valeurs t_a et t_b qui correspondent respectivement aux extrémités du chemin C . Dans ce cas l'intégrale est calculée par la formule suivante :

$$\int_C \vec{F}(\vec{r}) \cdot d\vec{r} \equiv \int_{t_a}^{t_b} \vec{F}(\vec{r}) \cdot \frac{d\vec{r}}{dt} dt$$

Si t est interprété comme un temps, on peut assimiler le vecteur $v \equiv d\vec{r}/dt$ à une vitesse. Le produit scalaire $\vec{F} \cdot \vec{v}$ peut alors être calculé dans un système de coordonnées adaptées de sorte que l'intégrale curviligne s'écrit encore une fois comme l'intégrale d'une fonction d'une seule variable.

Exemple. Soit à évaluer $\int_C \vec{F}(\vec{r}) \cdot d\vec{r}$ où C est un chemin du plan Oxy dont les équations paramétriques sont $x(u) = 1 - u$, $y(u) = u^2$ avec u qui varie de 0 à 1, et \vec{F} représente le champ de vecteurs $\vec{F}(x, y) = (x + y)\vec{e}_x - y\vec{e}_y$.

Comme $(\partial_u x, \partial_u y) = (-1, 2u)$ et $(F_x, F_y) = (1 - u + u^2, -u^2)$, on en déduit :

$$\int_C \vec{F}(\vec{r}) \cdot d\vec{r} = \int_0^1 \vec{F}(\vec{r}) \cdot \frac{d\vec{r}}{du} du = \int_0^1 (-1 + u - u^2 - 2u^3) du = -\frac{4}{3}$$

En éliminant u des expressions donnant $x(u)$ et $y(u)$, on trouve facilement que la trajectoire est un arc de parabole d'équation $y = (1 - x)^2$. En outre, x varie de 1 à 0 et y de 0 à 1 lorsque u varie de 0 à 1. On

2. Dans cette section, il doit être clair que A_x ne désigne pas une dérivée partielle (\vec{A} est un vecteur!), mais la projection du vecteur \vec{A} sur \vec{e}_x : $A_x \equiv \vec{A} \cdot \vec{e}_x$

aura donc aussi

$$\begin{aligned} \int_{\mathcal{C}} \vec{F}(\vec{r}) \cdot d\vec{r} &= \int_1^0 F_x(x, y) dx + \int_0^1 F_y(x, y) dy \\ &= \int_1^0 (x + (1-x)^2) dx + \int_0^1 (-y) dy = -\frac{4}{3} \end{aligned}$$

On remarquera que la valeur négative obtenue s'explique par le fait que le "mobile" remonte les lignes du champ.

En général la valeur de $\int_{\mathcal{C}} \vec{A}(\vec{r}) \cdot d\vec{r}$ dépend du chemin \mathcal{C} ; nous considérerons plus loin le cas important des intégrales curvilignes dont le résultat ne dépend pas du chemin d'intégration.

3. Flux et intégrales surfaciques

L'intégrale la plus importante est du type

$$\int_S \vec{F}(\vec{r}) \cdot d\vec{S},$$

où S est une surface régulière (ou au moins régulière par morceaux), qui correspond au flux du vecteur \vec{F} à travers la surface S .

Commençons par préciser que $d\vec{S} \equiv \vec{n}dS$, où dS est la mesure de l'élément de surface et \vec{n} un vecteur normalisé conventionnellement orienté vers l'extérieur si la surface est fermée, et obéissant à la règle du tire-bouchon si la surface est ouverte (et orientée). Si S est connue sous la forme paramétrique $\vec{r}(u, v) = x(u, v)\vec{e}_x + y(u, v)\vec{e}_y + z(u, v)\vec{e}_z$ où u et v sont des paramètres variant dans un domaine déterminé \mathcal{D} du plan uv , les 2 vecteurs $\partial_u \vec{r} du$ et $\partial_v \vec{r} dv$ forment les 2 côtés de l'élément de surface $d\vec{S}$ qui est donc donnée par le produit vectoriel $d\vec{S} \equiv (\partial_u \vec{r} \times \partial_v \vec{r}) dudv$, on obtient donc

$$\int_S \vec{F}(\vec{r}) \cdot d\vec{S} \equiv \iint_{\mathcal{D}} \vec{F}(x(u, v), y(u, v), z(u, v)) \cdot (\partial_u \vec{r} \times \partial_v \vec{r}) dudv.$$

On remarquera que le vecteur \vec{n} , normal à S s'écrit $\vec{n} = (\partial_u \vec{r} \times \partial_v \vec{r}) / \|\partial_u \vec{r} \times \partial_v \vec{r}\|$. Dans le cas du calcul d'une fonction scalaire f sur une surface, on a de même

$$\int_S f(\vec{r}) dS \equiv \iint_{\mathcal{D}} f(x(u, v), y(u, v), z(u, v)) \|\partial_u \vec{r} \times \partial_v \vec{r}\| dudv.$$

Exercice 9.3 Montrer que si la surface est connue par une équation du type $z = h(x, y)$, l'intégrale de surface est définie par la relation :

$$\int_S f(x, y, z) dS \equiv \iint_{\mathcal{D}} f(x, y, h(x, y)) [1 + (\partial_x h)^2 + (\partial_y h)^2]^{1/2} dx dy$$

où \mathcal{D} est le domaine de variation des variables x et y décrivant la surface S .

Utiliser cette formule pour montrer que la surface latérale d'un cône centré sur l'axe Oz d'équation $z = \sqrt{x^2 + y^2}$ est égale à $\sqrt{2}D$ où D est l'aire de la base du cône à la cote z .

La formule donnée dans l'exercice précédent est une expression générale qui peut prendre des formes beaucoup plus simples si on intègre sur des surfaces parallèles à des plans (par exemple sur les faces d'un cube); ainsi dans le cas d'une surface parallèle au plan Oxy , l'équation de la surface s'écrira $z = cte$ de sorte que l'intégrale se réduit à $\iint_{\mathcal{D}} f(x, y, cte) dx dy$.

Exemple. Il est bien connu que la surface de la sphère unité peut être paramétrée par la donnée de deux angles θ et φ tels que $x = \sin \theta \cos \varphi$, $y = \sin \theta \sin \varphi$, $z = \cos \theta$. Dans ce cas, le domaine \mathcal{D} de variations des paramètres est tel que $\mathcal{D} = [0, \pi] \times [0, 2\pi]$. Comme

$$\partial_\theta \vec{r} \times \partial_\varphi \vec{r} = \begin{vmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ x_\theta & y_\theta & z_\theta \\ x_\varphi & y_\varphi & z_\varphi \end{vmatrix} = \sin^2 \theta (\cos \varphi \vec{e}_x + \sin \varphi \vec{e}_y) + \sin \theta \cos \theta \vec{e}_z,$$

on en déduit que la norme de ce vecteur vaut $\sin \theta$ (le vérifier !) et donc

$$dS = \sin \theta d\theta d\varphi,$$

comme il se doit. On pourra vérifier également que le vecteur :

$$\vec{n} = \sin \theta (\cos \varphi \vec{e}_x + \sin \varphi \vec{e}_y) + \cos \theta \vec{e}_z$$

est bien normal en tout point à la surface de la sphère.

9.3 Opérateurs différentiels

Les opérateurs différentiels agissent comme leur nom l'indique par dérivation sur des champs scalaires ou vectoriels (voire tensoriels). Ces opérateurs sont dits du 1er ordre si leur définition ne met en jeu que des dérivées partielles du 1er ordre, du second ordre si apparaissent des dérivées secondes ..etc..

1. Premières définitions des opérateurs différentiels

Nous rappelons la définition usuelle des opérateurs *gradient*, *divergence* et *rotationnel* en coordonnées cartésiennes par rapport à la base canonique $(\vec{e}_x, \vec{e}_y, \vec{e}_z)$:

$$\begin{aligned} \vec{\text{grad}} f(\vec{r}) &\equiv \vec{\nabla} f(\vec{r}) \equiv f_x \vec{e}_x + f_y \vec{e}_y + f_z \vec{e}_z, \\ \text{div } \vec{F}(\vec{r}) &\equiv \vec{\nabla} \cdot \vec{F}(\vec{r}) \equiv \partial_x F_x + \partial_y F_y + \partial_z F_z, \\ \vec{\text{rot}} \vec{F}(\vec{r}) &\equiv \vec{\nabla} \times \vec{F}(\vec{r}) \equiv (\partial_y F_z - \partial_z F_y) \vec{e}_x + (\partial_z F_x - \partial_x F_z) \vec{e}_y + (\partial_x F_y - \partial_y F_x) \vec{e}_z, \end{aligned}$$

L'opérateur différentiel du second ordre le plus courant est le *laplacien*, noté Δ ou ∇^2 , défini par la relation

$$\Delta f(\vec{r}) \equiv \text{div} [\vec{\text{grad}} f(\vec{r})] \equiv \vec{\nabla} \cdot \vec{\nabla} f(\vec{r}) \equiv \nabla^2 f(\vec{r}) \equiv f_{xx} + f_{yy} + f_{zz}$$

Dans l'étude des phénomènes ondulatoires, on est également amené à introduire le laplacien d'un champ vectoriel :

$$\Delta \vec{F}(\vec{r}) \equiv \Delta F_x(\vec{r}) \vec{e}_x + \Delta F_y(\vec{r}) \vec{e}_y + \Delta F_z(\vec{r}) \vec{e}_z,$$

et l'opérateur de l'équation des ondes, ou d'*alembertien*, noté \square , qui agit sur les fonctions, scalaires ou vectorielles, des variables \vec{r} et t :

$$\square \equiv \Delta - \partial_{tt}^2.$$

Dans toutes ces définitions, il convient de noter le rôle permanent de l'opérateur $\nabla \equiv (\partial_x, \partial_y, \partial_z)$ (on dit "nabla"). La notation la plus répandue des opérateurs différentiels utilise cet opérateur, et non pas les notations grad, div, rot qui ne sont guère utilisées qu'en France. Telles quelles, ces définitions ne valent qu'en coordonnées cartésiennes, et il est intéressant de donner des définitions équivalentes - indépendantes du système de coordonnées - qui soulignent en outre le lien profond entre ces opérateurs et les notions de circulation et de flux déjà rencontrées.

2. Secondes définitions des opérateurs différentiels

Les opérateurs différentiels étudiés ici sont des opérateurs locaux ; on entend par là que les opérateurs sont définis en tout point de l'espace. Soit donc un point quelconque de l'espace, et considérons autour de ce point un volume $d\Omega$ (dont la forme pourra dépendre des symétries du problème étudié) limité par une surface fermée S . Une autre définition possible du gradient est donnée par la relation

$$\vec{\text{grad}} f(\vec{r}) \equiv \lim_{d\Omega \rightarrow 0} \frac{\oint_S f(\vec{r}) d\vec{S}}{d\Omega}.$$

En effet prenons comme volume $\Delta\Omega$ le cube de côtés $\Delta x \vec{e}_x, \Delta y \vec{e}_y, \Delta z \vec{e}_z$ situé à l'extrémité du vecteur \vec{r}

de sorte que l'on peut écrire :

$$\begin{aligned} \vec{\text{grad}} f(\vec{r}) &= \lim_{\Delta\Omega \rightarrow 0} \frac{1}{\Delta x \Delta y \Delta z} \left[\vec{e}_z \int_{\Delta x \Delta y} (f(x, y, z + dz) - f(x, y, z)) dx dy \right. \\ &+ \vec{e}_y \int_{\Delta x \Delta z} (f(x, y + dy, z) - f(x, y, z)) dx dz \\ &\left. + \vec{e}_x \int_{\Delta y \Delta z} (f(x + dx, y, z) - f(x, y, z)) dy dz \right] \end{aligned}$$

Un développement de Taylor à l'ordre le plus bas ($f(x, y, z + dz) - f(x, y, z) \approx dz \partial_z f$ pour la composante z par exemple), conduit à la contribution correspondante

$$\lim_{\Delta x \Delta y \rightarrow 0} \int_{\Delta x \Delta y} (f(x, y, z + dz) - f(x, y, z)) dx dy = (\Delta x \Delta y \Delta z) \partial_z f$$

En procédant de même sur les 2 autres axes, on obtient l'expression attendue.

La divergence d'un champ vectoriel peut être défini par la relation suivante :

$$\text{div } \vec{F}(\vec{r}) \equiv \lim_{\Delta\Omega \rightarrow 0} \frac{\oint_S \vec{F}(\vec{r}) \cdot d\vec{S}}{\Delta\Omega}$$

On vérifiera par des manipulations comparables à celles effectuées sur la définition du gradient, qu'on retrouve bien, à partir de cette définition, l'expression de la divergence en coordonnées cartésiennes. Mais il y a plus ; cette formulation est en fait indépendante du système de coordonnées utilisé (cf. exercices).

De la définition même de la divergence, il apparaît que les notions de divergence et de flux sont intimement liées puisque *la divergence d'un champ de vecteurs en un point P de l'espace est égale au flux sortant par les surfaces d'un volume infinitésimal situé autour de P, par unité de volume.*

Exercice 9.4 Utiliser la définition de l'opérateur divergence appliquée à un volume élémentaire construit sur les longueurs ($d\rho \vec{e}_\rho, \rho d\varphi \vec{e}_\varphi, dz \vec{e}_z$) pour établir la formule donnant la divergence d'un champ vectoriel en coordonnées cylindriques :

$$\text{div } \vec{F}(\vec{r}) = \frac{1}{\rho} \partial_\rho (\rho F_\rho) + \frac{1}{\rho} \partial_\varphi F_\varphi + \partial_z F_z.$$

Le lien entre circulation et rotationnel est tout aussi important ; selon la définition usuelle du rotationnel, la composante selon Oz du rotationnel s'écrit $(\vec{\nabla} \times \vec{F})_z = (\partial_x F_y - \partial_y F_x) \vec{e}_z$. Considérons à présent la petite surface rectangulaire orthogonale à l'axe Oz dont les côtés sont $\Delta x \vec{e}_x$ et $\Delta y \vec{e}_y$. Calculons la circulation de \vec{F} le long du contour \mathcal{C}_{xy} limitant la surface $\Delta x \Delta y$ et parcouru dans le sens trigonométrique :

$$\begin{aligned} \oint_{\mathcal{C}_{xy}} \vec{F} \cdot d\vec{r} &= \int_x^{x+\Delta x} F_x(x, y) dx + \int_y^{y+\Delta y} F_y(x + \Delta x, y) dy \\ &+ \int_{x+\Delta x}^x F_x(x, y + \Delta y) dx + \int_{y+\Delta y}^y F_y(x, y) dy \end{aligned}$$

Dans limite d'un élément de surface infiniment petit on obtient

$$\begin{aligned} \lim_{\Delta x \Delta y \rightarrow 0} \oint_{\mathcal{C}_{xy}} \vec{F} \cdot d\vec{r} &= F_x(x, y) \Delta x + F_y(x + \Delta x, y) \Delta y - F_x(x, y + \Delta y) \Delta x - F_y(x, y) \Delta y \\ &\approx (\partial_x F_y - \partial_y F_x) \Delta x \Delta y \end{aligned}$$

On en déduit donc la relation importante pour la composante du rotationnel selon Oz

$$\left(\vec{\nabla} \times \vec{F}\right)_z \equiv \lim_{\Delta x \Delta y \rightarrow 0} \frac{\oint_{C_{xy}} \vec{F} \cdot d\vec{r}}{\Delta x \Delta y}$$

On a évidemment le même genre de relation selon les deux autres directions, et d'une façon plus intrinsèque, la composante du rotationnel orthogonale à la surface quelconque $d\vec{S}$ peut être obtenue par la relation :

$$\left(\vec{\nabla} \times \vec{F}\right) \cdot d\vec{S} = \oint_{C_S} \vec{F} \cdot d\vec{r}, \quad (dS \rightarrow 0)$$

On peut donc dire que *la composante du rotationnel d'un champ de vecteurs, en un point P de l'espace, dans la direction arbitraire \vec{n} , est égal à la circulation le long d'un contour infinitésimal situé autour de P dans un plan orthogonal à \vec{n} , par unité de surface.*

9.4 Formule de Stokes

L'un des résultats fondamentaux relatif à l'intégration des fonctions d'une variable est la formule

$$\int_a^b f'(x) dx = f(b) - f(a).$$

Cette relation admet une généralisation en dimension supérieure connue en Mathématique sous le nom de formule de Stokes. On verra que les rôles de l'intervalle $[a, b]$ et du couple de point (a, b) seront tenus par une variété (volume, surface) et son bord (surface fermée, contour fermé), tandis que le rôle de f' sera tenu par les opérateurs différentiels divergence et rotationnel³. Bien qu'une formulation unifiée des théorèmes utilisés en Physique, (théorèmes de la divergence (ou de Gauss) et de Stokes) soit possible, elle nécessite l'introduction de notions mathématiques qui rallongerait l'exposé, et nous donnerons donc dans la suite deux énoncés distincts pour les théorèmes dits de Gauss et de Stokes.

1. Théorème de Stokes

Théorème 9.4.1 *Soit S une surface orientée, régulière par morceaux, et limitée par une courbe fermée C . Si $\vec{F}(\vec{r})$ possède des dérivées partielles continues dans S , alors*

$$\iint_S \left(\vec{\nabla} \times \vec{F}\right) \cdot d\vec{S} = \oint_C \vec{F} \cdot d\vec{r}$$

Sans donner une preuve rigoureuse de ce théorème, on peut en comprendre l'origine en revenant à la définition (locale) du rotationnel donné à la section précédente :

$$\left(\vec{\nabla} \times \vec{F}\right) \cdot d\vec{S} = \oint_{C_S} \vec{F} \cdot d\vec{r},$$

où dS est ici une surface élémentaire infiniment petite ($dS \rightarrow 0$) qui entoure le point situé en \vec{r} . Pour obtenir le théorème de Stokes qui s'applique à une surface S finie, on décompose cette surface en une infinité de surfaces infinitésimalement petites, et on en additionne les contributions. Toutes les intégrales curvilignes *intérieures* à S vont s'annuler puisque les segments correspondants seront parcourus en sens inverse (cf. Figure 9.1). La seule contribution restante est celle du contour *extérieur*, ce qui est le contenu du théorème.

2. Théorème de Gauss

Théorème 9.4.2 *Soit S une surface orientée, régulière par morceaux, qui contient le volume V . Si $\vec{F}(\vec{r})$ possède des dérivées partielles continues dans V , alors*

$$\iiint_V \left(\vec{\nabla} \cdot \vec{F}\right) dV = \oint_S \vec{F} \cdot d\vec{S}$$

3. en fait la dérivée extérieure d'une forme différentielle.

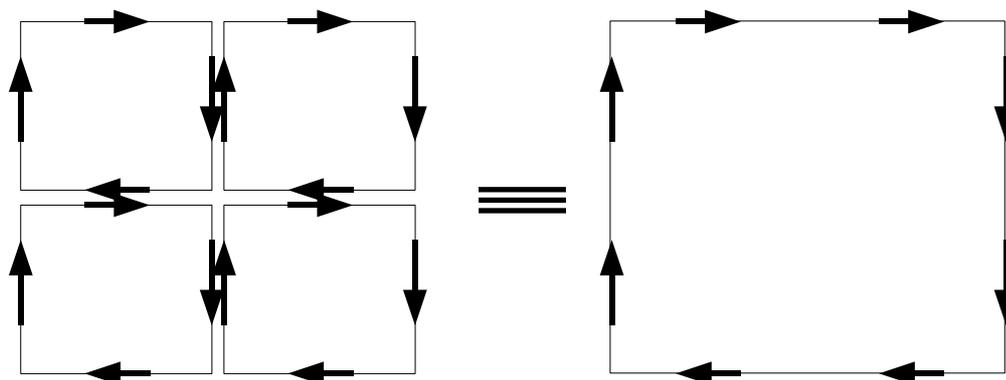


FIGURE 9.1 – Théorème de Stokes

L'idée de la démonstration est la même que pour celle du théorème de Stokes. La définition de l'opérateur divergence :

$$(\vec{\nabla} \cdot \vec{F}) dV = \oint_S \vec{F} \cdot d\vec{S}$$

s'applique à un volume élémentaire $dV \rightarrow 0$ (c'est une définition locale!). Pour obtenir le théorème de Gauss étendu à un volume fini V , on somme les contributions des cubes infinitésimaux dV . Le résultat est obtenu en prenant conscience que les contributions des flux à travers les surfaces des cubes adjacents s'annulent mutuellement, et que seule reste la contribution du flux à travers la surface extérieure S .

Exercice 9.5 Soit S une région du plan Oxy limité par la courbe C , P et Q 2 fonctions C^1 des variables x, y , montrer que le théorème de Green-Riemann dans le plan

$$\oint_C [P(x, y) dx + Q(x, y) dy] = \iint_S (Q_x - P_y) dx dy,$$

est une conséquence directe du théorème de Stokes.

Exercice 9.6 Montrer que la relation vectorielle

$$\vec{\nabla} \cdot (u \vec{\nabla} v) = u \Delta v + (\vec{\nabla} u) \cdot (\vec{\nabla} v),$$

associée au théorème de Gauss, conduit aux 2 formules de Green

$$\begin{aligned} \int_V [u \Delta v + (\vec{\nabla} u) \cdot (\vec{\nabla} v)] dV &= \oint_S u \vec{\nabla} v \cdot d\vec{S} \\ \int_V (u \Delta v - v \Delta u) dV &= \oint_S (u \vec{\nabla} v - v \vec{\nabla} u) \cdot d\vec{S} \end{aligned}$$

où u et v sont deux champs scalaires suffisamment réguliers.

9.5 Singularités

On aura noté l'importance des hypothèses de régularité exigées des champs de vecteurs pour appliquer les théorème de Stokes. Les lieux de l'espace où ces conditions ne sont pas vérifiées (ou a fortiori où les champs ne sont pas définis) ne sont pas moins importants. Ces points particuliers, ou *singularités* sont en général intrinsèquement associés aux caractéristiques physiques du problème étudié.

1. *Point*

Le champ électrique $\vec{E}(\vec{r})$ créé par une charge ponctuelle q située à l'origine des coordonnées, ainsi qu'il résulte de la loi expérimentale de Coulomb, est donné par l'expression

$$\vec{E}(\vec{r}) = \frac{q}{4\pi\epsilon_0} \frac{\vec{e}_r}{r^2}.$$

Il convient de noter sur cette expression que le champ électrique est défini en tout point de \mathbb{R}^3 excepté l'origine : les charges électriques sont des exemples de *singularités ponctuelles* pour le champ électrique.

On se propose de calculer le flux du champ \vec{E} à travers une sphère de rayon R . Considérons d'abord le cas où la sphère ne contient pas l'origine. Alors, \vec{E} et ses dérivées premières sont définies et continues dans tout le volume V de la sphère, de sorte que le théorème de Gauss s'écrit :

$$\oint_{S_R} \vec{E} \cdot d\vec{S} = \int_V (\vec{\nabla} \cdot \vec{E}) dV = \frac{q}{4\pi\epsilon_0} \iiint_V \left(\vec{\nabla} \cdot \frac{\vec{e}_r}{r^2} \right) dV$$

On se convaincra aisément (le faire!) que $\vec{\nabla} \cdot \frac{\vec{e}_r}{r^2} = 0$ (en fait $\vec{\nabla} \cdot \frac{\vec{e}_r}{r^n} = 0$ seulement pour $n = 2$). Le flux du champ électrique à travers une sphère qui ne contient pas l'origine est donc identiquement nul.

Considérons à présent le cas où la sphère contient l'origine. Dans ce cas, le champ n'est pas défini (ni a fortiori C^1) en tout point du volume V puisque le champ est singulier à l'origine. Il n'est donc pas possible d'appliquer le théorème de Gauss dans cette situation. Considérons cependant le volume compris entre les 2 sphères centrées sur l'origine de rayons $\epsilon < R$ et R . Puisque ce volume ne contient pas l'origine, il est possible d'appliquer le théorème de Gauss qui prend la forme :

$$\oint_{S_\epsilon} \vec{E} \cdot d\vec{S} + \oint_{S_R} \vec{E} \cdot d\vec{S} = 0$$

où S_ϵ et S_R désignent les surfaces des 2 sphères et où on a utilisé que la divergence du champ électrique est nulle dans le volume considéré. Le calcul du flux sur la sphère de rayon ϵ est immédiat :

$$\oint_{S_\epsilon} \vec{E} \cdot d\vec{S} = -\frac{q}{4\pi\epsilon_0} \int_0^\pi \frac{2\pi\epsilon^2 \sin\theta d\theta}{\epsilon^2} = -\frac{q}{4\pi\epsilon_0} (4\pi) = -\frac{q}{\epsilon_0},$$

où le signe provient du fait que le vecteur $d\vec{S}$ pointe vers l'extérieur du volume étudié, i.e. dans la direction opposée au vecteur unitaire \vec{e}_r .

