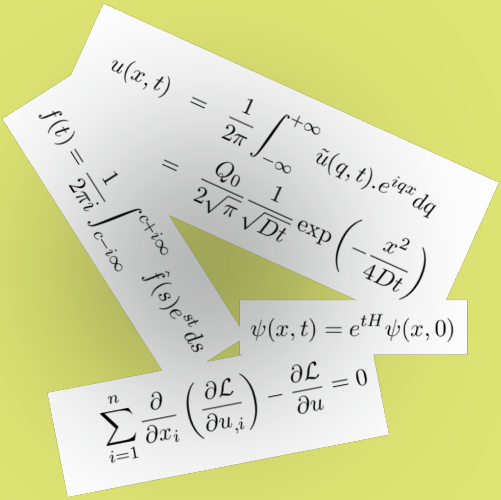
Mathématiques pour la Physique

Bahram Houchmandzadeh

UNIVERSITÉ Grenoble Alpes

Mathématiques pour la Physique.



Bahram Houchmandzadeh

Remerciements :
Je remercie sincérement Youssef Ben Miled et Mathias Legrand pour leur lecture attentive du manuscrit et leur très (très) nombreuses corrections et suggestions. Grace à leurs efforts, ce manuscrit a un aspect beaucoup plus présentable.

web	:	www-liphy.ujf-grenoble.fr/pagesperso/bahram/Math/math.htm
courriel	:	bahram.houchmandzadeh à univ-grenoble-alpes.fr
Première version	:	Septembre 2008
Version présente	:	January 12, 2018

Table des matières

1 Introduction

Durant les deux premières années à l'université, on apprend les bases essentielles des mathématiques : calcul différentiel et intégral, algèbre linéaire, équations différentielles linéaires, etc. L'objet de ce cours est d'utiliser ce corpus pour introduire les méthodes mathématiques dites supérieures utilisées pour résoudre les problèmes classiques de la physique.

Les mathématiques ne sont pas une collection de méthodes juxtaposées et sans relation : il existe des concepts extrêmement généraux qui nous permettent de porter le même regard sur des notions a priori disparates. Le concept qui reviendra tout au long de ce cours est celui de l'espace vectoriel. Ainsi, tourner un vecteur du plan d'un angle quelconque ou appliquer un opérateur intégrodifférentiel à une fonction sont fondamentalement la même chose ; de même que trouver les valeurs propres d'une matrice ou résoudre une équation à dérivée partielle linéaire. C'est bien pour cela que l'étudiant apprend un tel volume d'algèbre linéaire dans les cours de mathématiques élémentaires.

Le plan du cours est le suivant : Après une introduction (un rappel) des espaces vectoriels, nous verrons que les fonctions elles mêmes peuvent être considérées comme des points (des vecteurs) dans un grand espace des fonctions, et que nous pouvons définir des bases orthogonales dans cet espace presque comme on le fait dans l'espace tridimensionnel.

Le chapitre suivant est consacré aux séries de Fourier, le premier exemple pratique que nous verrons de bases dénombrables dans l'espace des fonctions sur un intervalle *fini*. Nous verrons entre autre comment cette base nous permet de résoudre les équations classique de la physique comme celle de diffusion de la chaleur ou des cordes vibrantes.

Nous avons souvent affaire à des fonctions définies sur des intervalles *infinis*. Les transformées de Fourier nous permettent de disposer de bases pour l'espace de ces fonctions. Comme souvent cependant, les infinis posent des problèmes particuliers et nous auront allors à définir les *distributions*, une généralisation des fonctions qui introduit en mathématique le concept de charge (ou force) ponctuelle si cher aux physiciens. Nous verrons alors le nombre incroyable de problèmes que ces nouvelles méthodes nous permettent d'aborder : de la résolution des équations différentielles à celle d'équations stochastiques (comme le mouvement brownien) en passant par la diffraction par les cristaux etc.

Le cousin germain des transformées de Fourier est la transformée de Laplace : nous verrons comment l'utiliser pour tous les problèmes où les conditions initiales sont importantes.

Finalement, un complément utile à tous ces outils est la méthode de Green (ou fonctions de Green) qui à nouveau a à voir avec la généralisation des charges ponctuelles : si on connaît l'effet d'une charge (ou d'une force ou ...) ponctuelle, on peut alors facilement calculer l'effet d'une distribution de charge (de force ...).

Nous allons revenir sur le concept général d'opérateur intégrodifférentiel. Une rotation ou une homothétie transforment de façon *linéaire* un vecteur dans un autre. Un opérateur différentiel linéaire comme $(\partial_t - D\partial_x^2)$ fait la même chose pour les fonctions, considérées comme des vecteurs d'un grand espace. Nous savons qu'étudier une application linéaire est toujours beaucoup plus simple dans sa base propre. La même chose est vrai pour les vecteurs et valeur propres des opérateurs. Les transformées de Fourier nous fournissaient une base très particulière bien adapté à certains problèmes de physique, nous verrons d'autres bases comme celle des polynômes orthogonaux et nous généraliserons le calcul des opérateurs.

Quelques chapitres sont consacrés aux notions plus avancées qui devront néanmoins être connues des étudiants à la fin de leur Master. Nous abordons le calcul des perturbations, outil indispensable dès que nous tentons la résolution de "vrai" problèmes, c'est à dire ceux qui s'écartent un peu des exemples classiques que nous savons résoudre. Par exemple, nous savons résoudre une équation différentielle d'une certaine forme, le calcul de perturbation nous permettra d'obtenir une solution approchée quand la forme change légèrement.

Un chapitre est consacré aux calculs des variations qui est une généralisation des problèmes d'extremum à l'espace des fonctions et des fonctionnelles qui y agissent. La plupart des problèmes de physique sont formulée dans ce langage ou gagne à être formulé dans ce langage.

Nous aborderons également la théorie des formes différentielles. Souvent ces objets sont enseignés dans le cadre de la théorie des tenseurs et vue comme des tenseurs alternés. Il est cependant beaucoup plus simple de les aborder directement et en donner une image géométrique, surtout que ce sont des objets très simple à manipuler et qui donnent de la cohérence aux divers opérateurs différentiels comme le gradient, rotationnel et divergence. Nous verrons comment certaines lois de la physique comme les équations de Maxwell acquiert une signification géométrique intuitive.

La théorie des tenseurs sera également dans un chapitre. Nous nous contenterons essentiellement des tenseurs dans un espace euclidien où il n'y a pas à faire de différence entre les vecteurs et les covecteurs.

Enfin un petit chapitre est consacré aux nombres. Nous les manipulons depuis si longtemps que nous avons oublié comment on les a construit. Ce chapitre tente de remédier à cet oubli.

Bon, suffisamment digressé, voyons du concret.

2 Éléments d'analyse fonctionnelle.

Les espaces vectoriels jouent un rôle unificateur fondamental en mathématiques. Peut-être cela rappelle au lecteur des souvenirs de matrices, de bases et de ses changements, ... Nous allons revoir tout cela de façon assez légère mais surtout appliquée à l'ensemble extrêmement vaste des *fonctions*. Nous allons voir que nous pouvons décrire les fonctions comme des *vecteurs* dans des espaces de dimensions infinies, en utilisant pratiquement les mêmes outils que pour des vecteurs à trois dimensions. Cela s'appelle *analyse fonctionnelle* et a été formalisé par Hilbert au début des années 1900. Le reste de ce cours s'appuie constamment sur les résultats de ce chapitre dont la lecture est indispensable.

2.1 Les espaces vectoriels.

Qu'est ce qu'un espace vectoriel ? Nous connaissons déjà certains ensemble célèbre comme celui des nombres réels $\mathbb{R}$ ou complexes $\mathbb{C}$. On les appellera dans la suite indifféremment l'ensemble des scalaires $\mathbb{S}$. Un espace vectoriel est un ensemble $\mathcal{E}$ où l'opération $+$ a un sens. Pas n'importe quel sens d'ailleurs, mais ce qu'on associe instinctivement [1] à cette opération : (i) si a et b appartiennent à notre ensemble, alors $a+b$ aussi ; (ii) $a+b=b+a$. En plus, multiplier un vecteur par un scalaire a un sens, qui plus est, lui aussi "très naturel" : si $s, s_1, s_2 \in \mathbb{S}$ et $a, b \in \mathcal{E}$, alors : (i) $sa \in \mathcal{E}$; (ii) $(s_1+s_2)a = s_1a + s_2a$; (iii) $s(a+b) = sa+sb$; (iv) $\mathcal{E}$ possède un élément *zéro*, qu'on notera $\mathbf{0}$ [2]tel que $a+\mathbf{0}=a,\ a, \mathbf{0} \in \mathcal{E}$. N'oublions pas que quand on parle de sa, on parle bien d'un *vecteur* et non d'un *scalaire*. L'ensemble des maisons d'une ville par exemple n'a pas vraiment une structure d'espace vectoriel. Par contre, l'ensemble des vecteurs dans un plan, l'ensemble des polynômes, ou l'ensemble des fonctions définies sur $[0,1]$ ont une structure d'espace vectoriel. L'intérêt majeur est que tout [3]ce que l'on peut affirmer pour l'un de ces ensembles (en rapport avec son caractère vectoriel) pourra être généralisé aux autres.

Bases d'espace vectoriel. Une base est l'ensemble de certains éléments de notre espace $\mathcal{E}$ qui nous permet de décrire tout les autres. Pour être plus rigoureux, supposons que $e_1, e_2, e_3, ..., e_i \in \mathcal{E}$, soit une base. Dans ce cas, pour n'importe quel élément a de

1. un instinct forgé par une douzaine d'année d'étude.
2. En caractère gras pour ne pas le confondre avec le 0 des scalaires.
3. Bon, il faut, de temps en temps, prendre des précautions.

$\mathcal{E}$, on peut trouver des scalaires (des chiffres donc) s_i tel que $a = \sum_i s_i e_i$. On dit que a est une combinaison linéaire des vecteurs e_i. Bien sûr, il faut prendre le minimum de e_i qui rende cette description faisable. Pour cela, il suffit d'exiger qu'aucun des e_i ne puisse être une combinaison linéaire des autres (on dit alors que ces vecteurs sont linéairement indépendant). Les scalaire s_i qu'on aura trouvé pour la description de a sont alors unique. On les appelle les $composantes$ du vecteur a dans la base $\{e\}$.

Le grand intérêt des bases est qu'elles nous permettent de manipuler les vecteurs comme des collections de chiffres. Pour les vecteurs dans le plan, nous ne sommes pas obligé de faire des dessins, nous pouvons les représenter par des doublets (x_1, x_2) si nous nous sommes fixés à l'avance deux vecteurs de références. A partir du moment où on peut représenter les objets par des chiffres, on peut pratiquement tout faire (hem).

Pour l'espace vectoriel des polynômes, les polynômes $1, X, X^2, ...$ constituent une base. Une autre serait $\{1, (1-X), (1-X)^2, ...\}$. Bien sûr, le choix de la base n'est pas unique. On peut cependant remarquer que l'ensemble des vecteurs du plan est de dimension 2 (il suffit de deux vecteurs pour définir une base), tandis que l'ensemble des polynômes est de dimension infinie. Ce n'est pas une très grande infinie, le nombre d'éléments dans la base qui couvre les polynômes est le même que celui des nombres dans $\mathbb{N}$. On dit alors que c'est une infinie dénombrable [4].

Quand est il de l'espace des fonctions ? A priori, c'est un espace d'une très grande dimension. On verra par la suite que si on se donne quelques restrictions, on peut également définir une base *dénombrable* pour cet espace. C'est un des théorèmes les plus fascinants d'analyse.

Le produit scalaire. On peut enrichir la structure d'espace vectoriel en rajoutant d'autres opération que le $+$ et le produit par un scalaire. L'opération la plus utile à définir pour l'Analyse est le *produit scalaire* (qu'on appelle également le produit intérieur). Le produit scalaire est une opération qui, à deux *vecteurs*, associe un *scalaire* [5]. Nous noterons le produit scalaire de deux vecteurs (a, b). En physique, on a plus l'habitude de le noter par $\overrightarrow{a}.\overrightarrow{b}$, en mécanique quantique par $\langle a|\ b\rangle$.

Nous sommes assez habitués depuis les années du lycée avec ce concept. Un "bon" produit scalaire doit avoir ces quelques propriétés :

(i) $(sa, b) = s(a, b)$ où $s \in \mathbb{S}$, et $a, b \in \mathcal{E}$.

(ii) $(a + b, c) = (a, c) + (b, c)$. où $a, b, c \in \mathcal{E}$.

(iii) $(a, a) \in \mathbb{R}$ et $(a, a) > 0$ si $a \neq \mathbf{0}$ et $(a, a) = 0$ si $a = \mathbf{0}$.

Par exemple, dans l'ensemble des vecteurs du plan, on peut définir un produit scalaire par $(a, b) = \sum x_i y_i$ où x_i et y_i sont les composantes des deux vecteurs a et b.

4. C'est le plus petit des infinis. Sans rentrer dans les détails, l'infini qui ensuite est vraiment plus grande que $\mathbb{N}$ est celui de $\mathbb{R}$. L'ensemble de toutes les fonctions est une infinie encore plus grande.

5. $(.,.) : \mathcal{E} \times \mathcal{E} \to \mathbb{S}$.Si l'espace vectoriel est associé aux réels (complexes), le scalaire est un réel (complexe).

La propriété (iii) est très intéressante. Elle nous permet de définir la *longueur* d'un vecteur, qu'on appelle sa *norme* et que l'on note $\|a\|^2 = (a, a)$. L'intérêt de pouvoir disposer d'une norme est immense. On peut par exemple savoir si deux vecteurs a, b sont "proches" l'un de l'autre en regardant la norme de leur différence $\|a - b\|$, ce qui nous permet à son tour de définir la notion de limite (souvenez vous, les $\forall$blabla, $\exists$blabla tel que blablabla ...). Cela paraît évident si l'on parle des vecteurs d'un plan, ça l'est beaucoup moins quand on discute des espaces vectoriels plus riches comme celui des fonctions. Est ce que par exemple, on peut dire que la fonction $\sin(.)$ et $\log(.)$ sont proches ?

Nous avons besoin aussi de préciser la commutativité. Nous exigeons du produit scalaire :

(iv) $(a, b) = (b, a)$ si $\mathcal{E}$ est associé aux réels ;

(iv') $(a, b) = (b, a)^*$ si $\mathcal{E}$ est associé aux complexes.

Par exemple, pour les vecteurs de $\mathbb{C}^2$, on peut définir le produit scalaire de a, b par $\sum_i x_i y_i^*$ où x_i, y_i sont les composantes de a et b. Notez bien que l'on doit multiplier la composante de l'un par le complexe conjugué de l'autre si on veut respecter la propriété (iii) et disposer d'une norme [6]. La propriété (iv) ou (iv)', combinée à la propriété (i) nous donne :

(i') $(a, sb) = s(a, b)$ si $\mathcal{E}$ est associé aux réels ;

(i") $(a, sb) = s^*(a, b)$ si $\mathcal{E}$ est associé aux réels.

L'orthogonalité. Nous nous souvenons que pour les vecteurs dans $\mathbb{R}^n$, deux vecteurs ($\neq \mathbf{0}$) sont perpendiculaires (qu'on note $a \perp b$) ssi leur produit scalaire est nul. Nous acceptons cette définitions pour tout espace vectoriel. On appelle une base orthogonale une base telle que tout ses éléments soit perpendiculaire l'un à l'autre. Nous avons un avantage fantastique à utiliser des bases orthogonales. D'abord, si les vecteurs $e_1, e_2, ...$ sont orthogonale les uns aux autres, ils sont linéairement indépendant. Si notre espace vectoriel est de dimension n, il nous suffit donc de trouver n vecteurs tous $\perp$les uns aux autres et le tour est joué : nous disposons d'une base !

On peut exiger encore plus d'une base : qu'elle soit orthonormée, c'est à dire que la norme de tous ses éléments soit l'unité. Si nous disposons d'une base orthonormé, on peut trouver les composantes d'un vecteur quelconque de façon extrêmement simple : si a est un vecteur et $(e_1, ...e_n)$ une base orthonormée, alors $a = \sum (a, e_i) e_i$, c'est à dire que la composante de a selon e_i est (a, e_i). Comme exemple, prenez le cas des vecteurs dans $\mathbb{R}^n$.

6. Un exemple intéressant de "produit scalaire" qui ne *respecte* pas (iii) est donné par la relativité restreinte. On repère un *événement* par ses quatre coordonnées spatio–temporelles (x, y, z, t) et le produit scalaire de deux événement est défini par $x_1x_2 + y_1y_2 + z_1z_2 - t_1t_2$. Deux événement *distincts* peuvent donc être à distance nulle l'un de l'autre.

Exercices.

§ 2.1 Démontrer que les deux vecteurs $(1,0)$ $(0,1)$ forment une base pour l'espace vectoriel $\mathbb{C}^2$ associé à $\mathbb{C}$. Même chose pour les deux vecteurs $(i,0)$ et $(0,i)$.

§ 2.2 Démontrer que si $\|a-b\|=0$, alors $a=b$.

§ 2.3 Démontrer que pour l'espace des matrices $n\times n$, $\sum a_{i,j}b_{i,j}$ est un produit scalaire. Ce produit scalaire est souvent utilisé en analyse matricielle numérique pour l'évaluation de la stabilité des méthode itératives.

§ 2.4 Démontrer que si n vecteurs sont mutuellement orthogonaux, alors ils sont linéairement indépendants.

§ 2.5 Démontrer que si $\{e_1,...,e_n\}$ est une base orthonormée, alors n'importe quel vecteur a peut s'écrire sous la forme

$$a=\sum_{i=1}^{n}(a,e_i)e_i$$

Comment doit on modifier cette formule si la base est simplement orthogonal, mais pas orthonormée ?

§ 2.6 Inégalité triangulaire.En réalité, une norme pour pouvoir légalement porter ce nom, doit respecter l'inégalité triangulaire :

$$\|a+b\|\leq\|a\|+\|b\|$$

Démontrez que la norme définie par le produit scalaire vérifie cette inégalité. Pour cela il faut d'abord démontrer l'inégalité de Cauchy-Schwarz :

$$|(a,b)|^2\leq\|a\|.\|b\|$$

qu'on peut assez facilement démontrer en considérant le produit $(a+\lambda b,a+\lambda b)\geq 0$.

§ 2.7 Pouvez vous généraliser le produit scalaire dans $\mathbb{R}^n$ à l'espace des polynômes ? Et surtout démontrer qu'il respecte toutes les propriétés d'un produit scalaire ?

§ 2.8 Un opérateur linéaire est une fonction linéaire de l'espace vectoriel dans lui même : il prend un vecteur en entrée et produit un vecteur en sortie. La linéarité veut dire que si L est un opérateur linéaire, a,b deux membres quelconques de l'espace vectoriel et λ,μ deux scalaires, alors

$$L(\lambda a+\mu b)=\lambda L(a)+\mu L(b)$$

(N'oublions pas que $L(a)$ et $L(b)$ sont des vecteurs au même titre que a et b). Si on se donne une base $\{e_i\}$, l'opérateur peut être caractérisé par son action sur *chaque* vecteur de la base :

$$L(e_j)=\sum_i L_{ij}e_i$$

Les nombres L_{ij} sont les composantes de l'application L dans la base des $\{e_i\}$. En général, pour les représenter, on les dispose dans un tableau (appelé matrice) où la $i-$ème ligne et la $j-$ème

colonne contient le nombre L_{ij}.
Démontrez que les composantes de deux vecteurs quelconque a et b tel que $b = L(a)$ sont relié par la relation (noter l'ordre des sommations)

$$b_i = \sum_j L_{ij} a_j$$

§ 2.9 Démontrer alors que si la base est orthonormale,

$$L_{ij} = (e_i, L(e_j))$$

En langage claire, pour connaître la composante L_{ij} d'une matrice, il faut trouver d'abord le vecteur qui résulte de l'application de l'opérateur au $j-$ème vecteur de la base$p = L(e_j)$, et former le produit scalaire de ce vecteur avec le $i-$ème vecteur de la base.

§ 2.10 Soit deux bases $\{e_i\}$ et $\{f_i\}$ et P une application linéaire tel que

$$\begin{aligned} P(e_i) &= f_i \quad i = 1, 2, ..., n \\ P^{-1}(f_i) &= e_i \end{aligned}$$

P est couramment appelé l'application de *passage*. Soit A une application linéaire quelconque dont les éléments dans la base des $\{f_i\}$ sont données par la matrice a_{ij}. Soit maintenant l'application linéaire $P^{-1}AP$. Calculer ses éléments de matrice dans la base des $\{e_i\}$.

2.2 L'espace vectoriel des fonctions.

Manipuler des vecteurs dans l'espace $\mathbb{R}^n$ c'est bien, mais nous nous intéressons à un espace beaucoup plus vaste, celui des fonctions. Soit $\mathcal{F}$ l'ensemble des fonctions $\mathbb{R} \to \mathbb{R}$ définies sur un intervalle donnée I. Les fonctions sont en faite des boites noire qui prennent des chiffres en entrée et produisent des chiffres en sortie. La fonction $\sin(.)$ par exemple, à une valeur $x \in I$ associe le nombre $\sin x$. On peut voir les fonctions comme des points dans un espace immensément grand où en se baladant, on rencontrerai de temps en temps des fonctions connues comme $\log(.)$, $\exp(.)$, $\exp(2.)$ et la plupart de temps des fonctions qui n'ont pas de nom[7].

Le produit scalaire. Il est évident que $\mathcal{F}$ possède une structure d'espace vectoriel. On ne sait pas encore si nous pouvons étendre la notion de base à cet espace, mais on peut parfaitement définir des produits scalaires. Le produit scalaire que l'on utilisera abondamment est le suivant :

$$(f, g) = \int_{\mathcal{I}} f(x) g(x) dx$$

7. En faite, si on se baladait dans cet espace de façon aléatoire, on ne rencontrerai jamais des fonctions connues.

On démontrera dans un exercice que ce produit scalaire a toute les bonnes propriétés. Mais on peut noter que cette définition généralise la somme $\sum x_i y_i$du produit scalaire dans $\mathbb{R}^n$, quand $n \to \infty$ (souvenez vous de la définition de l'intégral).

Bien, nous disposons d'un produit scalaire, on peut donc définir la norme d'une fonction.

$$\|f\|^2 = \int_I [f(x)]^2 dx$$

Cette norme, appelée $\mathcal{L}_2$, est très populaire. voyons quelques exemples, pour l'intervalle $[0, 2\pi]$,

1. $\|\exp(.)\|^2 = \int_0^{2\pi} \exp^2(x)dx = (\exp 4\pi - 1)/2$.
2. $\|\sin(.)\| = \sqrt{\pi}$
3. $\|\log(.)\| =$??? à faire en exercice.
4. $\|1/(.)^n\| = \infty$ si $n > 1$.

On voit ici les premières bizarreries des ces grands espaces (de dimension infini) apparaître : un élément à priori sympathique peut avoir une norme infinie.

Le lecteur a remarqué que jusque là, nous avons utilisé une notation particulière pour distinguer une fonction (un point dans l'espace vectoriel des fonctions) de la valeur que prend cette fonction pour une entrée particulière : la première est notée $f(.)$ est la deuxième $f(x)$. Comme cette notation est quelque peu lourde et que nous espérons que le lecteur est maintenant habitué à cette distinction, nous emploierons à partir de maintenant indifféremment la notation $f(x)$ pour les deux notions. Le contexte détermine si on parle de la fonction ou de sa valeur.

Nous avons mentionné plus haut que disposer d'une norme nous permet de savoir si deux fonctions sont proches ou même identique si $\|f - g\| = 0$. Considérons alors les deux fonctions, définies sur $[0, 1]$: $f(x) = 1$ et $g(x) = 1$ si $x \neq 0$ et $g(x) = 0$ si $x = 0$. Au sens de notre norme $\mathcal{L}_2$, ces deux fonctions sont identiques [8] ! Grossièrement parlant, notre norme est une lunette pas trop précise et ne distingue pas les différences subtiles entre deux fonctions. Elle ne va retenir que les traits les plus importants [9]. Ainsi, quand $n \to \infty$, la suite des fonctions $f_n(x) = x^n$converge vers $f(x) = 0$ sur l'intervalle $[0, 1]$ au sens $\mathcal{L}_2$, mais ne converge pas au sens des convergences uniformes.

Notons enfin que si nous manipulons l'ensemble des fonctions qui associent à une valeur réelle un nombre *complexe*, *i.e.* $f : \mathbb{R} \to \mathbb{C}$, nous devons légèrement modifier la définition du produit scalaire :

$$(f, g) = \int_I f(x)g(x)^* dx$$

où le symbole $*$ désigne le complexe conjugué.

8. C'est même pire : Si la fonction g est définie par $g(x) = 0$ si $x \in \mathbb{Q}$ et $g(x) = 1$ sinon, au sens de notre norme, elle est identique à la fonction f. Bien sûr, on aurait besoin de redéfinir ce que l'on entend par une intégrale.

9. Il existe bien sûr des normes aux pouvoirs de résolutions beaucoup plus grande, comme celle utilisée pour la convergence uniforme des suites de fonctions.