On déduit donc de cet exemple que le champ électrique $\vec{E}(\vec{r})$ créé par une charge ponctuelle située à l'origine vérifie

$$\oint_{S_R} \vec{E} \cdot d\vec{S} = \int_V \vec{\nabla} \cdot \vec{E} dV = \begin{cases} +q/\epsilon_0 \\ 0 \end{cases},$$

selon que l'origine est contenue ou pas dans la sphère. Dans ce cas particulier, la dernière égalité est équivalente à l'équation de Maxwell $\epsilon_0 \vec{\nabla} \cdot \vec{E} = q \delta(\vec{r})$ qui est une des équations fondamentale de l'électromagnétisme (équation de Maxwell-Gauss).

2. *Ligne*

Le champ magnétique \vec{B} créé par un fil rectiligne infini placé selon Oz et parcouru par un courant d'intensité I est donné par l'expression :

$$\vec{B}(\vec{r}) = \frac{\mu_0 I}{2\pi} \frac{\vec{e}_\varphi}{r},$$

où r est la distance entre le point où est calculé le champ et le fil, et où $\vec{e}_\varphi = \vec{e}_z \times \vec{e}_r$. Cette expression est valable en tout point de l'espace, sauf sur le fil. Un courant filiforme est un exemple de *singularité linéique* pour le champ magnétique.

On se propose de calculer la circulation de \vec{B} le long d'un contour circulaire \mathcal{C} . Deux cas sont à considérer selon que \mathcal{C} entoure ou pas la singularité. Dans ce dernier cas, le plus simple est d'utiliser le théorème de Stokes, ce qui est licite puisque le champ est C_1 dans toute la surface S_R limitée par le contour. Or, on

vérifiera sans peine (en coordonnées cylindriques) que $\vec{\nabla} \times (\vec{e}_\varphi/r) = \vec{0}$. La circulation du champ magnétique sur un cercle n'entourant pas le courant est donc identiquement nulle.

Lorsque le contour entoure le courant, on peut procéder comme dans le cas des singularités ponctuelles, en appliquant le théorème de Stokes à la couronne de rayons ϵ et R . Plus simplement, on peut effectuer directement le calcul avec pour résultat :

$$\oint_C \vec{B} \cdot d\vec{r} = \frac{\mu_0 I}{2\pi\epsilon} \int_0^{2\pi} \epsilon d\varphi = \mu_0 I$$

Cette fois encore, le résultat dépend d'une façon cruciale, de la présence ou de l'absence de la singularité à l'intérieur du domaine considéré.

9.6 Différentielles exactes

En thermodynamique, on est souvent amené à s'interroger si la variation globale d'une grandeur physique au cours d'une transformation ne dépend que de ses valeurs initiale et finale, ou si la variation dépend de la façon dont la transformation a été opérée. Dans l'affirmative on parle de fonction d'état. Ainsi l'énergie interne par exemple est une fonction d'état tandis que le travail et la chaleur ne le sont pas. La notion mathématique correspondant à la notion physique de fonction d'état est celle de *différentielle exacte*.

On dit que la différentielle $\vec{A} \cdot d\vec{r}$ est une différentielle exacte si l'intégrale curviligne $\int \vec{A} \cdot d\vec{r}$ dépend uniquement des points de départ et d'arrivée, et non du chemin particulier suivi. Notons immédiatement que si la différentielle est exacte, alors on a aussi :

$$\oint \vec{A} \cdot d\vec{r} = 0,$$

puisque le point de départ et d'arrivée coïncident. Il est également clair qu'une condition suffisante pour que l'intégrale sur un chemin fermé soit nulle est que \vec{A} dérive d'un potentiel, i.e. qu'il existe une fonction φ telle que $\vec{A} \equiv \vec{\nabla} \varphi$, puisque dans ces conditions $\vec{A} \cdot d\vec{r} = \vec{\nabla} \varphi \cdot d\vec{r} = d\varphi$, d'où il résulte que

$$\oint \vec{A} \cdot d\vec{r} = \oint d\varphi = 0.$$

On peut ajouter que lorsque les dérivées partielles secondes de φ sont continues, si $A_x = \partial_x \varphi$, $A_y = \partial_y \varphi$, $A_z = \partial_z \varphi$, alors, l'égalité des dérivées croisées implique en particulier que $\partial_y A_x = \partial_{yx} \varphi = \partial_{xy} \varphi = \partial_x A_y$, soit encore $\partial_x A_y - \partial_y A_x = (\vec{\nabla} \times \vec{A})_z = 0$. On a donc montré les implications

$$\vec{A} \equiv \vec{\nabla} \varphi \quad \Rightarrow \quad \oint \vec{A} \cdot d\vec{r} = 0 \quad \Rightarrow \quad \vec{\nabla} \times \vec{A} = 0$$

Les réciproques résultent essentiellement du théorème de Stokes de sorte que l'on peut énoncer

Théorème 9.6.1 *Soit \vec{A} un champ vectoriel continu dont les dérivées partielles sont continues dans une région U simplement connexe, alors les résultats suivants sont équivalents*

1. $\vec{\nabla} \times \vec{A} = 0$ dans U (\vec{A} est dit irrotationnel).
2. $\oint_C \vec{A} \cdot d\vec{r} = 0$, avec $C \subset U$.
3. $\int_C \vec{A} \cdot d\vec{r}$ est indépendant de C avec $C \subset U$.
4. $\vec{A} \cdot d\vec{r}$ est une différentielle exacte.
5. Il existe une fonction φ telle que $\vec{A} = \vec{\nabla} \varphi$.

Rappelons que schématiquement une région simplement connexe est une région sans trou ; cette restriction est nécessaire pour éviter d'éventuelles singularités des champs.

Exercice 9.7 Pour tout $(x, y) \neq (0, 0)$ on considère le champ vectoriel

$$\vec{A} = -\frac{y}{x^2 + y^2} \vec{e}_x + \frac{x}{x^2 + y^2} \vec{e}_y$$

Montrer que ce champ est irrotationnel, mais que $\oint_C \vec{A} \cdot d\vec{r} \neq 0$ lorsque C est le cercle unité centré sur l'origine.

Pour quelle raison le théorème ne s'applique-t-il pas ?

Exercice 9.8 Montrer que la différentielle $(y - x) dx + (x + y) dy$ est exacte et déterminer le potentiel φ dont le champ associé dérive.

9.7 Théorèmes de Helmholtz

On a insisté plus haut sur l'importance des opérateurs différentiels, divergence et rotationnel, surtout en relation avec les notions de flux et de circulation. Par ailleurs, peut-être avez-vous été frappé par le fait que l'électromagnétisme à travers les relations de Maxwell met essentiellement en jeu les divergence et rotationnel des champs électriques et magnétiques. Les théorèmes de Helmholtz apportent une justification mathématique à cet état de fait, et établissent, sous des conditions que nous précisons plus bas, qu'un champ vectoriel est parfaitement déterminé dès lors que ses rotationnels et divergences sont eux-mêmes connus.

Théorème 9.7.1 Soit U une région simplement connexe de \mathbb{R}^3 . Tout vecteur \vec{V} de U , de classe \mathcal{C}^2 , est déterminé d'une façon unique par la connaissance de $\vec{\nabla} \cdot \vec{V}$, de $\vec{\nabla} \times \vec{V}$ et de la composante normale de \vec{V} sur le bord de U .

Pour établir l'unicité, considérons 2 vecteurs \vec{V}_1 et \vec{V}_2 satisfaisant les mêmes hypothèses et montrons que $\vec{W} \equiv \vec{V}_1 - \vec{V}_2 = \vec{0}$. Par hypothèse, $\vec{\nabla} \times \vec{V}_1 = \vec{\nabla} \times \vec{V}_2$, donc $\vec{\nabla} \times \vec{W} = \vec{0}$. Il existe donc une fonction Φ telle que $\vec{W} \equiv \vec{\nabla} \Phi$. On a également par hypothèse, $\vec{\nabla} \cdot \vec{V}_1 = \vec{\nabla} \cdot \vec{V}_2$, donc $\vec{\nabla} \cdot \vec{W} = 0$. Comme $\vec{\nabla} \cdot \vec{\nabla} \Phi = \Delta \Phi$, on a encore $\Delta \Phi = 0$. Enfin, l'égalité des composantes normales sur le bord de U conduit à la relation $W_n dS = \vec{W} \cdot d\vec{S} = \vec{\nabla} \Phi \cdot d\vec{S} = 0$.

L'utilisation de la formule de Green démontrée plus haut en exercice,

$$\int_V \left[\Phi \Delta \Phi + (\vec{\nabla} \Phi) \cdot (\vec{\nabla} \Phi) \right] dV = \oint_S \Phi \vec{\nabla} \Phi \cdot d\vec{S},$$

conduit donc au résultat :

$$\int_V (\vec{\nabla} \Phi)^2 dV = \int_V \vec{W}^2 dV = 0$$

Comme l'intégrand est ≥ 0 , on doit avoir $\vec{W} = \vec{0}$ en tout point, ce qui démontre le théorème.

Le deuxième théorème d'Helmholtz est également intéressant

Théorème 9.7.2 Soit $\vec{V}(\vec{r})$ un champ vectoriel de classe \mathcal{C}^2 tel que

$$\begin{aligned} \vec{\nabla} \cdot \vec{V}(\vec{r}) &= s(\vec{r}), \\ \vec{\nabla} \times \vec{V}(\vec{r}) &= \vec{c}(\vec{r}), \end{aligned}$$

où $s(\vec{r})$ et $\vec{c}(\vec{r})$ sont des fonctions connues qui s'annulent à l'infini.

Il existe un champ scalaire $\phi(\vec{r})$ et un champ vectoriel $\vec{A}(\vec{r})$ qui déterminent $\vec{V}(\vec{r})$ d'une façon unique par la relation

$$\vec{V} = -\vec{\nabla} \phi + \vec{\nabla} \times \vec{A},$$

où

$$\begin{aligned} \phi(\vec{r}) &= \frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{s(\vec{r}')}{\|\vec{r} - \vec{r}'\|} dV' \\ \vec{A}(\vec{r}) &= \frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{\vec{c}(\vec{r}')}{\|\vec{r} - \vec{r}'\|} dV' \end{aligned}$$

$s(\vec{r})$ et $\vec{c}(\vec{r})$ sont des densités de source et de circulation. Notons qu'on a $\vec{\nabla} \cdot \vec{V} = -\Delta\phi$ puisque la divergence d'un rotationnel est identiquement nulle. On obtient donc :

$$\vec{\nabla} \cdot \vec{V} = -\frac{1}{4\pi} \int_{\mathbb{R}^3} s(\vec{r}') \Delta \left(\frac{1}{\|\vec{r} - \vec{r}'\|} \right) dV' = s(\vec{r}),$$

où on a utilisé la relation $\Delta \left(\frac{1}{\|\vec{r} - \vec{r}'\|} \right) = -4\pi \delta(\vec{r} - \vec{r}')$.

De même, $\vec{\nabla} \times \vec{V} = \vec{\nabla} \left(\vec{\nabla} \cdot \vec{A} \right) - \Delta \vec{A}$ puisque le rotationnel d'un gradient est identiquement nul. On admettra que $\vec{\nabla} \left(\vec{\nabla} \cdot \vec{A} \right) = 0$ lorsque les champs décroissent suffisamment vite à l'infini. On a donc, $\vec{\nabla} \times \vec{V} = -\Delta \vec{A}$, et on termine la démonstration comme dans le cas de la divergence :

$$\vec{\nabla} \times \vec{V} = -\frac{1}{4\pi} \int_{\mathbb{R}^3} \vec{c}(\vec{r}') \Delta \left(\frac{1}{\|\vec{r} - \vec{r}'\|} \right) dV' = \vec{c}(\vec{r})$$

L'unicité est une conséquence du premier théorème d'Helmholtz (le bord du domaine est rejeté à l'infini, où \vec{V} s'annule puisque les sources sont nulles à l'infini par hypothèse).

Exercice 9.9 On rappelle que les équations de Maxwell pour les champs électrique \vec{E} et magnétique \vec{B} s'écrivent

$$\begin{aligned} \vec{\nabla} \cdot \vec{E}(\vec{r}) &= \frac{\rho(\vec{r})}{\epsilon_0}, \\ \vec{\nabla} \times \vec{E}(\vec{r}) &= -\vec{B}_t(\vec{r}), \end{aligned}$$

d'une part, et,

$$\begin{aligned} \vec{\nabla} \cdot \vec{B}(\vec{r}) &= 0, \\ \vec{\nabla} \times \vec{B}(\vec{r}) &= \mu_0 \left(\vec{j}(\vec{r}) + \epsilon_0 \vec{E}_t(\vec{r}) \right), \end{aligned}$$

d'autre part. Utilisez le deuxième théorème d'Helmholtz pour exprimer les champs \vec{E} et \vec{B} en fonction des densités de charges et des densités de courants. Interprétez les différentes contributions.

9.8 Théorème de Leibnitz

On considère les fonctions définies par une équation du type suivant :

$$y(t) \equiv \int_{a(t)}^{b(t)} f(x, t) dx,$$

où la dépendance dans la variable t apparaît dans les bornes et dans la fonction à intégrer. Le problème est de dériver cette fonction par rapport à t ce qui nous conduira au théorème de Leibnitz qui est à l'origine de l'un des théorèmes les plus importants de la mécanique des fluides : le théorème du transport de Reynolds.

Un premier résultat à rappeler est celui de la dérivée d'une intégrale par rapport à une de ces bornes. Plus précisément, on a

$$\frac{dh}{db} = f(b) \quad \text{si} \quad h(b) \equiv \int_a^b f(x) dx.$$

En effet, par définition de la dérivée

$$\frac{dh}{db} = \lim_{db \rightarrow 0} \frac{h(b+db) - h(b)}{db} = \lim_{db \rightarrow 0} \frac{1}{db} \int_b^{b+db} f(x) dx = \frac{f(b)db}{db} = f(b)$$

Notez que le résultat est indépendant de la borne inférieure a . On se convaincra sans peine (le faire!) que

$$\frac{dh}{da} = -f(a) \quad \text{si} \quad h(a) \equiv \int_a^b f(x)dx.$$

Avant d'arriver à l'énoncé le plus général, considérons le cas plus simple où la fonction intégrée ne dépend pas explicitement du temps :

$$g(t) \equiv h(a(t), b(t)) \quad \text{avec} \quad h(a, b) \equiv \int_a^b f(x)dx.$$

La fonction g ainsi définie est la composition de la fonction h et des fonctions a et b , on peut donc appliquer la chain rule :

$$\frac{dg}{dt} = h_a \frac{da}{dt} + h_b \frac{db}{dt} = -f(a) \frac{da}{dt} + f(b) \frac{db}{dt}$$

Le dernier résultat nécessaire est l'égalité

$$\frac{dS}{dt} = \int_a^b \partial_t f(x, t) dx \quad \text{si} \quad S(t) \equiv \int_a^b f(x, t) dx,$$

qui exige seulement que la fonction intégrée soit C^1 par rapport aux variables dans le cas des intégrales propres.⁴

Le théorème de Leibnitz résulte de la combinaison des différents cas que nous venons de détailler. Le résultat précis est le suivant

Théorème 9.8.1 Soient $f : (x, t) \mapsto f(x, t)$ une fonction de classe C^1 par rapport à (x, t) , $a : x \mapsto a(x)$ et $b : x \mapsto b(x)$ deux fonctions également de classe C^1 , alors

$$\frac{d}{dt} \int_{a(t)}^{b(t)} f(x, t) dx = \int_{a(t)}^{b(t)} \partial_t f(x, t) dx + f(b(t), t) \frac{db}{dt} - f(a(t), t) \frac{da}{dt}$$

Deux généralisations importantes de ce théorème concernent l'extension de ce résultat lorsqu'on intègre une fonction scalaire ou vectorielle sur un volume $V(t)$ dépendant de la variable t , limité par une surface (fermée) $S(t)$:

$$\begin{aligned} \frac{d}{dt} \int_{V(t)} f(\vec{r}, t) d\tau &= \int_{V(t)} \partial_t f(\vec{r}, t) d\tau + \oint_{S(t)} f(\vec{r}, t) \frac{d\vec{r}}{dt} \cdot d\vec{S}, \\ \frac{d}{dt} \int_{V(t)} \vec{F}(\vec{r}, t) d\tau &= \int_{V(t)} \partial_t \vec{F}(\vec{r}, t) d\tau + \oint_{S(t)} \vec{F}(\vec{r}, t) \left(\frac{d\vec{r}}{dt} \cdot d\vec{S} \right). \end{aligned}$$

On notera que dans l'intégrale de surface, le vecteur $d\vec{r}/dt$ estimé en des points de $S(t)$ correspond à la vitesse de déplacement du volume $V(t)$ (si on interprète t comme un temps). Si le volume est fixe, les bornes de l'intégrale ne dépendent plus de la variable t , $d\vec{r}/dt \equiv 0$ et l'intégrale de surface ne contribue pas à la dérivée.

4. dans le cas des intégrales généralisées, il faut en outre la convergence uniforme.

Chapitre 10

Approximer et calculer

Formule de Taylor
Approximation de Stirling
Distribution de Dirac
Intégrales gaussiennes
Développements asymptotiques

10.1 Formule de Taylor

Commençons par rappeler la forme du développement de Mac Laurin pour une fonction d'une seule variable :

$$f(x) = f(0) + x f'(0) + \frac{x^2}{2!} f''(0) + \dots + \frac{x^n}{n!} f^{(n)}(0) + R_n,$$

où le reste $R_n = x^{n+1} f^{(n+1)}(\xi)/(n+1)!$ avec $0 < \xi < x$.

Considérons maintenant une fonction de 3 variables $f(x, y, z)$ et définissons une nouvelle fonction F par la relation $F(t) = f(a + ht, b + kt, c + lt)$. Une application de la chain rule donne :

$$F'(t) = h f_x + k f_y + l f_z \equiv (h \partial_x + k \partial_y + l \partial_z) f(a + ht, b + kt, c + lt),$$

pour la dérivée première, et

$$F''(t) = h (h f_{xx} + k f_{xy} + l f_{xz}) + k (h f_{xy} + k f_{yy} + l f_{yz}) + h (h f_{xz} + k f_{yz} + l f_{zz}),$$

pour la dérivée seconde. Supposons que toutes les dérivées partielles soient continues, on peut alors écrire :

$$F''(t) = (h \partial_x + k \partial_y + l \partial_z)^2 f(a + ht, b + kt, c + lt),$$

et plus généralement

$$F^{(n)}(t) = (h \partial_x + k \partial_y + l \partial_z)^n f(a + ht, b + kt, c + lt).$$

Le développement de Taylor de f au voisinage du point $\vec{r} = (a, b, c)$ peut alors être obtenu à partir de la formule de Mac Laurin appliquée à F , pour $t = 1, h = dx, k = dy, l = dz$:

$$f(\vec{r} + d\vec{r}) = \sum_{k=0}^n \frac{1}{k!} \left(d\vec{r} \cdot \vec{\nabla} \right)^k f(\vec{r}) + R_n,$$

avec $R_n \equiv \left(d\vec{r} \cdot \vec{\nabla} \right)^{n+1} f(\vec{\xi})/(n+1)!$ et $\vec{\xi}$ quelque part sur le segment entre \vec{r} et $\vec{r} + d\vec{r}$.

Lorsque la fonction f est telle que $\lim_{n \rightarrow \infty} R_n = 0$, le développement de Taylor d'une fonction C^∞ devient une série que l'on peut écrire formellement

$$f(\vec{r} + d\vec{r}) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(d\vec{r} \cdot \vec{\nabla} \right)^k f(\vec{r}) \equiv e^{d\vec{r} \cdot \vec{\nabla}} f(\vec{r}).$$

Exercice 10.1 Montrer que les 2 premiers termes du développement en série de Taylor de la fonction $f : \vec{r} \in \mathbb{R}^2 \mapsto f(\vec{r}) \in \mathbb{R}$ au voisinage du point \vec{r}_0 peut s'écrire sous la forme

$$f(\vec{r}) = f(\vec{r}_0) + df(\vec{r}_0) + \frac{1}{2!} d^2 f(\vec{r}_0) + \dots,$$

où $d^2 f(\vec{r}_0) \equiv ((\vec{r} - \vec{r}_0), H(\vec{r}_0)(\vec{r} - \vec{r}_0))$ est la forme quadratique associée à la matrice (dite hessienne)

$$H(\vec{r}) \equiv \begin{pmatrix} f_{xx}(\vec{r}) & f_{xy}(\vec{r}) \\ f_{xy}(\vec{r}) & f_{yy}(\vec{r}) \end{pmatrix}$$

Exercice 10.2 Utiliser la formule de Taylor pour développer la fonction $(x, y) \mapsto \ln(x^2 + y^2)$ au voisinage de $(1, 1)$, jusqu'à l'ordre 2 en x et y .

10.2 Approximation de Stirling

L'approximation de Stirling est une formule approximative qui permet d'évaluer $x!$ ou $\ln x!$ pour x grand. Elle est particulièrement utilisée en physique statistique.

Pour établir cette formule, on part de la définition des factorielles à partir de la fonction Gamma d'Euler¹ :

$$\Gamma(n) \equiv \int_0^{\infty} x^{n-1} e^{-x} dx, \quad \text{où } n \in \mathbb{N}.$$

Il est évident que $\Gamma(1) = 1$ et on montre facilement par intégration par parties puis par récurrence sur n que :

$$\Gamma(n) = (n-1)\Gamma(n-1) = (n-1)(n-2) \cdots \Gamma(1) = (n-1)!$$

On en déduit donc que $n!$ peut être calculé à partir de l'intégrale

$$n! = \int_0^{\infty} e^{n \ln x - x} dx$$

Etudions l'intégrand $f(x) \equiv e^{n \ln x - x}$. Le calcul des 2 premières dérivées (le faire!) montre que cette fonction à la forme d'une courbe en cloche qui présente un maximum en $x = n$ et des points d'inflexion en $x/n = 1 \pm 1/\sqrt{(n)}$. Le maximum est donc d'autant plus piqué que n est grand.

Le développement de Taylor de la fonction $x \mapsto n \ln x - x$ au voisinage de $x = n$ s'écrit $n \ln x - x \approx n \ln n - n - (x-n)^2/2n$; si on l'introduit dans l'intégrale précédente, on est conduit au résultat

$$n! \approx e^{n \ln n - n} \int_0^{\infty} e^{-(x-n)^2/2n} dx = e^{n \ln n - n} \int_{-n}^{\infty} e^{-u^2/2n} du$$

On peut remplacer $-n$ par $-\infty$ dans la limite des $n \rightarrow \infty$ et utiliser le résultat sur l'intégrale gaussienne $\int_{\mathbb{R}} e^{-\alpha x^2} dx = (\pi/\alpha)^{1/2}$ pour $\alpha > 0$ qui conduit à l'approximation de Stirling :

$$n! \approx (2\pi n)^{1/2} n^n e^{-n} \Leftrightarrow \ln n! \approx n \ln n - n + \frac{1}{2} \ln(2\pi n), \quad n \gg 1.$$

Le tableau suivant (à vérifier avec votre calculatrice) montre que l'approximation de Stirling est bonne même pour de petites valeurs de n .

n	$(2\pi n)^{1/2} n^n e^{-n} / n!$	$[n \ln n - n] / \ln n!$	$[n \ln n - n + 0.5 \ln(2\pi n)] / \ln n!$
2	0.959502	-0.88539	0.940358
5	0.983493	0.63649	0.996523
20	0.995842	0.942815	0.999902
50	0.998335	0.980626	0.999989

1. la définition qui suit permet également de définir $x!$ pour x non forcément entier

10.3 Distribution de Dirac

Au cours des premiers développements de la mécanique quantique, le physicien anglais Dirac a introduit une "fonction", notée δ , définie par les propriétés suivantes :

$$\delta(x) = 0 \quad \text{pour } x \neq 0, \quad \text{et} \quad \int_{\mathbb{R}} \delta(x) dx = 1$$

Cette définition implique que δ soit très localisée autour de l'origine, ce qui suggère de l'utiliser pour décrire le phénomène de *percussion* - c'est à dire le fait d'une grandeur physique finie qui ne s'exprime de façon infiniment intense que pendant une durée infiniment courte. La décharge d'un condensateur contenant une charge finie, fournissant un courant très intense en un temps extrêmement court en fournit un exemple. Un autre exemple, emprunté à la mécanique, est donné par la densité volumique de masse correspondant à une masse ponctuelle de module m localisée à l'origine, qui peut s'écrire $m \delta(\vec{r})$.