L'orthogonalité. La notion d'orthogonalité se généralise immédiatement aux fonctions : f et g ($\neq \mathbf{0}$) sont orthogonales si $(f, g) = 0$. Ainsi, sur l'intervalle $[-1, 1]$, les fonctions 1 et x sont orthogonales. De même pour les fonction $\exp(-x/2)$ et $\exp(-x/2)(1-x)$ sur l'intervalle $[0, \infty]$.

Nous avons vu plus haut que la notion d'orthogonalité nous donne un sérieux coup de main pour trouver une base. En particulier, dans un espace de dimension n, il nous suffit de trouver n vecteurs orthogonaux pour avoir une base. Peut on généraliser ce résultat à des espaces de dimension infinie ? la réponse est oui si on prend des précautions. Les fonctions de normes infinies nous posent de sérieux problèmes. Nous allons donc restreindre notre espace de fonctions en nous contentant des fonctions de carré sommable, c'est à dire des fonction f tel que $\int_I |f(x)|^2 dx < \infty$. Nous avons alors le théorème fondamental suivant :

> Dans l'espace des fonctions de carré sommable, on peut trouver des ensembles *infini dénombrable* de fonctions orthogonaux qui chacun constitue une base.

Le lecteur peut méditer sur ce théorème : pour l'énoncer (sans le démontrer) nous avons pris de nombreux raccourcies sans même avoir précisé certains termes, encore moins leur donner un peu de rigueur et de décence. Nous allons dans la suite clarifier un peu mieux les choses, sans les démontrer. Mais avant cela, voyons le côté étrange de ce théorème.

Comme nous l'avons indiqué, l'infini dénombrable, celui des nombres entiers, est le plus petit des infinis. Il a cette particularité que pour un nombre donné, on peut indiquer celui qui est juste avant et celui qui est juste après. L'infini des nombres rationnels n'est pas vraiment plus grand, ni celui des nombres algébriques. Par contre, l'infini des nombre réels est vraiment plus grand. On peut dire grossièrement [10] que $\mathbb{R} = 2^{\mathbb{N}}$(bien sûr, on parle en faite du cardinal, de la taille, de ces ensembles) : pour représenter *un* nombre réel, nous avons absolument besoin de $\mathbb{N}$ nombre entier. L'ensemble des fonctions est beaucoup, beaucoup plus vaste. Imaginez que pour représenter *une seule* fonction, nous avons besoin de $\mathbb{R}$ nombre réel. Le théorème ci-dessus nous dit que si la fonction est de carré sommable, nous n'avons alors besoin pour la représenter que de $\mathbb{N}$ nombre réel ! Une exigence a priori anodine, que les fonctions soient de carré sommable, réduit sérieusement la taille de l'ensemble des fonctions.

Après ces digressions philosophiques, un peu de concret. D'abord, qu'est ce que ça veut dire une base dans ces espaces infinis ? intuitivement, ça doit être la même chose que les espaces de dimensions fini : un ensemble d'objet élémentaire qui nous permet de décrire tous les autres. Supposons que, dans l'espace des fonctions, $E = \{e_1, e_2, ...\}$ constitue une base orthonormée. Dans ce cas, une fonction quelconque f doit pouvoir

10. le cardinal de $\mathbb{N}$ est noté $\aleph_0$ (aleph zéro), celui de $\mathbb{R}$ $\aleph_1$ si on accepte l'axiome de choix. En pensant aux nombres réels entre 0 et 1 comme une succession (infinie) de bits 0 et 1 (comme en informatique), la relation $\aleph_1 = 2^{\aleph_0}$ paraît assez raisonnable. Nous devons tous ces résultats sur les infinis aux travaux de Georg Kantor, à la fin du dix-neuvième siècle.

s'écrire, de façon unique,

$$f(x) = \sum_{n=1}^{\infty} f_n e_n(x)$$

où les f_n sont des *scalaires* qu'on appelle les composantes de f sur la base $\{e_n\}$. Elle sont données, comme pour des espaces de dimensions finis, par la *projection* de f sur les vecteurs de base en utilisant le produit scalaire :

$$f_n = \int_I f(x) e_n(x) dx$$

Remarquez que f_n est bien un nombre, un scalaire. On peut définir une suite de fonctions $\phi_N(x) = \sum_{i=1}^{N} f_n\, e_n(x)$. Si l'ensemble E est bien une base, alors $\|f - \phi_N\| \to 0$ quand $N \to \infty$. Cela veut dire qu'on peut approximer une fonction par une somme finie de fonctions de base, et on peut rendre cette approximation aussi bonne qu'on le souhaite en prenant suffisamment de composante. Le lecteur est déjà partiellement habitué à cette idée : le développement de Taylor approxime une fonction par la combinaison des fonctions x^n. L'espace des fonctions qui peuvent être couvert par un développement de Taylor est cependant beaucoup plus petit que $\mathcal{L}_2$. Les mathématiciens ont été amené à trouver donc d'autres bases. Chaque base est bien adapté aux traitements d'un certain nombres de problèmes, essentiellement la résolution d'une certaine classe d'équations différentielles. La base la plus populaire, est de loin, est celle proposée par monsieur Fourier, préfet de l'Isère en son temps, au tout début du XIX$^{\text{ème}}$ siècle. Ce sera l'objet du prochain chapitre.

Exercices.

§ 2.11 Donner une définition précise de la convergence d'une suite au sens de la norme $\mathcal{L}_2$ dans l'espace des fonctions de carré sommable.

§ 2.12 montrer que la fonction $f(x) = x^n$ définie sur $[0, 1]$ converge vers la fonction $g(x) = 0$ au sens $\mathcal{L}_2$.

§ 2.13 Démontrer que la convergence uniforme implique la convergence au sens $\mathcal{L}_2$. L'exemple précédent montre bien sûr que le contraire n'est pas vrai.

§ 2.14 En algèbre linéaire, les formes bilinéaires généralisent le concept du produit scalaire. On peut suivre le même chemin et définir le produit scalaire entre deux fonctions par

$$(f, g) = \int_I w(x) f(x) g(x) dx$$

où la fonction $w(x)$est appelé le poids. Démontrer que cette définition possède les propriétés d'un produit scalaire. Que doit on imposer à la fonction poids ?

§ 2.15 On appel polynômes orthogonaux des polynômes $P_n(x)$ de degrés n, orthogonaux les uns aux autres au sens du produit scalaire défini plus haut. Trouver les trois premiers polynômes associés au poids $w(x) = 1$ et à l'intervalle $[-1, 1]$. On appelle ces polynômes les polynômes de Legendre.

§ 2.16 Démontrer que les polynômes de Legendre que vous avez trouvé obéissent à l'équation différentielle

$$(1 - x^2)y'' - 2xy' + n(n + 1)y = 0$$

En réalité, c'est souvent pour cela que l'on cherche les polynômes orthogonaux : ils sont solution d'équations différentielles intéressante pour la physique.

§ 2.17 Même question que 5 pour le poids $w(x) = e^{-x}$et l'intervalle $[0, \infty[$. Ces polynômes (de Laguerre) sont associés à la solution de l'équation de Schrödinger pour l'atome d'hydrogène.

§ 2.18 L'opération $D = d/dx$ est une opération linéaire dans l'espace des fonctions infiniment dérivable ($\mathcal{C}^\infty$) : (i) elle prend une fonction en entrée et donne une fonction en sortie ; (ii) elle fait cela de façon linéaire, *i.e.* $D(\lambda f + \mu g) = \lambda Df + \mu Dg$, où λ, μ sont des scalaires et f, g des fonctions. Supposons que des fonctions orthonormées $f_n(x)$ constituant une base obéissent à la relation $df_n(x)/dx = f_n(x) + a_n f_{n+1}(x)$. Pouvez-vous donner la représentation matricielle de D dans la base des f_n ?

2.3 Quelques digressions historiques.

Le concept d'espace vectoriel des fonctions a été proposé par Hilbert à la fin du dix-neuvième et début du vingtième siècle et a unifié de nombreux champs de recherches en mathématique. La résolution des équations intégrales pouvait par exemple être ramené à la recherche des valeurs propres de certaines matrices. La résolution d'un système d'équations différentielles de premier ordre pouvait être donné directement comme l'exponentiel d'une matrice, ... Nous n'épuiserons pas par quelques exemples l'approche profondément novatrices d'Hilbert. Ce champ de recherche était cependant méconnu des physiciens. Au début des années 1920, Heisenberg et ses collègues ont inventé une mécanique matricielle (qu' Einstein qualifia de cabalistique dans une lettre à Planck) pour expliquer les phénomènes observés à la petite échelle des atomes et des électrons. Quelques années plus tard, Schrödinger a proposé sa célèbre équation d'onde, qui elle aussi expliquait assez bien les phénomènes observés. C'est Von Neumann qui a démontré à la fin des années 1920 que ces deux approches étaient fondamentalement les mêmes (voir exercice 8 plus haut) : l'équation de Schrödinger utilise un opérateur dans l'espace des fonctions de carré sommable (qui transforme une fonction en une autre fonction, comme une matrice qui transforme un vecteur dans un autre vecteur), donc associé à une matrice infinie. Depuis, les grands succès de la mécanique quantique ont encouragé les physiciens à assimiler ces concepts dès leur plus tendre age, ce qui les aide à traiter de nombreux champs de recherches autres que la mécanique quantique par les même techniques.

3 Les séries de Fourier.

Nous allons dans ce chapitre étudier les Séries de Fourier. On ne peut pas sérieusement toucher un sujet de physique sans utiliser d'une manière ou d'une autre ces séries (ou leur généralisation, les transformées de Fourier). Nous en verrons de nombreux exemples à travers ce cours. Les séries de Fourier ont également joué un grand rôle dans le développement des mathématiques. Quand Joseph Fourier présenta la première fois le résultat de son analyse de l'équation de la chaleur à l'Académie des Sciences, l'accueil était loin d'être enthousiaste et beaucoup, parmi les plus grands (Laplace et Lagrange) s'y sont violemment opposé : Comment la somme d'une suite de fonctions toutes continues peut être "égale" à une fonction discontinue ? Le pragmatisme a fait avancer l'usage des ces suites bizarres jusqu'à ce que d'autres mathématiciens comme Lebesgue (pour justifier ces pratiques un peu sales) redéfinissent la théorie de la mesure et fassent faire un bond à l'analyse mathématique. De tout cela, on ne parlera pas ici. Notre approche sera beaucoup plus pratique : Qu'est ce qu'une série de Fourier, à quoi elle sert, comment on fait pour l'obtenir.

3.1 Introduction.

Les premiers travaux sur la décomposition en série de Fourier viennent en faite du grand Lagrange lui même dans les années 1780 et son étude de l'équation des cordes vibrantes. Supposons une corde tendu entre 0 et L qu'on déforme à l'instant initial et que l'on relâche. Soit $y(x,t)$ l'écart à l'équilibre à la position x et à l'instant t. On démontre alors que

$$\frac{\partial^2 y}{\partial t^2} - v^2 \frac{\partial^2 y}{\partial x^2} = 0 \tag{3.1}$$

où v est un coefficient qui dépend de la densité et de la tension de la ligne. Cherchons la solution de cette équation sous la forme $y = A_{k,\omega} \cos \omega t . \sin kx$. En injectant cette forme dans l'équation (3.1), on trouve que cette forme ne peut être une solution que s'il existe une relation entre ω et k : $\omega = vk$. Ensuite, la fonction y doit satisfaire les conditions aux bords $y(0,t) = y(L,t) = 0$. La première condition est automatiquement satisfaite. La deuxième condition impose $\sin kL = 0$, c'est à dire $k = n\pi/L$, où $n = 0, 1, 2, 3, ...$On déduit de tout cela que les fonctions $f_n(x,t) = A_n \cos(n\pi vt/L) \sin(n\pi x/L)$ sont solution de notre équation d'onde *avec* ses conditions aux bords. On les appelle les modes propres de vibration. Le principe de superposition nous dit (le démontrer) que si f et g sont solution, alors $f + g$ l'est aussi. La

solution générale de l'équation d'onde (3.1) est donc de la forme

$$y = \sum_{n=1}^{\infty} A_n \cos(n\pi vt/L) \sin(n\pi x/L)$$

Jusque là, nous n'avons rien dit des coefficients A_k, puisqu'ils ne peuvent pas être obtenus de l'équation d'onde directement. Ils doivent sortir de la condition $y(x,0) = y_0(x)$, c'est à dire de la déformation originale que nous avons imprimé à notre corde à l'instant $t = 0$. Nous devons donc avoir :

$$y_0(x) = \sum_{n=1}^{\infty} A_n \sin(n\pi x/L)$$

Est-il possible de trouver des coefficient A_n pour satisfaire cette équation ? Nous verrons la réponse plus bas. A priori, trouver la réponse paraît assez compliquée. Notons que si y_0 a une forme simple, on peut trouver une solution. Par exemple, si $y_0(x) = 4\sin(11\pi x/L)$, alors $A_{11} = 4$ et tous les autres A_n sont nul.

3.2 Les séries de Fourier.

Nous allons étudier maintenant de façon approfondie les fonctions $\sin$ et $\cos$, puisqu'elles peuvent constituer une base. Plus précisément,

Théorème. Dans l'espace vectoriel $\mathcal{L}_2[0, L]$, c'est à dire celui des fonctions de carré sommable définies sur l'intervalle $[0, L]$, les fonctions

$$1, \sin(2\pi x/L), \cos(2\pi x/L), \ldots \sin(2\pi nx/L), \cos(2\pi nx/L), \ldots$$

constituent une base orthogonale.

Nous accepterons ce théorème sans démonstration [1], et allons plutôt contempler quelques uns de ses aspects. D'abord, l'orthogonalité. Puisque

$$(1, \sin n(.)\,) = \int_0^L \sin(2\pi nx/L)dx = -\frac{L}{2\pi n}\left[\cos(2\pi nx/L)\right]_0^L = 0$$

la fonction 1 est orthogonale à tous les sinus et tous les cosinus. Ensuite, comme

$$2\sin(2\pi nx/L)\sin(2\pi mx/L) = \cos(2\pi(n-m)x/L) - \cos(2\pi(n+m)x/L)$$

1. La démonstration est due à Weierstrass dans les années 1880. Elle ne pose pas de difficulté majeure. Disons que pour qu'une suite f_n de vecteurs orthogonaux constitue une base, il faut que si un élément g est orthogonal à tous les f_n, alors $g = 0$. C'est pour cela par exemple que la suite des $\sin(.)$ seul ne peut constituer une base : on peut trouver toujours des $\cos(.)$ qui soit orthogonaux à tous les $\sin(.)$.

les fonctions $\sin n(.)$ et $\sin m(.)$ sont orthogonales, sauf si $n = m$, auquel cas, $\| \sin n(.) \| = \| \cos n(.) \| = L/2$.

Ensuite, une fonction f quelconque de $\mathcal{L}_2[0, L]$ peut s'écrire sous la forme

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(2\pi nx/L) + b_n \sin(2\pi nx/L)$$

et comme notre base est orthogonale, les coefficient a_n et b_n sont donnés par le produit scalaire de f par les éléments de la base :

$$a_0 = (1/L) \int_0^L f(x)dx \tag{3.2}$$

$$a_n = (2/L) \int_0^L f(x) \cos(2\pi nx/L)dx \tag{3.3}$$

$$b_n = (2/L) \int_0^L f(x) \sin(2\pi nx/L)dx \tag{3.4}$$

Notons que le coefficient a_0 est la moyenne de la fonction f sur l'intervalle donnée.

Exemple 3.1 Prenons la fonction $f(x) = x$, $x \in [0, 1]$.

Le coefficient a_0 s'obtient facilement en utilisant l'eq.(3.2) : $a_0 = 1/2$. Pour les autres coefficients, nous avons besoin d'une intégration par partie :

$$b_n = 2 \int_0^1 x \sin(2\pi nx)dx = \frac{-1}{\pi n} \left[x \cos(2\pi nx) \right]_{x=0}^{x=1} + \frac{1}{\pi n} \int_0^1 \cos(2\pi nx)dx = -\frac{1}{\pi n}$$

$$a_n = 2 \int_0^1 x \cos(2\pi nx)dx = \frac{1}{\pi n} \left[x \sin(2\pi nx) \right]_{x=0}^{x=1} + \frac{1}{\pi n} \int_0^1 \sin(2\pi nx)dx = 0$$

Nous pouvons donc écrire

$$x = \frac{1}{2} - \sum_{n=1}^{\infty} \frac{1}{\pi n} \sin(2\pi nx) \quad x \in [0, 1] \tag{3.5}$$

La figure 3.1 montre la fonction x, ainsi que ses approximations successives en prenant de plus en plus de termes de la série de Fourier. . Nous pouvons constater plusieurs choses : (i) évidemment, plus on prend de terme, plus l'approximation est bonne , mais nous avons des oscillations de plus en plus violentes sur les bords, dont l'amplitude décroît ; (ii) l'approximation prend les mêmes valeurs aux deux bords, ce qui n'est pas le cas de la fonction originale ; (iii) cette valeur est $1/2$ dans le cas présent, ce qui est la moyenne des valeurs que prend la fonction originale aux deux bords.

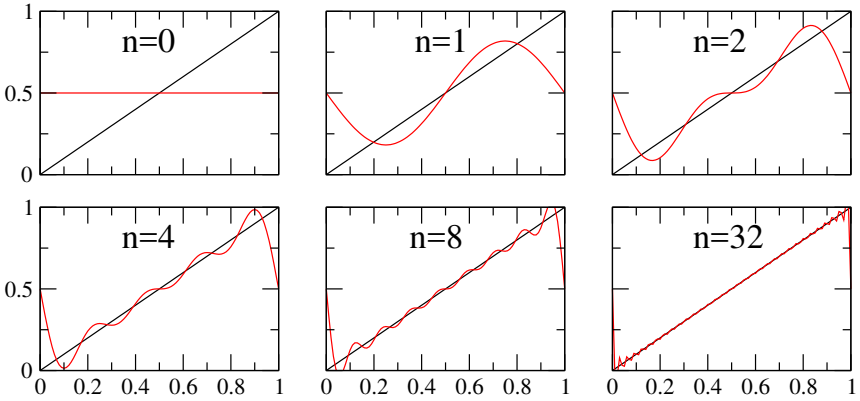


FIGURE 3.1 – Approximations successive de la fonction x par les séries de Fourier. En noir, la fonction original, en rouge l'approximation par la série géométrique.

Le point (ii) est dû à la périodicité de nos fonctions sin et cos : chaque fonction dans la somme, étant de période au moins 1, prend obligatoirement la même valeur sur les deux bords, donc la somme doit également prendre la même valeur sur les deux bords. Le point (iii) est plus troublant : la fonction originale $f(x) = x$ prend la valeur 0 en $x = 0$ et 1 en $x = 1$. La somme par contre, prend la valeur $1/2$ sur les deux bords : la somme ne converge donc pas en tout point vers la fonction originale (adieu la convergence uniforme ou point par point), mais seulement pour la *majorité* des points sur l'intervalle. Ils se trouvent que cette majorité est largement suffisante : si on prend une infinité de terme dans la somme, alors la somme et la fonction originale ne diffèrent qu'en deux points. Deux comparé à la taille de $\mathbb{R}$ donne tous son sens à la notion de majorité. On dit que la différence entre la fonction originale et la série *est de mesure nulle*.

Tout ce que nous avons dit ci-dessus se généralise immédiatement aux intervalles quelconques $[a, b]$. Il suffit simplement dans les formules, poser $L = b-a$ qui représente comme avant la longueur de l'intervalle.

Exemple 3.2 Prenons cette fois la même fonction $f(x) = x$, mais sur l'intervalle $[-1/2, 1/2]$.

Le même calcul que précédemment nous mène à

$$x = -\sum_{n=1}^{\infty} \frac{(-1)^n}{\pi n} \sin(2\pi n x) \quad x \in [\frac{-1}{2}, \frac{1}{2}] \tag{3.6}$$

Nous voyons que les coefficients dépendent également de l'intervalle sur lequel la fonction est définie.

Notons enfin qu'en prenant des valeurs de x particulières, nous disposons d'un moyen intéressant de trouver la limite de certaines sommes. Dans l'équation (3.5) par exemple, si on pose $x = 1/4$, nous trouvons que

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2n-1} = 1 - \frac{1}{3} + \frac{1}{5} - ... = \frac{\pi}{4}$$

Égalité de Parceval. En utilisant la notion d'orthogonalité et de produit scalaire, il est facile de démontrer que

$$\frac{1}{L}\int_0^L f(x)^2 dx = a_0^2 + \frac{1}{2}\sum_{n=1}(a_n^2 + b_n^2) \tag{3.7}$$

Cela veut dire qu'au lieu de calculer explicitement l'intégrale du carré de la fonction, nous pouvons simplement sommer le carré de ses coefficients de Fourier. A priori, la démarche paraît absurde, puisque pour calculer les coefficient, on a du déjà effectuer des intégrales. Mais nous allons voir dans la suite que dans de nombreuses applications, notamment celles liées à la solution d'équation à dérivée partielle, nous calculons directement les coefficients de Fourier de la fonction recherchée. Le coté gauche de l'équation (3.7) désigne souvent l'énergie stocké dans un volume, par exemple si f désigne la hauteur d'une corde tendu ou le champs électrique. L'égalité de Parceval nous permet alors d'accéder à cette quantité.

Exercices.

§ 3.1 Décomposez les fonction $f(x) = x^2$ et $f(x) = \exp(x)$ sur l'intervalle $[0, 1]$. Pour ce dernier, si le produit scalaire vous pose problème, noter que $\cos(x) = (e^{ix} + e^{-ix})/2$. Profiter de la décomposition de x^2 pour trouver la limite de $\sum 1/n^2$. C'était une des fierté d'Euler, dans les années 1730, que d'avoir pu déterminer cette somme.

§ 3.2 Soit la fonction "palier" f sur $[0, 1]$ tel que $f(x) = -1/2$ si $x < 1/2$ et $f(x) = 1/2$ si $x \geq 1/2$. Trouver sa décomposition en série de Fourier.

§ 3.3 Même question que précédemment, mais la fonction f est définie par $f(x) = 0$ si $x < 1/2$ et $f(x) = 1$ si $x \geq 1/2$. En comparant ce résultat au résultat précédent, pouvez en tirer des conclusions générales ?

§ 3.4 Décomposer la fonction triangle, $f(x) = 1 - 2|x|$ sur l'intervalle $[-1/2, 1/2]$.

§ 3.5 Trouver la série de Fourier de $\sin(x/2)$ sur l'intervalle $[0, 2\pi]$.

§ 3.6 Démontrer qu'à part le coefficient a_0, les deux fonctions $f(x)$ et $g(x) = f(x) + C$ ont les mêmes coefficients de Fourier. Que vaut le coefficient a_0 pour ces deux fonctions ?

§ 3.7 Démontrer que si $f(x)$ est une fonction $L-$périodique, sa décomposition en série de Fourier sur l'intervalle $[h, L+h]$ ne dépend pas de h. **Help :** vous pouvez utiliser la propriété des intégrales $\int_h^{L+h} = \int_h^0 + \int_0^L + \int_L^{L+h}$.

§ 3.8 Une fonction paire est tel que $f(-x) = f(x)$, c'est à dire que l'axe y constitue un axe de symétrie. Démontrer alors que sur un intervalle $[-L, L]$, les coefficients b_n de la série de Fourier sont nuls. On peut généraliser cette affirmation : si la fonction f, *sur l'intervalle* $[0, L]$, est symétrique par rapport à son milieu, c'est à dire telle que $f(L-x) = f(x)$, alors ses coefficient de Fourier b_n sont nuls (à démontrer bien sûr).

§ 3.9 Que peut on dire des coefficient de Fourier d'une fonction impaire ? En vous inspirant des deux précédents problèmes, que pouvez vous dire des coefficients de Fourier, sur $[0, L]$, d'une fonction telle que $f(L-x) = C - f(x)$?

§ 3.10 Quelle est la condition pour que la fonction x^α appartiennent à $\mathcal{L}_2[0, 1]$. Démontrer alors que si elle n'appartient pas à $\mathcal{L}_2[0, 1]$, elle ne peut pas se décomposer en série de Fourier. Pouvez vous généraliser ce résultat ?

§ 3.11 Démontrer l'égalité de Parceval.

§ 3.12 Qu'elle est la représentation matricielle de l'opérateur $D = d/dx$ dans la base de Fourier ? Et celle de $D^2 = d^2/dx^2$?

3.3 Pourquoi les séries de Fourier sont intéressantes ?

La base de Fourier est en fait très bien adapté pour étudier les phénomènes oscillatoires, et ces derniers sont abondants dans la vie de tous les jours : le mouvement d'un camion sur un pont génère des vibrations dans toute la structure de ce dernier ; le mouvement des pistons dans le moteur met la voiture en vibration ; une onde électromagnétique provoque l'oscillation des électrons à la surface du métal,...Nous ne pourrions pas traiter ces problèmes si nous ne disposions pas de la base de Fourier. Voyons cela ce plus près.

Soit une fonction f périodique, de période L. Cela veut dire que $f(x+L) = f(x)$. Pour connaître cette fonction, il suffit donc de connaître sa valeur seulement sur un intervalle de longueur L. Or, les fonctions $\sin(2\pi nx/L)$ et $\cos(2\pi nx/L)$ sont également périodiques, et de la même période L. Donc, si une somme de ces fonctions égale la fonction f sur une intervalle de longueur L, elle égale la fonction f partout !

Comme nous le savons, si ces fonctions sont de carré sommable sur une période, elle peuvent se décomposer en série de Fourier. Représenter leurs coefficients de Fourier a_n et b_n en fonction de n est aussi bien que de les représenter en fonction de x. Cette représentation est appelé le spectre d'une fonction, et ses composantes de Fourier les *harmoniques*. C'est un vocabulaire issue de la musique. La figure 3.2 représente le spectre de la fonction $f(x) = x$ sur $[-1/2, 1/2]$. En réalité, pour des raisons que nous verrons plus tard, on n'est même intéressé qu'aux coefficients $s_n = \sqrt{a_n^2 + b_n^2}$

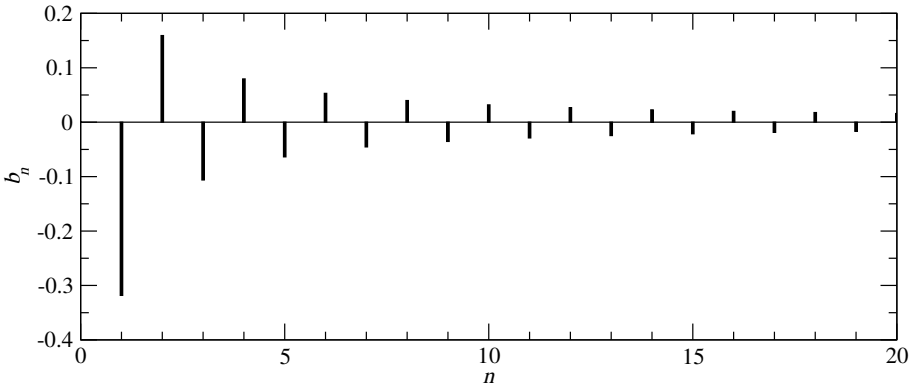


FIGURE 3.2 – Le spectre de la fonction x sur [-1/2,1/2]

et c'est ce dernier qu'on appelle couramment le spectre. Ce qui distingue le Do 240 Hz du piano de la même note d'une flûte n'est pas leurs fréquence de base, mais l'amplitude de leurs harmoniques (la forme de leur spectre). On appelle cela le timbre d'un instrument. L'oreille à une capacité fantastique à distinguer les timbres, et les meilleurs synthétiseur doivent sommer plus d'une vingtaine d'harmonique pour pouvoir imiter une note d'un vrai instrument.

L'oreille humain d'ailleurs ne fait que décomposer les sons en leurs harmoniques : la forme intérieure de l'oreille fait qu'une vibration pure (sinusoïdale) de fréquence donnée met une région particulière de l'oreille en oscillation, ce qui stimule les neurones à cet endroit. Chaque harmonique d'un son excite donc une région différente de l'oreille, et la carte de ces régions permet au cerveau de déterminer très précisément la nature du son et distinguer ainsi le piano du violon.

La compression JPEG des photos numériques utilise le principe des transformées de Fourier : une image est divisée en plusieurs régions et les composantes de Fourier de chaque sous régions sont calculées, mais seulement l'amplitude des premières harmoniques sont conservées, puisque l'œil humain n'est pas sensible aux petits détails.

Exercices.

Représenter le spectre des fonctions décomposées plus hauts.

3.4 Un peu de généralisation.

Nous avons manipulé des sinus et des cosinus séparément. Mais ces deux fonctions ne sont que des combinaisons d'exponentiel d'arguments imaginaires. Il serait donc tout aussi logique de choisir comme base les fonctions $\exp(2i\pi nx/L)$. Cela en plus

nous élargit un peu les horizons, puisqu'on peut alors étudier les fonctions $f : \mathbb{R} \to \mathbb{C}$ (pourvu qu'elles soient de carré sommable). N'oublions pas cependant que le produit scalaire s'obtient dans ce cas un intégrant une fonction qui multiplie le complexe conjugué de l'autre. Nous avons donc,

$$f(x) = \sum_{n=-\infty}^{+\infty} c_n \exp(2i\pi nx/L)$$

Tous ce que nous avons dit plus haut se généralise aisément.

Exercices.

§ 3.13 En vous inspirant du calcul des coefficient a_n et b_n, expliquer comment on calcule les coefficients c_n.

§ 3.14 Démontrer que si $f(x) \in \mathbb{R}$, alors $c_{-n} = c_n^*$.

§ 3.15 Trouver la relation entre a_n, b_n et c_n.

§ 3.16 Énoncer la relation de Parceval avec les coefficients c_n.

3.5 Les séries de sinus et de cosinus.

Nous avons vu que sur l'intervalle $[0, L]$, les fonctions $1, \sin(n2\pi x/L), \cos(n2\pi x/L)$ ($n \in \mathbb{N}$) constituent une base orthogonale qu'on appelle la base de Fourier : ces fonctions sont orthogonales les unes aux autres, et la suite est *complète*. Cette dernière assertion est équivalente à "il n'existe pas de fonction, à part celle uniformément nulle, qui soit orthogonale à toutes les fonctions de la base de Fourier". Est-il possible de trouver une autre base sur l'intervalle $[0, L]$ seulement composée de fonctions sinus ou seulement composée de fonction cosinus ? La réponse est oui : la suite des fonctions $\sin(n\pi x/L)$ (remarquer la disparition du coefficient 2 dans l'argument) également constitue une base, de même que la suite $\cos(n\pi x/L)$. Le choix d'une base plutôt que d'une autre est seulement dicté par le problème que nous avons à résoudre, nous en verrons un exemple plus bas.

Les séries de sinus. Prenons d'abord une fonction $f(x)$ définit sur $[-L, L]$ (un intervalle de longueur $2L$ donc). Sa série de Fourier est donnée par

$$f(x) = a_0 + \sum_{n=1} a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L}$$

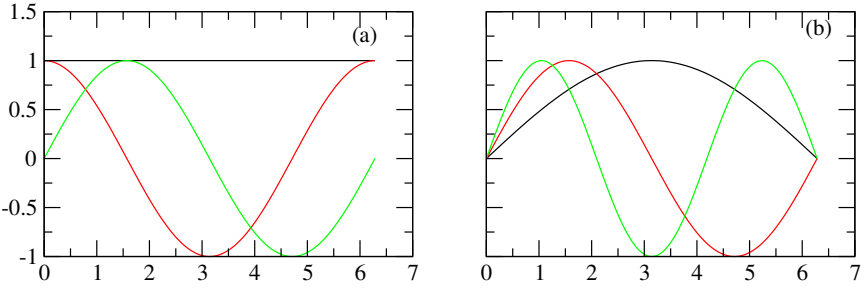


FIGURE 3.3 – Comparaison des trois premières fonction de la base de fourier (a) et celle des sinus.

et ses coefficients de Fourier sont

$$\begin{aligned} a_n &= (1/L)\int_{-L}^{+L} f(x)\cos(n\pi x/L)dx \\ b_n &= (1/L)\int_{-L}^{+L} f(x)\sin(n\pi x/L)dx \end{aligned}$$

Supposons maintenant que la fonction f est impaire, c'est à dire $f(-x) = -f(x)$. En écrivant les intégrales ci-dessus comme la somme de deux intégrales (de $-L$ à 0 et de 0 à L) il est alors évident que tous les coefficient a_n sont nuls, et les coefficients b_n sont données par

$$b_n = (2/L)\int_0^L f(x)\sin(n\pi x/L)dx \tag{3.8}$$

Considérons maintenant une fonction $f(x)$ définit sur $[0, L]$. On peut toujours trouver une extension g de f telle que sur l'intervalle $[0, L]$ les deux fonctions coïncident ($g(x) = f(x)$) et que sur l'intervalle $[-L, L]$, la fonction g est impaire. Il est donc évident que sur l'intervalle $[0, L]$, nous pouvons développer f en série de sinus seulement

$$f(x) = \sum_{n=1} b_n \sin\frac{n\pi x}{L}$$

où les coefficients b_n sont donnés par la relation (3.8). La figure (3.3) montre les trois premiers vecteurs de la base de Fourier et de la base des sinus.

Exercice.

§ 3.17 Développer de façon analogue le développement en série de cosinus pur d'une fonction sur l'intervalle $[0, L]$ en considérant les fonction *paires* sur l'intervalle $[-L, L]$.

§ 3.18 Démontrer que les fonction $\sin(n\pi x/L)$ sont orthogonales les unes aux autres sur l'intervalle $[0, L]$, et que l'on ne peut pas trouver une fonction $\cos(kx)$ qui soit orthogonale à toutes ces fonctions. Cela pourrait un peu plus nous convaincre de la complétude de cette suite.

3.6 Dérivation terme à terme des séries de Fourier.

3.6.1 Dérivation par rapport à la variable principale.

Soit la fonction continue et dérivable par morceau $f(x)$ dont la décomposition de Fourier est

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(2\pi nx/L) + b_n \sin(2\pi nx/L)$$

Dans quelles conditions nous pouvons la dériver terme à terme ? Décomposons la fonction $f'(x)$ en série de Fourier :

$$f'(x) = \alpha_0 + \sum_{n=1}^{\infty} \alpha_n \cos(2\pi nx/L) + \beta_n \sin(2\pi nx/L)$$

Nous avons alors pour le premier terme

$$\alpha_0 = \frac{1}{L}\int_0^L f'(x)dx = \frac{1}{L}(f(L) - f(0))$$

et pour les termes en cosinus

$$\begin{aligned} \alpha_n &= \frac{2}{L}\int_0^L f'(x)\cos(2\pi nx/L)dx \\ &= \frac{2}{L}(f(L) - f(0)) + \frac{2\pi n}{L}\frac{2}{L}\int_0^L f(x)\sin(2\pi nx/L)dx \\ &= \frac{2}{L}(f(L) - f(0)) + \frac{2\pi n}{L}b_n \end{aligned} \tag{3.9}$$

Finalement, pour les termes en sinus, on trouve, en suivant le même chemin,

$$\beta_n = -\frac{2\pi n}{L}a_n$$

Nous voyons donc que si $f(L) = f(0)$, nous pouvons dériver la série de Fourier terme à terme. Sinon, des termes additionnels apparaissent quand on dérive les termes en sinus dont il faut en tenir compte.

On peut généraliser ce résultat aux séries de sinus et de cosinus pures :

1. Si $f(x)$ est continue et dérivable par morceau, sa série de cosinus est dérivable terme à terme sans restriction.

2. Si $f(x)$ est continue et dérivable par morceau, sa série de sinus est dérivable terme à terme si $f(L) = f(0) = 0$! Sinon, des termes additionnelles apparaissent dans la série de sinus de la dérivée, qui sont de la forme $2(\,(-1)^n f(L) - f(0)\,)/L$.

Nous voyons maintenant dans quelles conditions nous avons pu dériver terme à terme la série de la corde vibrante. Comme nous avions la condition $u(0,t) = u(L,t) = 0$, nous pouvions dériver la série de sinus pour obtenir une série de cosinus. Nous avons pu ensuite dériver cette dernière encore une fois sans restriction particulière.

Une hérésie saute aux yeux dans l'équation (3.9) : la suite α_n comporte un terme qui ne dépend pas de n et ne tend pas vers 0. Cela n'est pas du plus bel effet pour la convergence de la série ! La réponse est dans le coefficient b_n, qui doit forcément avoir un terme en $-\Delta/\pi n$ pour annuler exactement le terme constant. C'est ce que l'on va voire plus bas dans un cas particulier.

En pratique, au lieu de dériver terme à terme et prendre en compte les termes additionnels dus aux conditions au bords, il est préférable de régulariser les conditions aux bords pour ne pas avoir de termes additionnels du tout. Supposons par exemple que les conditions aux bords soit $u(0,t) = a$, $u(L,t) = b$. Il est alors préférable d'utiliser la fonction

$$w(x,t) = u(x,t) - \frac{b-a}{L}x - a \tag{3.10}$$

Les dérivées temporelles et les dérivées secondes spatiales de u et w coïncident évidement. $w(x,t)$ est satisfait donc à la même équation que u s'il s'agit d'une équation de chaleur ou de corde vibrante. Par ailleurs, de par sa construction, $w(0,t)= w(L,t) = 0$. Nous pouvons donc d'abord rechercher $w(x,t)$ et ensuite retrouver u à l'aide de la relation (3.10). Nous verrons un exemple plus bas.

3.6.2 Dérivation par rapport à une autre variable.

Qu'en est il pour la dérivation par rapport à une autre variable ? En écrivant

$$u(x,t) = \sum_{n=1}^{\infty} b_n(t)\sin(n\pi x/L)$$

nous avons concentré toute la dépendance temporelle dans l'amplitude des harmoniques $b_n(t)$. Pour avoir le droit de dériver par rapport à t sous la somme, nous devons pouvoir écrire (le démontrer)

$$\frac{\partial}{\partial t}\int_0^L u(x,t)dx = \int_0^L \frac{\partial u}{\partial t}dx$$

L'échange de l'intégrale (sur x) et de la dérivation (par rapport au temps) est permis si $\partial u/\partial t$ existe sur l'intervalle $[0,L]$, est continue et si *elle est bornée*. Nous supposerons dans la suite que pour les fonctions que nous considérons (qui représentent des hauteurs de cordes, des pressions ou des températures) ces conditions sont toujours vérifiées.

3.7 Vibration d'une corde.

Les problèmes de vibration sont une des raisons du succès de l'analyse de Fourier. Nous traitons le cas de la corde vibrante, mais beaucoup d'autres problèmes ondulatoires comme le champs électromagnétique ou les ondes élastiques se traitent par exactement les mêmes méthodes. Revenons à l'équation d'une corde vibrante que nous avons mentionné au début de ce chapitre, et traitons le à l'aide de l'artillerie que nous avons développé. Comme nous l'avons dit plus haut, nous considérons une corde élastique dont (i) les deux extrémités sont fixées aux points $x = 0$ et $x = L$; (ii) est soumise à une tension T ; (iii) a une densité linéaire ρ. Si nous repérons la hauteur de la courbe par $y(x)$, l'équation d'évolution de y s'écrit

$$\frac{\partial^2 y}{\partial t^2} = v^2 \frac{\partial^2 y}{\partial x^2} \tag{3.11}$$

Nous ne justifions pas cette équation. Notons simplement que le membre de droite est l'accélération d'un élément infinitésimal de la corde (dans le sens vertical) au point x, et que le membre de droite est proportionnel à la force exercée sur cet élément par ses voisins. L'équation ci-dessus est simplement la relation de la dynamique $a = f/m$ pour un matériau continu. Le paramètre v a les dimensions d'une vitesse. Les seuls paramètres intrinsèques du problème étant la tension et la densité, il n'existe qu'une seule façon (dimentionnellement parlant) de former une vitesse et v^2 doit être proportionnelle à T/ρ. Nous supposons enfin qu'initialement, la corde est maintenue dans une certaine forme (pincée par exemple) $y_0(x)$ et qu'on la relâche à l'instant $t = 0$.

Nous pouvons à chaque instant, représenter la fonction $y(x;t)$ à l'aide de sa série de Fourier ou de sinus. Pour la représenter à tous les instants, il suffit donc que les *coefficients* de la série soient fonction du temps :

$$y(x;t) = \sum_{n=1} b_n(t) \sin \frac{n\pi x}{L} \tag{3.12}$$

Notons que nous faisons ici le choix de rechercher la solution sous la forme de fonction de sinus, puisque chaque terme de la série respecte les conditions aux bords $y(0,t) = y(L,t) = 0$. En injectant (3.12) dans (3.11) est en identifiant terme à terme (puisque les fonctions sont orthogonales), nous trouvons une équation différentielle pour l'amplitude de chaque mode :

$$b_n^{''}(t) + n^2\omega^2 b_n(t) = 0$$

où $\omega = \pi v/L$. Comme la corde est relâché avec une vitesse nulle, nous avons simplement

$$b_n(t) = B_n \cos(n\omega t)$$

où les coefficient B_n sont les coefficient de la série des sinus de la fonction $y_0(x)$.

L'image est donc la suivante : la déformation initiale est la superposition d'un certain nombre de mode, chacun avec une amplitude b_n. Une fois qu'on relâche la corde, l'amplitude de chaque mode oscillera dans le temps. Remarquez cependant qu'il n'y a pas de transfert entre les modes : si un mode n'était pas présent dans la déformation initiale, il ne sera pas excité par la suite. Chaque mode se comporte comme un oscillateur indépendant, non couplé aux autres. En langage plus chic, on dira que la base de Fourier est une base propre pour l'opérateur Laplacien : dans cette base, la représentation matricielle de cet opérateur est diagonale.

Nous avons, lors de cette résolution, inversé l'ordre des opérations de dérivation et de sommation. Nous savons qu'il existe des conditions très contraignantes pour pouvoir effectuer cette inversion, et elles sont loin d'être réunies à priori (voir ci-dessous). Notons enfin que la différence entre un clavecin et un piano, qui excitent pratiquement les même cordes tendues, est dans la façon de former la déformation initiale, donc de produire des coefficients B_n différents.

3.8 Équation de la chaleur.

La physique mathématique possède quelques équations "star". Ce sont les équations d'onde (rencontrées plus haut), l'équation de Laplace et l'équation de la chaleur. Cette dernière décrit les phénomènes de diffusion (de la chaleur, de la concentration, ...) et est de la forme [2]

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} \tag{3.13}$$

où u représente la température, la concentration, etc... Si x désigne l'espace et t le temps, nous voyons que D doit avoir la dimension d'une longueur au carré par unité de temps $[L^2/T]$.

Remarquez la différence avec l'équation d'onde (3.11), où la dérivée par rapport au temps est de deuxième ordre. Nous voulons traiter le cas d'une barre (comme d'habitude, de longueur L) dont les extrémités sont maintenues à deux températures *différentes* , disons 0 et T. Nous avons donc les deux conditions aux limites

$$u(0,t) = 0 \tag{3.14}$$
$$u(L,t) = T \tag{3.15}$$

et nous supposons la condition initiale

$$u(x,0) = f(x) \tag{3.16}$$

Voilà, le problème est maintenant bien posé. Avant de commencer son traitement total, voyons voir si il existe une solution stationnaire, c'est à dire une solution tel que $\partial_t u =$

2. Voir le chapitre "les équations de la physique" pour la signification de cette équation.

0. Dans les processus diffusifs, c'est la solution qui est atteinte au bout d'un temps plus ou moins long et correspond à l'équilibre. Il est évident ici que $u_s(x) = T(x/L)$ satisfait parfaitement aux équations (3.13-3.15) et est la solution stationnaire recherchée. Après quelques lignes de calculs, nous trouvons que sa série de sinus est donnée par

$$u_s(x) = \frac{2T}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(\frac{n\pi}{L}x)$$

Bon, revenons maintenant à la solution générale. Nous décomposons la fonction u en série de sinus avec des coefficients dépendant du temps :

$$u(x,t) = \sum b_n(t) \sin(n\pi x/L) \tag{3.17}$$

Nous dérivons une première fois par rapport à x. Mais attention, la fonction u prend des valeurs différentes sur les deux bords, il faut donc ajouter des termes

$$\frac{2}{L}\left((-1)^n u(L,t) - u(0,t)\right) = \frac{2T}{L}(-1)^n$$

à la série dérivée :

$$\partial_x u(x,t) = \frac{T}{L} + \sum \left[b_n(t)(\frac{n\pi}{L}) + \frac{2T}{L}(-1)^n \right] \cos(n\pi x/L)$$

Nous sommes maintenant en présence d'une série de cosinus, que nous dérivons encore une fois par rapport à x

$$\partial_x^2 u(x,t) = -\sum \left[b_n(t)(\frac{n\pi}{L}) + \frac{2T}{L}(-1)^n \right] (\frac{n\pi}{L}) \sin(n\pi x/L)$$

La dérivation par rapport au temps nous donne une série de sinus dont les coefficients sont $b_n'(t)$. En égalant terme à terme, nous obtenons une équation de premier ordre pour les coefficients

$$b_n'(t) = -D\left(\frac{n\pi}{L}\right)^2 b_n(t) - D\frac{2T}{L}\frac{n\pi}{L}(-1)^n$$

dont la solution est

$$b_n(t) = B_n \exp\left[-n^2(\pi^2 D/L^2)t\right] + \frac{2T}{\pi}\frac{(-1)^{n+1}}{n} \tag{3.18}$$

Notons d'abord que quand $t \to +\infty$, les coefficients b_n tendent vers les coefficients de sinus de la solution stationnaire : au bout d'un temps assez long, la distribution de température dans le barreau devient linéaire. Ensuite, les amplitudes des harmoniques sont de la forme $\exp(-n^2 t/\tau)$, où $\tau \sim L^2/D$. L' harmonique d'ordre n "disparaît" donc sur

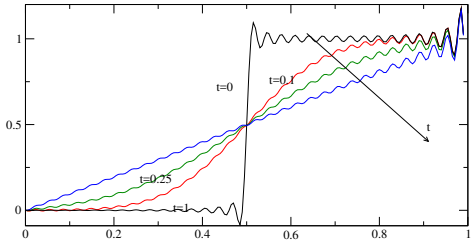


FIGURE 3.4 – Solution de l'équation de la chaleur par les séries de sinus avec $L = 1$, $u(0,t) = 0$; $u(1,t) = 1$, et condition initiale une fonction pallier. La solution est tracée pour 4 temps différents. Nous avons représenté la série par *simplement* ses 64 premières harmoniques. Notez les oscillations artificielles : les séries de sinus ont du mal à reproduire les discontinuités.

une échelle de temps $\sim L^2/(n^2 D)$. (i) Plus L est grand, plus le temps de "thermalisation" est grand. Si on multiplie par 2 la longueur du barreau, on doit multiplier par 4 le temps nécessaire à la thermalisation [3] ; (ii) Plus le coefficient de diffusion D est grand, plus la thermalisation est rapide : le cuivre est plus rapide à thermaliser que le verre ; (iii) plus l'ordre d'une harmonique est important, plus il disparaît rapidement, et ceci est proportionnel au *carré* de l'ordre. Très rapidement, il ne restera pratiquement que l'harmonique d'ordre 1, qui sera le plus lent à "mourir" (voir figure 3.4).

Nous avons réussi à nous en sortir même quand la dérivation sous la somme posait problème. Mais était-il vraiment nécessaire de faire appel à une telle artillerie lourde, qui numériquement n'est pas entièrement satisfaisant ? Et si au lieu de chercher la fonction $u(x,t)$ qui satisfait aux équations (3.13-3.16), nous cherchions la fonction

$$\phi(x,t) = u(x,t) - u_s(x) \tag{3.19}$$

Cette fonction obéit bien sûr à l'équation (3.13). Ces conditions aux limites sont

$$\phi(0,t) = u(0,t) - u_s(0) = 0 \tag{3.20}$$
$$\phi(L,t) = u(L,t) - u_s(L) = T - T = 0 \tag{3.21}$$

La fonction ϕ obéit donc (contrairement à u) à des conditions aux limites *continues*, et ne pose donc aucun problème lors de ses dérivations par rapport à x. Sa condition initiale est donnée par

$$\phi(x,0) = f(x) - u_s(x)$$

Les équations pour ϕ sont donc maintenant bien posées, et nous pouvons les résoudre par la technique habituelle des séries de sinus sans la complications des termes additionnels. Une fois ϕ trouvée, nous avons évidemment trouvé u. La figure (3.5) montre l'avantage (entre autre numérique) de cette dernière méthode.

3. Sachant qu'un gigot de 1 kg cuit en une heure au four, quel est le temps de cuisson d'un gigot de 2 kg ?

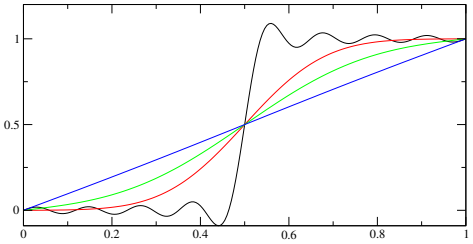


FIGURE 3.5 – La solution de la même équation de la chaleur mais où on a cherché d'abord la fonction ϕ. Nous n'avons pris ici que 16 termes harmoniques. Nous voyons qu'au temps $t = 0$, des oscillations artificielles sont toujours présentes à cause de la discontinuité de la condition initiale. Mais les oscillations artificielles pour des temps ultérieurs de la figure (3.4) ont disparu, puisqu'on a résolu le problème de la discontinuité des conditions aux limites.

3.9 Problèmes avancés.

Problème 3.1 Vibration d'une corde tendu.

L'exemple de la corde vibrante que nous avons considéré plus haut peut être complété de bien des façons. Si la corde est soumise à un frottement visqueux, il faut ajouter un terme en $-\lambda\partial y/\partial t$ (proportionnel à la vitesse locale) à droite de l'équation (3.11). Si la corde est en plus soumise à une force par unité de longueur $f(x,t)$, il faut également l'ajouter à droite. Résoudre l'équation de la corde vibrante (i) en présence d'un frottement (ii) en présence de la force de gravité $f = -\rho g$ (iii) en présence d'une force de rappelle harmonique $f = -ky$. Les conditions aux bords sont toujours les mêmes : corde fixée à ses deux extrémités et avec une déformation initiale $y_0(x)$.