La relation $\int_{\mathbb{R}} \delta(x) dx = 1$ montre cependant que δ ne peut absolument pas être une fonction usuelle puisqu'un théorème de la théorie de l'intégration dû à Lebesgue, affirme que l'intégrale d'une fonction dont le support² est réduit à un point est forcément nulle. Il est apparu cependant que cette pseudo-fonction lorsqu'elle était manipulée prudemment avec certaines règles empiriques définies par Dirac pouvait s'avérer extrêmement utile, voire incontournable dans certaines formulations de la théorie quantique.

Un cadre mathématique rigoureux d'étude de ces "fonctions", connues désormais sous le nom de fonctions généralisées ou de *distributions*, a été initié par le mathématicien français Laurent Schwartz, juste après la seconde guerre mondiale. Cette théorie s'est avérée extrêmement féconde, avec de nombreuses ramifications en Analyse, et incontournable en particulier pour l'étude des équations aux dérivées partielles. Bien que récente, il existe de nos jours de bonnes vulgarisations de cette théorie accessibles aux ingénieurs et physiciens (à commencer par certains ouvrages rédigés par Schwartz lui-même). Faute de temps, on ne trouvera pas dans ce qui suit un exposé selon ces lignes, mais on jouera le jeu pédagogiquement dangereux d'effectuer des calculs faux (!) - c'est-à-dire en appliquant les règles de calculs relatifs aux fonctions à des grandeurs qui n'en sont pas - pour justifier des résultats que l'on sait par ailleurs démontrer dans un cadre mathématique rigoureux...

1. Propriété fondamentale

La propriété fondamentale la plus importante est la relation

$$\int_{\mathbb{R}} \delta(x) f(x) dx = f(0),$$

valable pour toute fonction continue f . Cette propriété résulte intuitivement du fait que seul le voisinage immédiat de l'origine importe pour le calcul de l'intégrale. On aurait donc³ $\int_{\mathbb{R}} \delta(x) f(x) dx = f(0) \int_{\mathbb{R}} \delta(x) dx = f(0)$, par définition de δ .

La même propriété translatée de l'origine au point $x = a$ s'écrit

$$\int_{\mathbb{R}} \delta(x - a) f(x) dx = f(a).$$

Ainsi la valeur $f(a)$ d'une fonction au point $x = a$ est obtenue en multipliant f par $\delta(x - a)$ et en intégrant. Compte tenu de la définition de δ il n'est même pas nécessaire d'intégrer sur tout l'axe réel, mais seulement sur un voisinage du point où δ ne s'annule pas, ainsi

$$\int_A \delta(x - a) f(x) dx = f(a), \quad \text{mais} \quad \int_B \delta(x - a) f(x) dx = 0,$$

où A est n'importe quel intervalle contenant $\{a\}$, et B n'importe quel intervalle ne le contenant pas.

2. le support d'une fonction est le lieu des points où la fonction ne s'annule pas.
 3. Insistons encore une fois, que d'après le théorème de Lebesgue cité plus haut, cette intégrale ne peut qu'être identiquement nulle si δ était une fonction.

2. *Dérivée des fonctions discontinues (!)*

Il existe un lien profond entre δ et la dérivée des fonctions discontinues. Considérons par exemple la fonction de Heaviside H définie par

$$H(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases}$$

Il est clair que cette fonction est dérivable, de valeur nulle, en tout point de \mathbb{R} , sauf en $x = 0$ où la dérivée est infinie. Le comportement de H' est donc comparable à celui de δ . Ce point peut être précisé, en calculant l'intégrale suivante

$$\int_{-a}^{+a} H'(x)f(x) dx,$$

où a est n'importe quel nombre fini positif. Une intégration par partie abusive⁴ conduit au résultat :

$$\int_{-a}^{+a} H'(x)f(x) dx = [H(x)f(x)]_{x=-a}^{x=+a} - \int_{-a}^{+a} H(x)f'(x) dx = f(0)$$

En utilisant la propriété fondamentale, on écrirait donc

$$H' = \delta$$

Cette égalité doit être comprise au sens des distributions, c'est-à-dire que pour toute fonction continue f , lorsque H' est considérée comme une distribution, on a l'égalité

$$\int_{-a}^{+a} H'(x)f(x) dx = \int_{-a}^{+a} \delta(x)f(x) dx.$$

En fait - et c'est l'un des grands avantages de cette théorie - dans l'espace des distributions, *toutes* les distributions sont dérivables!

3. *Autres propriétés*

Donnons quelques autres propriétés, commodes dans les calculs, et qui découlent de la propriété fondamentale.

$$\begin{aligned} \delta(-x) &= \delta(x), \\ x \delta(x) &= 0, \\ \delta(ax) &= \frac{1}{|a|} \delta(x), \quad a \in \mathbb{R}, \\ f(x) \delta(x-a) &= f(a) \delta(x-a) \end{aligned}$$

A titre d'exemple, démontrons (si l'on peut dire!) l'étrange égalité $x \delta(x) = 0$. Comme on a

$$\int_{\mathbb{R}} (x\delta(x)) f(x) dx = \int_{\mathbb{R}} \delta(x) (xf(x)) dx = (xf(x))_{x=0} = 0,$$

pour n'importe quelle fonction f continue, l'égalité est "démontrée". Incidemment cette relation ouvre des horizons nouveaux. Alors que dans l'espace des fonctions, la seule fonction qui satisfait $xf(x) = 0$ pour tout x est la fonction identiquement nulle $f \equiv 0$, dans l'espace des distributions, la solution de cette équation est la distribution $f = c\delta$ où c est n'importe quel nombre réel!

Nous définirons dans les chapitres suivants la transformée de Fourier au sens des fonctions, et nous verrons que la définition même de la transformée de Fourier implique que les fonctions soient intégrables (ou de carré intégrables). Une constante ne satisfaisant pas cette condition, il n'est pas possible de définir la transformée de Fourier de la fonction $f \equiv 1$ par exemple. Cela est cependant possible au sens des distributions, et

4. l'application de l'intégration par parties impose que les fonctions soient dérivables dans l'intervalle d'intégration. Cette intégrale est en fait identiquement nulle au sens des fonctions.

nous admettrons que la transformée de Fourier de 1 est précisément δ , ce qui revient encore à dire que la représentation intégrale de δ est donnée par la notation abusive⁵

$$\delta(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{iqx} dq.$$

Nous terminerons par une importante propriété qui nous est suggérée par le calcul, effectué plus haut, du flux du champ électrique créé par une charge ponctuelle placée à l'origine, à travers une sphère de rayon R . Nous avons alors démontré :

$$\int_V \left(\vec{\nabla} \cdot \frac{\vec{e}_r}{r^2} \right) dV = \begin{cases} +4\pi \\ 0 \end{cases},$$

selon que l'origine est contenue ou pas dans la sphère. Comme $\vec{e}_r/r^2 = -\vec{\nabla}(1/r)$, et $\vec{\nabla} \cdot \vec{\nabla} = \Delta$, on a aussi

$$\int_V \Delta \left(\frac{1}{r} \right) dV = \begin{cases} -4\pi \\ 0 \end{cases}$$

Cette dernière égalité peut s'écrire entre distributions sous la forme :

$$\Delta \left(\frac{1}{r} \right) = -4\pi \delta(\vec{r}),$$

où l'on reconnaît, en passant, la relation de Poisson de l'électrostatique pour une charge ponctuelle unité.

Insistons une dernière fois en rappelant que toutes les relations intégrales ci-dessus qui contiennent la distribution δ ne doivent être comprises que comme des notations commodes, mais qu'il est possible de donner un sens mathématique rigoureux à toutes ces expressions.

Exercice 10.3 Justifiez les propriétés suivantes de la distribution δ :

$$\begin{aligned} \delta(-x) &= \delta(x), \\ \delta(ax) &= \frac{1}{|a|} \delta(x), \quad a \in \mathbb{R}, \\ f(x) \delta(x-a) &= f(a) \delta(x-a) \end{aligned}$$

10.4 Intégrales Gaussiennes

Partons de l'intégrale de Gauss :

$$\int_{\mathbb{R}} e^{-\pi x^2} dx = 1,$$

qu'il est bon de mémoriser. Pour retrouver ce résultat fondamental, l'astuce consiste à calculer le carré de cette intégrale et à passer en coordonnées polaires :

$$\left(\int_{\mathbb{R}} e^{-\pi x^2} dx \right)^2 = \int_{\mathbb{R}^2} e^{-\pi(x^2+y^2)} dx dy = \int_0^{\infty} e^{-\pi r^2} (2\pi r dr) = 1.$$

Pour les applications qui nous intéressent en physique statistique et probabilités, la fonction à intégrer est plutôt $x \mapsto e^{-a x^2/2}$ où a est un réel positif. Un simple changement de variables montre (faites-le) que :

$$\int_{\mathbb{R}} e^{-\frac{1}{2} a x^2} dx = \sqrt{\frac{2\pi}{a}}.$$

5. encore une fois, cette intégrale n'existe pas au sens des fonctions !

Cette formule admet une généralisation en dimension plus élevée sous la forme :

$$\int_{\mathbb{R}^n} e^{-\frac{1}{2}(\vec{x}, A\vec{x})} d\vec{x} = \sqrt{\frac{(2\pi)^n}{\det A}},$$

où \vec{x} est maintenant un vecteur de \mathbb{R}^n et A une matrice $n \times n$ réelle et symétrique. Dans cette formule, $(\vec{x}, \vec{y}) \equiv \sum_{i=1}^n x_i y_i$, représente le produit scalaire des 2 vecteurs \vec{x} et \vec{y} dans \mathbb{R}^n .

La justification de ce résultat est la suivante. Comme A est diagonalisable, il existe une transformation orthogonale⁶ P telle que $P^{-1}AP = D$ où D est une matrice diagonale dont les éléments sont les valeurs propres λ_i . On en déduit que

$$\frac{1}{2}(\vec{x}, A\vec{x}) = \frac{1}{2}(\vec{x}, PDP^{-1}\vec{x}) = \frac{1}{2}(P^T\vec{x}, DP^T\vec{x}) = \frac{1}{2}(\vec{y}, D\vec{y}),$$

où on a posé $\vec{y} = P^T\vec{x}$. Puisque le jacobien de la transformation $|\det P| = 1$, on obtient donc :

$$\int_{\mathbb{R}^n} e^{-\frac{1}{2}(\vec{x}, A\vec{x})} d\vec{x} = \int_{\mathbb{R}^n} e^{-\frac{1}{2}(\vec{y}, D\vec{y})} d\vec{y} = \prod_{i=1}^n \int_{\mathbb{R}} e^{-\frac{1}{2}\lambda_i y_i^2} dy_i = \prod_{i=1}^n \left(\frac{2\pi}{\lambda_i}\right)^{1/2} = \frac{(2\pi)^{n/2}}{\sqrt{\det A}},$$

où on a utilisé le fait que $\det A = \prod_i \lambda_i$.

Les résultats précédents peuvent être utilisés pour définir des densités de probabilité distribuées selon la loi de Gauss. A une dimension, la distribution de probabilité normalisée $P(x)$ est définie par

$$P(x) \equiv \frac{e^{-\frac{1}{2}ax^2}}{\int_{\mathbb{R}} e^{-\frac{1}{2}ax^2} dx} = \sqrt{\frac{a}{2\pi}} e^{-\frac{1}{2}ax^2}.$$

De même, la densité de probabilité normalisée $P(\vec{x})$ dans \mathbb{R}^n s'écrirait :

$$P(\vec{x}) = \frac{e^{-\frac{1}{2}(\vec{x}, A\vec{x})}}{\int_{\mathbb{R}^n} e^{-\frac{1}{2}(\vec{x}, A\vec{x})} d\vec{x}} = \sqrt{\frac{\det A}{(2\pi)^n}} e^{-\frac{1}{2}(\vec{x}, A\vec{x})}.$$

Signalons pour finir deux intégrales souvent rencontrées dans les calculs. Par dérivations successives par rapport à a , on obtient (à vérifier!)

$$\begin{aligned} \int_{\mathbb{R}} x^{2n} e^{-\frac{1}{2}ax^2} dx &= \frac{(2n-1)!!}{a^n} \sqrt{\frac{2\pi}{a}}, & \forall n \geq 1, \\ \int_{\mathbb{R}^+} x^{2n+1} e^{-\frac{1}{2}ax^2} dx &= \frac{2^n n!}{a^{n+1}}, & \forall n \geq 0, \end{aligned}$$

où $(2n-1)!! \equiv 1.3.5 \cdots (2n-1)$. Pour des raisons évidentes de parité on a bien sûr

$$\int_{\mathbb{R}} x^{2n+1} e^{-\frac{1}{2}ax^2} dx = 0, \quad \forall n \geq 0.$$

6. Une transformation orthogonale est une transformation telle que $P^{-1} = P^T$. Comme $\det P = \det P^T$, on en déduit que $(\det P)^2 = 1$.

Chapitre 11

Minimiser

Extrémalisation sans contraintes
Multiplicateurs de Lagrange
Transformation de Legendre
Calcul des Variations

11.1 Extrémalisation sans contraintes

Rappelons que pour les fonctions d'une seule variable, la fonction $f : x \mapsto f(x)$ admet un extremum au point $x = a$ si $f'(a) = 0$ et $f''(a) \neq 0$, et plus précisément, un maximum si $f''(a) < 0$ et un minimum si $f''(a) > 0$. Si $f''(a) = 0$, il faut étudier les dérivées d'ordre plus élevé.

On se pose la même question dans le cas des fonctions de plusieurs variables, et on se limitera à énoncer le résultat dans le cas de 2 variables. On dira que la fonction $f : (x, y) \mapsto f(x, y)$ admet un extremum local (maximum ou minimum) au voisinage d'un point particulier (a, b) , selon que la fonction est au-dessous ou au-dessus de son plan tangent dans un voisinage du point (a, b) . Ainsi, par exemple, f admet un maximum en (a, b) si $f(a + h, b + k) < f(a, b)$, pour tous les h et k tels que $|h| < \delta$ et $|k| < \delta$ avec δ suffisamment petit. En particulier, si on fait $k = 0$, la fonction $x \mapsto f(x, b)$ est une fonction d'une seule variable, qui sera donc maximale en $x = a$ si $\partial_x f = 0$ en (a, b) . Le même argument appliqué à la fonction $y \mapsto f(a, y)$, montre que l'on doit également avoir $\partial_y f = 0$ en (a, b) . D'une façon plus générale, les points où les égalités $\partial_x f = \partial_y f = 0$ sont vérifiées sont appelés *les points critiques* de la fonction f ¹.

On peut donc conclure de ce qui précède qu'une condition *nécessaire* pour qu'une fonction admette un extremum en un point est que ce point soit un point critique. Cette condition n'est cependant pas *suffisante* comme le montre le simple exemple de la fonction $f(x, y) = x^2 - y^2$. Le point $(0, 0)$ est clairement un point critique, mais tel que la fonction $x \mapsto f(x, 0) = x^2$ est minimale en $(0, 0)$, tandis que la fonction $y \mapsto f(0, y) = -y^2$ est maximale au même point. Ce type de point critique est appelé *point selle* pour des raisons évidentes. On perçoit donc qu'un point critique ne sera un extremum que s'il existe des contraintes sur les courbures dans les directions x et y . Ce critère doit donc mettre en jeu les dérivées secondes partielles ; en effectuant un développement de Taylor au voisinage d'un point (a, b) , on établit le théorème suivant :

Théorème 11.1.1 *Soit $(x, y) \mapsto f(x, y)$ une fonction continue telle que toutes ses dérivées secondes partielles soient elles-mêmes continues dans un voisinage du point (a, b) . On définit le déterminant $D(a, b)$ (dit déterminant Hessien de f) par la relation :*

$$D(a, b) \equiv \begin{vmatrix} f_{xx}(a, b) & f_{xy}(a, b) \\ f_{xy}(a, b) & f_{yy}(a, b) \end{vmatrix}$$

Alors, si (a, b) est un point critique, i.e.

$$f_x = f_y = 0, \quad \text{en } x = a, y = b$$

1. (a, b) est un maximum local de f si $f_{xx}(a, b) < 0$ et $D(a, b) > 0$

1. la même terminologie est utilisée en dimension quelconque.

2. (a, b) est un minimum local de f si $f_{xx}(a, b) > 0$ et $D(a, b) > 0$
3. (a, b) est un point selle de f si $D(a, b) < 0$
4. les dérivées d'ordre plus élevées doivent être étudiées si $D(a, b) = 0$.

Exercice 11.1 Déterminer la nature des points critiques des fonctions quadratiques du type :

$$f(x, y) = ax^2 + bxy + cy^2,$$

en fonction du signe de a et du déterminant hessien associé.

Pouvez-vous donner des exemples de fonctions quadratiques correspondant aux surfaces représentées sur la figure 11.1.



FIGURE 11.1 – Fonctions quadratiques du type $f(x, y) = ax^2 + bxy + cy^2$

Exercice 11.2 Montrer que la fonction $f : (x, y) \in \mathbb{R}^2 \mapsto f(x, y) = x^3 + y^3 - x - 6y + 10$ admet 4 points critiques dont on déterminera la nature.

Les résultats sont-ils confirmés par la figure 11.2 ?

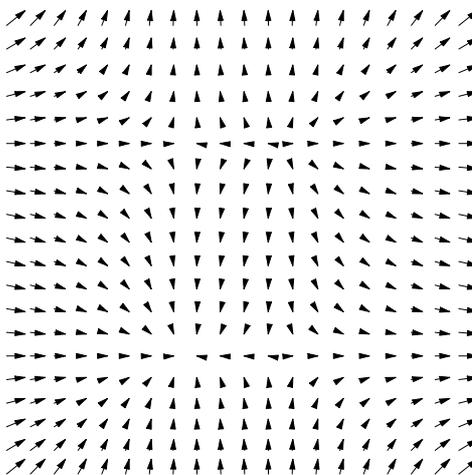


FIGURE 11.2 – Champs de gradients de la fonction $f(x, y) = x^3 + y^3 - x - 6y + 10$

11.2 Multiplicateurs de Lagrange

Considérons une fonction f de n variables indépendantes : $f(x_1, \dots, x_n)$. Sa différentielle au point $x \equiv (x_1, \dots, x_n)$ s'écrit

$$df = \frac{\partial f}{\partial x_1} dx_1 + \dots + \frac{\partial f}{\partial x_n} dx_n. \quad (11.1)$$

Une condition nécessaire pour que la fonction f possède un extremum (maximum ou minimum) au point x est que $df = 0$, soit :

$$\frac{\partial f}{\partial x_1} = \dots = \frac{\partial f}{\partial x_n} = 0. \quad (11.2)$$

Supposons maintenant qu'il existe une contrainte entre les variables x_1, \dots, x_n ; autrement dit, il existe une certaine relation fonctionnelle

$$g(x_1, \dots, x_n) = 0$$

entre ces variables. Cette relation montre qu'une des variables, disons x_n dépend des $n - 1$ autres. Il n'y a donc plus n variables indépendantes, mais seulement $n - 1$. Les conditions d'extrémalisation données par (11.2) ne sont donc pas valables en présence d'une contrainte.

Une première solution évidente consiste, lorsque cela est possible, à exprimer une des variables en fonctions des $n - 1$ autres, à substituer son expression dans f , puis à écrire les conditions d'extrémalisation (11.2) sur les $n - 1$ variables restantes.

Exemple. Soit à déterminer le rectangle d'aire maximale pour un périmètre donné. Avec des notations évidentes, on a $L(x, y) = 2(x + y)$ et $S(x, y) = xy$. En éliminant y entre ces 2 équations, on obtient $S(x) = x(L/2 - x)$, x étant maintenant une variable non contrainte. La condition $S_x = 0$ conduit à la relation $L/2 - 2x = 0$, et donc à la solution $x = y = L/4$, c'est-à-dire à un carré.

Un autre procédé plus général a été proposé par Lagrange. L'idée consiste à absorber la contrainte dans la définition d'une nouvelle fonction. Soit en effet la fonction $f \pm \lambda g$, où λ est un paramètre appelé *multiplicateur de Lagrange*. Il s'agit d'une fonction des n variables indépendantes (x_1, \dots, x_n) puisqu'aucune contrainte extérieure ne s'applique désormais sur les variables. On peut donc écrire directement :

$$d(f \pm \lambda g) = 0 \quad \Rightarrow \quad \frac{\partial f}{\partial x_i} \pm \lambda \frac{\partial g}{\partial x_i} = 0, \quad i = 1, \dots, n$$

Ces n équations et la contrainte $g = 0$ déterminent donc les $n + 1$ inconnues x_1, \dots, x_n, λ .

Exemple. Reprenons l'exemple précédent. On a $f(x, y) = xy$ et $g(x, y) = L - 2(x + y) = 0$. L'extremum de f est déterminé par les 2 relations $f_x + \lambda g_x = 0$ et $f_y + \lambda g_y = 0$, soit $y - 2\lambda = x - 2\lambda = 0$, qui conduit bien à la solution $x = y$.

Théorème 11.2.1 Soient f et g 2 fonctions \mathcal{C}^1 définies sur \mathbb{R}^n . On suppose en outre $\partial g / \partial x_i \neq 0$, pour $i = 1, \dots, n$.

Un extremum (maximum ou minimum) de la fonction f soumise à la contrainte $g(x_1, \dots, x_n) = 0$ est déterminé (s'il existe) par les $(n + 1)$ équations :

$$\begin{aligned} \frac{\partial f}{\partial x_i} - \lambda \frac{\partial g}{\partial x_i} &= 0, \quad i = 1, \dots, n \\ g(x_1, \dots, x_n) &= 0, \end{aligned}$$

où λ est un paramètre réel appelé *multiplicateur de Lagrange*.

Exercice 11.3 Représenter sur un schéma le plan décrit par l'équation $x + y + z = 1$, et utiliser la méthode de Lagrange pour déterminer la plus courte distance entre ce plan et l'origine.

En introduisant autant de multiplicateurs qu'il y a de contraintes, on généralise ce procédé au cas des fonctions soumises à plusieurs contraintes.

Exercice 11.4 On considère un système isolé de N particules identiques dont chaque particule est soit dans l'état d'énergie E_1 , soit dans l'état d'énergie E_2 (avec $E_1 \neq E_2$). Le nombre de particules N_1 et N_2 dans chaque état d'énergie est obtenu par extrémalisation de l'entropie $S(N_1, N_2)$ du système,

$$S(N_1, N_2) = N \ln N - N - (N_1 \ln N_1 - N_1) - (N_2 \ln N_2 - N_2),$$

soumises aux 2 contraintes :

$$\begin{aligned} N_1 + N_2 &= N, \\ N_1 E_1 + N_2 E_2 &= E, \end{aligned}$$

où N et E sont des constantes.

Utilisez la méthode des multiplicateurs de Lagrange pour déterminer N_1 et N_2 .

11.3 Transformation de Legendre

Il est bien connu qu'une courbe dans le plan $y = f(x)$ est parfaitement définie dès lors que sa fonction dérivée f' est connue en tout point x . On rencontre un certain nombre de situations en physique ou en mathématique où la donnée naturelle du problème étudié est la dérivée $f'(x)$ plutôt que la variable x elle-même. La transformation de Legendre est un moyen systématique qui permet de définir, sans perte d'informations, une nouvelle fonction dont la variable indépendante sera la pente de la fonction f au point x , c'est-à-dire $f'(x)$.

Définition 11.3.1 Soit $f : \mathbb{R} \mapsto \mathbb{R}$, une fonction convexe (i.e. $f''(x) > 0$).

La transformée de Legendre de f est la fonction g définie par la relation :

$$g(p) \equiv \max_x (px - f(x)) \quad \Leftrightarrow \quad g(f'(x)) \equiv xf'(x) - f(x)$$

Le contenu géométrique de la première définition est clair. $g(p)$ est la distance maximale entre la droite de pente p qui passe par l'origine et la fonction f . La solution est obtenue pour x_p tel que $p = f'(x_p)$, c'est-à-dire que la distance entre la droite et la fonction est maximale au point x_p où la tangente à f est égale à la pente de la droite (cf. figure). On a donc aussi $g(p) \equiv px_p - f(x_p)$, soit encore $g(f'(x_p)) = f'(x_p)x_p - f(x_p)$ qui correspond à la deuxième définition. Les variables x et $f'(x) \equiv p$ sont dites *variables conjuguées* par rapport au couple de fonctions f et g . Si f est concave, on définit la transformée de Legendre par la relation : $g(f'(x)) \equiv f(x) - xf'(x)$.

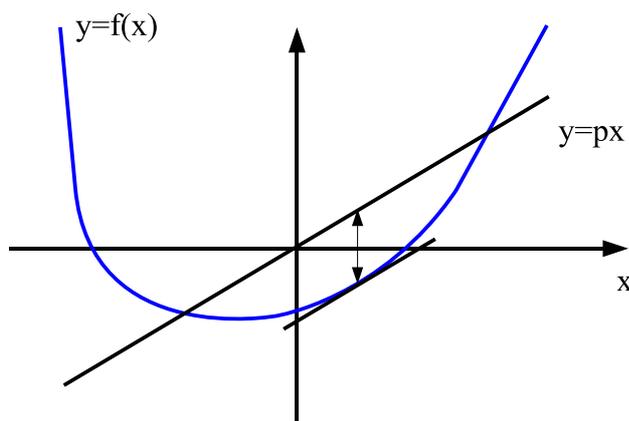


FIGURE 11.3 – Transformée de Legendre.

Exemple. Montrons que la transformée de Legendre de la fonction $f(x) = mx^2/2$ est la fonction $g(p) = p^2/2m$. En effet, comme $p \equiv f'(x) = mx$, on a par définition $g(f'(x)) \equiv x(mx) - mx^2/2 = mx^2/2$, donc $g(p) = p^2/2m$. Cet exemple simple montre que la transformée de Legendre peut être utilisée pour obtenir la mécanique hamiltonienne à partir du formalisme lagrangien.

Voici quelques exemples supplémentaires de transformées de Legendre que l'on établira à titre d'exercices :

$f(x)$	$g(p)$
x^n	$(n-1) \left(\frac{p}{n}\right)^{\frac{n}{n-1}}$
e^x	$p \ln p - p$
$\ln x$	$\ln ep$

On vérifiera sur ces exemples que les dérivées de f et g sont des fonctions inverses l'une de l'autre, ce qui, du reste, est une autre définition de la transformée de Legendre d'une fonction.

Le passage de $f(x)$ à $g(f'(x))$ peut d'ailleurs être retrouvé directement à partir de la définition. Soit en effet une fonction g définie à partir de x , f et f' :

$$g(x, f(x), f'(x)) \equiv f(x) - f'(x)x.$$

En utilisant le fait que $df(x) = f'(x) dx$ (f est différentiable en x), on voit que la différentielle totale de la fonction g au point x est telle que

$$dg = df(x) - f'(x) dx - x df'(x) = -x df'(x)$$

Il en résulte que $\partial_x g = \partial_{f(x)} g = 0$ de sorte que g est bien une fonction différentiable de la seule variable indépendante $f'(x)$.

Cette transformation se généralise aux cas des fonctions de plusieurs variables où il devient possible de faire une transformation de Legendre associée à une partie (ou à la totalité) des variables. Considérons ainsi une fonction différentiable f des n variables x_1, \dots, x_n et supposons que l'on cherche à définir une nouvelle fonction g qui dépende des variables $x_1, \dots, x_r, u_{r+1}, \dots, u_n$ où les $u_i \equiv \partial f / \partial x_i$ ($i = r+1, \dots, n$) sont les variables conjuguées des variables initiales x_{r+1}, \dots, x_n . En généralisant la procédure à une dimension, il suffit de définir :

$$g = f - (u_{r+1}x_{r+1} + \dots + u_n x_n),$$

pour constater que :

$$dg = u_1 dx_1 + \dots + u_r dx_r - x_{r+1} du_{r+1} - \dots - x_n du_n,$$

où on a utilisé le fait que $df = \sum_{i=1}^n u_i dx_i$. g est donc une fonction différentiable des n nouvelles variables $x_1, \dots, x_r, u_{r+1}, \dots, u_n$.

On notera en outre une belle propriété de symétrie de la transformation de Legendre. L'écriture même de dg montre qu'on a

$$u_i = \frac{\partial f}{\partial x_i} \quad \text{et} \quad x_i = -\frac{\partial g}{\partial u_i} \quad \text{pour} \quad i = r+1, \dots, n$$

Exercice 11.5 *Le premier principe de la thermodynamique s'écrit dans les cas les plus simples sous la forme :*

$$dU = T dS - p dV + \mu dN,$$

où l'énergie interne U est une fonction différentiable des variables S, V et N . On notera que les couples (T, S) , $(-p, V)$ et (μ, N) sont des variables conjuguées par rapport à U .

Déterminer par transformation de Legendre, une fonction différentiable des nouvelles variables N, V, T . Effectuer une seconde transformation de Legendre qui conduise à une fonction différentiable des variables μ, V, T . Interpréter les résultats obtenus.

Exercice 11.6 *Considérons un système mécanique à 2 degrés de liberté. Les équations du mouvement peuvent être obtenues à partir du lagrangien*

$$L(q_1, \dot{q}_1, q_2, \dot{q}_2, t)$$

où les q_i et \dot{q}_i représentent respectivement les coordonnées et vitesses généralisées.

Montrer qu'il est possible de définir par transformation de Legendre des fonctions $H(q_1, p_1, q_2, p_2, t)$ et $R(q_1, p_1, q_2, \dot{q}_2, t)$ où les p_i sont les impulsions généralisées, grandeurs conjuguées des \dot{q}_i par rapport au lagrangien L .

Reconnaissez-vous les fonctions H et R ?

Quatrième partie

Algèbre linéaire

Chapitre 12

Eléments d'algèbre linéaire

Déterminants

Matrices

Système d'équations linéaires

Diagonalisation

12.1 Déterminants

Cette section comprend quelques rappels sur le calcul et la manipulation des déterminants.

Définition 12.1.1 *Considérons un tableau carré, A , de nombres réels ou complexes arrangés en n lignes et n colonnes.*

Le déterminant correspondant, d'ordre n , noté $\det A$ ou $|A|$, est tel que :

$$\det A \equiv \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{\sigma \in G_n} \epsilon_\sigma a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

où G_n est le groupe des permutations de $(1, 2, \dots, n)$ (il y en a $n!$) et où ϵ_σ est la signature de la permutation considérée qui vaut 1 si on effectue un nombre pair d'inversions pour retrouver l'ordre naturel $(1, 2, \dots, n)$, et -1 dans le cas contraire.

Cette écriture un peu formelle montre qu'un déterminant est un cas particulier de forme multilinéaire alternée. Dans le cas des déterminants d'ordre 2, il n'y a que 2 permutations de l'ensemble $(1, 2)$: soit $(1, 2)$ avec $\epsilon = +1$, soit $(2, 1)$ avec $\epsilon = -1$. On obtient donc aussitôt dans ce cas :

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = +a_{11}a_{22} - a_{12}a_{21}$$

De même, il est facile de se convaincre que pour les déterminants d'ordre 3, il existe $3! = 6$ permutations qui conduisent au résultat :

$$\begin{aligned} & \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= +a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \\ &= +a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \end{aligned}$$

La dernière factorisation montre qu'un déterminant d'ordre 3 peut-être développé le long d'une ligne en faisant apparaître des déterminants d'ordre 2. Cette procédure est générale. On peut développer un déterminant d'ordre n selon une ligne ou une colonne en pondérant chaque déterminant d'ordre $(n - 1)$ obtenu en supprimant la ligne et la colonne correspondante. La procédure est la suivante :

Définition 12.1.2 Le déterminant mineur, noté M_{ij} , d'un déterminant d'ordre n est le déterminant d'ordre $n - 1$ obtenu en supprimant la i ème ligne et la j ème colonne.

Le cofacteur, noté A_{ij} , est tel que :

$$A_{ij} = (-1)^{i+j} M_{ij}$$

Le développement du déterminant de A selon une ligne ou une colonne s'écrit :

$$\det A = \sum_{j=1}^n a_{ij} A_{ij} = \sum_{i=1}^n a_{ij} A_{ij}$$

Exemple

$$\begin{vmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & -2 & 1 \end{vmatrix} = -0 \begin{vmatrix} -1 & 1 \\ -2 & 1 \end{vmatrix} + 3 \begin{vmatrix} 2 & 1 \\ 2 & 1 \end{vmatrix} - (-1) \begin{vmatrix} 2 & -1 \\ 2 & -2 \end{vmatrix} = 3(0) + (-2) = -2$$

A partir de ces définitions, on peut établir les propriétés suivantes utiles pour les calculs :

Théorème 12.1.1 – si une ligne (ou une colonne) ne comprend que des éléments nuls, alors le déterminant est nul.

- si chaque élément d'une ligne (ou d'une colonne) est multiplié par une constante, alors le déterminant est multiplié par cette constante.
- si l'on remplace une ligne (ou une colonne) par une combinaison linéaire de lignes (ou de colonnes), alors le déterminant n'est pas modifié.

Une conséquence importante de ces propriétés est la suivante : si 2 lignes (ou 2 colonnes) sont proportionnelles, alors le déterminant est nul. Si l'on identifie les lignes (ou les colonnes) avec des vecteurs, cela revient à dire que le déterminant d'une famille de vecteurs est non nul si les vecteurs de cette famille sont indépendants.

12.2 Matrices

Cette section comprend quelques rappels sur les opérations applicables aux matrices ainsi que la définition et le calcul de l'inverse d'une matrice.

12.2.1 Opérations

Commençons par rappeler quelques définitions relatives aux opérations sur les matrices. L'addition de 2 matrices s'effectue terme à terme :

$$C = A + B \Leftrightarrow c_{ij} = a_{ij} + b_{ij}, \quad \forall i, j$$

(la matrice nulle, dont tous les éléments sont nuls, joue le rôle d'élément neutre pour l'addition).

La multiplication par un scalaire s'obtient en multipliant chaque élément par le scalaire :

$$C = \lambda A, \quad \lambda \in \mathbb{C} \Leftrightarrow c_{ij} = \lambda a_{ij}, \quad \forall i, j$$

L'opération la plus importante est bien sûr la multiplication entre matrices définie par les relations :

$$C = AB \Leftrightarrow c_{ij} = \sum_k a_{ik} b_{kj}, \quad \forall i, j$$

Le premier indice étant (conventionnellement) l'indice de ligne et le second, l'indice de colonne, on constate que la multiplication entre matrices est définie en multipliant terme à terme les éléments d'une ligne de A par les éléments de la colonne de B . Il est donc possible d'effectuer des produits de matrices de taille différentes pour peu que le nombre de colonnes de la 1ère coïncide avec le nombre de lignes de la seconde.

Prendre garde que le produit matriciel n'est pas commutatif¹ en général,

$$AB \neq BA \quad \text{bien que} \quad \det(AB) = \det(BA) \quad \text{et} \quad \text{tr}(AB) = \text{tr}(BA)$$

où tr désigne la trace, c'est-à-dire la somme des éléments diagonaux de la matrice.

L'élément unité pour la multiplication des matrices est la matrice identité, noté I , qui ne comprend que des 1 sur la diagonale et des 0 partout ailleurs (ses éléments sont représentés par le symbole de Kronecker, δ_{ij}) :

$$AI = IA = A \quad \Leftrightarrow \quad \sum_k a_{ik} \delta_{kj} = \sum_k \delta_{ik} a_{kj} = a_{ij}, \quad \forall i, j$$

12.2.2 Inversion

Venons-en à la définition et au calcul de l'inverse des matrices. L'inverse d'une matrice A , que nous noterons, A^{-1} est telle que son produit avec A redonne l'identité :

$$AA^{-1} = A^{-1}A = I$$

L'inverse de la matrice A peut être calculé très facilement à l'aide du déterminant de A et des cofacteurs de la transposée A^T (la transposée d'une matrice est la matrice obtenue en permutant ligne et colonne)²

$$a_{ij}^{-1} = \frac{A_{ij}^T}{\det A}, \quad \forall i, j$$

Ce résultat implique donc qu'une matrice de déterminant nul n'a pas d'inverse (pour les nombres réels ou complexes, 0, de même n'a pas d'inverse). Une telle matrice est dite *singulière*.

D'un point de vue pratique, on procède comme suit. Dans le cas des matrices d'ordre 2 par exemple :

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \quad A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Dans le cas d'une matrice d'ordre 3, traitons à titre d'exemple un cas particulier.

$$\text{Soit } A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 2 & 2 & 0 \end{pmatrix} \quad \text{alors } A^T = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 1 & 2 \\ 0 & 1 & 0 \end{pmatrix}$$

Le déterminant de A peut être calculé par développement selon la première ligne : $\det A = 1(1 \cdot 0 - 2 \cdot 1) + 2(-1 \cdot 1 - 1 \cdot 0) = -4$, d'où l'expression de A^{-1} :

$$A^{-1} = \frac{1}{-4} \begin{pmatrix} +(1 \cdot 0 - 1 \cdot 2) & -(-1 \cdot 0 - 2 \cdot 0) & +(-1 \cdot 1 - 1 \cdot 0) \\ -(0 \cdot 0 - 2 \cdot 1) & +(1 \cdot 0 - 2 \cdot 0) & -(1 \cdot 1 - 0 \cdot 0) \\ +(0 \cdot 2 - 2 \cdot 1) & -(1 \cdot 2 - 2 \cdot (-1)) & +(1 \cdot 1 - 0 \cdot (-1)) \end{pmatrix}$$

soit encore,

$$A^{-1} = \begin{pmatrix} +1/2 & 0 & +1/4 \\ -1/2 & 0 & +1/4 \\ +1/2 & 1 & -1/4 \end{pmatrix}$$

1. Une autre façon d'exprimer cette propriété de non-commutativité (utilisée en particulier en mécanique quantique) consiste à introduire les commutateurs : $[A, B] \equiv AB - BA$. La relation de non-commutation s'écrit donc simplement sous la forme $[A, B] \neq 0$.

2. On peut au moins le vérifier :

$$\sum_j a_{ij}^{-1} a_{jk} = \sum_j \frac{A_{ij}^T}{\det A} a_{jk} = \frac{1}{\det A} \sum_j A_{ji} a_{jk}$$

Si $k = i$, $\sum_j A_{ji} a_{jk} = \sum_j A_{ji} a_{ji} = \det A$ et on trouve bien 1, tandis que si $k \neq i$, $\sum_j A_{ji} a_{jk} = 0$ car la matrice A a deux colonnes identiques.

12.3 Système d'équations linéaires

Une des applications importantes du calcul matriciel est la solution des systèmes d'équations algébriques.

Définition 12.3.1 *Un système d'équations linéaires $m \times n$ est une famille de m équations linéaires avec n inconnues, x_1, x_2, \dots, x_n :*

$$\begin{array}{rcl} a_{11}x_1 + \dots + a_{1n}x_n & = & b_1 \\ \dots\dots\dots\dots\dots\dots\dots & \dots & \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n & = & b_m \end{array}$$

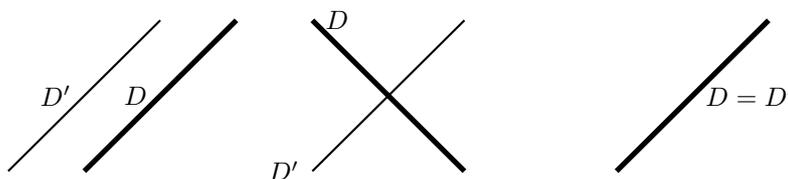
Les coefficients a_{ij} sont des nombres réels ou complexes. Une forme matricielle équivalente s'écrit $Ax = b$ où A est une matrice rectangulaire $m \times n$, x et b sont des vecteurs colonnes respectivement à n et m termes.

Si $b \equiv 0$, le système est dit homogène ou sans second membre. Le système est dit sous-déterminé si $n > m$ (plus d'inconnues que d'équations) et sur-déterminé dans le cas opposé.

La question centrale est de savoir si un tel système possède des solutions et si oui, combien.

Par exemple, dans le cas des systèmes de 2 équations à 2 inconnues :

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 & = & b_1 \\ a_{21}x_1 + a_{22}x_2 & = & b_2, \end{array}$$



Ce système correspond à 2 droites dans le plan (x_1, x_2) . 3 cas sont donc possible : aucune solution si les droites sont parallèles, 1 solution si les droites se coupent, une infinité si elles sont confondues. Dans le cas des systèmes 3×3 , les équations mettent en jeu 3 plans dans \mathbb{R}^3 qui conduisent également aux 3 cas possibles (aucune, une seule ou une infinité). Cette situation est en fait générale dès que le corps auquel appartiennent les coefficients est infini (ce qui est le cas de \mathbb{R} et \mathbb{C}) :

Théorème 12.3.1 *Un système d'équations linéaires dont les coefficients sont réels ou complexes possède : une seule solution, aucune solution ou une infinité de solutions.*

La notion de "rang d'une matrice" permet de trancher entre ces 3 possibilités.

Définition 12.3.2 *Le rang d'une matrice A , que nous noterons $r(A)$ est l'ordre de la plus grande sous-matrice carrée de A dont le déterminant ne s'annule pas.*

Par exemple la matrice rectangulaire 2×3 définie par

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 0 & 0 \end{pmatrix}$$

a 3 sous-matrices 2×2 :

$$\begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix} \quad \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

Il est clair que le déterminant des 2 dernières est nul mais le déterminant de la première vaut -6, donc $r(A) = 2$.

Le théorème fondamental est le suivant :

Théorème 12.3.2 Soit le système linéaire $m \times n$, $Ax = b$. On note $A|b$ la matrice augmentée $m \times (n + 1)$ obtenue en ajoutant b comme colonne supplémentaire à A . Alors,

- si $r(A) = r(A|b) = n$, le système admet une seule solution.
- si $r(A) = r(A|b) < n$, le système admet une infinité de solutions.
- si $r(A) < r(A|b)$, le système admet aucune solution.

Ce théorème couvre les cas les plus généraux. Deux cas particuliers sont importants :

1. les systèmes inhomogènes carrés ($m = n$).

Si $\det A \neq 0$, $r(A|b) = n$: la première condition du théorème est vérifiée, et le système $Ax = b$ admet une solution unique donnée par $x = A^{-1}b$ (A^{-1} existe puisque A n'est pas singulière).

2. les systèmes homogènes ($b \equiv 0$).

Si $\det A \neq 0$, $r(A|b) = r(A) = n$: la première condition du théorème est vérifiée, et le système $Ax = 0$ admet une solution unique (triviale) donnée par $x = 0$. Pour les systèmes homogènes, on obtient des solutions non nulles si et seulement si $\det A = 0$.

Exemple Soit à résoudre le système d'équations :

$$\begin{aligned} x_1 + x_2 + x_3 &= -2 \\ x_1 - x_2 + x_3 &= +2 \\ -x_1 + x_2 - x_3 &= -2 \end{aligned}$$

La matrice augmentée s'écrit sous la forme :

$$A|b = \left(\begin{array}{ccc|c} 1 & 1 & 1 & -2 \\ 1 & -1 & 1 & 2 \\ -1 & 1 & -1 & -2 \end{array} \right)$$

La 2ème et la 3ème ligne de A étant proportionnelle, on a $\det A = 0$, ce qui prouve qu'il existe d'autres solutions que la solution triviale nulle pour le système homogène. Par combinaisons linéaires, on trouve $x_2 = 0$ et $x_3 = -x_1$. Pour le système inhomogène, on trouve $r(A|b) = r(A) = 2 < 3$: il y a donc une infinité de solutions. Plus précisément, pour le système inhomogène, on trouve $x_2 = -2$ et $x_1 = -x_3$.

12.4 Diagonalisation

Définition 12.4.1 Une matrice carrée A est diagonalisable s'il existe une matrice inversible P telle que la matrice $D = P^{-1}AP$ soit diagonale.

Comme nous le verrons dans la suite, le fait qu'une matrice A soit diagonalisable ou pas est important en tant que tel comme caractérisation de l'opérateur associée à la matrice A . Lorsqu'une matrice est diagonalisable, tous les calculs algébriques sont également simplifiés. En effet, soit à calculer A^n par exemple, comme $PP^{-1} = I$ alors $P^{-1}A^nP = P^{-1}AIAI \dots IAP = P^{-1}AP \dots P^{-1}AP = D^n$ qui se calcule aisément puisque D est diagonale. A^n est ensuite obtenue par le produit $A^n = PD^nP^{-1}$. Ainsi, tout polynôme en A se calcule aisément lorsque A est diagonalisable.

Les vecteurs qui se trouvent seulement contractés ou dilatés dans leur direction sous l'action de l'opérateur linéaire associé à la matrice A jouent un rôle spécifique pour la diagonalisation. Ces vecteurs sont appelés vecteurs propres :

Définition 12.4.2 Les vecteurs non nuls $|v\rangle$ tels que :

$$A|v\rangle = \lambda|v\rangle$$

sont appelés les vecteurs propres associés à A . λ qui peut être réel ou complexe est appelée la valeur propre correspondante (c'est le coefficient de dilatation ou de contraction de $|v\rangle$ dans sa propre direction).

D'après cette définition, déterminer le problème aux valeurs propres (i.e. trouver les vecteurs et valeurs propres) associé à une matrice A revient à résoudre le système linéaire homogène :

$$(A - \lambda I) |v\rangle = 0$$

Comme on l'a vu plus haut, la seule solution non triviale (i.e. telle que $|v\rangle \neq |0\rangle$) est telle que :

$$\det(A - \lambda I) = 0$$

Pour A d'ordre n connue, il s'agit d'un polynôme en λ d'ordre n , que nous noterons $P(\lambda)$ et que l'on appelle *le polynôme caractéristique*. Les valeurs propres λ sont donc les racines de ce polynôme :

$$P(\lambda) = \det(A - \lambda I) = 0$$

Exemple Considérons la matrice 2×2 :

$$A = \begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix}$$

Le polynôme caractéristique est $P(\lambda) = (1-\lambda)^2 + 4$, dont les 2 racines sont $1 \pm 2i$. Le vecteur propre correspondant à $1 + 2i$ est $(i\alpha, \alpha)$ tandis que celui correspondant à $1 - 2i$ est $(-i\alpha, \alpha)$ avec $\lambda \neq 0$.

Dans le cas des matrices d'ordre 2, il est facile de montrer que le polynôme caractéristique s'exprime simplement en fonction de la trace et du déterminant de A :

Exercice 12.1 Soit A une matrice 2×2 .

1. Montrer que le polynôme caractéristique ne dépend que de $\text{tr}A$ et $\det A$.
2. Déterminer les 2 valeurs propres λ_1 et λ_2 et montrer que :

$$\begin{aligned} \text{tr}A &= \lambda_1 + \lambda_2, \\ \det A &= \lambda_1 \lambda_2. \end{aligned}$$

3. En déduire qu'il existe 3 types de valeurs propres possibles pour les systèmes linéaires d'ordre 2.

Le nombre de valeurs propres possibles dépend du corps, typiquement \mathbb{R} ou \mathbb{C} dans lequel on cherche les racines de l'équation $P(\lambda) = 0$. Dans \mathbb{R} , tout est possible, depuis aucune jusqu'à n solutions. Dans \mathbb{C} , le théorème de D'Alembert (cf. chapitre "Nombres Complexes"), nous garantit l'existence de n valeurs propres, éventuellement multiples. Dans ce dernier cas (racine multiple), on dit que *la valeur propre est dégénérée*.

Une fois trouvées les valeurs propres, la résolution des systèmes linéaires associés à chaque valeur propre permet d'obtenir les vecteurs propres correspondants. Compte tenu de la forme linéaire $A|v\rangle = \lambda|v\rangle$, il est clair que les vecteurs propres sont déterminés à un facteur multiplicatif près (si $|v\rangle$ est vecteur propre, $\alpha|v\rangle$ avec $\alpha \in \mathbb{C}$ l'est aussi).

Deux cas peuvent alors se présenter :

- où l'ensemble des vecteurs propres $(|v_1\rangle, |v_2\rangle, \dots, |v_n\rangle)$ sont indépendants (ils constituent alors une base).
- où l'ensemble des vecteurs propres $(|v_1\rangle, |v_2\rangle, \dots, |v_n\rangle)$ ne sont pas indépendants.

Considérons la matrice P dont les colonnes sont formées par les vecteurs propres de la matrice A . Alors, puisque $A|v_i\rangle = \lambda_i|v_i\rangle$, le produit matriciel AP est une matrice dont les colonnes sont les vecteurs $\lambda_i|v_i\rangle$. Cette dernière matrice s'écrit elle-même comme le produit de P par la matrice diagonale D dont les éléments sont les valeurs propres λ_i . En clair, la matrice P construite avec les vecteurs propres satisfait la relation :

$$AP = PD$$

Si l'ensemble des vecteurs propres sont indépendants, $\det P \neq 0$, la matrice P est donc inversible et on peut donc écrire :

$$D = P^{-1}AP$$

où, rappelons-le, P a pour colonnes les vecteurs propres de A . Si l'ensemble des vecteurs propres ne sont pas indépendants, $\det P = 0$, la matrice P n'est donc pas inversible et on doit en rester là. On peut montrer qu'il n'existe pas d'autre matrice P qui satisfasse la relation $D = P^{-1}AP$: la matrice A n'est donc pas diagonalisable. Le théorème précis et fondamental est le suivant :

Théorème 12.4.1 Une matrice carrée $n \times n$ est diagonalisable si et seulement si elle possède n vecteurs propres linéairement indépendants.