Problème 3.2 Vibration d'une barre élastique.

L'équation de vibration d'une barre élastique est donnée par

$$\frac{\partial^2 y}{\partial t^2} = \alpha \frac{\partial^4 y}{\partial x^4}$$

Discuter les solutions de cette équation. Que pensez vous des conditions initiales ?

Problème 3.3 Équation de Schrödinger dans un puits rectangulaire.

Dans un puits de potentiel rectangulaire est très profond, à une dimension, l'équation de Schrödinger s'écrit :

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2}$$

avec les condition aux limites $\psi(-L,t) = \psi(L,t) = 0$ et $\psi(x,0) = f(x)$. Discuter de la solution de cette équation en suivant l'exemple de la corde vibrante.

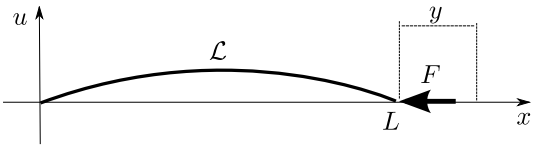


FIGURE 3.6 – Flambage sous l'effet d'une force.

Problème 3.4 Équation de la Chaleur I.

Soit une barre thermiquement isolée, c'est à dire $\partial_x u(0,t) = \partial_x u(L,0) = 0$. En effet, le flux de chaleur à travers une section est proportionnel à la pente de u en cet endroit, et que "isolée" (pas de flux de chaleur de/vers l'extérieur), impose les conditions ci-dessus. En partant d'une distribution de température initiale parabolique $u(x,0) = Tx(L-x)/L^2$, résolvez l'équation de la chaleur. A t'on plutôt intérêt à prendre des séries de cosinus ou des séries de sinus ?

Problème 3.5 Équation de la chaleur II.

Si une source de chaleur est présente dans le milieu (sous forme de résistance électrique ou de particule radioactive ...), l'équation de la chaleur prend la forme

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + Q(x)$$

où Q est la quantité de chaleur produite en x. Prenez une source constante localisée autour de $L/2$, et résolvez alors l'équation de la chaleur.

Problème 3.6 Équation de la chaleur III.

Nous nous intéressons à la distribution de température dans une barre dont l'un des cotés est maintenu à température constante et l'autre extrémité à une température qui varie dans le temps selon une loi connue. Résolvez

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2}$$

avec les conditions aux limites $u(0,t) = 0$; $u(L,t) = g(t)$. Nous supposons une condition initiale homogène $u(x,0) = 0$ et bien sûr g(0)=0. Notez que vous avez intérêt à chercher plutôt la fonction $\phi(x,t) = u(x,t) - g(t)x/L$ pour régulariser les conditions aux limites. Si vous n'aimez pas cette façon de faire, il faut faire attention aux dérivations terme à terme (qui donneront, bien sûr, la même réponse).

Problème 3.7 flambage d'une poutre

Posez une règle à la verticale et appuyez avec un doigt dessus : au bout d'une certaine pression, la règle fléchit soudain. On appelle cette transition brusque le "flambage". Le flambage est par exemple la raison principale qui limite la hauteur des immeubles : au delà d'une certaine taille, l'immeuble flambe sous son propre poids. Nous allons étudier ce phénomène à l'aide des séries de Sinus.

0. Minimum. Pour une fonction $f(z_1, ..., z_n) = a_1 z_1^2 + ... + a_n z_n^2$, le point $z_1 = z_2 = ... = z_n = 0$ est un minimum si $a_i > 0 \; \forall i$. En mécanique, Un point constitue un équilibre stable si c'est un minimum de l'énergie potentielle.

0'. Définitions. Nous considérons une barre de longueur $\mathcal{L}$ qui sous l'effet d'une force F "flambe". La flèche de la barre est repéré par la fonction $u(x)$. La barre est contraintes de garder ses extrémités à $u = 0$. La différence entre la longueur de la barre et sa longueur projeté L est notée $y = \mathcal{L} - L$ (Fig. 3.6).

Soit la fonction $u(x)$ sur l'intervalle $[0, L]$ telle que $u(0) = u(L) = 0$. Nous notons sa décomposition en série de sinus par

$$u(x) = \sum_{n=1}^{\infty} b_n \sin(n\pi x/L)$$

1. Parseval. En utilisant l'orthogonalité des sinus, démontrer l'égalité de Parseval :

$$\int_0^L u(x)^2 dx = (L/2) \sum_{n=1}^{\infty} b_n^2$$

Obtenez alors

$$\int_0^L \left(u'(x)\right)^2 dx \text{ et } \int_0^L \left(u''(x)\right)^2 dx$$

en fonction des harmoniques b_n.

2. Énergie Potentielle. L'énergie potentielle de la barre s'écrit

$$\mathcal{E} = \int_0^L (1/2) B \left(u''(x)\right)^2 dx - Fy$$

où B est le module de rigidité de courbure et F la force appliquée sur la barre à son extrémité. Nous considérons le début de l'instabilité où la flèche est faible ($u'(x) \ll 1$). Dans ce cas

$$\begin{aligned} y = \mathcal{L} - L &= \int_0^L \left(\sqrt{1 + u'(x)^2} - 1\right) dx \\ &= (1/2) \int_0^L u'(x)^2 dx \end{aligned}$$

Obtenez alors l'énergie potentielle en fonction de l'amplitude des harmoniques b_n.

3. Instabilité. La position de la barre droite (non flambée) correspond à $b_n = 0 \; \forall n$. A partir de quelle valeur F_c de la force cette équilibre devient instable ? Quelle est le mode qui devient instable en premier ? Comment cela correspond à votre expérience de poussée sur une règle ?

4. Hauteur. Discutez comment vous pourriez utiliser ce résultat pour calculer la hauteur limite d'un bâtiment.

Problème 3.8 Énergie d'un corps noir.

Reprenons le cas de la corde vibrante fixée à ses deux extrémités. Quelle est son énergie libre F quand elle est maintenue à une température T[4] ? Ce problème a joué un rôle majeur dans

4. Nous mesurons la température en échelle d'énergie : $T = K_B \theta$, où θ est la température habituelle. Cela nous évite de trainer la constante de Boltzmann.

l'évolution de la physique au tournant du XIX$^{\text{ème}}$siècle et a donné lieu à la première formulation de la mécanique quantique.

Avant d'attaquer de front ce problème, quelques rappels sur un cas simple. Prenons un oscillateur harmonique (une boule au bout d'un ressort) dont l'amplitude à un instant est $x(t)$. L'énergie élastique emmagasinée dans l'oscillateur est $E(x) = (k/2)x^2$. L'énergie libre est une sorte de moyenne pondérée par la température de toutes les énergies disponible :

$$Z = e^{-F/T} = \sum_{\{x\}} e^{-E(x)/T} \tag{3.22}$$

où la somme s'entend au sens de toutes les configurations possibles. Comme l'amplitude x est une variable continue, la somme peut être transformée en intégrale et [5]

$$Z = \int_{-\infty}^{+\infty} e^{-kx^2/2T} dx = C.\sqrt{T/k}$$

L'énergie élastique moyenne de l'oscillateur est

$$\langle E \rangle = \sum_{\{x\}} E(x) e^{-E(x)/T} / Z = T/2$$

Si à la place d'un oscillateur, nous avions deux oscillateurs indépendants (non couplés), x_1 et x_2 de raideur k_1 et k_2, l'énergie serait

$$E = (1/2)(k_1 x_1^2 + k_2 x_2^2) \tag{3.23}$$

et la somme (3.22) pour obtenir l'énergie serait cette fois une double intégrale qui se calcule tout aussi facilement. Un calcul élémentaire nous montre alors que l'énergie élastique moyenne est

$$\langle E \rangle = \langle E_1 \rangle + \langle E_2 \rangle = T$$

Pour N oscillateurs, nous aurions $\langle E \rangle = NT/2$, ce qu'on appelle en physique statistique, l'équipartition.

Revenons à notre corde vibrante, dont la hauteur à l'abscisse x est repérée par $u(x)$. L'énergie élastique emmagasiné dans la corde, pour une configuration donnée $u(x)$ vaut [6]

$$E[u(x)] = \int_0^L k \left(\frac{\partial u}{\partial x} \right)^2 dx \tag{3.24}$$

Au lieu de représenter la corde par $u(x)$, nous pouvons le représenter sur la base de Fourier :

$$u(x) = \sum_{n=-\infty}^{+\infty} c_n \exp[(2i\pi n/L)x]$$

5. Le résultat s'obtient facilement en effectuant le changement de variable $x \to x\sqrt{T/k}$; la constante C est juste une intégrale définie qui vaut $\sqrt{2\pi}$

6. Pour la signification de cette expression, voir les deux chapitres sur le calcul variationnel et le sens des équations de la physique.

et une application simple du théorème de Parceval nous montre que

$$E = \int_0^L k \left(\frac{\partial u}{\partial x} \right)^2 dx = \sum_{n=-\infty}^{\infty} k \left(\frac{2\pi n}{L} \right)^2 |c_n|^2$$

Remarquez l'expression de droite et comparez le à (3.23). C'est comme si nous étions en présence de N ($N \to \infty$) oscillateurs harmonique indépendants, l'oscillateur n, d'amplitude c_n ayant une constante de ressort $k(2\pi n/L)^2$. L'énergie moyenne emmagasinée dans la corde est

$$\langle E \rangle = \sum_{n=-\infty}^{+\infty} T/2$$

Aïe .. Une simple corde à température non nulle emmagasine une énergie infinie. La raison principale pour cette divergence est la forme (3.24) de l'énergie de la corde : elle n'est valable que pour des petites déformations de la corde. Les modes à grand n (les hautes fréquences) imposent cependant de très fortes déformations. Il est évident que vous ne pouvez pas plier en 10000 une corde de 1m. Il doit donc exister une sorte de longueur minimum qui limiterait les hautes fréquences, et on parle alors de longueur de *cut-off*.

Par contre, la forme (3.24) décrit parfaitement l'énergie du champ électrique (poser $E = \nabla u$) dans une cavité. Bien sûr, il faut prendre un champs électrique tridimensionnel et prendre en compte les diverses polarisations ; cela est légèrement plus long à calculer mais c'est exactement le même genre de calcul. Ce problème que l'on appelle divergence ultra-violet (pour les hautes fréquence spatiale) a été résolu par Planck et Einstein en supposant que l'énergie d'un mode n ne pouvait pas prendre des valeurs continues mais varie par palier discret. Ceci a été la naissance de la mécanique quantique.

Problème 3.9 Le mouvement Brownien.

Considérons une particule sur un réseau discret unidimensionnel de pas a, c'est à dire une particule qui peut sauter de site en site. Appelons α la probabilité de saut par unité de temps autant vers la gauche que vers la droite. Cela veut dire que la probabilité pour que la particule saute à gauche ou à droite pendant un temps infinitésimal dt est αdt (et donc, la probabilité de rester sur place est $1 - 2\alpha dt$). On cherche à déterminer $P(n, t)$, la probabilité pour qu'au temps t, la particule se trouve sur le site n, sachant qu'à $t = 0$, la particule se trouvait à $n = 0$ [7]. Pour que la particule soit en n au temps $t + dt$, il faut qu'il ait été en $n \pm 1$ au temps t, et qu'il ait effectué un saut vers n pendant l'intervalle dt. Ce phénomène enrichit la probabilité d'être en n. Par ailleurs, si la particule se trouvait en n au temps t, il a une probabilité de sauter à gauche ou à droite pendant l'intervalle dt, ce qui appauvrit la probabilité d'être en n. En prenant en compte ces deux phénomènes, on obtient une équation qu'on appel maîtresse :

$$\frac{1}{\alpha} \frac{dP(n)}{dt} = P(n-1) + P(n+1) - 2P(n) \tag{3.25}$$

7. Le mouvement de petites graines de poussière dans l'eau, étudié par Brown à la fin du dix-neuvième siècle, est une version continue de ce mouvement. En 1905, Einstein a donné l'explication de ce mouvement ératique en supposant la nature moléculaire de l'eau. Cet article est celui le plus cité dans le monde scientifique.

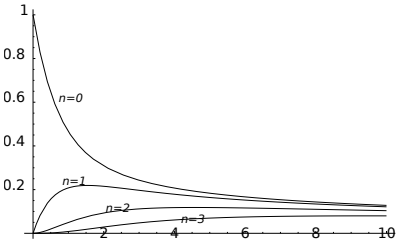


FIGURE 3.7 – Les fonctions $e^{-x}I_n(x)$ pour $n = 0, ..., 3$. Pour $z \ll 1$, $I_n(2z) \approx z^n/n!$, et pour $z \to \infty$, $I_n(z) \approx e^z/\sqrt{2\pi z}$.

Ceci est en faite une infinité d'équations différentielles de premier ordre, avec la condition initiale $P(n = 0, t = 0) = 1$,$P(n \neq 0, t = 0) = 0$. La série de Fourier est devenue une célébrité en résolvant en particulier ce genre d'équation. La méthode que l'on va suivre est une méthode extrêmement générale. Supposons que les $P(n)$ sont les coefficients de Fourier d'une fonction $\phi(s, t)$

$$\phi(s,t) = \sum_{n=-\infty}^{+\infty} P(n,t)\exp(ins) \tag{3.26}$$

et essayons de voir à qu'elle équation doit obéir ϕ. Notons tout d'abord que ϕest 2π-périodique en s. La fonction ϕ est appelée la fonction génératrice des probabilités et caractérise entièrement le processus stochastique. Par exemple, on obtient la moyenne très facilement :

$$\langle n\rangle = \sum nP(n) = -i\left.\frac{\partial\phi}{\partial s}\right|_{s=0}$$

Question : comment on obtiendrai la variance ?

En multipliant l'eq(3.25) par $\exp(ins)$ et en sommant sur les n, on obtient pour ϕ :

$$\frac{\partial\phi}{\partial t} = -2\alpha(1-\cos s)\phi$$

Ce qui nous donne, en intégrant l'équation différentielle par rapport à t :

$$\phi = A(s)\exp[-2\alpha(1-\cos s)t]$$

Et la condition initiale $\phi(s, t = 0) = 1$ (pourquoi ?) nous impose $A(s) = 1$. On connaît donc ϕ et on peut par exemple calculer que la variance $\sim 2\alpha t$, qui est un des résultats important du mouvement brownien. Mais on peut pousser l'analyse plus loin, et calculer explicitement les probabilité. Il existe une classe de fonctions qu'on appelle des fonctions de Bessel d'ordre n que l' on étudiera plus tard dans de ce cours. Ces fonctions (voir Fig.3.7) sont définies par des intégrales :

$$I_n(z) = \frac{1}{\pi}\int_0^\pi e^{z\cos\theta}\cos(n\theta)d\theta$$

Cela nous donne directement les $P(n, t)$:

$$P(n,t) = \exp(-2\alpha t)I_n(2\alpha t)$$

On peut à partir de là effectuer une analyse beaucoup plus fine du mouvement Brownien.

Problème 3.10 fluctuation de mort et naissance.

Soit un système (comme par exemple un ensemble de bactéries) dont la taille varie par palier discret de façon aléatoire. Supposons que la probabilité par unité de temps pour qu'il passe de la taille n à $n \pm 1$ soit proportionnelle à n (notez la différence avec le cas brownien). montrez alors que la probabilité $P(n,t)$ pour que le système ait la taille n à l'instant t obéit à l'équation :

$$dP(n)/dt = (n-1)P(n-1) + (n+1)P(n+1) - 2nP(n) \tag{3.27}$$

Nous supposons qu'à l'instant initial $t = 0$, le système a une taille 1 : $P(1,0) = 1$, $P(n \neq 1, 0) = 0$. Ceci constituera notre condition initiale. Résoudre cette équation n'est pas à priori facile. Posons $u(x,t) = \sum P(n,t) exp(inx)$, et cherchons si on peut trouver une équation pour cette fonction. Notons tout de suite que de par sa définition, u est 2π périodique. On appelle $u(x,t)$ la *fonction génératrice*. Si on réussit à trouver u, on voit alors que les $P(n,t)$ sont simplement les coefficients de la transformée de Fourier de cette fonction. Quelle est la condition initiale pour u, c'est à dire $u(x,0) =?$ En multipliant les deux côtés de l'équation (3.27) par $\exp(inx)$ et en sommant sur n, démontrez que

$$\frac{\partial u}{\partial t} = 2(\cos x - 1)\frac{\partial u}{\partial x}$$

Démontrez que la solution de cette dernière (nous verrons plus tard [8] comment résoudre ces équations) est donnée par :

$$u(x,t) = \frac{1 - i(1+2t)\tan x/2}{1 + i(1-2t)\tan x/2}$$

Calculez la moyenne et la variance en fonction du temps. Avec un peu de vigueur en plus, démontrez que

$$\begin{aligned} P(0,t) &= t/(1+t) \\ P(n,t) &= t^{n-1}/(1+t)^{n+1} \end{aligned}$$

Discutez ce résultat.

Problème 3.11 Propagation du son dans un cristal (les phonons).

Mise en place du problème. Nous allons considérer un cristal uni-dimensionnel de maille élémentaire a. La position d'équilibre de l'atome n est $x_n = na$. Chaque atome n'interagit qu'avec avec ses deux plus proches voisins. Nous voulons savoir comment une perturbation des atomes par rapport à leur positions d'équilibre se propage dans le cristal. Soit u_n l'écart de l'atome n par rapport à sa position d'équilibre (Fig. 3.8). La force $f_{n+1\to n}$ exercée par l'atome $n+1$ sur son voisin n est une fonction de la distance entre les deux : $f_{n+1\to n} = C(u_{n+1} - u_n)$ où C est un coefficient qui dépend de la nature de la liaison chimique (grand pour les cristaux ioniques ou covalents plus petit pour les cristaux moléculaires). Soit m la masse d'un atome ; son équation de mouvement ($\gamma = F/m$) s'écrit, en considérant les forces exercées par ses deux voisins,

8. Voir le chapitre sur les équations à dérivées partielles de premier ordre.

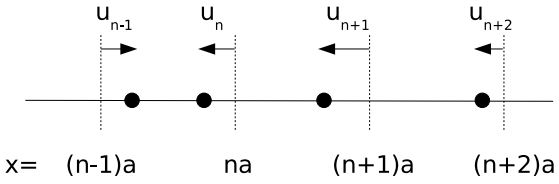


FIGURE 3.8 – Une portion du cristal montrant 4 atomes

$$\frac{d^2u_n}{dt^2} = (C/m)(u_{n+1} - u_n) + (C/m)(u_n - u_{n-1}) \tag{3.28}$$

$$(C/m)\left(-2u_n + u_{n+1} + u_{n-1}\right) \tag{3.29}$$

Dorénavant, nous poserons $\omega_0^2 = C/m$.

Résolution. Nous considérerons les quantités $u_n(t)$ comme les coefficients de Fourier d'une fonction caractéristique $\phi(q,t)$ où la variable q appartient à l'intervalle $[-\pi/a, \pi/a]$ (la longueur de l'intervalle est donc $L = 2\pi/a$). Le paramètre q est souvent appelé "nombre d'onde". ϕ s'écrit donc

$$\phi(q,t) = \sum_{n=-\infty}^{+\infty} u_n(t) \exp\left(iaqn\right) \tag{3.30}$$

1. En multipliant l'équation (3.28) pour l'atome n par $\exp(iaqn)$ et en sommant sur toutes les équations, démontrer que la fonction ϕ obéit à l'équation

$$\frac{\partial^2\phi}{\partial t^2} + 2\omega_0^2(1 - \cos(aq))\,\phi = 0 \tag{3.31}$$

[Help : Utilisez et justifiez $\sum_{n=-\infty}^{+\infty} u_{n\pm1} \exp(iaqn) = \sum_{n=-\infty}^{+\infty} u_n \exp\left(iaq(n \mp 1)\right)$].

2. Vérifiez que la solution de cette équation différentielle linéaire homogène de seconde ordre en t est donnée par

$$\phi(q,t) = A(q)\exp(i\omega_q t) + B(q)\exp(-i\omega_q t)$$

où $\omega_q = 2\omega_0 \sin(aq/2)$. [Help : arranger d'abord l'équation (3.31) en exprimant $\cos(2\theta)$ en fonction de $\sin\theta$]. $A(q)$ et $B(q)$ sont des fonctions du paramètre q, mais pas du paramètre t qui peuvent être déterminées à l'aide des conditions initiales, mais nous n'avons pas besoin de les expliciter ici.

3. (i) Donnez enfin la forme des $u_n(t)$. Comme vous pouvez le constater, les u_n peuvent être considérés comme des superpositions d'ondes planes $\exp\left[i(\omega_q t \pm qna)\right]$. (ii) La vitesse du propagation d'onde est définie par $v_q = d\omega_q/dq$. Donnez l'expression de v_q et tracez la en fonction de q. Que vaut cette vitesse pour $q \to 0$ (grande longueur d'onde) et $q \approx \pm\pi/a$?

4. Généralisation. Pouvez-vous indiquer, en suivant rapidement le même chemin que pour les questions précédentes, ce qui changerait si un atome interagissait avec ses 4 plus proches

voisins ?

$$\begin{aligned} \frac{d^2 u_n}{dt^2} &= \omega_0^2 \left(-2u_n + u_{n+1} + u_{n-1}\right) \\ &+ \alpha\omega_0^2 \left(-2u_n + u_{n+2} + u_{n-2}\right) \end{aligned}$$

4 Les transformations de Fourier.

4.1 Entrée en matière.

Nous avons vu plus haut que si une fonction était définie sur un intervalle *fini* de taille L, elle pouvait être approximée aussi précisément qu'on le veuille par les séries de Fourier $\exp(2i\pi nx/L)$. On peut, pour clarifier la notation, poser $q = 2\pi n/L$ et écrire pour notre fonction :

$$f(x) = \sum_q \widetilde{f}_q \exp(iqx) \tag{4.1}$$

où q varierait par pallier discret de $1/L$. Notez également que les coefficients sont donnés de façon fort symétrique

$$\widetilde{f}_q = (1/L) \int_0^L f(x) \exp(-iqx) dx \tag{4.2}$$

C'est en quelque sorte une formule d'inversion que nous devons à l'orthogonalité. Cela est fort sympathique, mais si on voulait approcher notre fonction sur toute l'intervalle $]-\infty, \infty[$? La réponse simple (simpliste) serait que q deviendrait alors une variable continue, la somme se transforme en une intégrale, et nous avons :

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \widetilde{f}(q) \exp(iqx) dq \tag{4.3}$$

et par analogie avec (4.2), on doit avoir

$$\widetilde{f}(q) = \int_{-\infty}^{\infty} f(x) \exp(-iqx) dx \tag{4.4}$$

On appelle les équations (4.3,4.4) des transformations de Fourier : prenez votre fonction, multipliez par $\exp(-iqx)$, intégrez sur $]-\infty, +\infty[$, et hop, c'est prêt. Notez la beauté de la symétrie : si $\widetilde{f}(q) = \text{TF}[f(x)]$ alors $f(-x) = (1/2\pi)\text{TF}[\widetilde{f}(q)]$. Par convention, on met un petit chapeau sur la TF pour distinguer la fonction originale de la fonction transformée. La variable q est la variable réciproque de x. Pourquoi ce facteur 2π doit apparaître dans la TF inverse ? En faite, si nous définissions la TF par [multiplier par $\exp(2i\pi qx)$ et intégrer] la TF et son inverse seraient parfaitement symétrique, et certaines personnes préfèrent cette dernière définition. Cela nous obligerait cependant à

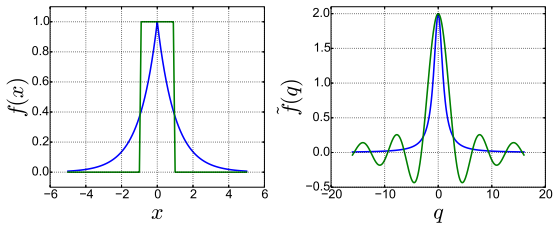


FIGURE 4.1 – Quelques exemples de fonctions et de leurs transformées de Fourier. En bleu, la fonction $\exp(|x|)$; en vert, la fonction $\Pi(x)$.

traîner des 2π lors des dérivations et des changements de variables et nous préférons donc la définition (4.4).

La signification de la TF est la suivante : une fonction $f(x)$ peut être considérée comme la superposition d'oscillations pures $\exp(iqx)$, chaque oscillation ayant un poids $\tilde{f}(q)$. Cette signification, comme nous l'avons dit, est juste une généralisation des séries de Fourier. La TF est un exemple d'opérateur linéaire, c'est à dire une boite noire qui prend une fonction en entrée et produit une nouvelle fonction en sortie, et fait cela de façon linéaire, c'est à dire :

$$\mathrm{TF}[\lambda f(x) + \mu g(x)] \quad = \quad \lambda \mathrm{TF}[f(x)] + \mu \mathrm{TF}[g(x)]$$

Voyons pour l'instant quelques exemples (figure 4.1).

Exemple 4.1 $f(x) = e^{-k|x|}$

Il est facile de démontrer que $\tilde{f}(q) = 2k/(k^2 + q^2)$. La formule d'inversion est un peu plus compliqué à démontrer, et nécessite quelques éléments de la théorie des variables complexes.