La matrice diagonale D a pour éléments les valeurs propres de A , et la matrice P telle que $P^{-1}AP = D$ a pour colonnes les vecteurs propres correspondants.

Suivant les exigences de ce théorème, on doit s'attendre à ce que beaucoup de matrices ne soient pas diagonalisables. C'est bien le cas. Il existe cependant certaines matrices fréquemment rencontrées dans les applications pour lesquelles la diagonalisation est garantie. Il s'agit :

- des matrices ayant n valeurs propres distinctes,
- des matrices symétriques ou hermitiennes.

Le fait que les matrices ayant des valeurs propres distinctes soient diagonalisables résultent du théorème 12.4.1 et du théorème suivant :

Théorème 12.4.2 Une matrice ayant des valeurs propres distinctes a nécessairement des vecteurs propres indépendants.

La démonstration se fait par l'absurde³. Si certaines valeurs propres sont dégénérées (valeurs propres confondues), les vecteurs propres peuvent être indépendants ou pas. Le système doit être étudié au cas par cas.

Dans le cas des matrices symétriques ou hermitiennes, les vecteurs propres sont linéairement indépendants, même si les valeurs propres sont dégénérées. Rappelons qu'une matrice symétrique est égale à sa matrice transposée ($a_{ij} = a_{ji}$) et qu'une matrice hermitienne (notée A^\dagger) est égale à sa matrice transposée conjuguée ($a_{ij} = \overline{a_{ji}}$). Une matrice symétrique est donc un cas particulier de matrice hermitienne dont les éléments sont réels. Le formalisme de la mécanique quantique met précisément en jeu des opérateurs hermitiques. On dispose du théorème suivant :

Théorème 12.4.3 Les valeurs propres d'une matrice hermitienne sont réelles.

Les vecteurs propres d'une matrice hermitienne sont mutuellement orthogonaux (et donc indépendants).

En effet, utilisant la notation $\langle v | \equiv |v\rangle^\dagger$, on a d'une part $\langle v_j | A | v_i \rangle = \lambda_i \langle v_j | v_i \rangle$ et d'autre part $\langle v_j | A^\dagger | v_i \rangle = \overline{\lambda_j} \langle v_j | v_i \rangle$. Comme la matrice est hermitienne par hypothèse, $A = A^\dagger$, en faisant la différence de ces 2 expressions on obtient : $(\lambda_i - \overline{\lambda_j}) \langle v_j | v_i \rangle = 0$. Si $i = j$, on en déduit que $\lambda_i = \overline{\lambda_i}$ puisque $\langle v_i | v_i \rangle > 0$, tandis que si $i \neq j$ et $\lambda_i \neq \lambda_j$, on obtient la condition d'orthogonalité $\langle v_j | v_i \rangle = 0$ (si $\lambda_i = \lambda_j$, $|v_j\rangle$ et $|v_i\rangle$ ne sont pas automatiquement orthogonaux mais peuvent être rendus tels).

En utilisant les résultats précédents, on a donc obtenu le théorème suivant :

Théorème 12.4.4 Les matrices ayant des valeurs propres distinctes, et les matrices symétriques ou hermitiennes, sont diagonalisables.

Exemple Considérons 2 particules identiques de masse m reliées à 3 ressorts identiques et contraintes de se déplacer selon une seule direction Ox . x_1 et x_2 désignant les déplacements des particules par rapport à leurs positions d'équilibre, les équations du mouvement s'écrivent :

$$\begin{aligned} m \ddot{x}_1(t) &= k(x_2 - 2x_1), \\ m \ddot{x}_2(t) &= k(x_1 - 2x_2) \end{aligned}$$

En notation matricielle, le système différentiel s'écrit :

$$m \begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = k \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

3. Supposons que $(|v_1\rangle, \dots, |v_p\rangle)$ soit des vecteurs liés. Un vecteur propre particulier, disons $|v_1\rangle$ s'exprime donc en fonction des $(p-1)$ autres supposés indépendants : $|v_1\rangle = \sum_{i=2,p} c_i |v_i\rangle$ avec c_i non tous nuls. Mais $(A - \lambda_1)|v_1\rangle = \sum_{i=2,p} c_i (\lambda_i - \lambda_1)|v_i\rangle$. Comme $(|v_2\rangle, \dots, |v_p\rangle)$ sont indépendants, il en résulte que $c_i (\lambda_i - \lambda_1) = 0, \forall i = 2, p$. Mais $\lambda_i \neq \lambda_1, \forall i = 2, p$ par hypothèse, donc $c_i = 0, \forall i = 2, p$ en contradiction avec le fait que les c_i étaient supposés non tous nuls.

On cherche les solutions de ce système différentiel sous la forme $x_1(t) = u_1 e^{i\omega t}$ et $x_2(t) = u_2 e^{i\omega t}$. Le système différentiel prend alors la forme suivante :

$$-m\omega^2 \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} -2k & k \\ k & -2k \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

ce que l'on peut écrire sous la forme d'un problème aux valeurs propres :

$$A \mathbf{u} = -m\omega^2 \mathbf{u}$$

Les 2 valeurs propres de A sont $-k$ et $-3k$, auxquelles correspondent les 2 fréquences caractéristiques

$$\omega_1 = \sqrt{k/m} \quad \text{et} \quad \omega_2 = \sqrt{3k/m}$$

Les 2 vecteurs propres correspondants sont respectivement $\begin{pmatrix} \alpha \\ \alpha \end{pmatrix}$ et $\begin{pmatrix} \alpha \\ -\alpha \end{pmatrix}$ où α est un nombre réel non nul. La matrice étant symétrique, les 2 vecteurs propres sont bien orthogonaux. La solution générale du système différentiel s'écrit donc (comme attendu) sous la forme :

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = A_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{i\omega_1 t} + A_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{i\omega_2 t}$$

où A_1 et A_2 sont des constantes fixées par les conditions initiales.

Les 2 vecteurs propres correspondent aux 2 modes "normaux" du mouvement des ressorts, respectivement en phase ou en opposition de phase.

Chapitre 13

Scalars, vecteurs, tenseurs

Introduction

Notation et changement de base

Scalars, vecteurs et tenseurs

Produits et contractions

Base réciproque

Formes et espace dual

On présente dans ce chapitre quelques notions d'algèbre tensorielle utiles, en particulier, pour la compréhension des cours d'élasticité et de cristallographie. Comme on le découvrira dans l'introduction, les tenseurs généralisent la notion familière de vecteurs et d'opérateurs linéaires en les plaçant dans le cadre mathématique plus large de l'algèbre multilinéaire.

13.1 Introduction

Soit E un espace vectoriel réel (le corps de base est donc \mathbb{R}) de dimension finie N . Tout ensemble de N vecteurs linéairement indépendants forment une base, et il en existe une infinité. Soit $\{\vec{e}_i\}$ et $\{\vec{e}'_i\}$ deux bases différentes de E . Notons P et P^{-1} les matrices de passage¹ d'une base à l'autre, soit :

$$\vec{e}'_i = \sum_{j=1}^N P_{ji} \vec{e}_j, \quad \text{et} \quad \vec{e}_j = \sum_{i=1}^N P_{ij}^{-1} \vec{e}'_i \quad (13.1)$$

Un vecteur quelconque $\vec{x} \in E$ peut s'écrire indifféremment dans l'une ou l'autre base de E :

$$\vec{x} = \sum_{j=1}^N x_j \vec{e}_j = \sum_{i=1}^N x'_i \vec{e}'_i$$

En utilisant, les formules (13.1), on est conduit aux relations suivantes entre les composantes d'un vecteur, après changement de base :

$$x'_i = \sum_{j=1}^N P_{ij}^{-1} x_j, \quad \text{et} \quad x_j = \sum_{i=1}^N P_{ji} x'_i \quad (13.2)$$

Considérons maintenant un opérateur linéaire $A : \vec{x} \in E \mapsto A\vec{x} \in E$. L'opérateur A est représenté dans la base $\{\vec{e}_i\}$ par une matrice telle que sa j ème colonne soit le vecteur $A\vec{e}_j$:

$$A\vec{e}_j = \sum_{k=1}^N A_{kj} \vec{e}_k$$

1. Par définition, la matrice de passage P de la base $\{\vec{e}_i\}$ à la base $\{\vec{e}'_i\}$ est telle que sa i ème colonne est formée des composantes de \vec{e}'_i par rapport à la base $\{\vec{e}_i\}$.

Rappelons quelle est l'expression de A dans la base $\{\vec{e}'_i\}$. Par définition on a $A\vec{e}'_j = \sum_k A'_{kj} \vec{e}'_k$, mais on a aussi :

$$A\vec{e}'_j = A \left(\sum_{l=1}^N P_{lj} \vec{e}_l \right) = \sum_{l,m=1}^N P_{lj} A_{ml} \vec{e}_m = \sum_{k=1}^N \left(\sum_{l,m=1}^N P_{lj} A_{ml} P_{km}^{-1} \right) \vec{e}'_k,$$

soit encore la formule bien connue :

$$A'_{kj} = \sum_{m,l=1}^N P_{km}^{-1} A_{ml} P_{lj}, \quad (13.3)$$

Réécrivons (13.2) et (13.3) sous forme matricielle :

$$\vec{x}' = P^{-1} \vec{x} \quad \text{et} \quad A' = P^{-1} A P.$$

Un vecteur met donc en jeu une seule matrice de passage (P^{-1}) lors d'un changement de base, tandis qu'il en faut 2 (P et P^{-1}) pour transformer un opérateur linéaire. Un nombre réel, invariant par changement de base, n'en met aucune en jeu. Le calcul tensoriel systématise ce point de vue en considérant des êtres mathématiques nouveaux : les tenseurs, qui se transforment lors d'un changement de base en mettant en jeu un nombre quelconque de matrices de passage P et P^{-1} . L'algèbre (et l'analyse) tensorielle constitue une généralisation de l'algèbre (et de l'analyse) linéaire élémentaire que les mathématiciens appellent algèbre multilinéaire.

13.2 Notations et changement de base

En plus d'un approfondissement des structures linéaires, le calcul tensoriel introduit un ensemble de notations commodes et cohérentes indispensables dans les calculs.

Commençons par les vecteurs. Soit \vec{x} un vecteur quelconque de l'espace vectoriel E . On écrira désormais :

$$\vec{x} = \sum_{i=1}^N x^i \vec{e}_i \equiv x^i \vec{e}_i$$

où (x^1, \dots, x^N) sont les composantes du vecteur \vec{x} dans la base $\{\vec{e}_i\}$. La position des indices dans x^i et \vec{e}_i peut sembler arbitraire pour l'instant, mais sera justifiée par la suite. La deuxième égalité correspond à la convention de sommation d'Einstein : tout indice répété 2 fois, une fois en position inférieure et une fois en position supérieure, implique une sommation sur cet indice.

Considérons maintenant une nouvelle base de E : $\{\vec{e}'_i\}$. Chacun des vecteurs de cette nouvelle base peut être décomposé sur l'ancienne base $\{\vec{e}_i\}$. On passe donc d'une base à l'autre par une matrice de changement de base que l'on notera désormais α :

$$\vec{e}'_i = \alpha_i^j \vec{e}_j$$

Dans cette expression, pour utiliser la convention d'Einstein, nous avons écrit les composantes de la matrice sous la forme α_i^j où l'indice supérieur j désigne une ligne et l'indice inférieur i une colonne.

De ce point de vue, comme les vecteurs peuvent être considérés comme des matrices ligne ou colonne, on a de façon plus explicite :

$$\vec{x} = x^i \vec{e}_i = \vec{e}_i x^i = (\vec{e}_1 \cdots \vec{e}_N) \begin{pmatrix} x^1 \\ \vdots \\ x^N \end{pmatrix}$$

(le produit matriciel est défini comme un produit ligne \times colonne : on somme sur les colonnes de la première matrice et sur les lignes de la seconde matrice).

Inversement, notons $\beta \equiv \alpha^{-1}$, l'inverse de la matrice α . On a donc (toujours en utilisant la convention d'Einstein) :

$$\beta_k^i \vec{e}'_i = \beta_k^i \alpha_i^j \vec{e}_j = \alpha_i^j \beta_k^i \vec{e}_j = \delta_k^j \vec{e}_j = \vec{e}_k,$$

où δ est la matrice identité. En résumé, on a donc obtenu :

$$\vec{e}'_i = \alpha_i^j \vec{e}_j \quad \text{et} \quad \vec{e}_k = \beta_k^i \vec{e}'_i,$$

qui n'est rien d'autre qu'une réécriture des relations (13.1).

Exercice 13.1 *Ecrire $\vec{e}'_i = \alpha_i^j \vec{e}_j$ et $\vec{e}_k = \beta_k^i \vec{e}'_i$ sous forme de produits matriciels explicites.*

13.3 Scalaires, vecteurs et tenseurs

Un vecteur existe indépendamment de sa représentation dans une base particulière. On peut donc écrire :

$$\vec{x} = x^k \vec{e}_k = x'^i \vec{e}'_i$$

(la somme est sur un indice muet que l'on peut noter à notre convenance). Comme $\vec{e}_k = \beta_k^i \vec{e}'_i$, on trouve que la relation entre les composantes du vecteur dans les 2 bases est :

$$x'^i = \beta_k^i x^k,$$

c'est-à-dire, qu'à la différence des vecteurs de bases, qui se transforment via la matrice α ($\vec{e}'_i = \alpha_i^j \vec{e}_j$), les composantes des vecteurs se transforment via la matrice inverse β . Pour cette raison, les composantes x^i du vecteur \vec{x} sont dites *contravariantes*.

Exercice 13.2 *On considère l'espace $E = \mathbb{R}^3$ et $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ une base orthonormée.*

1. Déterminer la matrice α , \vec{e}'_i étant obtenu à partir de \vec{e}_i par une rotation d'angle θ autour de l'axe portant \vec{e}_3 .
2. Vérifier que α^{-1} est égale à la transposée de la matrice α .
3. Un vecteur admet les composantes $(0, -1, +1)$ dans la base $\{\vec{e}'_i\}$. Quelles étaient ses anciennes composantes ?
4. Un vecteur admet les composantes $(0, +1, 0)$ dans la base $\{\vec{e}_i\}$. Quelles étaient ses composantes dans la base $\{\vec{e}'_i\}$?

Tout espace vectoriel E défini sur un corps commutatif peut être muni d'une structure euclidienne dès lors qu'on y définit un produit scalaire². On parle alors d'*espace vectoriel euclidien*. Soient \vec{x} et \vec{y} des vecteurs quelconques de E , le produit scalaire de ces 2 vecteurs sera noté indifféremment :

$$g(\vec{x}, \vec{y}) = \langle \vec{x} | \vec{y} \rangle = \vec{x} \cdot \vec{y}$$

Compte tenu des propriétés du produit scalaire, on a encore :

$$g(\vec{x}, \vec{y}) = \langle x^i \vec{e}_i | y^j \vec{e}_j \rangle = x^i y^j \langle \vec{e}_i | \vec{e}_j \rangle = g_{ij} x^i y^j,$$

où on a posé $g_{ij} \equiv g(\vec{e}_i, \vec{e}_j)$. g est appelée *la métrique* ou la forme fondamentale, et g_{ij} les coefficients de la métrique. Noter la propriété $g_{ij} = g_{ji}$ puisque le produit scalaire est symétrique par définition.

Etudions comment les g_{ij} se transforme lors d'un changement de base :

$$g'_{ij} \equiv \langle \vec{e}'_i | \vec{e}'_j \rangle = \alpha_i^k \alpha_j^l \langle \vec{e}_k | \vec{e}_l \rangle = \alpha_i^k \alpha_j^l g_{kl}$$

La métrique g constitue un exemple de tenseur (euclidien) de rang 2, c'est-à-dire un exemple d'être mathématique dont les composantes, caractérisées par 2 indices, se transforment lors d'un changement de base selon la relation précédente (ici deux matrices α).

Il est possible de définir des composantes du vecteur \vec{x} qui se transforment, lors d'un changement de base, comme les vecteurs de base. De telles composantes seront notées avec un indice en bas x_i , et naturellement appelées *covariantes*. Soit en effet, par définition :

$$x_i \equiv \langle \vec{e}_i | \vec{x} \rangle = g_{ij} x^j,$$

2. Un produit scalaire est une forme bilinéaire, symétrique et définie positive.

c'est-à-dire la projection orthogonale de \vec{x} sur \vec{e}_i si la base $\{\vec{e}_i\}$ est normée. Lors du changement de base $\{\vec{e}_i\}$ en $\{\vec{e}'_i\}$, on obtient :

$$x'_i = \langle \vec{e}'_i | \vec{x} \rangle = \alpha_i^j \langle \vec{e}_j | \vec{x} \rangle = \alpha_i^j x_j,$$

qui correspond bien à la même transformation que les vecteurs de bases.

Les 2 relations de transformations des composantes d'un vecteur : $x'^i = \beta_k^i x^k$ et $x'_i = \alpha_i^j x_j$, qui ne dépendent que d'un seul indice, sont caractéristiques des transformations des composantes d'un tenseur de rang 1 lors des changements de base.

Exercice 13.3 1. Montrer que la norme du vecteur \vec{x} s'écrit sous la forme

$$\|\vec{x}\| = \sqrt{g_{ij}x^i x^j},$$

et retrouver la formule usuelle dans le cas où la base est orthonormée.

2. Etablir le résultat : $\langle \vec{x} | \vec{y} \rangle = x_i y^i = x^i y_i$

Considérons pour finir le produit scalaire de 2 vecteurs \vec{x} et \vec{y} lors d'un changement de base. On trouve

$$\langle \vec{x} | \vec{y} \rangle = x^i y_i = (\alpha_k^i x'^k) (\beta_i^l y'_l) = (\beta_i^l \alpha_k^i) x'^k y'_l = \delta_k^l x'^k y'_l = x'^k y'_k$$

Le produit scalaire est donc invariant lors d'un changement de base.

Le produit scalaire, qui ne dépend d'aucun indice, est un exemple de tenseur de rang 0.

Remarque Il apparaît clairement dans ce qui précède que le caractère invariant du produit scalaire lors d'un changement de base résulte de la présence dans le produit $x^i y_i$ du terme contravariant x^i (se transformant donc avec $\beta = \alpha^{-1}$), et du terme covariant y_i (qui se transforme avec α). Cette situation est générale. Toutes les expressions qui comprennent des sommations sur un même indice une fois en position contravariante et une autre fois en position covariante sont des invariants. Les produits scalaires ($x^i y_i$ et $g_{ij}x^i x^j$), les vecteurs ($x^i \vec{e}_i$), et plus généralement les tenseurs, quels que soient leurs rangs, sont des exemples d'invariants.

On peut maintenant se donner une définition (opératoire) des tenseurs, ici de rang 5 pour fixer les idées.

Définition 13.3.1 Soit E un espace vectoriel euclidien de dimension N . Les nombres réels $t_{ij}{}^k{}_l{}^m$, où $i, j, k, l, m = 1 \dots N$, sont les composantes mixtes d'un tenseur t de rang 5, 3 fois covariants et 2 fois contravariants, si leurs lois de transformation dans un changement de base s'écrit :

$$t'_{ij}{}^k{}_l{}^m = \alpha_i^r \alpha_j^s \beta_u^k \alpha_l^v \beta_w^m t_{rs}{}^u{}_v{}^w,$$

où α et β désignent respectivement la matrice de changement de base et son inverse.

13.4 Produits et contractions

1. Produit tensoriel

On peut naturellement définir la somme de 2 tenseurs de même nature. Par exemple, si t_{ij} et u_{ij} sont les composantes des tenseurs doublement covariants t et u , on définit le tenseur $r = t + u$ de telle façon que ses composantes $r_{ij} \equiv t_{ij} + u_{ij}$. Le tenseur obtenu est évidemment doublement covariant par construction. On définit de même le tenseur λt , où $\lambda \in \mathbb{R}$, par la relation $(\lambda t)_{ij} = \lambda t_{ij}$.

Plus intéressante est la définition du produit de 2 tenseurs.

Théorème 13.4.1 Soient u un tenseur de rang p , et v un tenseur de rang q .

Les produits de composantes d'un type donné de u par les composantes d'un type donné de v définissent un tenseur d'ordre $(p + q)$, noté $u \otimes v$.

La démonstration est évidente. $u \otimes v$ s'appelle le produit tensoriel des tenseurs u et v .

Exercice 13.4 Soit 2 vecteurs \vec{x} et \vec{y} .

Montrer qu'il est possible de définir 4 tenseurs différents à partir du produit d'une composante (covariante ou contravariante) de \vec{x} par une composante (covariante ou contravariante) de \vec{y} .

Préciser la loi de transformation de ces tenseurs.

2. *Produit contracté*

Considérons maintenant la notion de produit contracté.

Définition 13.4.1 Soient u un tenseur de rang p , et v un tenseur de rang q .

On appelle produit contracté des tenseurs u et v , le tenseur d'ordre $(p + q - 2)$ obtenu en égalant dans le produit tensoriel $u \otimes v$, un indice covariant et un indice contravariant de chacun des 2 tenseurs.

Le fait d'égaliser un indice supérieur et un indice supérieur implique une sommation sur l'indice commun, et abaisse ainsi de 2 le rang du produit tensoriel initial. Par exemple, partant du tenseur de composantes $u_{ij}{}^k v_p{}^q$, on peut par exemple contracter sur le couple d'indice k et p , puis éventuellement sur j et q :

$$u_{ij}{}^k v_p{}^q \longrightarrow u_{ij}{}^k v_k{}^q \longrightarrow u_{iq}{}^k v_k{}^q$$

Il est clair que par contractions successives, on aboutit aux composantes d'un vecteur (un seul indice libre) si le tenseur initial était de rang impair, et sur un scalaire (pas d'indice libre) dans le cas contraire.

Il est également possible de procéder à une auto-contraction sur un tenseur qui ne s'exprime pas comme un produit tensoriel. Par exemple,

$$t^{ij}{}_k \longrightarrow t^{ik}{}_k$$

Exercice 13.5 Montrer que la trace d'un tenseur de rang 2, définie comme la somme de ses composantes diagonales mixtes, est un tenseur de rang 0, donc un invariant.

La réciproque de la règle du produit contracté permet d'obtenir le critère de tensorialité suivant que nous ne démontrerons pas.

Théorème 13.4.2 Soient $t_{ij}{}^k$ (pour fixer les idées) une liste de nombre réels, et soit u un tenseur quelconque.

Si le produit contracté de t par u est un tenseur, alors t est un tenseur.

Par exemple, on peut montrer de cette façon que les nombres $\delta^i{}_j$ tels que :

$$\delta^i{}_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

sont les composantes mixtes d'un tenseur (tenseur de Kronecker). Pour cela, considérons le tenseur de composantes $x_p y^q$ où \vec{x} et \vec{y} sont des vecteurs (tenseurs de rang 1). Formons le produit $\delta^i{}_j x_p y^q$ et contractons sur le couple d'indices i et p . On obtient $\delta^i{}_j x_i y^q = x_j y^q$, qui constituent bien les composantes mixtes du tenseur $x \otimes y$. D'après le critère de tensorialité, on en déduit donc que δ est bien un tenseur.

3. *Tenseurs symétriques et antisymétriques*

Définition 13.4.2 Un tenseur est dit symétrique (respectivement antisymétrique) par rapport à un couple d'indices de même variance si ses composantes sont inchangées (respectivement changées de signe) par permutation de ces 2 indices.

Un tenseur est dit complètement symétrique (ou antisymétrique) s'il est symétrique (ou antisymétrique) par rapport à tout couple d'indices de même variance.

Il est facile de se convaincre (le faire!) que la nature symétrique (ou antisymétrique) d'un tenseur se conserve par changement de base : la symétrie (ou l'antisymétrie) est une propriété intrinsèque.

L'égalité :

$$t^{ij} = \frac{t^{ij} + t^{ji}}{2} + \frac{t^{ij} - t^{ji}}{2},$$

montre que tout tenseur d'ordre 2 peut s'écrire comme la somme d'un tenseur symétrique et d'un tenseur antisymétrique.

Les tenseurs symétriques (ou antisymétriques) possèdent un nombre restreint de composantes indépendantes non nulles. On vérifiera, que dans un espace de dimension n , ce nombre vaut $n(n + 1)/2$ dans le cas d'un tenseur symétrique, et $n(n - 1)/2$ dans le cas d'un tenseur antisymétrique.