Exemple 4.2 $f(x) = \Pi(x)$

La fonction $\Pi(x)$, appelée *porte*, est définie par $\Pi(x) = 0$ si $|x| > 1$ et $\Pi(x) = 1$ si $|x| \leq 1$. Sa TF est donnée par $\tilde{f}(q) = 2\sin(q)/q$. Cette fonction joue un rôle important dans la théorie de la diffraction.

Nous n'avons pas énoncé dans quelle condition les TF existent. Une condition suffisante évidente serait que f doit être sommable. Nous ne rentrons pas plus dans le détail, disons simplement qu'en général, les résultats obtenus sont radicalement aberrants si on a violé les limites permises.

Exercices.

§ 4.1 Démontrer la formule donnée pour la TF de $k\exp(-k|x|)$. Que devient cette transformée si $k \to +\infty$?

§ 4.2 Calculer la TF de la fonction $f(t) = H(t)\exp(-\nu t)$. $H(t)$ est appelé la fonction d'Heaviside (Physicien anglais de la fin du XIXème). Elle est nulle pour $t < 0$ et vaut 1 pour $t \geq 0$.

§ 4.3 Calculer la transformée de Fourier de la fonction $a^{-1}\Pi(x/a)$. Que devient cette transformée quand $a \to 0$?

§ 4.4 Sachant que $\int_{-\infty}^{+\infty} \exp(-x^2)dx = \sqrt{\pi}$, calculez la TF de $\exp(-x^2/2)$. Pour cela, notez que $x^2 + 2iqx = (x + iq)^2 + q^2$. Le résultat d'intégration reste valable si on parcours un axe parallèle à l'axe réel.

§ 4.5 Calculez la TF de $a^{-1}\exp[-x^2/2a^2]$. Que devient cette transformée quand $a \to 0$?

§ 4.6 Calculer la TF d'un "train d'onde", $f(x) = \Pi(x/a)\exp(ik_0x)$. k_0^{-1}est la période de l'onde et a son extension spatial. Discutez les divers limites.

§ 4.7 La fonction d'Airy que l'on rencontre souvent dans les problèmes du calcul des bandes de certain semi conducteur est définie par une intégrale :

$$Ai(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left(i(xt + t^3/3)\right) dt$$

Démontrez que sa Transformée de Fourier est donnée par

$$\tilde{Ai}(k) = e^{ik^3/3}$$

[Help : Vous devez connaître les *distributions*, traités au prochain chapitre. Pour cet exercice, il peut être judicieux d'échanger l'ordre d'intégration pour le calcul de la TF].

4.2 Les opérations sur les TF.

Si les TF sont si intéressante, c'est en partie parce qu'elles nous permettent de manipuler les équations différentielles comme des équations algébriques. Ceci est un peu l'analogue de l'invention des logarithme par monsieur Neper au début des années 1600. Il est difficile de multiplier deux grands nombres. On prend alors le log de chacun (en consultant une table), on *additionne* (au lieu de multiplier) les deux nombres obtenus, et on prend l'antilog du résultat. C'est presque la même chose pour les équations différentielles et les TF : on prend la TF des équations (parfois en consultant une table), on résout l'équation *algébrique* correspondante, on prend la TF inverse du résultat. Pour cela, nous devons connaître quelques règles de manipulation des TF, l'équivalent des règles comme $\log(ab) = \log(a) + \log(b)$ pour les logarithme. Voyons cela de plus près.

Translation. Si TF$[f(x)] = \tilde{f}(q)$ alors TF$[f(x-a)] = \exp(-iqa)\tilde{f}(q)$

Translater dans l'espace direct revient à multiplier par un facteur de phase dans l'espace réciproque. Le changement de variable $x \to x + a$ (si on remplace x par $x + a$) nous donne la démonstration :

$$\int f(x-a)e^{-iqx}dx = e^{-iqa}\int f(x)e^{-iqx}dx$$

Inversion. Si TF$[f(x)] = \tilde{f}(q)$, alors TF$[f(-x)] = \tilde{f}(-q)$

Changement d'échelle. Si TF$[f(x)] = \tilde{f}(q)$, alors TF$[f(x/a)] = a\tilde{f}(qa)$

La dilatation d'échelle dans l'espace direct conduit à la contraction d'échelle dans l'espace réciproque. Cela se démontre par le changement de variable $x \to ax$, comme ce que vous avez fait dans les exercices 1 et 3 ci-dessus. Nous avons en réalité supposé que $a > 0$. Dans le cas général, on doit écrire TF$[f(x/a)] = |a|\tilde{f}(qa)$.

Dérivation. Si TF$[f(x)] = \tilde{f}(q)$, alors TF$[df(x)/dx] = iq\tilde{f}(q)$.

Dériver dans l'espace direct revient à multiplier par iq dans l'espace réciproque. C'est là le grand avantage qui permet de transformer les équadifs en équation algébrique dans l'espace réciproque. Pour démontrer cela, il faut simplement effectuer une intégration par partie, et noter que puisque f est sommable, $f(x) \to 0$ quand $x \to \pm\infty$.

4.3 Transformée de Fourier Rapide.

Une des raisons qui a grandement popularisé les TF est la disponibilité, depuis le début des années 1960, des algorithmes qui permettent de les calculer efficacement. La première étape pour traiter numériquement un signal est de l'échantillonner, c'est à dire de le mesurer et de l'enregistrer tous les pas de temps Δt. Le son sur un CD est par exemple échantillonné à 48 KHz, c'est à dire 48000 enregistrement de l'amplitude par seconde. Nous sommes alors en possession de N nombres (qui sont les $f_n = f(n\Delta t)$). Normalement, si on voulait calculer la TF, on devrait effectuer N^2 opérations (de multiplications et d'addition). Les transformés de Fourier Rapide (ou FFT, pour *Fast Fourier Transform* en anglais) n'effectuent pour ce calcul que $N \log N$ opérations. La différence est énorme en temps de calcul. Par exemple, en supposant que notre ordinateur effectue un milliard d'opérations par seconde, la TF d'une seconde d'un CD prendrait environ 2 secondes, tandis que sa TFR ne prendrait que 0.5 ms. C'est cette différence qui permet d'analyser le signal en "temps réel".

4.4 Manipulation et utilisation des TF.

Filtrage.

Un filtre ne laisse passer que certaines fréquences. Par exemple, pour la réception radio de France Info, on règle un circuit électrique pour ne laisser passer que le 105.5 MHz. En optique, on fait souvent un filtrage spatial pour "nettoyer" un faisceau laser et enlever les *speckles*. Le principe est toujours le même : nous avons un signal $x(t)$ en entrée et un signal $y(t)$ en sortie. Dans le cas d'un circuit RLC, ils sont reliés par une

équation différentielle

$$\frac{d^2y}{dt^2} + \alpha\frac{dy}{dt} + \omega_0^2 y = x(t)$$

Une habitude veut que la variable réciproque soit notée q (ou k) quand la variable directe est x, et ω (ou ν) quand la variable directe est t. En prenant la TF des deux côtés de l'équation, on obtient

$$\tilde{y}(\omega) = \frac{\tilde{x}(\omega)}{(-\omega^2 + i\alpha\omega + \omega_0^2)} \tag{4.5}$$

Le signal en entrée $x(t)$ est la superposition d'oscillations pures $\exp(i\omega t)$, chaque oscillation ayant un poids $\tilde{x}(\omega)$. L'équation (4.5) montre comment le poids de chacune de ces oscillations est modifié en sortie. Le signal (temporel) en sortie est la superposition de ces oscillations avec le poids $\tilde{y}(\omega)$. L'amplitude du poids de la fréquence ω en entrée est donc divisée par $[(\omega^2-\omega_0^2)^2+\alpha^2\omega^2]^{1/2}$. Chaque composante de sortie subit également un déphasage $\phi = \arctan[(\omega_0^2 - \omega^2)/\alpha\omega]$.

Il existe bien sûr autant de filtre que de problème à traiter. Les images issues de la microscopie électronique sont souvent brouillées par des pixels aléatoires. Pour nettoyer ces images, on filtre les hautes fréquences : on prend la TF de l'image (c'est une TF à deux dimensions) et on "coupe" les hautes fréquences, en multipliant la TF par une fonction d'Heaviside $H(q_0 - q)$ où q_0 est la fréquence (spatiale) de coupure. On prend alors la TF inverse et l'image résultantes a été nettoyé du bruit aléatoire. Bien sûr, dans l'opération, on a aussi perdu peut-être quelques informations. L'opération peut-être résumé comme suit : $I_n(\mathbf{x}) = \mathrm{TF}^{-1}[H(q_0 - q)\mathrm{TF}[I(\mathbf{x})]\,]$.

TF de $H(t)\cos(\omega_0 t)$.

Nous souhaitons calculer la TF de la fonction $f(t) = H(t)\cos(\omega_0 t)$ où $H(t)$ est la fonction d'Heaviside. Cela paraît à priori problématique, la fonction $\cos(\omega_0 t)$ ne tendant pas vers zéro pour $t \to +\infty$. Calculons plutôt la TF de la fonction $f_\nu(t) = H(t)\exp(-\nu t)\cos(\omega_0 t)$. Pour $\nu > 0$, cette fonction converge rapidement vers zéro et son intégrale est très bien définie. Donc,

$$\begin{aligned}\tilde{f}_\nu(\omega) &= \frac{1}{2}\int_0^\infty \left(e^{-(\nu - i\omega_0)t} + e^{-(\nu+i\omega_0 t)}\right) e^{-i\omega t}dt \\ &= \frac{\nu + i\omega}{(\nu + i\omega)^2 + \omega_0^2}\end{aligned}$$

Maintenant, si on prend la limite $\nu \to 0$, nous voyons que $f_\nu(t) \to f(t)$ (pas uniformément), et que la transformée de Fourier tend également vers une limite bien définie. Nous posons donc :

$$\tilde{f}(\omega) = i\frac{\omega}{\omega_0^2 - \omega^2}$$

Bien sûr, si on voulait prendre la TF inverse, on aurait à nouveau des problèmes pour l'intégration autour des singularités $\omega = \pm\omega_0$. On s'en sort en prenant la *valeur principale* des intégrales. Quelques connaissances de la théorie d'intégration dans le plan complexe nous montre alors qu'on trouve bien le bon résultat [1]. Le lecteur peut démontrer, en suivant une démarche analogue, que

$$\mathrm{TF}\left[H(t)\sin(\omega_0 t)\right] = \frac{\omega_0}{\omega_0^2 - \omega^2}$$

Théorie de la diffraction de la lumière et de la formation d'image.

Considérons un rayon de lumière qui se propage d'un point A à un point B . Si la phase du champs au point A est $\exp(i\omega_0 t)$, elle est de $\exp(i\omega_0 t + \phi)$ au point B. ω_0 est (à 2π près) la fréquence de la lumière (de l'ordre de $10^{14}\mathrm{s}^{-1}$ pour la lumière visible) et ϕ est le déphasage dû au temps que la lumière met pour aller de A à B (distant de l) :

$$\phi = \omega_0 \Delta t = 2\pi f \frac{AB}{c} = 2\pi \frac{l}{\lambda}$$

où λ est la longueur d'onde de la lumière (entre 0.3 et 0.8 micron pour la lumière visible). Le lecteur connaît sans doute tout cela depuis le premier cycle universitaire.

Chaque point d'un objet recevant une onde lumineuse peut être considéré comme une source secondaire. Si $a\exp(i\omega_0 t)$ est le champs qui arrive au point P, le champs émis est $ar\exp(i\omega_0 t + i\phi)$. Le coefficient r (≤ 1) désigne l'absorption de la lumière au point P. Le coefficient ϕ est le déphasage induit au point P si par exemple en ce point, le matériaux a un indice différent de son environnement. Le coefficient complexe $T = r\exp(i\phi)$ est le coefficient de transmission du point P. Un objet est donc caractérisé par une fonction complexe $f(\mathbf{x})$ qui est son coefficient de transmission pour chacun de ses points $\mathbf{x}$.

Considérons maintenant une onde plane arrivant sur un objet (qui pour plus de simplicité, nous considérons unidimensionnel) et un point P à l'infini dans la direction θ (Fig. 4.2(a)). Le champ reçu en ce point est la somme des champs secondaire émis par les divers points de l'objet. Par rapport au rayon OP que l'on prend comme référence, le rayon AP aura un déphasage de $\phi = -2\pi AA'/\lambda = -(2\pi/\lambda)x\sin(\theta)$. En appelant $q = (2\pi/\lambda)\sin(\theta)$, et en appelant $f(x)$ la fonction de transmission de l'objet, nous voyons que le champs au point P vaut

$$\int f(x)\exp(-iqx)dx$$

qui n'est rien d'autre que la TF de la fonction f.

Mettons maintenant une lentille à une distance U en face de l'objet, une image se formera dans un plan à distance V de la lentille (Fig. (b)). Les rayons qui partaient dans

1. Cela est en dehors du champs de ce cours.

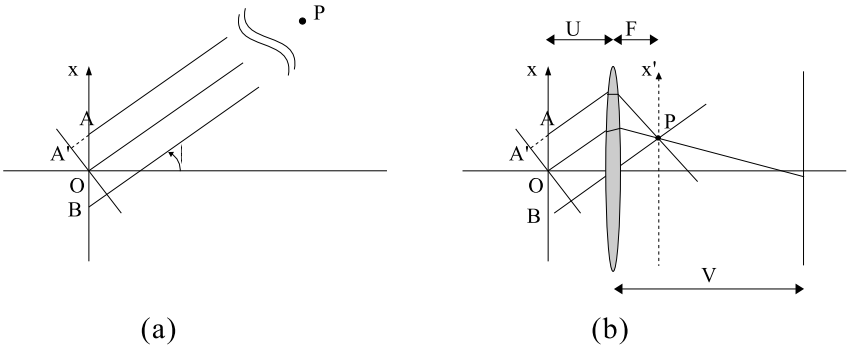


FIGURE 4.2 – Formation d'image vu comme une double transformée de Fourier.

la direction θ vont maintenant se focaliser dans le plan focal arrière de la lentille (distant de F) en un point P dont la coordonnée x' vaut $F \tan\theta \approx F \sin\theta$ tant que l'angle θ n'est pas trop important. Nous en déduisons l'intensité du champ $g(x')$ dans le plan focal arrière de la lentille :

$$g(x') = \int f(x) \exp\left(-i(\frac{2\pi}{\lambda F})x'.x\right) dx = \tilde{f}(\frac{2\pi}{\lambda F}x')$$

Il n'est pas trop difficile de démontrer que l'image formée est la TF du plan focal arrière, nous laissons cela au soin du lecteur. La formation d'image peut donc être vu comme une double transformation de Fourier. Cela ouvre de grandes perspectives pour effectuer des opérations de filtrage directement dans le plan focal arrière (pfa) d'une lentille. Voir des objets transparents, comme par exemple des cellules dans l'eau n'est pas possible en microscopie classique. Zerniké, dans les années 1950, a inventé une technique appelée contraste de phase, qui consiste à introduire des filtres dans le pfa de l'objectif et permet la visualisation des objets transparents sous microscope.

Énergie injecté dans un oscillateur.

Soit une particule dans un puits harmonique ($E_p = (1/2)kx^2$) soumis à une force extérieure $F(t)$. Nous désirons savoir quelle énergie cette force transfert à la particule. L' équation du mouvement s'écrit :

$$\frac{d^2x}{dt^2} + \omega_0^2 x = (1/m)F(t) \tag{4.6}$$

où $\omega_0^2 = k/m$ est la fréquence propre d'oscillation de la particule. Nous supposons qu'au temps T_1du début, l'oscillateur est au repos. L'énergie totale transférée à l'oscillateur est donc la somme de son énergie cinétique et potentielle au bout d'un temps T_2 (que nous prendront égale à $+\infty$ par la suite).

Notons tout de suite que la gauche de l'équation (4.6) peut s'écrire $(d/dt - i\omega_0)(d/dt + i\omega_0)x$. Comme nous allons voir, cette décomposition a son utilité. En mécanique quantique, on appellerai l'analogue de ces termes des opérateurs de création et d'annihilation qui sont fréquemment utilisé. Par ailleurs, H, L'énergie totale [2] du système (cinétique + potentielle), s'écrit :

$$\begin{aligned}(2/m)H &= (dx/dt)^2 + \omega_0^2 x^2 \\ &= (dx/dt - i\omega_0 x)(dx/dt + i\omega_0 x)\end{aligned}$$

Si on pose $z = dx/dt + i\omega_0 x$, nous aurons alors $(2/m)H = zz^*$, et l'équation (4.6) se transforme en

$$dz/dt - i\omega_0 z = (1/m)F(t) \tag{4.7}$$

L'énergie transférée à l'oscillateur est $\Delta E = H(T_2) - H(T_1) = H(T_2)$. Multiplions maintenant les deux cotés de l'équation (4.7) ci-dessus par $\exp(-i\omega_0 t)$ et intégrons entre T_1et T_2

$$\int_{T_1}^{T_2} (dz/dt - i\omega_0 z)e^{-i\omega_0 t}dt = (1/m)\int_{T_1}^{T_2} F(t)e^{-i\omega_0 t}dt$$

Il nous suffit maintenant d'effectuer une intégration par partie du côté gauche de l'intégrale et d'utiliser le fait que l'oscillateur est au repos à l'instant T_1pour trouver que ce côté vaut $z(T_2)\exp(-i\omega_0 T_2)$. Comme en plus l'oscillateur est au repos avant T_1, on peut étendre l'intégrale à $-\infty$. Quand $T_2 \to +\infty$, le côté droit devient égale à la TF de F évaluée pour la fréquence ω_0, et nous avons

$$\Delta E = \frac{1}{2m}\tilde{F}(\omega_0)\tilde{F}^*(\omega_0)$$

Pour connaître l'énergie totale transférée à l'oscillateur, nous n'avons pas à résoudre l'équation différentielle de second ordre avec second membre, évaluer simplement la TF de la Force appliquée à la fréquence propre de l'oscillateur nous suffit.

Exercices. Vous pouvez donc facilement calculer l'énergie transférée dans les cas suivants :

§ 4.8 $F(t) = f_0 e^{-t/t_0}$ si $t \geq 0$; sinon, $F(t) = 0$.

§ 4.9 $F(t) = f_0 \Pi(t/t_0)$

§ 4.10 $F(t) = f_0$ si $t \geq 0$; sinon, $F(t) = 0$

§ 4.11 $F(t) = f_0 \cos(\omega_1 t)$

Dans les cas 4.8 et 4.9, discutez le transfert d'énergie en fonction du temps t_0. Pour résoudre le cas 4.10 et 4.11, vous aurez besoin des résultats sur les distributions disponible dans les prochains chapitres.

2. En mécanique analytique, on appelle Hamiltonien l'énergie totale du système, d'où le H. Comme le système n'est pas isolé, H n'est pas une constante du mouvement : $H = H(t)$

4.5 Relation entre les séries et les transformés de Fourier.

Nous avons indiqué au début du chapitre, sans le démontrer, que l'on passe des séries au transformés de Fourier en laissant la longueur de l'intervalle L tendre vers l'infini. Revoyons ce passage avec quelques détails maintenant. Considérons une fonction $f(x)$ sur l'intervalle $[-L/2, L/2]$. Ses coefficients de Fourier (complexe) sont donnés par

$$c_q = \frac{1}{L}\int_{-L/2}^{L/2} f(x)e^{-iqx}dx$$

où pour plus de simplicité, nous notons $q = 2\pi n/L$. Désignons par $I(q)$ l'intégrale ci-dessus (sans le facteur $1/L$ donc). Par définition, nous avons pour $f(x)$:

$$f(x) = \sum_{q,2\pi/L} (1/L)I(q)e^{iqx}$$

où dans la somme, l'indice q varie par pas discret $dq = 2\pi/L$. Quand $L \to \infty$, $dq \to 0$ et par définition de l'intégrale de Riemann, la somme ci-dessus tend vers

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} I(q)e^{iqx}dq$$

Par ailleurs, il est évident que quand $L \to \infty$, $I(q)$ tend vers $\tilde{f}(q)$ donnée par l'équation (4.4).

4.6 Approfondissement : TF à plusieurs dimensions.

Nous nous occupons dans ce cours essentiellement des fonctions à une seule variable $f(x)$. Ces concepts cependant se généralisent sans problème à plusieurs variable. Si nous notons les variables de façon vectorielle

$$\mathbf{x} = (x_1, x_2, ..., x_d)$$

alors la TF est définie comme

$$\tilde{f}(\mathbf{k}) = \int_{\mathbb{R}^d} f(\mathbf{x})e^{-i\mathbf{k}.\mathbf{x}}d\mathbf{x} \tag{4.8}$$

où $\mathbf{k}.\mathbf{x} = k_1x_1 + ...k_dx_d$ désigne le produit scalaire. De même, la TF inverse est

$$f(\mathbf{x}) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} f(\mathbf{k})e^{i\mathbf{k}.\mathbf{x}}d\mathbf{k}$$

Étudions deux cas particuliers que nous rencontrons souvent.

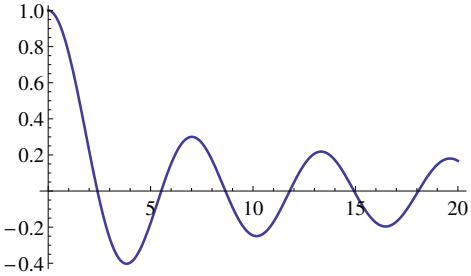


FIGURE 4.3 – La fonction $J_0(x)$

4.6.1 Symétrie cylindrique.

Dans le premier cas, $d = 2$ et la fonction est à symétrie cylindrique : $f(\mathbf{x}) = f(r)$ où $r^2 = x_1^2 + x_2^2$. La fonction $f(x, y) = 1/(x^2 + y^2)$ en est un bon exemple. Il est évident que dans un tel cas, nous avons intérêt à utiliser les coordonnées polaire. Dans ce cas,

$$\tilde{f}(\mathbf{k}) = \int_0^\infty \int_0^{2\pi} f(r) e^{-i\mathbf{k}.\mathbf{x}} r dr d\theta$$

Le problème consiste à calculer le facteur $\exp(-i\mathbf{k}.\mathbf{x})$. La coordonnées θ désigne l'angle entre le vecteur $\mathbf{x}$ et l'axe des x. Cependant, si nous tournons l'axe des x pour l'aligner sur l'axe de $\mathbf{k}$, la fonction $f(\mathbf{x})$ ne change pas, puisqu'elle est de symétrie cylindrique. Dans ce cas, θ désigne l'angle entre le vecteur $\mathbf{x}$ et le vecteur $\mathbf{k}$ et

$$\mathbf{k}.\mathbf{x} = kr\cos\theta$$

où $k = \sqrt{k_1^2 + k_2^2}$

L'intégration sur θ nous donne

$$\int_0^{2\pi} e^{-ikr\cos\theta} d\theta = 2\pi J_0(kr)$$

La fonction $J_0(z)$ est appelée la fonction de Bessel d'ordre 0 et apparaît en physique mathématique à de nombreux endroits. Les fonctions de Bessel jouent le rôle des fonctions trigonométrique en coordonnées polaires. Nous avons donc

$$\tilde{f}(\mathbf{k}) = 2\pi \int_0^\infty r f(r) J_0(kr) dr \tag{4.9}$$

Il est évident que la fonction $\tilde{f}(\mathbf{k})$ est également de symétrie cylindrique et la TF inverse est

$$f(r) = \frac{1}{2\pi} \int_0^\infty k \tilde{f}(k) J_0(kr) dk$$

4.6.2 Symétrie sphérique.

Étudions maintenant le cas $d = 3$ où la fonction est à symétrie sphérique. En répétant les arguments ci-dessus pour les coordonnées sphériques, nous aboutissons à

$$\tilde{f}(\mathbf{k}) = \int_0^\infty \int_0^\pi \int_0^{2\pi} f(r) e^{-i\mathbf{k}.\mathbf{x}} r^2 \sin\theta dr d\theta d\phi$$

L'intégration sur ϕ nous donne juste un facteur 2π. Par ailleurs,

$$\int_0^\pi e^{-ikr\cos\theta} \sin\theta d\theta = 2\frac{\sin(kr)}{kr}$$

et donc finalement

$$\tilde{f}(\mathbf{k}) = \frac{4\pi}{k} \int_0^\infty r f(r) \sin(kr) dr$$

En faisant le même chemin pour la TF inverse, nous avons

$$f(r) = \frac{-1}{\pi r} \int_0^\infty k \tilde{f}(k) \sin(kr) dk$$

§ 4.12 Calculer la Transformée de Fourier de

$$f(\mathbf{x}) = \frac{1}{\sigma^{d/2}} \exp\left(-\frac{r^2}{4\sigma}\right)$$

aux dimensions $d = 1, 2, 3$.

5 Les distributions.

5.1 Ce qu'il faut savoir.

Les transformées de Fourier nous posent quelques problèmes quant à la définition d'une base orthogonale. Nous avons vu que, sur l'intervalle $]-\infty,+\infty[$, la fonction $f(x)$ peut être représentée comme la superposition des fonctions $\exp(iqx)$ avec le poids $\tilde{f}(q)$. Pour en revenir à notre image de base dans l'espace des fonction, les fonctions $e^{iq(.)}$ ($q \in \mathbb{R}$) forment une base, et les coefficients $\tilde{f}(q)$ sont les projections du vecteur $f(.)$ sur les vecteurs de cette base.

Nous avions aux sections précédentes basé nos démonstrations sur le concept d'orthogonalité. Mais peut on dire que $\exp(iq_1x)$ et $\exp(iq_2x)$ sont orthogonales ? Le produit scalaire a t' il encore un sens ? En effet, comment définir la valeur de

$$\int_{-\infty}^{+\infty} \exp(iq_1x)\exp(-iq_2x)dx \tag{5.1}$$

qui au sens normal de la théorie d'intégration, n'a aucun sens ?

Il faut comprendre Le produit scalaire (5.1) dans le sens suivant d'un passage à la limite :

$$\lim_{L\to\infty} \frac{1}{L}\int_{-L/2}^{+L/2} \exp(iq_1x)\exp(-iq_2x)dx \tag{5.2}$$

Quand $q_1 \neq q_2$, l'intégrale est au plus de l'ordre de 1 et le $(1/L)$ fait tout tendre vers zéro. Par contre, si $q_1 = q_2$, l'intégrale vaut L et l'expression (5.2) vaut 1. Le produit scalaire (5.1) est donc de l'ordre de L (avec $L \to \infty$) pour $q_1 = q_2$ et de l'ordre de 1 sinon. Nous noterons ce genre d'objet $\delta(q_1 - q_2)$ et nous l'appellerons le *delta* de Dirac, du nom du physicien qui a établi les règles de manipulation de ces objets dans son livre sur la mécanique quantique en 1930.

Pour un physicien, le concept de la *fonction* δ est très intuitif et généralise le concept de charge ou de masse ponctuel. Supposez que vous ayez des charges répartit continuellement dans l'espace avec une densité $\rho(\mathbf{x})$ et que vous voulez calculer la charge totale contenue dans une sphère de rayon R autour d'un point. Rien de plus simple,il suffit d'intégrer la densité autour $C = \int_V \rho(\mathbf{x})d\mathbf{x}$. Supposez maintenant que $R \to 0$, c'est à dire que vous prenez des sphères de plus en plus petite autour de votre point. Il est évident qu'il y aura de moins en mois de charge à l'intérieur et que $C \to 0$. C'est vrai, sauf si vous avez placé une charge *ponctuelle* au point considéré. Pour une charge

ponctuelle Q, quel que soit la taille de la sphère autour, la quantité totale de la charge à l'intérieur reste constante. En gros, pour une charge ponctuelle placée en $\mathbf{x}_0$, la densité de charge est *nulle* partout, sauf en $\mathbf{x}_0$ où elle est infinie ! Ce genre de densité infinie en un point, nulle partout et dont l'intégrale est *finie* est justement un "delta" de Dirac. Les mathématiciens nous exécuteraient si on appelait ces objets des fonctions et nous obligent à les nommer des *distributions*. La propriété du delta de Dirac est la suivante :

$$\int_I \delta(x)dx = 1 \qquad \forall I \ni 0 \tag{5.3}$$

Du moment que l'intervalle I contient 0, l'intégrale vaut 1, sinon elle *vaut* zéro. L'objet de ce chapitre est de se familiariser avec les distributions, et en particulier avec la distribution de Dirac.

On peut voir $\delta(x)$ comme un processus de limite. Prenons le cas de la fonction

$$f_a(x) = \frac{1}{a\sqrt{\pi}} e^{-(x/a)^2}$$

C'est une gaussienne centrée sur 0, et son intégrale vaut 1. Quand $a \to 0$, elle devient de plus en plus piquée, avec une extension de moins en moins large, mais l'intégrale reste constante. On peut dire la même chose de la fonction $g_a(x) = (1/2a)\Pi(x/a)$ ou en fait de n'importe quelle fonction qui, lors d'un processus de passage à la limite, réduit son extension, augmente l'amplitude de son pique, et garde son intégrale constante. La distribution $\delta(x)$ est la limite de ce genre de fonction.

La définition (5.3) nous permet quelques généralisations. Par exemple, on peut définir $3\delta(x)$, la distribution dont l'intégrale vaut 3 sur des intervalles contenant 0. On peut même définir $f(x)\delta(x)$ où on suppose $f(x)$ continue en 0. L'intégrale vaut :

$$\int_{-\infty}^{+\infty} f(x)\delta(x)dx = \lim_{\epsilon \to 0} \int_{-\epsilon}^{+\epsilon} f(x)\delta(x) = f(0)$$

Vous pouvez démontrer cela facilement en utilisant la définition de la continuité d'une fonction. Mais intuitivement, cela paraît évident : comme $\delta(x)$ est nulle partout sauf en 0, le multiplier par $f(x)$ revient simplement à le multiplier par $f(0)$. En fait, on utilise cela comme la définition de la distribution $\delta(x)$:

$$\int_I f(x)\delta(x)dx = f(0) \qquad \forall I \ni 0 \tag{5.4}$$

$\delta(x)$ est une distribution centrée sur 0. $\delta(x - x_0)$ est une distribution centrée sur x_0 et $\int \delta(x - x_0)f(x) = f(x_0)$. Finalement, les règles pour manipuler les δ ne sont pas vraiment compliquées.

Une dernière chose à savoir sur $\delta(x)$ est sa transformée de Fourier (on pose dorénavant $R =]-\infty, +\infty[$:

$$\tilde{\delta}(q) = \int_R \delta(x) \exp(-iqx)dx = 1$$

La TF de $\delta(x)$est la fonction constante 1. Cela veut dire que $\delta(x)$ est la superposition, à poids égal, de toutes les modulations $\exp(iqx)$! Cela n'est pas vraiment étonnant : comme $\delta(x)$ varie vraiment très rapidement, toutes les modulations doivent y être présent. Inversement,

$$\frac{1}{2\pi}\int_R \exp(iqx)dq = \delta(x) \tag{5.5}$$

Exercice : Démontrer que la fonction δ *définie* par (5.5) est bien une δ de Dirac, c'est à dire que $\int_R \delta(x)f(x) = f(0)$.

La dimension de la distribution δ. En mathématique, nous manipulons essentiellement des chiffres, c'est à dire des nombre sans dimensions. En physique cependant, les quantités que nous manipulons représentent des grandeurs telles que des longueurs, énergies, temps, vitesses, etc. Ces grandeurs ont des dimensions. Les physiciens attachent beaucoup d'importance à cette question pour plusieurs raisons. Une de ces raisons est purement gramaticale et permet de vérifier la cohérence des divers étapes d'un calcul. Prenons le cas d'une équation différentielle du genre $d^2y/dt^2 + \omega_0^2 y = f_0$ et supposons que nous avons trouvé $y = f_0 \sin(\omega_0 t)$. Si y représente une quantité de dimension $[y]$, alors $[\omega_0] = T^{-1}$ et $[f_0] = [y]T^{-2}$. Ceci est nécessaire si nous voulons que les deux cotés de l'équation aient la même dimension. La dimension du côté gauche de la solution est $[y]$. Comme la fonction $\sin$ n'a pas de dimension, le coté droit de la solution a la dimension de $[y]T^{-2}$! Les deux cotés de la solution n'ont pas la même dimension et nous nous sommes manifestement trompé à une étape de la résolution. Ces vérifications peuvent (et doivent) être effectuées à chaque étape du calcul.

Comme nous manipulerons pas mal les distributions par la suite, nous avons besoin de connaître leurs dimensions. Les fonctions $\sin(x)$ et $\exp(x)$ n'ont pas de dimensions. Qu'en est-il de la distribution $\delta(x)$? Pour cela, il faut d'abord répondre à la question de la dimension de $\int ydx$. Le signe $\int$ n'est qu'une généralisation de l'opération addition. La dimension d' "une pomme + une pomme" est toujours "une pomme". Le signe d signifie "une très petite quantité de" et la dimension d'une "très petite quantité de pomme" est toujours "une pomme". Nous en déduisons de tout cela que $[\int ydx] = [y][x]$

Nous pouvons maintenant utiliser la propriété de la distribution δ : $\int \delta(x)f(x)dx = f(0)$. Il est alors aisé de voir que $[\delta(x)] = [x]^{-1}$! Nous aurions pu bien sûr arriver au même résultat en utilisant l'expression par passage à la limite $\delta(x) = (1/a)\exp(-x^2/a^2)$ quand $a \to 0$. Voilà, il faut avoir cela en tête à chaque fois que l'on veut vérifier la cohérence des équations qui impliquent des δ.

5.2 Un peu de décence.

Du point de vue du mathématicien, ce que nous avons raconté plus haut est, en restant poli, mal propre. Laurent Schwarz, dans les années 1950, a rendu rigoureux la théorie des distributions. Il est utile de connaître les grandes lignes de sa construction. Il est

parti de la constatation que les distributions ne sont utilisées en pratique que sous le signe intégral, comme en (5.4).

Une fonctionnelle est une généralisation d'une fonction. Elle prend *une fonction* en entrée et produit *un scalaire* en sortie. C'est une fonction de fonction en quelque sorte. Notons que la TF *n'est pas* une fonctionnelle, puisqu'elle produit une fonction en sortie.

Appelons $\mathcal{E}$ l'espace des fonctions [1], et $\mathcal{F}$ l'espace des fonctionnelles *linéaires* (ou dit plus sérieusement, des formes linéaires définies sur $\mathcal{E}$). Un exemple de fonctionnelle est $\mathcal{L}_{\exp(-x^2)}$ qui prend une fonction en entrée, calcul son produit scalaire avec la fonction $\exp(-x^2)$, et produit ce chiffre en sortie :

$$\mathcal{L}_{\exp(-x^2)}[f] = \int_R \exp(-x^2) f(x) dx$$

Nous pouvons généraliser cet exemple : à chaque fonction $g \in \mathcal{E}$, nous pouvons associer une fonctionnelle $\mathcal{L}_g \in \mathcal{F}$ tel que

$$\mathcal{L}_g[f] = \int_R f(x) g(x) dx$$

Et nous pouvons démontrer facilement que ce $\mathcal{L}_g$ est bien une fonctionnelle linéaire. Nous pouvons trouver beaucoup d'autres fonctionnelles linéaires. Par exemple, la fonctionnelle δ_{x_0} est définie par

$$\delta_{x_0}[f] = f(x_0)$$

Voilà, le tour est joué. Cette fonctionnelle est bien le delta de Dirac $\delta(x - x_0)$ définie plus haut. Noter bien l'opération : on peut identifier une partie de l'espace $\mathcal{F}$ avec l'espace $\mathcal{E}$ via ces $\mathcal{L}_g$ que nous avions construit : à chaque élément de $\mathcal{E}$ nous pouvons faire correspondre un élément de l'espace $\mathcal{F}$. Mais l'espace $\mathcal{F}$ est plus vaste, et quelques uns de ses éléments supplémentaires constituent les distributions inhabituelles. C'est un peu comme enrichir l'ensemble des nombres rationnels $\mathbb{Q}$ pour arriver à l'ensemble des nombres réel $\mathbb{R}$.

On peut définir des opérations sur les distributions. Il est toujours plus simple de partir des distributions du genre $\mathcal{L}_g$ dont le sens est familier pour définir ensuite les mêmes opérations sur les distributions du genre δ. Par exemple, que veut dire $\mathcal{L}_{g'}$? En intégrant par partie, on trouve

$$\mathcal{L}_{g'}[f] = \int_R g'(x) f(x) dx = -\int_R g(x) f'(x) dx = -\mathcal{L}_g[f']$$

(N'oublions pas que comme f et g sont au moins sommable, elle tendent vers zéro pour $x \to \infty$). On peut donc définir :

$$\delta'_{x_0}[f] = -\delta_{x_0}[f'] = -f'(x_0)$$

1. En réalité, l'espace des fonctions à support borné et infiniment dérivable, mais nous ne sommes pas à notre premier délit.

ou dans le langage moins élégants des physiciens,

$$\int \delta^{'}(x-x_0)f(x)dx = -f'(x_0)$$

De même, nous pouvons démontrer que pour la fonction d'Heaviside $H(x)$, $H'(x) = \delta(x)$. Nous ne continuerons pas plus le développement formel des distributions. Mais la constructions de Schwarz est extrêmement élégante et nous conseillons au lecteur de voir au moins une fois les bases rigoureuses de cette construction.

Nous voyons cependant que l'espace plus large des distributions nous permet de manipuler aisément des objets qui nous semblaient interdit. Une force ponctuelle a un sens. Une discontinuité également. En physique, une fonction ne *peut* pas être discontinue. La densité de l'eau ne saute pas de ρ_l à ρ_v à l'interface liquide–solide, il existe une couche d'épaisseur petite (très petite devant les autres échelle de longueur) où la densité varie continuellement d'une valeur à une autre. La lumière réfléchit par un miroir pénètre sur une petite longueur *dans* le miroir où son intensité décroît exponentiellement et ainsi de suite. Nous pouvons donc caractériser les discontinuité des fonctions par des distributions. Soit la fonction $f(x) = g(x) + \Delta H(x-x_0)$, où la fonction g est une fonction continue et dérivable en x_0. La fonction f par contre, saute de la valeur $g(x_0)$ à x_0^-à $\Delta + g(x_0)$ à x_0^+. Au sens des distributions, la dérivé de f est donnée par $f'(x) = g'(x) + \delta(x-x_0)$. Imaginez donc f' comme une fonction normale, avec une flèche positionnée en x_0.

Exercices.

§ 5.1 En utilisant la définition (5.4), démontrer que l'expression (5.1) égale $\delta(q_1 - q_2)$.

§ 5.2 Que valent $\delta'(x)$ et $\int \delta(x)$?

§ 5.3 Soit une fonction $L-$périodique f. Que vaut sa TF (au sens des distributions) ?

§ 5.4 Une peigne de Dirac est défini par $\Xi(x) = \sum_{n=-\infty}^{+\infty} \delta(x-n)$. C'est comme si nous avions posé un delta de Dirac sur chaque nombre entier. Quelle est la TF de $\Xi(x/a)$?

§ 5.5 Démontrer que $\delta(x+a) = \delta(x) + a\delta'(x) + (1/2)a^2\delta''(x) + ...$ On peut faire un développement de Taylor des δ comme pour les fonctions usuelles. Pour pouvoir démontrer cette égalité, appliquer les deux côtés de l'égalité à une fonction f

§ 5.6 Démontrer que $\delta(-x) = \delta(x)$ et $\delta(ax) = (1/|a|)\delta(x)$.

§ 5.7 Considérons une fonction $g(x)$ avec un zéro simple en $x_0 : g(x_0) = 0, g'(x_0) \neq 0$. Prenons un intervalle $I = [x_0 - a, x_0 + a]$ autour de x_0(on peut supposer a aussi petit que l'on veut). En développant g autour de sa racine à l'ordre 1, démontrez que

$$\int_I \delta(g(x))f(x)dx = \frac{1}{|g'(x_0)|}f(x_0)$$

§ 5.8 En supposant que la fonction $g(x)$ n'a que des racines simples, et en utilisant le résultat ci-dessus, démontrer :

$$\delta\left(g(x)\right) = \sum_i \frac{1}{|g'(x_i)|}\delta(x - x_i)$$

où les x_i sont les racines simples de $g(x)$. Donner l'expression de $\delta(x^2 - a^2)$.

§ 5.9 En vous inspirant du résultat de la question 5, pouvez-vous indiquer pourquoi dans la question 7, nous pouvions nous restreindre à un développement d'ordre 1 ?

5.3 Manipulation et utilisation des distributions.

Oscillateur soumis à une force périodique. Il obéit à l'équation $d^2x/dt^2 + \omega_0^2 x = A\exp(i\omega_1 t)$. En prenant la TF des deux côtés, nous avons :

$$\tilde{x}(\omega) = \frac{2\pi A\delta(\omega - \omega_1)}{\omega_0^2 - \omega^2}$$

comme $x(t) = (1/2\pi)\int \tilde{x}(\omega)\exp(i\omega t)d\omega$, nous trouvons

$$x(t) = \frac{A\exp(i\omega_1 t)}{\omega_0^2 - \omega_1^2}$$

Nous connaissions ce résultat depuis l'exercice sur le filtrage.

Oscillateur amorti soumis à une force impulsionnelle. Un oscillateur amorti soumis à une force $F(t)$ obéit à l'équation

$$m\frac{d^2y}{dt^2} + \lambda\frac{dy}{dt} + ky = F(t)$$

Nous souhaitons connaître la réponse de l'oscillateur à une force impulsionnelle $F(t) = F_0\delta(t)$. Ceci est l'idéalisation d'un coup de marteau très bref et très puissant sur l'oscillateur. Pour simplifier le problème, nous supposons dans un premier temps que la masse est négligeable (que les forces d'inertie sont petites devant les forces de frottement) et que l'oscillateur est au repos. En renormalisant nos coefficients, l'équation prend la forme :

$$\frac{dy}{dt} + \nu y = f_0\delta(t) \tag{5.6}$$

et en prenant la TF des deux cotés, nous trouvons que $\tilde{y}(\omega) = f_0/(\nu + i\omega)$. Il suffit maintenant de prendre la TF inverse. Il se trouve que dans ce cas, si l'on se souvient de l'exercice (4.1 :4.2), nous pouvons directement écrire

$$y(t) = f_0 H(t)\exp(-\nu t) \tag{5.7}$$

Ce résultat est représenté sur la figure (5.1). Nous suggérons au lecteur de discuter les limites $\lambda \to 0$ et $\lambda \to \infty$.

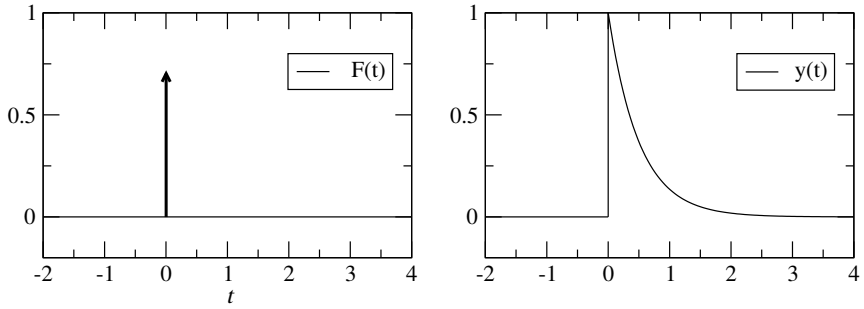


FIGURE 5.1 – Reponse d'un oscillateur amortie à une force impulsionnelle. La distribution δ est représentée par une flèche verticale.

Équation de la chaleur avec une source ponctuelle. Une goutte d'encre extrêmement concentrée, déposée en un point de l'espace va se diluer par diffusion. La même chose est valable pour un pulse ponctuel de chaleur. Comme nous l'avons vu précédemment, les phénomènes de diffusion sont gouvernés par l'équation de la chaleur :

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + Q(x,t) \tag{5.8}$$

où u désigne la température ou la concentration et Q est un terme de source. Dans le problème qui nous intéresse ici, $Q(x,t) = Q_0\delta(x)\delta(t)$. En prenant la TF par rapport à la variable d'espace x, nous avons :

$$\partial_t \tilde{u}(q,t) + Dq^2\tilde{u}(q,t) = Q_0\delta(t) \tag{5.9}$$

Mais cette équation est exactement eq.(5.6), celle qu'on a écrit pour l'oscillateur amorti. C'est bien une équation différentielle *ordinaire* par rapport à la variable temps, et q peut être considérée comme une constante : Pour chaque mode q, nous avons une EDO indépendante. La solution est donc analogue à (5.7), et s'écrit :

$$\tilde{u}(q,t) = Q_0 H(t)\exp(-Dq^2 t)$$

Il nous suffit maintenant de prendre la TF inverse pour obtenir la solution dans l'espace direct :

$$\begin{aligned} u(x,t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\tilde{u}(q,t).e^{iqx}dq \\ &= \frac{Q_0}{2\sqrt{\pi}}\frac{1}{\sqrt{Dt}}\exp\left(-\frac{x^2}{4Dt}\right) \end{aligned}$$

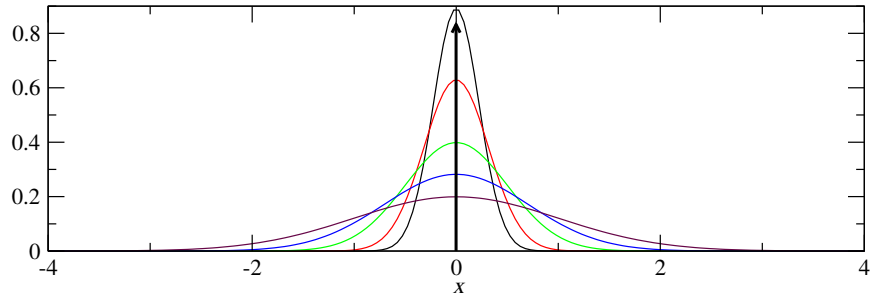


FIGURE 5.2 – profil de concentration en fonction de x, à différent temps $t = 0.1, 0.2, 0.5, 1, 2$. Ici, $D = 1/4$. La distribution originale, en $\delta(x)$, est représenté par une flèche verticale.

La dernière intégrale s'obtient facilement par les techniques que nous avons déjà utilisé. L'évolution de $u(x)$ pour différente valeur de t est représentée sur la figure (5.2).

Extension (difficile) : si la source n'est pas ponctuelle dans le temps, mais seulement dans l'espace, *i.e.* $Q(x) = Q_0\delta(x)$, quel est le comportement de la solution ? Help : Essayez comme avant d'obtenir une expression pour $\tilde{u}(q,t)$. Cette expression est trop compliquée pour inverser, mais $\partial_t\tilde{u}(q,t)$ l'est beaucoup moins. En changeant alors l'ordre des opération TF^{-1} et ∂_t, vous pouvez obtenir une expression pour $\partial_t u(x,t)$. Il vous suffit alors d'évaluer

$$u(x,\tau) = \int_0^\tau \partial_t u(x,t)dt$$

Il n'est pas difficile alors d'obtenir le comportement asymptotique de u pour $t \to \infty$.

Équation d'onde avec source ponctuelle. Considérons une corde tendue infinie et au repos à l'instant initial. A l'instant $t = 0$, on la soumet à une force ponctuelle dans le temps et dans l'espace (l'idéalisation d'un marteau de piano tapant sur la corde). l'équation d'onde s'écrit

$$\frac{\partial^2 u}{\partial x^2} - v^2\frac{\partial^2 u}{\partial x^2} = \gamma\delta(x)\delta(t) \tag{5.10}$$

En suivant la même démarche que ci-dessus, on peut obtenir la propagation de l'onde. on peut également montrer que l'extension du domaine ou $u \neq 0$ croit à la vitesse v.

Vitesse de phase, vitesse de groupe. Donnons nous un signal $u(x,t)$ qui se propage (Fig.5.3). Comment devrait on définir la vitesse du signal ? On pourrait par exemple repérer le maximum de u et suivre ce point en fonction de temps ; ceci n'est pas très bon

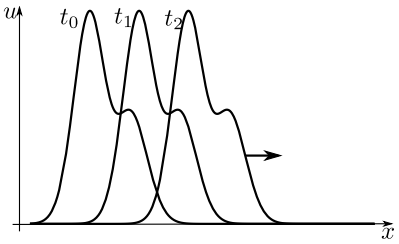


FIGURE 5.3 – un signal $u(x,t)$ en fonction de x à trois temps t différents, se propageant vers la droite.

cependant, puisque le signal peut se déformer et notre maximum disparaître ou d'autres maxima apparaître. Nous devons définir la vitesse en prenant en compte l'ensemble du signal. Une bonne définition est par exemple de suivre le barycentre du signal, ou même mieux, le barycentre du carré du signal pour éviter les compensations de signe :

$$\bar{x}(t) = \int_I x u^2(x,t) dx$$

Notons que dans la plupart des exemples physique, le *carré* du signal est relié au concept de l'énergie. Par la suite, sans perte de généralité, nous supposons notre signal normé : $\int_I u^2(x,t)dx = 1$. Supposons par exemple que notre signal se propage sans se déformer $u(x,t) = u_0(x-ct)$ et nous avons alors

$$\begin{aligned} \bar{x}(t) &= \int_I x u_0^2(x) dx + ct \int_I u_0^2(x) dx \\ &= \bar{x}_0 + ct \end{aligned}$$

ce qui correspond bien à notre intuition de la vitesse d'un signal.

En utilisant la définition des TF et de la distribution $\delta(x)$, il est facile de démontrer que (cf exercice 5.17) :

$$\bar{x}(t) = \frac{i}{2\pi} \int_I \left(\frac{\partial \tilde{u}(q,t)}{\partial q} \right) \tilde{u}^*(q,t) dq \tag{5.11}$$

où $\tilde{u}(q,t)$ est la TF de $u(x,t)$ par rapport à x. Reprenons à nouveau notre signal qui se propage sans se déformer : $u(x,t) = u_0(x-ct)$. Par la règle des manipulation des TF, nous savons qu'une translation dans l'espace direct revient à multiplier par une exponentielle complexe dans l'espace réciproque :

$$\tilde{u}(q,t) = \tilde{u}_0(q).e^{-iqct}$$

En remplaçant dans l'expression (5.11), nous voyons que cela nous donne

$$\begin{aligned}\bar{x}(t) &= \bar{x}_0 + \frac{ct}{2\pi}\int_I \tilde{u}_0(q)\tilde{u}_0^*(q)dq \\ &= \bar{x}_0 + ct\end{aligned}$$

puisque, par la relation de Parseval (cf §5.18), $(1/2\pi)\int_I \tilde{u}_0(q)\tilde{u}_0^*(q)dq = \int_I u_0^2(x)dx$.