Exercice 13.6 Soit \vec{x} et \vec{y} 2 vecteurs de \mathbb{R}^3 . Le tenseur t , appelé produit extérieur (ou bivecteur), noté $\vec{x} \wedge \vec{y}$, est défini par l'égalité :

$$t \equiv \vec{x} \wedge \vec{y} = \vec{x} \otimes \vec{y} - \vec{y} \otimes \vec{x}$$

- (a) *Ecrire explicitement les composantes 2 fois contravariantes de t en fonction des composantes contravariantes de \vec{x} et \vec{y} .*
- (b) *Utiliser le critère de tensorialité pour vérifier que les t^{ij} sont bien les composantes d'un tenseur de rang 2.*
- (c) *Vérifier que le produit extérieur est un tenseur antisymétrique.*
- (d) *Montrer que t n'a que 3 composantes indépendantes que l'on notera u^1, u^2 et u^3 .*
- (e) *Montrer que le triplet de réels (u^1, u^2, u^3) , qui correspondent aux composantes du produit vectoriel usuel, ne constitue pas les composantes d'un vecteur.*
On dit que le produit vectoriel est un pseudo-vecteur.

13.5 Base réciproque

On se pose la question de savoir s'il existe une base $\{\vec{e}^i\}$ de E telle que les composantes du vecteur \vec{x} sur cette base soient les composantes covariantes x_i précédemment définies, soit :

$$\vec{x} = x_i \vec{e}^i$$

(le caractère invariant de \vec{x} suggère la nature contravariante des vecteurs \vec{e}^i , ce que l'on vérifiera plus loin).

En utilisant la définition $x_i = g_{ij} x^j$ et l'égalité $\vec{x} = x^j \vec{e}_j$ on obtient par identification

$$\vec{e}_j = g_{ji} \vec{e}^i$$

Ecrivons la relation inverse sous la forme :

$$\vec{e}^i \equiv g^{ik} \vec{e}_k$$

On obtient alors $\vec{e}_j = g_{ji} g^{ik} \vec{e}_k$ et puisque $\vec{e}_j = \delta_j^k \vec{e}_k$, on en déduit que la contraction de g^{ik} avec le tenseur g_{ji} donne le tenseur de Kronecker,

$$g_{ji} g^{ik} = \delta_j^k$$

D'après le critère de tensorialité, les g^{ik} sont donc les composantes 2 fois covariantes d'un tenseur.

Notons enfin que la relation d'orthogonalité suivante :

$$\langle \vec{e}^k | \vec{e}_i \rangle = g^{kj} \langle \vec{e}_j | \vec{e}_i \rangle = g^{kj} g_{ji} = \delta_i^k,$$

fournit une méthode de construction des \vec{e}^k . Ceux-ci sont orthogonaux à tous les \vec{e}_i pour $i \neq k$, et tels que la projection orthogonale de \vec{e}^k sur \vec{e}_k vaut exactement \vec{e}_k , si la base $\{\vec{e}_i\}$ est normée.

On notera que c'est dans le seul cas où la base est orthonormée, que les bases directe et réciproque ainsi que les composantes covariantes ou contravariantes des tenseurs sont identiques.

Exercice 13.7 *Etablir les égalités :*

1. $\langle \vec{e}^i | \vec{e}^j \rangle = g^{ij}$,
2. $\vec{e}^i = \beta_j^i \vec{e}^j$.

Exercice 13.8 *On se place dans l'espace euclidien à 2 dimensions rapporté à une base (\vec{e}_1, \vec{e}_2) telle que*

$$\|\vec{e}_1\| = a, \quad \|\vec{e}_2\| = b, \quad \langle \vec{e}_1 | \vec{e}_2 \rangle = ab \cos \theta,$$

avec $0 < \theta < \pi/2$.

1. *Calculer les 4 composantes deux fois covariantes de la métrique g .*
2. *Calculer la métrique inverse.*
3. *Déterminer la base réciproque (\vec{e}^1, \vec{e}^2)*
4. *Représenter sur un même schéma les bases directe et réciproque, lorsque $a = b = 1$ et $\theta = 30^\circ$.*

Exercice 13.9 *Étant donné une base quelconque $\{\vec{a}, \vec{b}, \vec{c}\}$ de \mathbb{R}^3 , les cristallographes définissent la base réciproque par les relations :*

$$\vec{a}^* = \frac{\vec{b} \times \vec{c}}{(\vec{a}, \vec{b}, \vec{c})} \quad + \text{permutations circulaires,}$$

où $(\vec{a}, \vec{b}, \vec{c})$ désigne le produit mixte des 3 vecteurs et $\vec{b} \times \vec{c}$ le produit vectoriel de \vec{b} et \vec{c} .

Vérifier que cette définition correspond avec celle du cours.

Cinquième partie

Analyse de Fourier

Chapitre 14

Séries de Fourier

Des problèmes historiques

Définitions des séries de Fourier

Convergence des séries de Fourier

Dérivation et intégration terme à terme

Phénomène de Gibbs

Formule de Plancherel

Séries de fonctions orthogonales

14.1 Des problèmes historiques

L'idée de représenter une fonction par une série de sinus et cosinus de différentes fréquences remonte au problème des cordes vibrantes résolu par l'encyclopédiste d'Alembert en 1747. Ce point de vue sera repris par Fourier en 1808 dans son mémoire sur la propagation de la chaleur.

Dans ce travail, Fourier montre que la température $T(\vec{r}, t)$ obéit à l'équation aux dérivées partielles :

$$T_t = T_{xx} + T_{yy} + T_{zz}.$$

La situation stationnaire ($T_t = 0$), correspond à une équation de Laplace : $\Delta T = 0$. Fourier commence par considérer un cas stationnaire où il y a invariance par translation selon $0z$: $T(\vec{r}, t) = T(x, y, t)$, et cherche une solution dans le domaine défini par $y > 0$ et $0 < x < 1$. Le problème est de déterminer une solution bornée avec les conditions aux limites particulières $T(0, y) = T(1, y) = 0$ et $T(x, 0) = 1$.

Fourier cherche une solution sous forme séparable $T(x, y) = f(x)g(y)$ qui le conduit au système d'équations différentielles

$$\frac{f''(x)}{f(x)} = -\frac{g''(y)}{g(y)} = -C^2,$$

où C est une constante. On a donc comme solution générale de ces équations :

$$\begin{aligned} f(x) &= A \sin Cx + B \cos Cx, \\ g(y) &= A' e^{+Cy} + B' e^{-Cy} \end{aligned}$$

Comme $T(0, y) = 0$ et que la solution recherchée doit être bornée pour $y \rightarrow \infty$, les solutions sont toutes de la forme $T(x, y) = \sin Cx e^{-Cy}$; la condition aux limites $T(1, y) = 0$ fixe la valeur de C qui vaut $n\pi$ où n est un entier positif. Une solution générale s'écrit donc sous la forme de la série : $\sum_{n=1}^{\infty} a_n e^{-n\pi y} \sin n\pi x$, où les coefficients a_n sont contraints par la dernière relation $T(x, 0) = 1$.

En utilisant la propriété $\int_0^1 \sin n\pi x \sin m\pi x dx = \delta_{mn}/2$, on trouve $a_n = 4/(n\pi)$ pour n impair et $a_n = 0$ pour n pair. Les a_n étant bornés, la série obtenue est convergente, et on a donc trouvé une solution sous la forme d'une série de sinus qui comprend toutes les harmoniques impaires :

$$T(x, y) = \sum_{n=1,3,5,\dots}^{\infty} \frac{4}{n\pi} e^{-n\pi y} \sin n\pi x, \quad (x, y) \in [0, 1] \times [0, +\infty[$$

Partant de là, il était naturellement tentant de savoir si le procédé était généralisable, et quelles conditions devait satisfaire une fonction pour qu'on puisse la représenter sur un intervalle par une série trigonométrique. Autrement dit, $a_n(f)$ et $b_n(f)$ étant des coefficients dépendant de la fonction f , quand peut-on écrire

$$f(x) = \sum_{n=0}^{\infty} (a_n(f) \sin n\pi x + b_n(f) \cos n\pi x) \quad ?$$

14.2 Définition des séries de Fourier

Définition 14.2.1 Soit f une fonction de période T , définie sur \mathbb{R} et à valeurs dans \mathbb{R} . On pose $\omega = 2\pi/T$. La série de Fourier associée à la fonction f est la série définie par la relation

$$b_0(f) + \sum_{n=1}^{\infty} b_n(f) \cos n\omega t + \sum_{n=1}^{\infty} a_n(f) \sin n\omega t \equiv \sum_{n=-\infty}^{+\infty} c_n e^{+in\omega t},$$

avec $c_n = (b_n - i a_n)/2$.

Les coefficients de Fourier a_n, b_n et c_n , lorsqu'ils existent, sont définis par les relations

$$\begin{aligned} b_0(f) &= \frac{1}{T} \int_0^T f(t) dt, \\ b_n(f) &= \frac{2}{T} \int_0^T f(t) \cos n\omega t dt, \\ a_n(f) &= \frac{2}{T} \int_0^T f(t) \sin n\omega t dt, \\ c_n(f) &= \frac{1}{T} \int_0^T f(t) e^{-in\omega t} dt. \end{aligned}$$

Pour que les coefficients existent, il n'est pas nécessaire que f soit continue, il suffit que f soit intégrable sur l'intervalle période. On remarquera en particulier que $c_0 = b_0$ et que c_{-n} est la quantité conjuguée de c_n : $c_{-n} = \bar{c}_n$. La définition de ces coefficients est suggérée (en cas de convergence) par les relations d'orthogonalité (à démontrer) :

$$\begin{aligned} \int_0^T \sin m\omega t \sin n\omega t dt &= \int_0^T \cos m\omega t \cos n\omega t dt = \frac{T}{2} \delta_{m,n} \quad (m \neq 0), \\ \int_0^T \sin m\omega t \cos n\omega t dt &= 0, \quad \forall m, n, \in \mathbb{N} \end{aligned}$$

Le développement en séries d'exponentielles complexes peut apparaître plus économique. Il est cependant préférable d'utiliser les développements en sinus ou cosinus lorsque les fonctions ont une parité définie. En effet, comme les intégrations peuvent être prises sur l'intervalle symétrique $[-T/2, +T/2]$, les coefficients a_n ou b_n sont nuls selon que la fonction est paire ou impaire. La série de Fourier d'une fonction impaire ne contient donc que des sinus (fonction impaire), et la série de Fourier d'une fonction paire est une série de cosinus (fonction paire).

Il est bon de mentionner qu'on parle parfois de *développement de Fourier sur un intervalle (a, b) pour une fonction non périodique*. Il s'agit en fait du développement de la fonction f_1 , de période $(b - a)$, qui coïncide avec f sur l'intervalle (a, b) . Si la série converge vers la fonction, le développement en série de f n'est pas en général valable en dehors de (a, b) .

14.3 Convergence des séries de Fourier

Il ne suffit malheureusement pas que les coefficients de Fourier existent pour que la série de Fourier converge vers sa fonction. Kolmogorov a même exhibé un exemple de fonction sommable dont la série de Fourier diverge en tout point de l'intervalle période !

En physique, on a souvent affaire à des fonctions assez régulières, et un premier théorème garantissant la convergence dans des conditions pas trop contraignantes est donné par l'énoncé suivant :

Théorème 14.3.1 *Si f est de carré sommable ou si f est de classe C^2 sur l'intervalle période I , la série de Fourier converge (normalement) vers f dans I .*

Si f est C^2 , en intégrant 2 fois par parties, on établit facilement que $|c_n| \leq \sup |f''(t)|/(4\pi^2 n^2)$, ce qui montre que $\sum_n |c_n| < \infty$. On peut justifier que la somme de la série de Fourier construite avec les c_n s'identifie bien avec f .

Il ne suffit pas qu'une fonction soit continue pour que sa série de Fourier converge. Là encore, les mathématiciens ont pu exhiber des exemples (pathologiques) de fonctions continues dont la série de Fourier diverge au moins en un point de l'intervalle période. On donne maintenant un résultat plus fin dans le cas de fonctions éventuellement discontinues, mais régulières par morceaux.

Définition 14.3.1 *Soit I un intervalle de \mathbb{R} . Une fonction f définie sur I et à valeurs dans \mathbb{R} est dite régulière par morceaux sur I , si f est bornée et n'a qu'un nombre fini de discontinuités et d'extrema dans l'intervalle I .*

On dit également d'une fonction régulière par morceaux, qu'elle vérifie les conditions de Dirichlet. Par exemple, la fonction $x \mapsto \sin 1/x$ est bornée mais a un nombre infini d'extrema dans tout intervalle qui contient 0 : elle n'est donc pas régulière par morceaux.

Théorème 14.3.2 *Soit f une fonction de période T définie sur \mathbb{R} et à valeurs dans \mathbb{R} . On pose $\omega = 2\pi/T$. Si f est régulière par morceaux sur $[0, T]$, alors*

– aux points où f est continue :

$$\sum_{n=-\infty}^{+\infty} c_n e^{+in\omega t} = f(t)$$

– aux points de discontinuité :

$$\sum_{n=-\infty}^{+\infty} c_n e^{+in\omega t} = \lim_{\epsilon \rightarrow 0} \frac{f(t+\epsilon) + f(t-\epsilon)}{2},$$

Ce théorème montre qu'à la différence des séries de Taylor, il est possible de représenter une fonction discontinue par une série de Fourier (sous les conditions de Dirichlet). On notera en outre que la convergence de la série est uniforme aux points de continuité, mais seulement simple aux points de discontinuité. Cet aspect sera discuté de façon plus approfondie lorsque nous traiterons du phénomène de Gibbs.

14.4 Dérivation et Intégration terme à terme

L'intégration, terme à terme, de la série de Fourier : $\sum_n c_n e^{+in\omega t}$, s'écrit

$$\sum_{n=-\infty}^{+\infty} c_n \int_a^b e^{+in\omega t} dt = \sum_{n=-\infty}^{+\infty} \frac{c_n}{in\omega} e^{+in\omega t} \Big|_a^b$$

Le résultat de l'intégration est d'introduire une division par n de chaque coefficient. L'intégration améliore donc la convergence. En conséquence, *une série de Fourier convergente peut toujours être intégrée terme par terme, la série obtenue convergeant vers l'intégrale de la fonction.*

Il est clair au contraire, que la dérivée de la série introduit une multiplication par n de chaque coefficient de Fourier. La convergence s'en trouve dégradée et peut même être perdue. En conséquence, *on ne peut pas, en général, dériver terme à terme, une série de Fourier convergente.*

On dispose cependant du théorème suivant

Théorème 14.4.1 *Soit f une fonction continue de période T définie sur \mathbb{R} et à valeurs dans \mathbb{R} . On pose $\omega = 2\pi/T$. Si f' est régulière par morceaux sur $[0, T]$, alors*

- *La série de Fourier de f' peut être obtenue en dérivant la série de Fourier de f terme à terme.*
- *aux points où f' est continue :*

$$\sum_{n=-\infty}^{+\infty} (in\omega c_n) e^{+in\omega t} = f'(t)$$

- *aux points où f' est discontinue :*

$$\sum_{n=-\infty}^{+\infty} (in\omega c_n) e^{+in\omega t} = \lim_{\epsilon \rightarrow 0} \frac{f'(t + \epsilon) + f'(t - \epsilon)}{2},$$

14.5 Phénomène de Gibbs

Le phénomène de Gibbs concerne l'étude du développement en série de Fourier au voisinage d'un point de discontinuité.

Pour le mettre en évidence, considérons la fonction de période 2, définie sur $] -1, 1[$ par

$$f(x) = \begin{cases} -1 & -1 < x < 0 \\ +1 & 0 < x < 1, \end{cases}$$

On montrera en TD que le créneau carré admet pour développement en série de Fourier

$$f(x) = \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{\sin(2n-1)\pi x}{2n-1},$$

éventuellement valable sur tout \mathbb{R} si on demande en outre que $f(0) = f(1) = 0$.

La figure (14.1) représente ce développement lorsqu'on retient 10 et 50 termes. On constate que loin du point de discontinuité (pour $x = 0.5$, par exemple), un nombre plus grand de termes améliore la convergence (ici vers 1), tandis qu'un phénomène de rebond apparaît au voisinage du point de discontinuité. L'amplitude du rebond ne semble pas se réduire lorsqu'on augmente le nombre de termes de la série. Ce phénomène fut observé par Michelson lors de la mise au point d'un analyseur harmonique et expliqué par J. W. Gibbs.

La raison mathématique de ce comportement surprenant tient dans le fait que la convergence *uniforme* de la série de Fourier vers la fonction est garantie aux points où la fonction est continue, mais perdue aux points de discontinuités.

Considérons la somme finie à N termes correspondant à la série de Fourier

$$g(x, N) \equiv \frac{4}{\pi} \sum_{n=1}^N \frac{\sin(2n-1)\pi x}{2n-1}.$$

Il est facile de montrer (le faire!) que :

$$\frac{\partial g}{\partial x} = 2 \frac{\sin 2N\pi x}{\sin \pi x}$$

Le premier extremum de g (un maximum, en fait) apparaît donc en $x = 1/2N$. Calculons la valeur de g en ce point. Comme $g(0, N) = 0$, on a

$$g\left(\frac{1}{2N}, N\right) = 2 \int_0^{1/2N} \frac{\sin 2N\pi x}{\sin \pi x} dx = \frac{2}{\pi} \int_0^{\pi} \frac{\sin z}{2N \sin(z/2N)} dz.$$

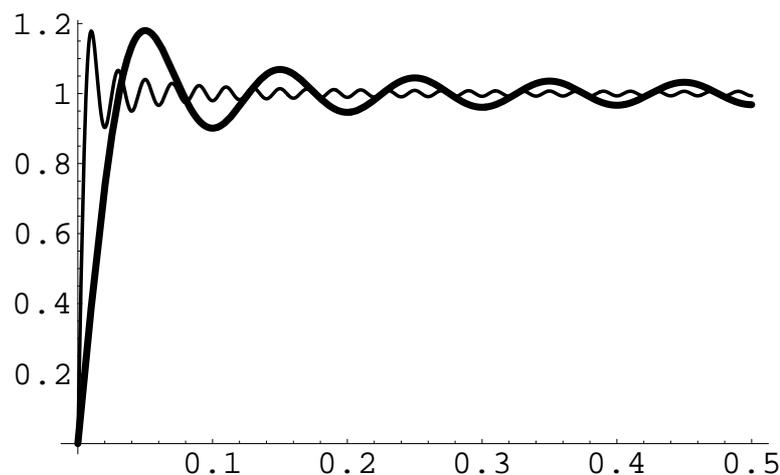


FIGURE 14.1 – Série de Fourier du créneau carré avec 10 termes (trait épais) et 50 termes (trait fin).

Ainsi, même lorsque $N \rightarrow \infty$, la somme dépasse systématiquement la valeur attendue (1) puisque :

$$\lim_{N \rightarrow \infty} g\left(\frac{1}{2N}, N\right) = \frac{2}{\pi} \int_0^{\pi} \frac{\sin z}{z} dz \approx 1.17898,$$

en accord avec le résultat numérique.

On notera qu'il n'en reste pas moins vrai que la série converge (simplement, donc) vers la demi-somme des valeurs à droite et à gauche du point de continuité, c'est-à-dire ici vers 0.

14.6 Exercices

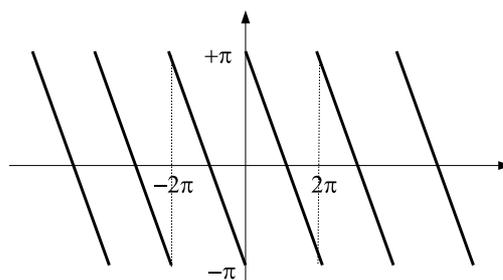
Exercice 14.1 On considère la fonction définie sur \mathbb{R} par :

$$f(t) = |\sin \alpha t|,$$

où α est une constante positive.

1. Représenter schématiquement cette fonction et préciser sa parité.
2. Le développement en série de Fourier contiendra-t-il toutes les harmoniques en sinus et cosinus ?
3. Donner le développement en série de Fourier de la fonction f .

Exercice 14.2 On considère la fonction f , de période 2π , égale à $f(x) = \pi - x$ si $0 < x < 2\pi$.



1. Quelle est la parité de cette fonction ?

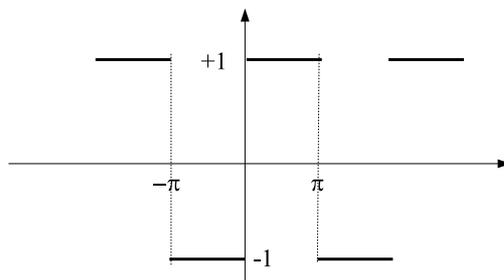
2. Montrer que le développement de f en série de Fourier s'écrit :

$$f(x) = 2 \sum_{n=1}^{\infty} \frac{\sin nx}{n}.$$

Préciser pour quelles valeurs de x ce développement est-il valable.

3. Comment faudrait-il compléter la définition de f pour que le développement précédent soit valable sur tout \mathbb{R} .

Exercice 14.3 Calculer le développement en série de Fourier de la fonction 2π périodique qui vaut -1 pour $t \in]-\pi, 0[$, et $+1$ pour $t \in]0, \pi[$.

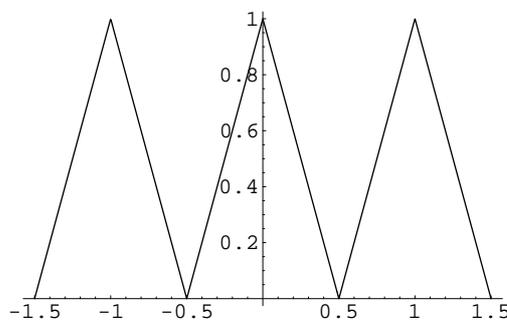


Exercice 14.4 Montrer que le développement en série de Fourier de la fonction $x \mapsto \cos ax$ entre $-\pi$ et $+\pi$, avec a réel non entier s'écrit :

$$\cos ax = \frac{2a \sin \pi a}{\pi} \left(\frac{1}{2a^2} - \frac{\cos x}{a^2 - 1^2} + \frac{\cos 2x}{a^2 - 2^2} - \dots \right).$$

Exercice 14.5 On considère la fonction f définie sur \mathbb{R} , de période $T = 1$, telle que $f(t) = 1 - 2|t|$ pour $t \in [-1/2, +1/2]$.

1. Quelle est sa parité ?
2. Déterminer le développement en série de Fourier de f . On précisera le domaine sur lequel ce développement est valable.
3. Représenter la fonction dérivée f' .
4. Peut-on obtenir le développement en série de Fourier de f' par dérivation terme à terme du développement en série de Fourier de f ? Justifier votre réponse.
5. Déterminer le développement en série de Fourier de f' .



Chapitre 15

Transformation de Fourier au sens des fonctions

Motivations

Transformée de Fourier dans $L^1(\mathbb{R})$

Dérivation et Inversion

Convolution

Transformée de Fourier dans $L^2(\mathbb{R})$

Transformée de Fourier dans $S(\mathbb{R})$

Transformées à plusieurs variables

Applications aux équations différentielles à coefficients constants

15.1 Motivations

Les phénomènes physiques ne se restreignent évidemment pas aux phénomènes périodiques, et il est très naturel de se poser la question de la représentation des phénomènes physiques non périodiques par un analogue des séries de Fourier. La transformation de Fourier joue précisément ce rôle. La transformée de Fourier de la fonction f est la fonction \hat{f} , définie pour $\xi \in \mathbb{R}$ par :

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-i2\pi\xi x} f(x) dx, \quad (15.1)$$

Sous certaines conditions que l'on précisera plus loin, on dispose de la formule d'inversion qui exprime la fonction f comme une superposition de contributions en "fréquences" ξ :

$$f(x) = \int_{\mathbb{R}} e^{i2\pi\xi x} \hat{f}(\xi) d\xi. \quad (15.2)$$

$\hat{f}(\xi)$ représente donc l'amplitude de la sinusoïde de fréquence ξ dans f .

Il est important de réaliser que la transformation de Fourier est une application qui associe une fonction f à une autre fonction \hat{f} : c'est-à-dire *une application fonctionnelle*. Il est bon de s'interroger sur les types de fonctions que le physicien est amené à utiliser le plus souvent. En physique des milieux continus, par exemple, il est commode d'introduire des densités des grandeurs physiques considérées ; ainsi on parle fréquemment de densité de masses, de quantité de mouvement ou d'énergie¹. Par exemple, la relation :

$$\int_{\mathbb{R}^3} \frac{1}{2} \epsilon_0 |\mathbf{E}(x)|^2 dx < \infty,$$

1. c'est-à-dire la masse, la quantité de mouvement ou l'énergie contenue dans le petit élément de volume dx

où $\mathbf{E}(x)$ représente le champ électrostatique, exprime le caractère fini de l'énergie électrique totale. Mathématiquement, cela revient à dire que la fonction $x \mapsto \mathbf{E}(x)$ doit être de carré sommable. Un autre exemple important du même genre de contrainte intervient en physique microscopique lorsqu'on adopte un point de vue probabiliste pour décrire une particule quantique dans son espace des phases. On associe à la particule, une "fonction d'onde" ψ , dont le carré représente une densité de probabilité. La condition de normalisation de la probabilité (la particule doit être quelque part) :

$$\int_{\mathbb{R}^3} |\psi(x)|^2 dx = 1$$

montre que la fonction d'onde ψ doit être de carré sommable.