En général le facteur qui multiplie le temps dans l'exponentiel complexe est appelé la fréquence angulaire ω, qui dans ce cas simple de signal se propageant sans déformation s'écrit

$$\omega = cq$$

et nous voyons que nous pouvons définir la vitesse comme

$$c = \frac{d\omega}{dq}$$

Prenons maintenant le cas plus général de signaux se déformant en se propageant. Il existe un cas très important appelé milieu dispersif, où la déformation du signal prend une forme simple *dans l'espace réciproque* :

$$\tilde{u}(q,t) = \tilde{u}_0(q)e^{-i\omega(q)t} \tag{5.12}$$

c'est à dire que le mode q est pondéré par un facteur de phase $\omega(q)t$ au temps t, avec une forme $\omega(q)$ quelconque [2], sans plus nécessairement être proportionnel au mode q. Dans le cas d'un cristal par exemple, on peut démontrer (voir le problème correspondant au chapitre sur les séries de Fourier) que $\omega(q) = A\sin(q)$. Nous pouvons néanmoins calculer la vitesse du barycentre du signal comme avant :

$$\bar{x}(t) = \bar{x}_0 + \frac{t}{2\pi}\int_I \left(\frac{d\omega}{dq}\right)\tilde{u}_0(q)\tilde{u}_0^*(q)dq$$

Si $\omega(q)$ varie de façon lente par rapport à $\tilde{u}_0(q)\tilde{u}_0^*(q)$, et que ce dernier possède un pic étroit en q_0, alors une bonne approximation pour la vitesse du barycentre serait

$$c = \left.\frac{d\omega}{dq}\right|_{q=q_0}$$

Ceci est ce qu'on appelle la vitesse du groupe. L'expression ω/q, ayant un sens pour les signaux se propageant sans déformation, s'appelle la vitesse de phase. Une mauvaise compréhension de ses formules peut parfois conclure à des transmissions surluminalles.

2. Notez que dans un milieu dispersif, la quantité $\int_I u^2(x,t)dx$ *se conserve*. Si cette quantité est proportionnelle à l'énergie, nous concluons que dans les milieux dispersif, l'énergie se conserve. Nous pouvons conclure à l'inverse que dans tout milieu non-dissipatif, la forme du signal change nécessairement selon la relation (5.12).

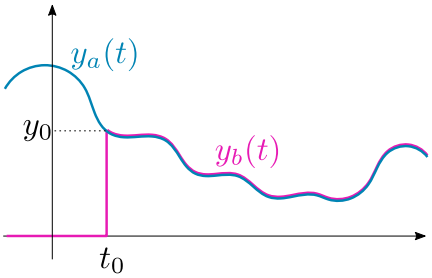


FIGURE 5.4 – La fonction de Green $y_b(t)$ d'une équation différentielle du type (5.13-5.14) comparée à sa solution exacte.

5.4 Les distributions et les conditions initiales des équations différentielles.

Les distributions, en particulier les $\delta(t)$ et leurs dérivées, sont un moyen d'intégrer les conditions initiales directement dans l'équation différentielle. La solution de ces dernières est alors appelée "la fonction de Green", auquel le chapitre 8 est consacré.

Considérons une équation différentielle de première ordre sur la fonction $y(t)$:

$$y' + F(y,t) = 0 \tag{5.13}$$
$$y(t_0) = y_0 \tag{5.14}$$

où $F(.,.)$ est une fonction quelconque. L'équation (5.14) constitue la condition initiale de l'équation différentielle (5.13). Comme vous le constatez, nous devons écrire l'équation sur deux lignes, précisant indépendamment la relation différentielle et la valeur à l'origine.

Appelons $y_a(t)$ la solution de l'équation (5.13) muni de la relation (5.14). Considérons maintenant une autre équation

$$y' + F(y,t) = y_0\delta(t - t_0) \tag{5.15}$$

où nous n'indiquons pas de condition initiale. Nous allons démontrer que l'équation (5.15) possède la même solution que les équations (5.13-5.14) ; en d'autres terme, cette équation a intégré la condition initiale sur sa première ligne.

En effet, soit la fonction (figure 5.4)

$$y_b(t) = H(t - t_0)y_a(t)$$

et donc, par une simple dérivation,

$$y_b' = \delta(t - t_0)y_a(t) + H(t - t_0)y_a'$$

Nous nous souvenons que $\delta(t-t_0)y(t) = y(t_0)\delta(t-t_0)$ et que $H(t-t_0)f(t) = f(t)$ pour $t \geq t_0$. Il est donc trivial de démontrer que

$$y_b' + F(y_b, t) = y_0\delta(t-t_0)$$

c'est à dire que la fonction $y_b(t)$ est bien solution de l'équation (5.15). La fonction $y_b(t)$ est appelé la fonction de Green de l'équation (5.13) et est souvent noté $G(t|y_0, t_0)$. Nous les verrons plus en détail au chapitre 8.

Notez cependant que nous avons cherché une solution $y_b(t)$ qui coïncide avec la fonction $y_a(t)$ pour $t \geq t_0$. C'est pour cela que nous avons appelé la condition (5.14) condition *initiale*. Cela est souvent le cas dans des problèmes physique que l'on rencontre, mais la condition (5.14) est juste une égalité en un point. Nous aurions pu chercher une solution $y_b(t)$ qui coïncide avec $y_a(t)$ pour $t \leq t_0$.

§ 5.10 Montrer que la fonction $y_b(t) = H(t_0-t)y_a(t)$ est solution de l'équation $y' + F(y,t) = -y_0\delta(t)$ pour $t \leq t_0$.

Ce que nous avons discuté se généralise à des fonctions de plusieurs variables. Considérons par exemple le mouvement brownien où $u(x, t|x_0, t_0)$ désigne la densité de probabilité de trouver la particule au point x au temps t, *sachant que* la particule se trouvait au point x_0 au temps t_0. La densité de probabilité obéit à l'équation [3]

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} \tag{5.16}$$

avec la condition initiale donc

$$u(x, t_0) = \delta(x - x_0) \tag{5.17}$$

D'après ce que nous avons dit, au lieu de résoudre les équations ci-dessus, nous pouvons résoudre l'équation

$$\frac{\partial u}{\partial t} - D\frac{\partial^2 u}{\partial x^2} = \delta(x-x_0)\delta(t-t_0)$$

Or, nous avons déjà résolu cette équation (voir la relation 5.8) et nous connaissons sa solution pour les temps $t \geq t_0$:

$$u(x, t|x_0, t_0) = \frac{1}{\sqrt{4\pi D(t-t_0)}} \exp\left(-\frac{(x-x_0)^2}{4D(t-t_0)}\right) \tag{5.18}$$

qui est souvent appelé *le propagateur* de l'équation de diffusion. Ce propagateur joue un rôle fondamental dans de nombreuse branche de la physique, et nous le rencontrerons régulièrement par la suite.

3. Ce n'est pas un hasard que l'équation du mouvement brownien, l'équation de la chaleur et l'équation de diffusion soient les mêmes.

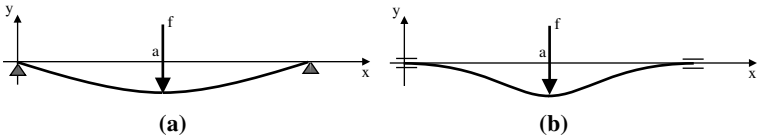


FIGURE 5.5 – la flèche d'un pont sous l'effet d'une force ponctuelle.

5.5 Exercices .

§ 5.11 Que valent les distributions $\delta(x)\cos(qx)$, $\delta(x)\sin(qx)$ et $\delta'(x)\sin(qx)$?

§ 5.12 En dérivant directement la fonction $y(t) = (f_0/\omega_0)H(t)\sin(\omega_0 t)$, démontrer qu'elle est la solution de $\ddot{y} + \omega_0^2 y = f_0\delta(t)$.

§ 5.13 Démontrer que $tH(t)$ est la primitive de $H(t)$. En utilisant une intégration par partie, trouver la primitive de $tH(t)$.

§ 5.14 Une particule initialement au repos de masse m soumise à une force impulsionnelle obéit à l'équation $m\ddot{y} = f_0\delta(t)$. En intégrant directement et en utilisant les conditions initiales, trouver la solution. Trouver la même solution en considérant la particule soumise à une force constante avec une certaine durée $f = (f_0/2T)\Pi(t/T - 1)$ et faire tendre ensuite la durée vers zéro.

§ 5.15 Intégrer directement l'équation $dy/dt + \nu y = f_0\delta(t)$ en utilisant la méthode de la variation des constantes.

§ 5.16 L'élasticité des barres est donnée par l'équation

$$Bd^4y/dx^4 = F(x)$$

où $F(x)$ est la densité de force (force par unité de longueur) appliquée au point x et B une constante qui donne l'amplitude de la rigidité de la barre et qu'on appelle module de courbure. C'est par exemple cette équation qui donne la *flèche* d'un pont sous l'effet d'une charge. Nous souhaitons connaître la flèche d'un pont de longueur L sous l'effet du mouvement d'un camion à la position a dessus. Comme les dimensions du camion sont petit par rapport au pont, on le modélise par une distribution de Dirac. En résolvant donc l'équation $y^{(4)} = f_0\delta(x - a)$ trouver la forme du pont. Nous utiliserons deux formes de conditions aux limites : (i) pont posé sur des piliers, $y(0) = y(L) = 0$; $y"(0) = y"(L) = 0$ (figure 5.5.a ; (ii) pont ancré aux deux bouts $y(0) = y(L) = 0$; $y'(0) = y'(L) = 0$ (figure 5.5.b) . Pour quelle valeur de a la flèche est maximum ?

§ 5.17 Démontrer que

$$\int_I xu^2(x)dx = \frac{i}{2\pi}\int_I \tilde{u}'(q)\tilde{u}^*(q)dq$$

où $\tilde{u}(q)$ est la TF de $u(x)$. Help : écrire $\tilde{u}'(q)$ et $\tilde{u}^*(q)$ par leurs définition des TF, former leurs produit et intégrer sur q. Il suffira juste de remarquer que $\int_I \exp(iq(x-y))dq = 2\pi\delta(x-y)$.

§ 5.18 Parseval
Démontrer de façon générale que

$$\int_I f(x)g^*(x)dx = \frac{1}{2\pi}\int_I \tilde{f}(q)\tilde{g}^*(q)dq$$

et en déduire le résultat de §5.17 à nouveau en utilisant les règles de manipulation des TF.

§ 5.19 Conditions initiales
Soit l'équation de second ordre

$$\begin{aligned} y'' + ay' + by &= 0 \\ y(0) = y_0 \quad &; \quad y'(0) = v_0 \end{aligned}$$

Démontrer que cette équation a la même solution, pour $t \geq 0$, que l'équation

$$y'' + ay' + by = (2v_0 + ay_0)\delta(t) + y_0\delta'(t)$$

5.6 Problèmes.

Problème 5.1 Peigne de Dirac
Il est évident que l'expression

$$\Psi(x) = \sum_{n=-\infty}^{n=+\infty} \exp(2i\pi nx) \tag{5.19}$$

n'a pas de signification au sens des fonctions. Nous allons voir par contre qu'au sens des distribution, elle est définie.

1. Expliquer simplement pourquoi l'expression (5.19) n'a pas de sens usuel de fonction.
2. A supposer que cette expression ait un sens, qu'elle est sa périodicité ?

Soit la distribution "peigne de Dirac" $W(x) = \sum_{n=-\infty}^{+\infty} \delta(x-n)$ où $\delta(x)$ est la delta de Dirac.

3. Démontrer que la période de $W(x)$ est 1. Représenter graphiquement $W(x)$.
4. Comme $W(x)$ est de période 1, la décomposer en *série* de Fourier sur l'intervalle $[-0.5, +0.5]$, c'est à dire trouver les coefficient a_n et b_n tels que sur cet intervalle,

$$W(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(2\pi nx) + b_n \sin(2\pi nx)$$

5. En déduire que $\sum_{n=-\infty}^{+\infty} \delta(x-n) = \sum_{n=-\infty}^{n=+\infty} \exp(2i\pi nx)$.
6. Que pouvez vous dire maintenant de la *transformée* de Fourier d'une peigne de Dirac ?

Problème 5.2 les sommes d'Abel, les noyaux de Dirichlet.
0. Introduction. La convergence des séries de Fourier a joué un grand rôle dans l'avancée de l'Analyse au début du XIXème, et certains concepts très proche des distributions inventés pour cela. Nous allons visiter quelques uns.

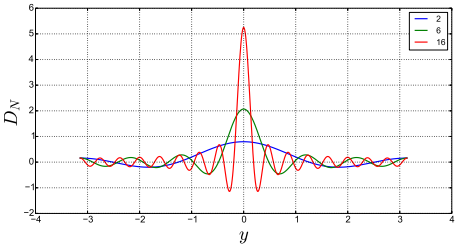


FIGURE 5.6 – Le noyau de Dirichlet $D_N(y)$ pour plusieurs valeurs de N.

soit la fonction 2π périodique $f(x)$, dont la série de Fourier s'écrit :

$$f(x) = a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$$

Soit la fonction $h_N(x)$ la somme partielle des N premiers termes :

$$h_N(x) = a_0 + \sum_{n=1}^{N}(a_n \cos nx + b_n \sin nx)$$

1. Coefficients de Fourier. Démontrer que

$$a_n \cos nx + b_n \sin(nx) = (1/\pi)\int_0^{2\pi} f(y)\cos\left(n(x-y)\right)dy$$

2. Intervalle. Soit une fonction $T-$périodique : $g(y+T) = g(y)$. Démontrer alors que

$$\int_a^{a+T} g(y)dy = \int_0^{T} g(y)dy$$

En déduire que

$$a_n \cos nx + b_n \sin(nx) = (1/\pi)\int_0^{2\pi} f(x-y)\cos(ny)dy$$

3. Noyau de Dirichlet. Soit la fonction (appelé le noyau de Dirichlet)

$$D_N(y) = 1/2\pi + (1/\pi)\left(\cos y + \cos 2y + ... + \cos Ny\right)$$

(figure 5.6). Déduire alors que

$$h_N(x) = \int_0^{2\pi} f(x-y)D_N(y)dy$$

En multipliant et divisant $D_N(y)$ par $\sin y/2$, déduire

$$D_N(y) = \frac{\sin\left((N+1/2)y\right)}{2\pi\sin(y/2)}$$

Pouvez dire vers quoi tend $D_N(y)$ quand $N \to \infty$ [Help : Vers quoi tend $h_N(x)$?]

4. Les sommes d'Abel. Certaines séries divergentes peuvent être régularisées par la procédure d'Abel : au lieu de considérer la série $S = \sum u_n$, on considère la série $S(r) = \sum r^n u_n$ où $0 \leq r < 1$; Une fois le calcul fait, on étudie la limite de $S(r)$ pour $r \to 1$, qu'on appelle la somme d'Abel de la série $\sum u_n$.

Sachant la somme d'une progression géométrique

$$1 + \lambda + ... + \lambda^n = (1 - \lambda^{n+1})/(1 - \lambda)$$

démontrer que

$$1 - 1 + 1 - 1 + ... = 1/2$$

au sens d'Abel.

5. Somme d'Abel du noyau de Dirichlet. Reprenons le noyau de Dirichlet pour $N \to \infty$

$$\pi D_\infty(y) = 1/2 + \cos x + ...$$

notons $\delta_r(y)$ sa somme d'Abel . En considérant $r^n \cos nx$ comme la partie réelle du nombre complexe z^n, démontrer que

$$\delta_r(y) = \frac{1}{2} \frac{1 - r^2}{1 - 2r \cos y + r^2}$$

Quelle est la limite de $\delta_r(y)$ quand $r \to 1$? Considérer séparément les deux cas $y = 0$ et $y \neq 0$. Discuter ce résultat.

6 Convolution et corrélation.

Deux concepts abondement utilisés en physique (et bien d'autres endroits) sont les convolutions et les corrélations. Les TF nous permettent de calculer ces choses de façon assez simple.

6.1 Les convolutions.

Le produit de convolution $f * g$ de deux fonctions f et g est définie par

$$h(x) = (f * g)(x) = \int_{-\infty}^{+\infty} f(s)g(x-s)ds$$

§ 6.1 démontrer que le produit est commutatif : $f * g = g * f$.

L'endroit où l'on rencontre fréquemment ce produit est quand on mesure un signal. Supposons que le signal qu'on mesure est l'intensité lumineuse sur un écran, $f(x)$. Pour mesurer ce signal, l'expérimentateur doit positionner son détecteur à un point x, et mesurer son intensité. Bien sûr, il va effectuer cette mesure en plusieurs points. Le détecteur est cependant un instrument réel, de taille finie, disons 2ℓ (et non infinitésimal). Quand l'instrument est positionnée en x, toute la lumière dans l'intervalle $[x-\ell, x+\ell]$ rentre dans le détecteur, et l'expérimentateur mesure donc en faite la moyenne de l'intensité sur une intervalle autour du point x, et non la valeur exacte de l'intensité en ce point. Évidemment, plus ℓ est petit, meilleure est la précision de l'appareil. En terme mathématique, l'expérimentateur enregistre le signal $h(x)$:

$$\begin{aligned} h(x) &= \int_{x-\ell}^{x+\ell} f(s)ds \\ &= \int_{-\infty}^{+\infty} f(s)\Pi(\frac{x-s}{\ell})ds \\ &= (f * \Pi_\ell)(x) \end{aligned}$$

Ici, $\Pi_l(x) = \Pi(x/\ell)$ est la *fonction de l'appareil.* Les fonctions d'appareil peuvent avoir des formes plus compliquées, comme par exemple une gaussienne. Le facteur limitant la précision du signal est le pouvoir de résolution ℓ de l'appareil qui lisse et rend flou le signal original. Par exemple, un objectif de microscope est un appareil de mesure dont le signal mesuré est l'image formée. Ernst Abbe, physicien de la compagnie Carl Zeiss

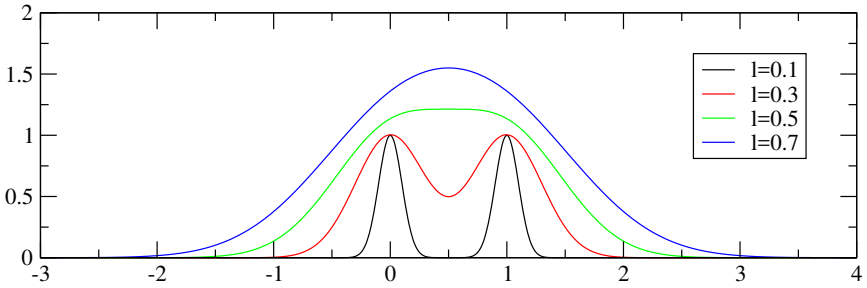


FIGURE 6.1 – La convolution du signal $\delta(x) + \delta(x-1)$ par des gaussiennes G_l de différente largeur.

dans les années 1890, a développé la théorie de la formation d'image et démontré que le pouvoir de résolution des objectifs et, au mieux, $\ell = \lambda/2NA$, où λ est la longueur d'onde utilisée et NA est l'ouverture de l'objectif (le sinus de l'angle maximum de capture de la lumière). Les microscopes optiques ne peuvent donc pas "voir" les échelles plus petites que 0.2 micron.

§ 6.2 soit le signal $f(x) = \delta(x) + \delta(x - x_0)$, c'est à dire deux piques de Dirac distant de x_0. Calculer et tracer le signal mesuré si la fonction de l'appareil est (i) Π_l ; (ii) $G_l = \exp(-x^2/2\ell^2)$. Traiter particulièrement les cas $x_0 \ll \ell$, $x_0 \gg \ell$ et $x_0 \approx \ell$ (voir figure 6.1). Pouvez vous déterminer, pour la Gaussienne, , à partir de quelle ℓ, nous ne pouvons plus distinguer deux piques séparées ?

Les transformées de Fourier nous permettent de calculer facilement les produits de convolution :

$$\mathrm{TF}[f * g] = \mathrm{TF}[f].\mathrm{TF}[g]$$

La transformée de Fourier du produit de convolution de deux fonctions est le produit (normal) de leurs transformée de Fourier. Soit $h(x) = (f * g)(x)$, alors

$$\begin{aligned}
\tilde{h}(q) &= \int_{-\infty}^{+\infty} dx\, e^{-iqx} \int_{-\infty}^{+\infty} ds\, f(s) g(x-s) \\
&= \int_{-\infty}^{+\infty} ds\, f(s) \int_{-\infty}^{+\infty} dx\, e^{-iqx} g(x-s) \\
&= \int_{-\infty}^{+\infty} ds\, f(s) e^{-iqs} \int_{-\infty}^{+\infty} dx\, e^{-iqx} g(x) \\
&= \tilde{f}(q)\tilde{g}(q)
\end{aligned}$$

Calculer numériquement le produit de convolution dans l'espace direct est de l'ordre de N^2, où N est le nombre de points d'échantillonnage des fonctions. Par contre, prendre

la TFR, effectuer une multiplication entre les TF et prendre une TFR inverse ne coûtera que $N \log N$ opérations.

Un autre endroit où l'on rencontre fréquemment les convolutions est la théorie des probabilités et le théorème central limite. Soit deux variables aléatoires continues X_1 et X_2 *indépendantes* de densité $f(x)$ et $g(x)$. Cela veut dire que la probabilité pour qu'une réalisation de X_1 "tombe" entre x et $x+dx$ est égale à $f(x)dx$: $\Pr(x < X_1 < x+dx) = f(x)dx$. Nous nous demandons maintenant si nous pouvons déterminer la densité de probabilité $h(z)$ de la variable $Z = X_1 + X_2$.

$$\begin{aligned} h(z)dz &= \Pr(z < X_1 + X_2 < z + dz) \\ &= \int_{x_1=-\infty}^{x_1=+\infty} \Pr(z - x_1 < X_2 < z - x_1 + dz)\Pr(x_1 < X_1 < x_1 + dx_1) \\ &= dz \int_{-\infty}^{+\infty} g(z - x_1)f(x_1)dx_1 \end{aligned}$$

Nous voyons donc que $h(z) = (f * g)(z)$.

§ 6.3 Démontrer que la densité de probabilité de la *moyenne* de deux variables aléatoires est donnée par $h(z) = 2(f * g)(2z)$.

§ 6.4 Démontrer que le produit de convolution de deux gaussiennes de largeur l et p est encore une gaussienne

$$\frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{l^2 + p^2}} \exp\left(-\frac{x^2}{2(l^2 + p^2)}\right)$$

pour vraiment apprécier les TF, faire le calcul d'abord dans l'espace direct, et ensuite à l'aide des TF. Une gaussienne de largeur l est la fonction

$$\frac{1}{\sqrt{2\pi}l} \exp(-x^2/2l^2)$$

Les résultats ci-dessus sont important. Supposons que nous ayons deux variables aléatoires gaussienne de largeur l. Leur moyenne est alors également une variable aléatoire gaussienne, mais de largeur $l/\sqrt{2}$. Ce résultat se généralise à N variables aléatoires : la moyenne est alors une gaussienne de largeur $l/\sqrt{N}$. La moyenne de N variables aléatoires est également une variable aléatoire, mais qui fluctue $\sqrt{N}$ fois *moins* que les variables originales. C'est pour cette raison par exemple qu'un expérimentateur, pour caractériser un phénomène physique, prend plusieurs mesures et calcule leur moyenne (voir les problèmes avancés).

Exercices :

§ 6.5 Calculer $\Lambda(x) = (\Pi * \Pi)(x)$, et représenter le graphiquement.

§ 6.6 Démontrer que la distribution δ est l'unité pour la convolution : $f * \delta = f$

§ 6.7 Que vaut $f * \delta'$?

§ 6.8 Démontrer que la translation $T_a[f(x)] = f(x-a)$ est la convolution de $\delta(x-a)$ avec la fonction f.

§ 6.9 Le Graal de l'expérimentateur est de déconvoluer son signal, c'est à dire connaissant le signal enregistré $h(t) = (f * A)(t)$ et la fonction d'appareil $A(t)$, déterminer $f(t)$. On pourrait se dire que pour connaître $f(t)$ il suffit de diviser la TF de h par la TF de A et de prendre la TF inverse du résultat. En pratique, ceci n'est pas une bonne solution, puisqu'on ne peut jamais enregistrer un signal pendant un temps infiniment long. Soit $H(t)$ le signal enregistré de $-T$ à $+T$. Mathématiquement parlant, $H(t) = h(t).\Pi(t/T)$. Montrer alors que

$$\tilde{H}(\omega) = 2T \int_{\omega-1/T}^{\omega+1/T} \tilde{f}(\nu)\tilde{A}(\nu)d\nu$$

On voit donc que l'intervalle de temps fini *mélange* les fréquences. Que trouve t'on à la limite $T \to \infty$?