Dans ce chapitre, on se limitera² à définir et à étudier les propriétés des fonctions de puissance p sommable ($p = 1, 2$), définies sur \mathbb{R}^n . Plus précisément :

Définition 15.1.1 Une fonction f , définie dans \mathbb{R} et à valeurs dans \mathbb{C} , est de puissance p -sommable, si l'on a

$$\int_{\mathbb{R}} |f(x)|^p dx < \infty$$

On note $L^p(\mathbb{R})$, l'espace des fonctions de puissance p -sommable.

On vient de présenter la transformation de Fourier, comme une représentation intégrale exprimée sur une base de fonctions sinusoidales. Le contenu physique d'une telle décomposition est clair, et l'on peut facilement imaginer que la transformation de Fourier jouera un rôle déterminant dans tous les phénomènes physiques mettant en jeu des périodicités spatiales ou temporelles, et plus généralement des longueurs ou temps caractéristiques.

Ainsi, pour donner quelques exemples, tous les phénomènes de diffusion de rayonnement (lumière, rayons X, mais aussi électrons ou neutrons via la dualité onde corpuscule) sont susceptibles d'une analyse reposant sur la transformation de Fourier ; plus précisément, l'amplitude diffusée³ $A(\mathbf{k})$:

$$A(\mathbf{k}) \propto \int_{\mathbb{R}^3} e^{-i\mathbf{k}\cdot\mathbf{r}} \rho(\mathbf{r}) d\mathbf{r},$$

où $\rho(\mathbf{r})$ représente la densité de particules diffusantes. De là, on déduit facilement que l'intensité s'exprime comme la transformée de Fourier d'un produit de convolution (cette notion sera définie plus loin).

La transformation de Fourier, continue ou discrète, est également l'outil de base pour *le traitement de l'information numérique*. Le traitement du signal analogique repose essentiellement sur l'utilisation de circuits électroniques qui fonctionnent mathématiquement comme des opérateurs linéaires diagonalisables par transformation de Fourier. Le traitement du signal discret a été rendu possible par la mise au point d'algorithmes rapides de calcul de transformées de Fourier discrètes. La transformation de Fourier est cependant limitée à une analyse globale soit en temps, soit en fréquence. Les signaux complexes doivent être analysés par des outils plus sophistiqués que les filtres linéaires. Le développement récent du traitement de l'information a conduit à une analyse utilisant des transformées de Fourier à fenêtre et aux célèbres "ondelettes".

Les équations décrivant les phénomènes physiques s'expriment bien souvent par des équations différentielles ou par des équations aux dérivées partielles. Dans le cas où les opérateurs associés sont linéaires et à coefficients constants, il est possible d'utiliser la transformation de Fourier pour diagonaliser ces opérateurs. Parmi celles-ci on peut citer :

– l'équation de Poisson :

$$\Delta\phi(x) = -\frac{\rho(x)}{\epsilon_0}$$

– l'équation de la diffusion :

$$\frac{\partial c(x, t)}{\partial t} - D \frac{\partial^2 c(x, t)}{\partial x^2} = 0$$

2. Un cadre plus confortable pour les besoins de la physique nécessiterait l'étude de la transformation de Fourier au sens des distributions.

3. dans le cadre de l'approximation de Fraunhofer.

– l'équation de Schrödinger libre :

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \Delta \psi(x, t)$$

– l'équation des ondes :

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

Par transformée de Fourier, on se ramène à l'étude d'équations algébriques, ce qui est, bien sûr, considérablement plus simple. Bien que le cadre confortable de travail pour l'étude par transformée de Fourier des équations de la Physique, soit celui, plus général, des distributions, on donnera quelques exemples d'applications à la fin de cette partie du cours.

15.2 Transformée de Fourier dans $L^1(\mathbb{R})$

Définition 15.2.1 Soit $f \in L^1(\mathbb{R})$, on appelle transformée de Fourier de f , la fonction, à valeurs complexes, notée \hat{f} ou $\mathcal{F}(f)$, définie pour tout $\xi \in \mathbb{R}$ par :

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-i2\pi\xi x} f(x) dx$$

on a bien sûr :

$$|\hat{f}(\xi)| = \left| \int_{\mathbb{R}} e^{-i2\pi\xi x} f(x) dx \right| \leq \int_{\mathbb{R}} |e^{-i2\pi\xi x} f(x)| dx = \int_{\mathbb{R}} |f(x)| dx < \infty$$

d'où l'on déduit que la transformée de Fourier d'une fonction sommable est une fonction bornée.

En physique, x représente soit une variable d'espace, soit le temps. La variable conjuguée ξ s'identifiera alors à un vecteur d'onde ou à une pulsation.

Il existe d'autres *conventions de définition*; par exemple :

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-i\xi x} f(x) dx, \quad \text{ou} \quad \hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\xi x} f(x) dx.$$

Avec ces définitions de la transformation de Fourier, il n'apparaît pas de facteur 2π dans la formule d'inversion (voir plus loin), mais il en apparaît soit dans la dérivation, soit dans la convolution.

Premiers exemples

Donnons tout de suite quelques exemples de transformées de Fourier de fonctions sommables couramment utilisées en Physique. Il s'agit des fonctions "porte" ou de ses généralisations et des exponentielles décroissantes à support limité ou non limité.

La fonction caractéristique $\mathbf{1}_{[a,b]}$ de l'intervalle $[a, b]$ est la fonction qui vaut l'unité dans l'intervalle et 0 ailleurs, soit :

$$\begin{aligned} \mathbf{1}_{[a,b]}(x) &= 1 \quad \text{pour } x \in [a, b] \\ &= 0 \quad \text{ailleurs.} \end{aligned}$$

On notera Π la fonction caractéristique de l'intervalle $[-1/2, +1/2]$: c'est la fonction "porte" couramment utilisée en théorie du signal. On notera également H , la fonction dite de Heaviside, c'est-à-dire la fonction caractéristique de l'intervalle $[0, \infty[$.

Exercice 15.1 Etablir les résultats suivants par un calcul direct :

$$\begin{aligned} \Pi(x) &\xrightarrow{FT} \frac{\sin \pi\xi}{\pi\xi}, \\ e^{-\alpha|x|} &\xrightarrow{FT} \frac{2\alpha}{\alpha^2 + 4\pi^2\xi^2} \quad \text{pour } \alpha > 0, \\ H(x) e^{-\alpha x} &\xrightarrow{FT} \frac{1}{\alpha + i2\pi\xi} \quad \text{pour } \alpha > 0, \end{aligned}$$

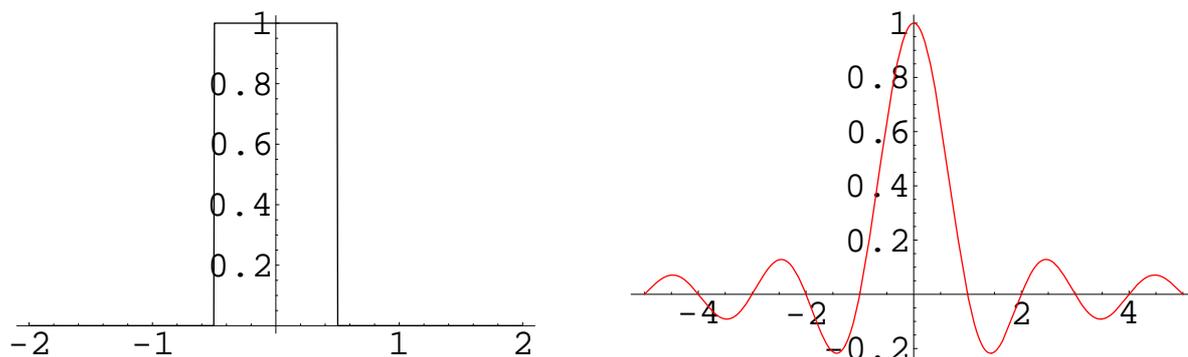
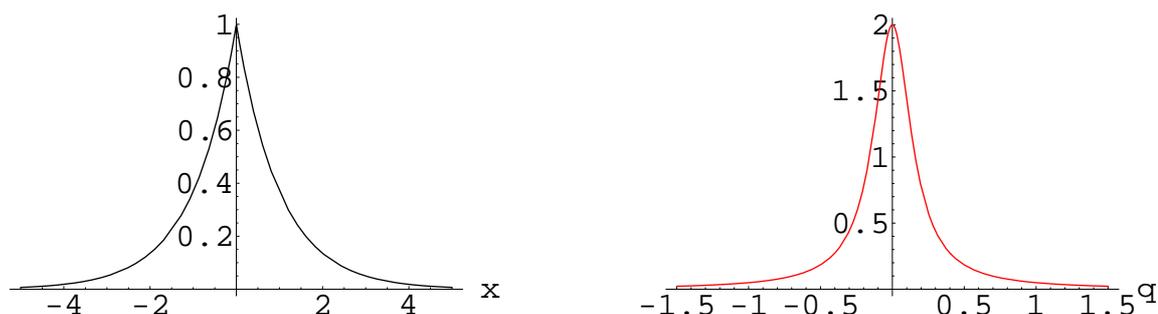


FIGURE 15.1 – Fonction porte et sa transformée de Fourier.

où $x \in \mathbb{R}$, Π la fonction indicatrice de l'intervalle $[-1/2, +1/2]$ et H est la fonction de Heaviside.

FIGURE 15.2 – Fonction $x \mapsto e^{-|x|}$ et sa transformée de Fourier.

Propriétés

Les propriétés qui suivent découlent immédiatement de la définition de la transformation de Fourier (TD).

1. Conservation de la parité.

si f est une fonction paire (resp. impaire), alors \hat{f} est une fonction paire (resp. impaire).

2. Echange de la réalité et de la symétrie hermitienne.

si f est une fonction réelle, alors $\hat{f}(-\xi) = \overline{\hat{f}(\xi)}$,

si $f(-x) = \overline{f(x)}$, alors \hat{f} est une fonction réelle.

3. Echange de la translation et de la modulation.

$$\begin{aligned} f(x - x_0) &\xrightarrow{FT} e^{-i2\pi x_0 \xi} \hat{f}(\xi), \\ e^{i2\pi x \xi_0} f(x) &\xrightarrow{FT} \hat{f}(\xi - \xi_0) \end{aligned}$$

4. Dilatation.

$$f(x/\lambda) \xrightarrow{FT} |\lambda| \hat{f}(\lambda \xi) \quad \text{pour } \lambda \text{ réel non nul}$$

La propriété de dilatation est l'une des propriétés les plus importantes de la transformée de Fourier ; elle traduit le fait que la transformée de Fourier d'une fonction large est une fonction étroite et réciproquement. Cela se manifeste très souvent en Physique ; en optique par exemple, la tache de diffraction créée par un diaphragme circulaire est d'autant plus étendue que le rayon de l'orifice est petit.

Exercice 15.2 1. Etablir la propriété :

$$e^{i2\pi x \xi_0} f(x) \xrightarrow{FT} \hat{f}(\xi - \xi_0).$$

En déduire la transformée de Fourier de la fonction $x \mapsto \cos(2\pi\beta x)e^{-\alpha|x|}$ ($\alpha, \beta > 0$) et la représenter sur un schéma.

2. Etablir la propriété (de dilatation) :

$$f(x/\lambda) \xrightarrow{FT} |\lambda| \hat{f}(\lambda\xi)$$

Une autre convention de définition de la transformation de Fourier est $\mathcal{F}[f(x)](\xi) = \int_{\mathbb{R}} dx f(x) e^{-i\xi x}$. Utilisez la propriété précédente pour calculer $\mathcal{F}[\Pi(x)](\xi)$ avec cette convention.

Le théorème fondamental de la transformée de Fourier est le théorème suivant dit *lemme de Riemann-Lebesgue*

Théorème 15.2.1 La transformée de Fourier d'une fonction sommable est bornée, continue et tend vers 0 à l'infini.

On a déjà montré plus haut que la transformée de Fourier était bornée. La continuité de \hat{f} résulte des égalités

$$\lim_{\eta \rightarrow 0} \hat{f}(\xi + \eta) = \int_{\mathbb{R}} \lim_{\eta \rightarrow 0} [f(x)e^{-i(2\pi\xi + \eta)x}] dx = \int_{\mathbb{R}} f(x)e^{-i2\pi\xi x} dx = \hat{f}(\xi),$$

où la permutation de la limite et du signe intégral est justifiée.⁴

Pour montrer que \hat{f} tend vers 0 à l'infini, on approche f par une fonction en escalier g qui satisfait $\int_{\mathbb{R}} |f(x) - g(x)| dx < \epsilon$, ϵ étant un nombre aussi petit qu'on le désire. g étant une fonction en escalier est une combinaison linéaire de fonctions portes translatées, on a donc : $|\hat{g}(\xi)| < \epsilon$, pour ξ assez grand. La linéarité de l'intégrale jointe à l'inégalité triangulaire permettent d'écrire :

$$|\hat{f}(\xi)| = |(\widehat{f-g})(\xi) + \hat{g}(\xi)| \leq |(\widehat{f-g})(\xi)| + |\hat{g}(\xi)| \leq \int_{\mathbb{R}} |f(x) - g(x)| dx + |\hat{g}(\xi)|,$$

qui peut donc être rendu aussi petit qu'on le désire pour ξ assez grand.

On notera en particulier que la transformation de Fourier est une opération "régularisante" puisqu'elle rend continue une fonction qui ne l'est pas forcément (le cas de la porte, par exemple).

15.3 Dérivation et Inversion

1. Dérivation

Les deux théorèmes de cette section traduisent les deux idées réciproques suivantes : plus une fonction est dérivable, plus sa transformée de Fourier décroît vite à l'infini, plus une fonction décroît vite à l'infini, plus sa transformée de Fourier est dérivable.

Commençons par un résultat sur la transformée de Fourier de la dérivée d'une fonction :

Théorème 15.3.1 Soit $f \in L^1(\mathbb{R})$, de classe C^1 dont la dérivée est sommable, alors

$$f'(x) \xrightarrow{FT} 2i\pi\xi \hat{f}(\xi)$$

et donc,

$$|\xi| \hat{f}(\xi) \rightarrow 0 \text{ pour } \xi \rightarrow \infty.$$

4. une application du théorème de la convergence dominée ..

Pour retenir ce résultat, il suffit d'intégrer par parties

$$\mathcal{F}(f'(x)) = \int_{\mathbb{R}} f'(x)e^{-i2\pi\xi x} dx = [f(x)e^{-i2\pi\xi x}]_{-\infty}^{+\infty} + i2\pi\xi \int_{\mathbb{R}} f(x)e^{-i2\pi\xi x} dx$$

qui donne donc le résultat puisque $f(\pm\infty) = 0^5$. Le comportement à l'infini est une conséquence du lemme de Riemann-Lebesgue appliqué à \hat{f}' .

Pour la dérivée de la transformée de Fourier d'une fonction, on a :

Théorème 15.3.2 Soit $f \in L^1(\mathbb{R})$ telle que $x \rightarrow xf(x)$ soit sommable, alors

$$-2i\pi x f(x) \xrightarrow{FT} \frac{d\hat{f}}{d\xi}(\xi)$$

et donc f est de classe C^1 .

Ce résultat découle des égalités

$$\frac{d\hat{f}}{d\xi}(\xi) = \int_{\mathbb{R}} \frac{d}{d\xi} (f(x)e^{-i2\pi\xi x}) dx = \int_{\mathbb{R}} (-i2\pi x f(x))e^{-i2\pi\xi x} dx = \mathcal{F}(-i2\pi x f(x)),$$

où là encore, il faudrait justifier de l'inversion de la dérivée et du signe intégral.

Ces 2 résultats se généralisent aisément aux dérivées d'ordre plus élevées.

Exercice 15.3 En s'inspirant des théorèmes 2.3.1 et 2.3.2, établir la formule donnant la dérivée d'ordre n de la transformée de Fourier d'une fonction, et celle donnant la transformée de Fourier de la dérivée d'ordre n d'une fonction.

On s'attachera à préciser les conditions requises pour l'application de ces théorèmes.

Exercice 15.4 (a) Montrer que la fonction $\xi \mapsto \hat{f}(\xi) \equiv \mathcal{F}(e^{-x^2})(\xi)$ vérifie l'équation différentielle :

$$\hat{f}'(\xi) + 2\pi^2\xi\hat{f}(\xi) = 0.$$

(b) Calculer $\hat{f}(0)$ puis déterminer la solution de l'équation différentielle.

(c) Utiliser la propriété de dilatation pour établir le résultat :

$$e^{-ax^2} \xrightarrow{FT} \sqrt{\frac{\pi}{a}} e^{-\pi^2\xi^2/a} \quad \text{pour } a > 0.$$

2. Inversion

Nous avons vu comment \hat{f} s'exprime en fonction de f à la seule condition que f soit sommable. On se pose maintenant la question de savoir s'il est possible de définir f en fonction de \hat{f} : c'est le problème de l'inversion de la transformation de Fourier. On a le résultat suivant :

Théorème 15.3.3 Soit $f \in L^1(\mathbb{R})$ telle que $\hat{f} \in L^1(\mathbb{R})$, on a alors

$$f(x) = \int_{\mathbb{R}} e^{+i2\pi x\xi} \hat{f}(\xi) d\xi$$

Ainsi, tant qu'on reste dans L^1 , l'inversion n'est possible que si f et \hat{f} sont toutes les deux sommables. On peut montrer que l'on a pas mieux concernant cette propriété d'inversion dans L^2 . Le recours aux distributions dites tempérées pour lesquelles on sait définir une transformation de Fourier lève en partie ces contraintes.

5. $f(x) = f(0) + \int_0^x f'(y) dy$, donc $f(\pm\infty)$ existe puisque f' est sommable. De plus $f(\pm\infty)$ doit être nul, faute de quoi f ne serait pas sommable.

15.4 Convolution

Dans cette section on définit le produit de convolution de deux fonctions sommables, et l'on montre que la transformation de Fourier échange le produit de convolution et la multiplication. Cette dernière propriété est d'une importance fondamentale dans les applications comme nous le verrons dans la suite.

Définition 15.4.1 Soient f et g des fonctions sommables, le produit de convolution, noté $f \star g$, est défini par

$$(f \star g)(x) = \int_{\mathbb{R}} f(x-y)g(y) dy$$

Pour bien saisir cette définition, on pourra essayer de convoluer 2 portes identiques : le résultat est une fonction ayant la forme d'un triangle.

Comme premier exemple, considérons le potentiel électrostatique créé en un point \mathbf{r} par une distribution de charge volumique $\rho(\mathbf{r})$ contenue dans un volume \mathcal{V} , on a :

$$V(r) = \int_{\mathcal{V}} \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' = \left(\frac{1}{r} \star \rho_{\mathcal{V}}\right)(\mathbf{r}),$$

où $\rho_{\mathcal{V}}$ est la fonction égale à ρ sur \mathcal{V} et nulle ailleurs. Cette formule peut être interprétée comme la réponse du milieu étudié à la perturbation électrique localisée $\rho(\mathbf{r})$. La fonction a dans ce cas est le potentiel électrique créé par une charge ponctuelle.

On pourra vérifier aisément que le produit de convolution des fonctions sommables possède les propriétés de commutativité, associativité et distributivité par rapport à l'addition, soit :

$$\begin{aligned} f \star g &= g \star f, \\ f \star (g \star h) &= (f \star g) \star h, \\ f \star (ag + bh) &= a(f \star g) + b(f \star h) \quad \text{pour } a, b \in \mathbb{C}. \end{aligned}$$

Le théorème le plus utile pour les applications est le théorème suivant qui montre que la transformation de Fourier remplace un produit de convolution par une multiplication.

Théorème 15.4.1 Soient f et g sommables, alors $f \star g$ est sommable et,

$$\widehat{f \star g} = \hat{f}\hat{g},$$

qui se démontre en écrivant les définitions de chaque membre. Si f et g sont deux fonctions sommables, leur produit n'est en général pas sommable, ce qui ne permet pas toujours de définir $\mathcal{F}(fg)$. Par contre, on verra plus loin qu'on peut démontrer une formule réciproque dans L^2 .

Le produit de convolution intervient fréquemment dans l'analyse des systèmes linéaires et homogènes. Une situation assez générique consiste en effet à exciter un système par une fonction extérieure, que nous noterons $x \mapsto e(x)$ et à mesurer la réponse $r(x)$ du système. Dans le cadre de notre étude, on supposera e et r sommables. Le système physique étudié est dit *linéaire et homogène* si l'on peut définir un opérateur $\mathcal{L} : e \mapsto r = \mathcal{L}(e)$, qui satisfait aux deux propriétés suivantes :

- $\mathcal{L}(\alpha e_1(x) + \beta e_2(x)) = \alpha \mathcal{L}(e_1(x)) + \beta \mathcal{L}(e_2(x))$ avec $\alpha, \beta \in \mathbb{R}$
- $\mathcal{L}(e(x - x_0)) = r(x - x_0)$

La convolution permet de construire aisément un opérateur linéaire et homogène. On montre en effet, que si la réponse du système s'écrit sous la forme $r = \mathcal{L}(e) = a \star e$, où a une fonction sommable, l'opérateur \mathcal{L} est bien linéaire et homogène. Toute l'information sur le système est alors contenue dans la fonction a qu'il importe de déterminer.

Le théorème sur la transformée de Fourier du produit de convolution permet de calculer la réponse $r = a \star e$ d'un système linéaire et homogène. En effet la transformée de Fourier donne $\hat{r} = \hat{a}\hat{e}$, c'est-à-dire que chaque composante $\hat{e}(\xi)$ du signal d'entrée se trouve multipliée (amplifiée ou atténuée) par un coefficient $\hat{a}(\xi)$. Le

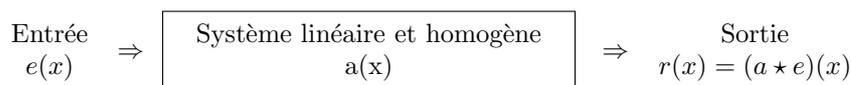


FIGURE 15.3 – Effet de filtre d'un système linéaire et homogène.

système fonctionne donc comme *un filtre fréquentiel*, la fonction a est appelée la fonction de transfert du filtre. Comme exemples concrets de systèmes physiques homogènes et linéaires, on peut penser aux circuits électriques constitués d'éléments passifs (résistance, bobine et capacité) qui, correctement combinés, ont effectivement des fonctions de filtres.

Exercice 15.5 On considère les gaussiennes de "largeur" $\ell_a = a^{-1/2}$ définies par :

$$f_a(x) = e^{-ax^2} \quad a > 0.$$

1. Après avoir rappelé l'expression de \hat{f}_a , montrer que l'on peut utiliser la formule d'inversion pour retrouver f_a à partir de \hat{f}_a .
2. a et b étant 2 réels positifs, calculer $f_a \star f_b$, et en déduire la relation :

$$f_a \star f_b = \sqrt{\frac{\pi}{a+b}} f_{\frac{ab}{a+b}}.$$

3. Etudier le cas où $\ell_a \gg \ell_b$ et en déduire l'approximation :

$$f_a \star f_b \approx \sqrt{\frac{\pi}{b}} f_a \quad \text{pour } \ell_a \gg \ell_b,$$

c'est-à-dire que la convolution d'une fonction large et d'une fonction étroite est une fonction large.

15.5 Transformée de Fourier dans $L^2(\mathbb{R})$

On a déjà signalé, aussi bien pour des raisons physiques que mathématiques, que le cadre des fonctions sommables est beaucoup trop étroit pour un usage commode des transformées de Fourier. Comme il n'existe pas de relation d'inclusion⁶ entre $L^1(\mathbb{R})$ et $L^2(\mathbb{R})$ on peut, par exemple, être amené à rencontrer des fonctions qui appartiennent à L^2 sans appartenir à L^1 . Dans ce cas, il n'est plus possible de définir la transformée de Fourier par une formule intégrale, puisque cette dernière expression n'a aucun sens. Un autre cas problématique est celui des fonctions sommables dont la transformée de Fourier n'est pas sommable mais de carré sommable. Dans ce cas, on ne peut utiliser la formule d'inversion, ce qui est très limitant dans la pratique.

Au début de ce siècle, Plancherel a pu étendre les résultats connus sur la transformation de Fourier des fonctions sommables aux fonctions de carré sommable, plus quelques autres que nous mentionnons dans la suite.

Sans rentrer dans le détail, contentons nous de dire que la transformée de Fourier-Plancherel d'une fonction de L^2 est définie comme une limite de transformées de Fourier de fonctions appartenant à $L^1 \cap L^2$.

On a signalé dans une section précédente, que l'on ne pouvait pas en général définir \widehat{fg} pour f et g dans L^1 . Mais dans L^2 , on a

Théorème 15.5.1 Soient f et g deux fonctions de $L^2(\mathbb{R})$. Alors :

$$\widehat{fg} = \hat{f} \star \hat{g}$$

A titre d'exercice, vous pouvez essayer de vérifier ce théorème dans le cas particulier où on peut utiliser la représentation intégrale.

Le résultat le plus important est la *formule de Parseval-Plancherel* :

6. Les exemples standards sont $x \mapsto e^{-x^2}/\sqrt{x} \in L^1 \not\subset L^2$, $x \mapsto e^{-x^2} \in L^1 \cap L^2$, mais $x \mapsto (x^2 + 1)^{-1/2} \in L^2 \not\subset L^1$.

Théorème 15.5.2 Soient f et g deux fonctions de $L^2(\mathbb{R})$. Alors :

$$\int_{\mathbb{R}} f(x)\overline{g(x)} dx = \int_{\mathbb{R}} \hat{f}(\xi)\overline{\hat{g}(\xi)} d\xi$$

En effet $\int_{\mathbb{R}} f(x)\overline{g(x)} dx \equiv \widehat{f\overline{g}}(0) = \hat{f}(0) \star \widehat{\overline{g}}(0)$ par le théorème précédent. Le résultat est obtenu en observant que $\widehat{\overline{g}}(-\eta) = \overline{\hat{g}(\eta)}$ (à vérifier!).

Dans le cas particulier où $f = g$, la formule de Parseval-Plancherel s'écrit

$$\int |f(x)|^2 dx = \int |\hat{f}(\xi)|^2 d\xi.$$

Lorsque f représente l'amplitude d'une onde ou d'une déformation, ce résultat exprime que l'énergie totale associée peut être calculée de façon équivalente dans l'espace des positions ou dans l'espace des fréquences.

Exercice 15.6 Utiliser la formule de Parseval-Plancherel pour calculer les intégrales suivantes :

$$\int_{\mathbb{R}} \left(\frac{\sin x}{x}\right)^n dx \quad \text{pour } n = 2, 3, 4.$$

Exercice 15.7 1. Calculer la transformée de Fourier de la fonction

$$x \rightarrow \frac{1}{1+x^2},$$

et en déduire la transformée de Fourier de la fonction

$$x \rightarrow \frac{x}{1+x^2}.$$

2. Après avoir étudié la parité des deux fonctions précédentes, calculer :

$$\int_0^{+\infty} \frac{\cos mx}{1+x^2} dx \quad \text{et} \quad \int_0^{+\infty} \frac{x \sin mx}{1+x^2} dx \quad \text{pour } m > 0.$$

3. Utiliser la formule de Parseval-Plancherel pour calculer les intégrales :

$$\int_{-\infty}^{+\infty} \frac{1}{(1+x^2)^2} dx \quad \text{et} \quad \int_0^{+\infty} \frac{x^2}{(1+x^2)^2} dx$$

15.6 Transformées de Fourier dans $\mathcal{S}(\mathbb{R})$

On introduit un nouvel espace de fonctions, que l'on appelle **l'espace de Schwartz**, et que l'on note $\mathcal{S}(\mathbb{R})$. Cet espace désigne *l'espace des fonctions indéfiniment dérivables à décroissance rapide*, c'est-à-dire des fonctions φ indéfiniment dérivables sur \mathbb{R} qui vérifient :

$$\forall n, p \in \mathbb{N}, \quad |x^n \varphi^{(p)}(x)| \leq M_{p,n}, \quad \text{avec } M_{p,n} \text{ réels.}$$

Du fait de la contrainte très forte de décroissance plus rapide que toute fonction en x^{-n} , on notera que les fonctions de $\mathcal{S}(\mathbb{R})$ sont, d'un point de vue pratique, très semblables aux fonctions indéfiniment dérivables à support bornés.

Donnons quelques propriétés de ces espaces.

Théorème 15.6.1 Les fonctions de \mathcal{S} ainsi que toutes leurs dérivées sont bornées et intégrables sur \mathbb{R} .

Soit $\varphi \in \mathcal{S}$. Par définition, toutes ses dérivées sont également dans \mathcal{S} . Donc, $\forall n, p, |x^n \varphi^{(p)}(x)| \leq M_{p,n}$, et en particulier que $|(1+x^2)\varphi^{(p)}(x)| \leq M_{p,0} + M_{p,2}$. Toutes les dérivées $\varphi^{(p)}$ sont donc bornées et majorées par des fonctions intégrables : elles sont donc elles-mêmes intégrables.

Ainsi, puisque $\mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R})$, toutes les définitions, propriétés et théorèmes vus pour les fonctions sommables s'étendent aux fonctions de l'espace de Schwartz. En particulier la transformée de Fourier d'une fonction $\varphi \in \mathcal{S}(\mathbb{R})$ est définie, pour $q \in \mathbb{R}$, par l'intégrale :

$$\hat{\varphi}(q) = [F\varphi](q) \equiv \int_{\mathbb{R}} \varphi(x) e^{-i2\pi qx} dx$$

Mais il y a plus : la formule d'inversion est valable pour toute fonction de \mathcal{S} (ce qui n'était assuré dans L^1 , que si la TF était elle-même dans L^1), et la formule de Parseval-Plancherel (valable dans L^2 mais pas dans L^1), s'étend aux fonctions de \mathcal{S} . Le théorème qui suit résume la situation :

Théorème 15.6.2 1. Soit $\varphi \in \mathcal{S}$. Sa transformée de Fourier $\hat{\varphi}$ est elle-même dans \mathcal{S} .

2. Toute fonction $\varphi \in \mathcal{S}$ est la TF inverse d'un élément $\hat{\varphi} \in \mathcal{S}$, telle que :

$$\varphi(x) = [F^{-1}\hat{\varphi}](x) \equiv \int_{\mathbb{R}} \hat{\varphi}(q) e^{+i2\pi qx} dq$$

3. Pour φ et ψ dans \mathcal{S} :

$$\int_{\mathbb{R}} \varphi(x) \overline{\psi(x)} dx = \int_{\mathbb{R}} \hat{\varphi}(q) \overline{\hat{\psi}(q)} dq,$$

et en particulier, pour $\varphi \in \mathcal{S}$:

$$\int_{\mathbb{R}} |\varphi(x)|^2 dx = \int_{\mathbb{R}} |\hat{\varphi}(q)|^2 dq.$$

La dernière égalité montre que la TF conserve la norme dans $\mathcal{S}(\mathbb{R})$.

15.7 Transformées à plusieurs variables.

$\xi \cdot x$ désignant le produit scalaire dans \mathbb{R}^n , soit $\xi \cdot x = x_1 \xi_1 + \dots + x_n \xi_n$ si $x = (x_1, \dots, x_n)$ et $\xi = (\xi_1, \dots, \xi_n)$, on a la définition.

Définition 15.7.1 Soit $f \in L^1(\mathbb{R}^n)$, on appelle transformée de Fourier de f , la fonction, à valeurs complexes, notée \hat{f} ou $\mathcal{F}(f)$, définie pour tout $\xi \in \mathbb{R}^n$ par :

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-i2\pi \xi \cdot x} f(x) dx$$

On dispose dans \mathbb{R}^n des mêmes résultats que ceux obtenus dans les sections précédentes pour les fonctions à une variable. On discute maintenant 2 cas importants pour lesquels on peut obtenir des formules explicites.

1. Fonctions séparables

Un cas particulièrement simple est le cas des fonctions séparables, qui s'écrivent comme un produit de fonctions à une variable : $f(x) \equiv \prod_{i=1}^n f_i(x_i)$. Si chaque fonction f_i est sommable dans \mathbb{R} , on a simplement :

$$\hat{f}(\xi) = \prod_{i=1}^n \hat{f}_i(x_i)$$

puisque $e^{-i2\pi \xi \cdot x} = \prod_{i=1}^n e^{-i2\pi \xi_i x_i}$.

2. Fonctions radiales

Les fonctions radiales, d'un usage très fréquent en Physique, sont des fonctions qui ne dépendent que de la distance à l'origine. On a donc dans ce cas :

$$f(x) = F(r) \quad \text{où} \quad r \equiv (x_1^2 + \dots + x_n^2)^{1/2}$$

Montrons d'abord que *si f est radiale alors \hat{f} aussi*. Si f est radiale cela veut dire que n'importe quelle rotation \mathcal{R} centrée sur l'origine est telle que $f(\mathcal{R}x) = f(x)$ pour tout x . Montrons alors que $\hat{f}(\mathcal{R}\xi) = \hat{f}(\xi)$.

$$\hat{f}(\mathcal{R}\xi) = \int_{\mathbb{R}^n} e^{-i2\pi\mathcal{R}\xi \cdot x} f(x) dx$$

Or, $\mathcal{R}\xi \cdot x = \xi \cdot \mathcal{R}^{-1}x$, de sorte qu'en effectuant le changement de variable $y = \mathcal{R}^{-1}x$, de jacobien égal à 1, on a

$$\hat{f}(\mathcal{R}\xi) = \int_{\mathbb{R}^n} e^{-i2\pi\xi \cdot y} f(\mathcal{R}y) dy = \hat{f}(\xi),$$

puisque $f(\mathcal{R}y) = f(y)$ par hypothèse

Posons $\rho \equiv (\xi_1^2 + \dots + \xi_n^2)^{1/2}$, il existe donc une fonction G telle que $\hat{f}(\xi) = G(\rho)$. On va montrer que G s'exprime en fonction de F . On peut toujours choisir un système d'axes orthonormés, de telle sorte que $\xi = (0, \dots, \xi_n)$ (ξ étant donné, cela revient à placer l'axe Ox_n dans la direction du vecteur ξ), alors, puisque $\rho = \xi_n$ et $f(x) = F(r)$:

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-i2\pi\xi_n x_n} f(x) dx = \int_{\mathbb{R}^n} e^{-i2\pi\rho x_n} F(r) dx = G(\rho)$$

La suite du calcul dépend de la dimension (le faire!) :

$$\begin{aligned} G(\rho) &= 2 \int_0^\infty \cos(2\pi\rho r) F(r) dr \quad \text{pour } n = 1, \\ G(\rho) &= 2\pi \int_0^\infty r J_0(2\pi\rho r) F(r) dr \quad \text{pour } n = 2, \\ G(\rho) &= \frac{2}{\rho} \int_0^\infty r \sin(2\pi\rho r) F(r) dr \quad \text{pour } n = 3, \end{aligned}$$

où on a introduit la fonction de Bessel $J_0(u)$ définie par la relation $J_0(u) \equiv \frac{1}{2\pi} \int_0^{2\pi} e^{-iu \cos \theta} d\theta$.

Exercice 15.8 Soit (x, y, z) un élément de \mathbb{R}^3 ; on pose $r = (x^2 + y^2 + z^2)^{1/2}$.

Calculer la transformée de Fourier de la fonction :

$$(x, y, z) \rightarrow \frac{e^{-\alpha r}}{r}, \quad \alpha > 0.$$

15.8 Applications aux EDO à coefficients constants

Dans certaines circonstances que nous allons préciser plus loin, la transformation de Fourier transforme une équation différentielle à coefficients constants en une équation algébrique. On dit parfois que la transformation de Fourier "diagonalise" les opérateurs à coefficients constants. Lorsque cette méthode est applicable, la transformée de Fourier permet de déterminer une solution particulière d'une équation différentielle ayant un "bon" comportement physique.

Considérons une équation différentielle linéaire à coefficients constants d'ordre m . Sa forme générale s'écrit

$$f^{(m)}(x) + \sum_{i=1}^{m-1} a_i f^{(i)}(x) + a_0 f(x) = h(x) \quad (15.3)$$

où h est une fonction donnée et f l'inconnue.

A l'équation sans second membre, on associe l'équation caractéristique :

$$P(\lambda) = \lambda^m + \sum_{i=1}^{m-1} a_i \lambda^i + a_0 = 0$$

Prenons la transformée de Fourier de l'équation différentielle ; alors, s'il est légitime d'appliquer le théorème sur la transformée d'une dérivée, on obtient $P(i2\pi\xi)\hat{f}(\xi) = \hat{h}(\xi)$. Si le polynôme caractéristique n'a aucune racine imaginaire pure, on peut procéder à la division et l'on obtient :

$$\hat{f}(\xi) = \frac{\hat{h}(\xi)}{P(i2\pi\xi)}$$

Si maintenant \hat{f} est elle-même sommable, on peut utiliser la formule d'inversion et l'on obtient comme solution :

$$f(x) = \int_{\mathbb{R}} e^{+i2\pi x\xi} \frac{\hat{h}(\xi)}{P(i2\pi\xi)} d\xi$$

Exemple

Un certain nombre de problèmes de Physique, dans des contextes variés, admettent une modélisation qui est celle de l'oscillateur harmonique amorti. La dynamique du système est régie par l'équation différentielle :

$$u''(t) + \gamma u'(t) + \omega_0^2 u(t) = f(t)$$

où ω_0 , réel positif, est la fréquence propre de l'oscillateur, $\gamma \in \mathbb{R}^+$, un coefficient d'amortissement et f une force extérieure convenablement normalisée. Pour des raisons de commodité évidente, on posera $\xi \equiv \omega/2\pi$ la variable conjuguée du temps t dans la suite du problème. Le problème est de trouver u pour f donnée.

Le polynôme caractéristique $P(i\omega) = -\omega^2 + i\gamma\omega + \omega_0^2$ admet pour racines :

$$\begin{aligned} \omega_{\pm} &= \frac{i\gamma}{2} \mp \omega_1, \\ \omega_1 &= \sqrt{\omega_0^2 - \frac{\gamma^2}{4}} \end{aligned}$$

ω_{\pm} ne peut être réel sauf si $\gamma = 0$ (dans ce cas $\omega = \omega_0$ annule $P(i\omega)$).

Plus précisément,

- si $\omega_0 > \gamma/2$, ω_1 est réel, ω_{\pm} a une partie réelle et une partie imaginaire, et on s'attend à ce que la solution oscille avec la fréquence ω_1 et s'atténue avec le temps caractéristique $\tau = 2/\gamma$.
- si $\omega_0 < \gamma/2$, ω_1 et ω_{\pm} sont imaginaires purs, le système est dit sur-amorti, et il apparaît 2 temps caractéristiques d'amortissement :

$$\frac{1}{\tau_{\pm}} = \frac{\gamma}{2} \left(1 \mp \sqrt{1 - \left(\frac{\omega_0}{\gamma/2}\right)^2} \right)$$

il est facile de voir qu'en cas de fort amortissement : $\gamma/2 \gg \omega_0$, $\tau_+ \approx \gamma/\omega_0^2$ et $\tau_- \approx 1/\gamma$; au temps longs devant τ_- , seul compte τ_+ ce qui revient à négliger le terme inertiel dans l'équation différentielle et à considérer

$$\gamma u'(t) + \omega_0^2 u(t) = f(t),$$

équation effectivement utilisée dès lors que les effets visqueux sont prépondérants.

Pour u et f dans $L^1(\mathbb{R})$, et $\gamma \neq 0$, on peut donc écrire :

$$\hat{u}(\omega) = \frac{\hat{f}(\omega)}{-\omega^2 + i\gamma\omega + \omega_0^2}$$

La fonction de réponse χ du système est définie par le rapport “effet/cause”, soit ici :

$$\chi(\omega) = \frac{\hat{u}(\omega)}{\hat{f}(\omega)} = \frac{1}{P(i\omega)} = \frac{1}{-\omega^2 + i\gamma\omega + \omega_0^2},$$

comme $\omega_1^2 = \omega_0^2 - \gamma^2/4$, on a

$$-\omega^2 + i\gamma\omega + \omega_0^2 = -\omega^2 + i\gamma\omega + \omega_1^2 + \frac{\gamma^2}{4} = \omega_1^2 - (\omega - i\gamma/2)^2$$

donc,

$$\chi(\omega) = \frac{1}{2\omega_1} \left[\frac{1}{\omega_1 - \omega + i\gamma/2} + \frac{1}{\omega_1 + \omega - i\gamma/2} \right]$$

Pour $\omega_0 > \gamma/2$ (ω_1 réel), on a encore :

$$\chi(\omega) = \frac{i}{2\omega_1} \left[\frac{1}{i(\omega + \omega_1) + \gamma/2} - \frac{1}{i(\omega - \omega_1) + \gamma/2} \right] = \frac{i}{2\omega_1} \mathcal{F} \left([e^{-i\omega_1 t} - e^{+i\omega_1 t}] H(t) e^{-\gamma t/2} \right),$$

et donc

$$\begin{aligned} \chi(t) &= \frac{1}{\omega_1} \sin(\omega_1 t) H(t) e^{-\gamma t/2}, \\ u(t) &= \chi(t) \star f(t) \end{aligned}$$

qui donne donc la solution pour toute force extérieure sommable, dans le cas ω_1 réel. On trouve bien, comme annoncé, une réponse du système qui oscille à la fréquence ω_1 tout en s'amortissant avec le temps caractéristique $\tau = 2/\gamma$.

Lorsque $\omega_0 < \gamma/2$, on a $\omega_1 = i\sqrt{\gamma^2/4 - \omega_0^2}$, et l'on peut écrire :

$$\begin{aligned} \chi(\omega) &= \frac{1}{2\omega'_1} \left[\frac{1}{i\omega + \tau_+^{-1}} - \frac{1}{i\omega + \tau_-^{-1}} \right] = \frac{1}{2\omega'_1} \mathcal{F} \left(H(t) [e^{-t/\tau_+} - e^{-t/\tau_-}] \right), \\ \omega'_1 &= \sqrt{\gamma^2/4 - \omega_0^2} \end{aligned}$$

Compte tenu des relations entre τ_+ et τ_- , on peut aussi écrire :

$$\begin{aligned} \chi(t) &= \frac{1}{\omega'_1} \sinh(\omega'_1 t) H(t) e^{-\gamma t/2}, \\ u(t) &= \chi(t) \star f(t), \end{aligned}$$

solution qui se comporte bien comme $e^{-\omega_0^2 t/\gamma}$ au temps longs.

Exercice 15.9 On considère le circuit RC série. $u(t)$ désigne la tension d'entrée, et on mesure la tension de sortie $v(t)$ aux bornes de la capacité C . On note $\tau = RC$ la constante de temps du circuit.

1. Montrer que les tensions satisfont l'équation différentielle

$$v'(t) + \frac{v(t)}{\tau} = \frac{u(t)}{\tau}$$

2. On applique une tension d'entrée dont on suppose seulement qu'elle correspond à une fonction sommable. Montrer que

$$v(t) = (a \star u)(t)$$

Donner l'expression de la fonction a .

3. On donne la forme suivante au signal d'entrée :

$$u(t) = u_0 H(t) e^{-\alpha t}$$

Quelle est la forme du signal de sortie :

- (a) pour $\alpha \neq \tau^{-1}$?
 - (b) à la résonance : $\alpha = \tau^{-1}$?
4. Pourrait-on traiter le problème par la même méthode si le signal d'entrée avait la forme $u(t) = u_0 H(t)$ (signal continu) ?

15.9 Conclusion

On a déjà fait remarquer qu'une définition de la transformation de Fourier limitée aux fonctions sommables (ou même de carré sommables) était beaucoup trop contraignante dans les applications. Ainsi, des fonctions aussi courantes que les fonctions constantes ou la fonction de Heaviside, n'admettent pas de transformées de Fourier *au sens des fonctions*, puisqu'elles n'appartiennent ni à $L^1(\mathbb{R})$, ni à $L^2(\mathbb{R})$. En s'appuyant sur l'espace de Schwartz des fonctions indéfiniment dérivables à décroissance rapide introduit plus haut, il est possible de définir beaucoup plus de TF (y compris celles de fonctions constantes) dans le cadre de la théorie des fonctions généralisées ou *distributions*.

Table des matières

I	Variables complexes	5
1	Rappels sur les nombres complexes	7
1.1	Motivations historiques	7
1.2	Le corps des nombres complexes	8
1.3	Exponentielle complexe	9
1.4	Représentation des nombres complexes	10
1.4.1	Représentation cartésienne	10
1.4.2	Représentation géométrique	10
2	Fonctions d'une variable complexe	13
2.1	Fonctions d'une variable complexe	13
2.2	Dérivation des fonctions d'une variable complexe	14
2.3	Intégration des fonctions d'une variable complexe	15
2.4	Développement en série	17
2.5	Méthode des résidus	18
3	Transformation de Laplace au sens des fonctions	21
3.1	Introduction	21
3.2	Définition de la transformée de Laplace	22
3.3	Holomorphie	23
3.4	Propriétés de la transformée de Laplace	24
3.5	Comportements asymptotiques	26
3.6	Inversion de la transformée de Laplace	27
3.7	Remarque	30
3.8	Exercices	30
II	Equations différentielles	35
4	Introduction	37
4.1	Mécanique.	37

4.2	Dynamique des Populations.	38
4.3	Equations aux Dérivées Partielles.	39
5	Rappels sur les équations différentielles	41
5.1	Terminologie	41
5.2	Quelques conséquences de la linéarité	42
5.3	Deux solutions explicites importantes	42
5.3.1	Equation différentielle linéaire du 1er ordre	43
5.3.2	Equations différentielles du second ordre à coefficients constants	44
5.4	Système d'équations différentielles du 1er ordre	46
5.5	Equation intégrale	48
5.6	Théorème d'existence et d'unicité	49
6	Systèmes Différentiels Linéaires	51
6.1	Exponentielle de matrice	51
6.2	Propagateur	52
6.3	Calcul pratique du propagateur	53
6.4	Equations différentielles en dimension 2	54
6.4.1	Diagonalisation des matrices 2×2	54
6.4.2	Forme explicite du propagateur en dimension 2	55
7	Analyse qualitative des équations différentielles	59
7.1	Exemple	59
7.2	Classification des points fixes (1 dimension)	60
8	Stabilité des systèmes différentiels	63
8.1	Stabilité des systèmes différentiels linéaires	63
8.1.1	Portraits de phase en dimension 2	63
8.2	Stabilité des systèmes différentiels non linéaires	65
8.2.1	Linéarisation	65
8.2.2	Cycles limites	67
III	Analyse dans \mathbb{R}^n	69
9	Différentiel et Intégrer	71
9.1	Dérivées et différentielles	71
9.2	Circulation et flux	73
9.3	Opérateurs différentiels	76
9.4	Formule de Stokes	78
9.5	Singularités	79

9.6	Différentielles exactes	81
9.7	Théorèmes de Helmholtz	82
9.8	Théorème de Leibnitz	83
10	Approximer et calculer	85
10.1	Formule de Taylor	85
10.2	Approximation de Stirling	86
10.3	Distribution de Dirac	87
10.4	Intégrales Gaussiennes	89
11	Minimiser	91
11.1	Extrémalisation sans contraintes	91
11.2	Multiplicateurs de Lagrange	92
11.3	Transformation de Legendre	94
IV	Algèbre linéaire	97
12	Eléments d'algèbre linéaire	99
12.1	Déterminants	99
12.2	Matrices	100
12.2.1	Opérations	100
12.2.2	Inversion	101
12.3	Système d'équations linéaires	102
12.4	Diagonalisation	103
13	Scalars, vecteurs, tenseurs	107
13.1	Introduction	107
13.2	Notations et changement de base	108
13.3	Scalars, vecteurs et tenseurs	109
13.4	Produits et contractions	110
13.5	Base réciproque	112
V	Analyse de Fourier	115
14	Séries de Fourier	117
14.1	Des problèmes historiques	117
14.2	Définition des séries de Fourier	118
14.3	Convergence des séries de Fourier	119
14.4	Dérivation et Intégration terme à terme	119
14.5	Phénomène de Gibbs	120

14.6 Exercices	121
15 Transformation de Fourier au sens des fonctions	123
15.1 Motivations	123
15.2 Transformée de Fourier dans $L^1(\mathbb{R})$	125
15.3 Dérivation et Inversion	127
15.4 Convolution	129
15.5 Transformée de Fourier dans $L^2(\mathbb{R})$	130
15.6 Transformées de Fourier dans $\mathcal{S}(\mathbb{R})$	131
15.7 Transformées à plusieurs variables.	132
15.8 Applications aux EDO à coefficients constants	133
15.9 Conclusion	136