6.2 Auto-corrélation.

Un outil indispensable en physique est le concept d'auto-corrélation. Cela joue un rôle important dans les processus stochastiques, la diffraction, ... Supposons que nous ayons une fonction $x(t)$. Pour plus de simplicité, nous considérons notre signal de moyenne nulle, c'est à dire

$$\lim_{T\to\infty} \frac{1}{T} \int_t^{t+T} x(t)dt = 0$$

Nous désirons savoir combien d'information nous pouvons avoir sur $x(t+\tau)$ si nous connaissons le signal en t. Cette quantité est contenu dans la fonction d'auto-corrélation

$$G(\tau) = \int_{-\infty}^{+\infty} x^*(t)x(t+\tau)dt$$

Le complexe conjugué est nécessaire si l'on veut que pour $\tau = 0$, $G(\tau)$ soit réelle. Dans beaucoup de cas, le signal est réel et le complexe conjugué dans l'espace réel n'a pas d'importance. Concrètement, nous prenons notre signal au temps t, nous le multiplions par le signal au temps $t+\tau$, nous répétons cette opération pour tous les temps t et ajoutons le résultat. Nous donnerons plus loin quelques exemples de la façon dont cette mesure est utilisée pour déterminer les caractéristiques de certains systèmes

physiques. Que vaut la TF de la fonction d'auto-corrélation ?

$$\begin{aligned}
\tilde{G}(\omega) &= \int d\tau \int dt\, x^*(t)x(t+\tau)\exp(-i\omega\tau) && (6.1)\\
&= \int dt\, x^*(t) \int d\tau\, x(t+\tau)\exp(-i\omega\tau) && (6.2)\\
&= \int dt\, x^*(t)\exp(+i\omega t) \int d\tau\, x(\tau)\exp(-i\omega\tau) && (6.3)\\
&= \tilde{x}^*(\omega)\tilde{x}(\omega) = |\tilde{x}(\omega)|^2 && (6.4)
\end{aligned}$$

Le résultat est d'une grande beauté : la TF de la fonction d'auto-corrélation est égale au module de la TF du signal au *carré*. Rappelons simplement que pour passer de (6.1) à (6.2), nous avons échangé l'ordre d'intégration ; pour passer de (6.2) à (6.3) nous avons effectué le changement de variable $\tau \to \tau - t$.

La fonction d'auto-corrélation reçoit des interprétation différentes dans différents contextes. Par exemple en probabilités, soit X_1 la valeur d'une fonction aléatoire au temps t, et X_2 la valeur de même fonction au temps $t+\tau$. En suivant la discussion sur les convolutions, on peut alors démontrer que l'autocorrélation $G(\tau)$ est la densité de probabilité de la variable aléatoire $X_2 - X_1$.

En physique de la matière condensée, on a coutume d'imager autrement la fonction d'auto-correlation. Supposez que vous ayez des particules distribuées dans l'espace. Quelle est la distribution des distances entre les particules ? Prenez n'importe quelles deux particules $i,\ j$ et calculer la distance r_{ij} entre les deux. Faites maintenant un histogramme de toutes les distances, et vous avez une fonction d'autocorrélation des concentrations. Nous avions vu, dans le chapitre sur les TF, que le champs $E(\mathbf{q})$ de lumière diffusé dans une direction $\mathbf{q}$ est la TF de la fonction de transmission local. En utilisant des rayons γ ou neutron à très petite longueur d'onde, la fonction de transmission devient proportionnelle à la concentration des molécules qui diffusent ces longueurs d'onde efficacement, c'est à dire : $E(\mathbf{q}) \sim TF[c(\mathbf{x})]$. Or, les plaques photographiques ou les senseurs de nos caméras ne mesurent pas le champ, mais l'intensité, c'est à dire $I(\mathbf{q}) = E(\mathbf{q})E^*(\mathbf{q})$. Les clichés de diffusion des Rayons γ sont donc une mesure directe de la fonction d'auto-corrélation des concentrations moléculaires.

§ 6.10 le démontrer.

6.3 Relation entre l'équation de diffusion et les convolutions.

Soit l'équation de diffusion

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}$$

avec la condition initiale

$$c(x,0) = f(x)$$

Nous savons depuis la section 5.4 que nous pouvons le résoudre par l'équation

$$\frac{\partial c}{\partial t} - D\frac{\partial^2 c}{\partial x^2} = \delta(t)f(x) \tag{6.5}$$

Pour résoudre ce dernier, nous pouvons utiliser diverses méthodes, comme les TF ou les fonction de Green (voir chapitre 8). Soit

$$G(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right) \quad t \geq 0$$

Cette fonction, appelée aussi une gaussienne, est de plus en plus large en x au fur et à mesure de l'écoulement de t (voir figure 5.2). Or, la solution de l'équation (5.2) est donnée par

$$c(x,t) = \int_I f(x)G(y-x,t)dy$$

en d'autre terme,

$$c(x,t) = \left(f(.) * G(.,t)\right)(x)$$

Pour obtenir la solution au temps t, nous convoluons la condition initiale $f(x)$ avec une gaussienne dont la largeur est donnée par t : La diffusion est une simple convolution de la condition initiale ; plus le temps coule, plus la gaussienne est large est plus les détails de la condition initiale sont gommées.

6.4 Problèmes avancés.

Problème 6.1 **Diffusion des corrélations.**

Soit une fonction (représentant par exemple une concentration ou une probabilité, ...) obéissant à l'équation de diffusion

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}$$

Et soit la fonction d'auto-corrélation spatiale

$$G(y;t) = \int_{-\infty}^{\infty} c(x;t)c(x+y;t)dx$$

Démontrer que G obéit également à une équation de diffusion, mais avec un coefficient de diffusion de $2D$. [indication : il suffit d'échanger soigneusement les dérivations et les intégrations]

Problème 6.2 **Ressort soumis au bruit thermique.**

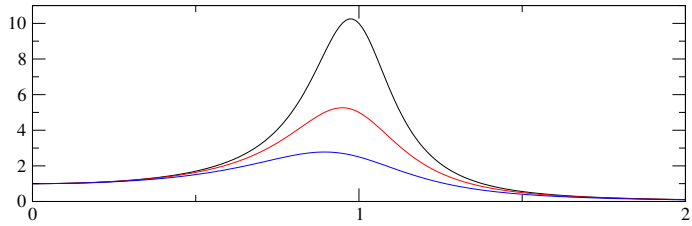


FIGURE 6.2 – $|\tilde{x}(\omega)|^2$ (eq. 6.9) en fonction de ω pour $\omega_0 = 1$ et $\nu = 0.1, 0.2, 0.4$.

(Discuter ergodicité). Supposons une particule dans un puits harmonique, soumis au bruit thermique. Son équation du mouvement s'écrit :

$$m\frac{d^2x}{dt^2} + \nu\frac{dx}{dt} + kx = f\xi(t) \tag{6.6}$$

m est la masse de la particule, ν est la force visqueuse et k la constante du ressort. Ceci constitue une équation différentielle stochastique, et le formalisme a été développé par Langevin vers 1910. La partie gauche de l'équation est celle du mouvement classique d'une particule attachée à un ressort. La partie droite tient compte des chocs aléatoires des molécules du fluide qui entourent la particule et qui font subir à cette dernière une force. La fonction ξ est une fonction aléatoire, c'est à dire qu'on ne connaît pas vraiment la valeur qu'elle peut prendre, mais seulement la *probabilité* qu'elle prenne une certaine valeur. Cela généralise le concept de variable aléatoire utilisée en calcul des probabilités. f est l'amplitude des chocs aléatoires et vaut K_BT/a, où a est la taille de la particule.

On suppose que la fonction ξ est de moyenne nulle, c'est à dire qu'il y a autant de chance, *en moyenne*, que les chocs mènent vers la gauche que vers la droite. De plus, on suppose que la connaissance de la valeur de $\xi(t)$ ne nous donne aucune information sur $\xi(t+\tau)$, quelque soit τ. On exprime cela par

$$\int_I \xi(t)\xi(t+\tau) = \delta(\tau) \tag{6.7}$$

où bien sûr, δ désigne le delta de Dirac. Cela n'est pas trop dur à imaginer : comme $\xi(t+\tau)$ est complètement indépendant de $\xi(t)$, il y a autant de chance qu'il soit de signe contraire que de même signe. A la longue, l'intégral doit tendre vers 0. Par contre, $\xi^2(t) > 0$, son intégrale tend donc vers l'infini (reportez vous à notre discussion sur ce genre d'objet au chapitre précédent). En prenant la TF (par rapport à τ) de l'éq.(6.7), on obtient :

$$\tilde{\xi}(\omega)\tilde{\xi}^*(\omega) = 1 \tag{6.8}$$

En notant $\omega_0^2 = k/m$ et en prenant la TF de l'équation (6.6), nous obtenons :

$$(\omega_0^2 - \omega^2 + i\nu\omega)\tilde{x}(\omega) = (f/m)\tilde{\xi}(\omega)$$

ce qui nous donne, grâce à la relation (6.8),

$$|\tilde{x}(\omega)|^2 = \frac{(f/m)}{(\omega_0^2 - \omega^2)^2 + \nu^2\omega^2} \tag{6.9}$$

Cette fonction présente un pique à $\omega \approx \omega_0$, comme on peut le constater sur la figure 6.2.

Soit la fonction d'autocorrélation des positions

$$G(\tau) = \int_I x(t)x(t+\tau)dt$$

et nous avons vu que par définition, $\tilde{G}(\omega) = |\tilde{x}(\omega)|^2$. La relation (6.9) nous donne donc directement la fonction d'autocorrélation des positions.

On peut faire beaucoup de chose à partir de là. En physique, on réalise souvent [1] des ressorts de taille micrométrique pour exercer des forces sur des bactéries où des molécules biologiques. Un problème majeur est celui de calibrer le ressort, c'est à dire trouver sa constante k. L'équation (6.9) nous montre qu'il existe une façon extrêmement robuste de trouver cette constante : (i) enregistrer la position $x(t)$ d'une particule au bout de ce ressort au cours du temps (ses fluctuations thermiques) ; (ii) prendre la TF de $x(t)$; (iii) élever le module de la TF au carré ; (iv) trouver pour quelle fréquence, cette dernière présente un maximum : nous avons la fréquence propre de l'oscillateur.

Problème 6.3 Somme de deux variables aléatoires et théorème central limite .

Une variable aléatoire X est une fonction qui produit un nombre aléatoire à chaque réalisation. On peut se donner l'image d'un boîtier électronique qui affiche un nombre à chaque fois qu'on appuie sur un bouton (une réalisation). C'est par exemple, le jeté d'un dés ; ou le temps entre l'arrivé de deux particules sur notre senseur ; ou la direction prise par une amibe au fond d'une boite de pétri quand on la photographie toute les 30 secondes ; ou le cours de la bourse à chaque seconde ; ...

On caractérise une variable aléatoire (que l'on suppose continue) par sa densité de probabilité $f(x)$: la probabilité d'observer une réalisation de X entre x et $x+dx$ est égale à $f(x)dx$. Cela veut dire concrètement que si on effectue par exemple 10^6 réalisations (mesurons l'arrivé d'un million de particule sur notre senseur), une proportion $f(x)dx$ des réalisations tomberont dans l'intervalle $[x, x+dx[$. D'après ce que nous venons de dire, $f(x) \geq 0$ et

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

Soit maintenant deux variables aléatoires *indépendantes* [2] X et Y de densité de probabilité $f(x)$ et $g(y)$. Quelle est la densité de la variable $Z = X + Y$ (comme par exemple la somme de deux dés)? En probabilité, le "et" d'événements indépendants se traduit par le produit de chacune des probabilités et le "ou" par l'opération somme des probabilités. Appelons $h(z)dz$ la probabilité d'observer Z dans l'intervalle $[z, z+dz[$. La probabilité d'observer un tel événement égale la probabilité d'observer X dans $[x, x+dx[$ et Y dans $[z-x, z-(x+dx)+dz[$ pour un x quelconque. Cet événement a la probabilité

$$f(x)dx.g(z-x)(dz-dx) = f(x)g(z-x)dxdz + O(dx^2)$$

1. Par des pinces optiques, magnétique, des micropipettes, ...
2. Une réalisation de l'une n'influe pas sur le résultat de la réalisation de l'autre.

pour une valeur x quelconque. Il faut donc ajouter la probabilité pour toutes les valeur possible de x pour obtenir $h(z)dz$, ce qui nous donne

$$h(z) = \int_{-\infty}^{+\infty} f(x)g(z-x)dx$$

La densité de probabilité de la somme de deux variables aléatoires égale le produit de convolution des densités de chaque variable.

La moyenne d'une variable aléatoire X de densité de probabilité $f(x)$ est notée $\langle X \rangle$ et est définie par

$$\langle X \rangle = \int_{-\infty}^{+\infty} xf(x)dx$$

De façon générale, pour une fonction quelconque V, on définit

$$\langle V(X) \rangle = \int_{-\infty}^{+\infty} V(x)f(x)dx$$

6.5 Exercices.

La suite des exercices suivantes vous entraîne à manipuler les probabilités. Si vous les suivez dans l'ordre jusqu'au bout (bravo), cela vous mènera à la démonstration du théorème central limite : quelque soit la densité de probabilité de la fonction X (pourvu qu'elle ait une variance finie), la densité de probabilité de la moyenne de N de ces variables est une gaussienne, de largeur $\sigma/\sqrt{N}$, où σ^2 est la variance de X. L'ensemble de ces exercices constitue un bon cours de probabilité.

§ 6.11 Démontrer que $\langle aX \rangle = a \langle X \rangle$ où a est un nombre réel. De façon générale, quelle est la densité de probabilité de $Z = aX$?

§ 6.12 Démontrer que $\langle X + Y \rangle = \langle X \rangle + \langle Y \rangle$. Que vaut la moyenne de la variable $Z = (X + Y)/2$? Soit $Z_N = (1/N) \sum_{i=1}^{N} X_i$ où les variables aléatoires X_i sont identiques. Que Vaut $\langle Z_n \rangle$?

§ 6.13 La variance d'une variable est définie par $Var(X) = \langle X^2 \rangle - \langle X \rangle^2$. Que vaut $Var(X + Y)$? Et $Var(Z_N)$?

§ 6.14 La fonction caractéristique $\phi_X(t)$ d'une variable aléatoire X de densité $f(x)$ est définie par

$$\phi_X(t) = \langle \exp(itX) \rangle$$

Quelle est la relation entre la densité de X et la fonction $\phi_X(t)$?

§ 6.15 Démontrer que $\phi_X(0) = 1$; $\phi_X^{'}(0) = i \langle X \rangle$; $\phi_X^{''}(0) = -\langle X^2 \rangle$; généraliser ce résultat. Vous pouvez obtenir ce résultat par le développement de Taylor de la fonction exponentielle.

§ 6.16 Que valent $\phi_{aX}(t)$ et $\phi_{X+Y}(t)$? Que vaut $\phi_{Z_n}(t)$?

§ 6.17 Démontrer que de façon générale, $\phi_X(t)$ a un maximum absolue à $t = 0$.

§ 6.18 On suppose que $\langle X \rangle = 0$ et $Var(X) = \sigma^2$. Développer $\phi_{Z_n}(t)$ à l'ordre 2 en t autour de son maximum, et démontrer qu'elle tend vers $\exp(-\sigma^2 t/2\sqrt{N})$. [Help : $(1+x/n)^n \to \exp(x)$]. En déduire la densité de probabilité de Z_n. Généraliser ce résultat au cas $\langle X \rangle \neq 0$.

6.6 Problèmes.

Problème 6.4 Fluctuation de la courbure des polymères.

D'abord, un peu de géométrie différentielle. Soit une courbe dans le plan. Nous pouvons par exemple la décrire par l'équation $y(x)$ ou par ses coordonnées paramétrique $x(t), y(t)$. Si nous appelons l'extrémité de la courbe A, la longueur d'arc à partir de A jusqu'à un point P est définie par

$$s = \int_0^t \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} dt$$

Appelons l'angle $\theta(s)$ l'angle que fait la tangente à la courbe au point P avec l'axe y. En faite, nous pouvons parfaitement définir la courbe par la donnée de la fonction $\theta(s)$. Par exemple, $\theta =$ Cte décrit une droite, $\theta = s/R$ décrit un cercle de rayon R. Cette description d'une courbe s'appelle semi-intrinsèque. La courbure de la courbe à la position s est donnée par $\kappa = (d\theta/ds)^2$. Nous pouvons également décrire une courbe dans le plan par la donnée de $\kappa(s)$ de façon totalement intrinsèque, sans référence à aucun système d'axe.

Soit maintenant un polymère (à deux dimensions) de longueur L ($L \to \infty$ à l'échelle moléculaire, comme l'ADN par exemple) baignant dans un bain à température T. L'énergie emmagasinée dans le polymère par unité de longueur dépend de la courbure de sa conformation et s'écrit

$$E = \int_0^L B\kappa^2(s) ds$$

où B est le module de rigidité du polymère. Quelle est la corrélation entre les tangentes à la courbe distantes de σ ? Plus exactement, démontrer que

$$\langle \mathbf{u}(s).\mathbf{u}(s+\sigma) \rangle = \exp(-\sigma/L_P)$$

où $\mathbf{u}(s)$ est le vecteur tangent à l'abscisse curviligne s et $L_p = B/KT$.

Ceci est loin d'être un calcul anodin : c'est comme cela que l'on mesure la rigidité des polymères biologiques comme l'actine, les microtubules ou l'ADN.

Problème 6.5 Théorème d'échantillonnage.

L'échantillonnage consiste a enregistrer un signal $f(t)$ seulement sur un nombre discret de points $f_n = f(n\tau)$ (Fig. 6.3). Le théorème d'échantillonnage de Shannon-Nyquist nous affirme que nous sommes capable de reconstituer *exactement* la fonction $f(t)$ à partir des f_n à deux conditions :

1. La fonction originale $f(t)$ ne contient pas de fréquences plus élevé qu'une certaine fréquence que nous notons ω_0. Précisément, cela veut dire que
$$\tilde{f}(\omega) = 0 \text{ si } \omega \notin [-\omega_0, \omega_0]$$
2. nous avons échantillonné le signal à au moins $\tau = \pi/\omega_0$.

Si ces deux conditions sont réalisées, alors nous pouvons reconstituer exactement la fonction à l'aide seulement des f_n de la façon suivante :

$$f(t) = \sum_{n=-\infty}^{+\infty} f_n sinc\left(\omega_0 (t - n\tau)\right) \tag{6.10}$$

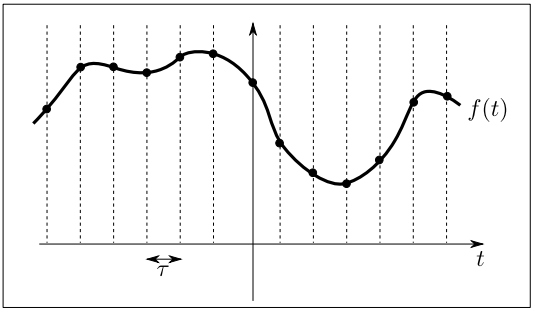


FIGURE 6.3 – Echantillonnage d'un signal $f(t)$ à intervalle régulier τ

où $sinc(u) = (\sin u)/u$. Ce théorème, énoncé vers 1953, joue un rôle crucial dans les communications radio et enregistrement des signaux ; les CD par exemple sont échantillonné à 48kHz, deux fois le seuil de l'audition humaine en fréquence.

1. Commentaires. Expliquez ce que vous comprenez par la formule (6.10). Aidez-vous d'un graphe.

2. Porte. Soit une fonction $f(t)$ quelconque. Montrez graphiquement les fonctions $f(t)$ et $2f(t)\Pi(t/a)$, où Π représente la fonction porte. Désignons par ω la variable conjuguée à t lors d'une TF. Que valent

$$\mathrm{TF}\left[\Pi(t/a)\right] \text{ et } \mathrm{TF}^{-1}\left[\Pi(\omega/\omega_0)\right]$$

où a et ω_0 sont des constantes ?

3. Peigne de Dirac. La peigne de Dirac est définie par

$$\Psi_a(t) = \sum_{n=-\infty}^{\infty} \delta(t - na)$$

Soit la distribution

$$\Psi_a(t).f(t) = \sum_n f(t)\delta(t - na)$$

Représenter graphiquement la distribution $\Psi_a(t)$ et la distribution $\Psi_a(t).f(t)$ où f est une fonction quelconque. Démontrer que

$$\Psi_a(t).f(t) = \sum_n f(na)\delta(t - na)$$

[Help : égalité entre distribution].

4. Convolution. Rappeler ce que vaut $(f * \delta_a)(t)$, le produit de convolution de la fonction quelconque $f(t)$ avec la distribution de Dirac $\delta(t-a)$. Soit maintenant la fonction $f(t)$ à support bornée : $f(t) = 0$ si $t \notin [-b, b]$. Représenter graphiquement $(f * \delta_{2a})(t)$ dans les cas où $b > a$, $b = a$ et $b < a$. En utilisant ces résultats, représenter graphiquement

$$(\Psi_{2a} * f)(t)$$

dans les trois cas précédent.

5. Linéarité. Démontrer que le produit de convolution est linéaire, c'est à dire

$$(\lambda f + \mu g) * h = \lambda(f * h) + \mu(g * h)$$

6. Résultat admis. Nous admettons le résultat suivant : la TF d'une peigne de Dirac est une peigne de Dirac (voir problème 5.1) :

$$\mathrm{TF}[\Psi_a(t)] = (2\pi/a)\Psi_{2\pi/a}(\omega)$$

De quelle distribution $\Psi_{\omega_0}(\omega)$ est elle la TF ?

7. Mise en place. Soit maintenant la fonction à support bornée $\tilde{f}(\omega)$, nulle en dehors de l'intervalle $[-\omega_0, \omega_0]$. En utilisant les résultats précédents, argumenter pourquoi

$$\tilde{f}(\omega) = 2\left(\tilde{f} * \Psi_{2\omega_0}\right)(\omega).\Pi(\omega/\omega_0)$$

Aidez vous d'un graphique.

8. TF inverse. En appliquant la TF inverse à l'équation précédente, démontrer que

$$f(t) = \left(\sum_{-n=\infty}^{\infty} f(t).\delta(t - n\pi/\omega_0)\right) * sinc(\omega_0 t)$$

[**Help** : $\mathrm{TF}^{-1}[(\tilde{f} * \tilde{g})(\omega)] = 2\pi f(t).g(t)$].

9. Fin. A partir de l'équation précédente, obtenir enfin l'équation (6.10) de Nyquist-Shannon.

Note. En général, le théorème de Nyquist est énoncée à l'aide des fréquences ν et non des fréquences angulaires ω comme nous l'avons fait ici. Sachant que $\omega = 2\pi\nu$, le théorème d'échantillonnage est souvent énoncé par la formulation $\tau < 1/2\nu_0$.

Problème 6.6 Corrélation dans le mouvement brownien.

Calculer la fonction d'autocorrélation pour un mouvement brownien $x(t)$.

7 Les transformées de Laplace.

7.1 Entrée en matière.

Les mathématiciens ont inventé de nombreux transformation intégrale d'une fonction, parmi lesquels nous avons vu les transformées de Fourier. Une autre transformation extrêmement utilisée est celle de Laplace. Les transformées de Laplace sont les cousins des transformées de Fourier. Leur relation est celle de la fonction exponentielle et de la fonction sinus ou cosinus. Comme vous vous souvenez, pour prendre la TF, on multiplie la fonction $f(t)$ par $\exp(i\omega t)$ et on intègre entre, notez le bien, $-\infty$ et $+\infty$. Pour les TL, on multiplie la fonction par $\exp(-st)$ et on intègre entre, cette fois, 0 et $+\infty$

$$\tilde{f}(s) = \mathrm{TL}[f(t)] = \int_0^\infty f(t)\exp(-st)dt$$

Les conventions veulent que la variable conjuguées à t s'appelle ω pour les TF et s pour les TL. La fonction $f(t)$ est appelée l'original, et sa TL son *image*. Dans la plupart des livres que vous consulterez, l'image est noté $F(s)$, mais nous maintenons ici la convention $\tilde{f}(s)$ ou $\hat{f}(s)$. Il existe de nombreux avantages et désavantages à utiliser les TL à la place des TF. D'un point de vue pratique, toutes les deux transforment des équations différentielles linéaires en des équations algébriques. Mais il est difficile d'intégrer les conditions initiales dans les TF, tandis qu'elles s'introduisent naturellement dans les TL, comme nous en verrons des exemples plus bas. Prenons le cas d'un signal temporel $x(t)$. Pour les transformées de Fourier, ce signal a toujours existé (depuis $t = -\infty$) et existera toujours. Pour les Transformée de Laplace, le signal ne commence son existence qu'à un temps fini ($t = 0$).

Un autre (grand) avantage des TL est que nos exigences sur le comportement de $f(t)$ quand $t \to \infty$ sont beaucoup plus légères : comme la fonction $\exp(-st)$ décroît très rapidement à l'infini (pour $Re(s) > 0$), la transformée de Laplace de la plupart des fonctions usuelles existera. Voyons quelques exemples.

Exemple 7.1 $\mathrm{TL}[1] = 1/s$

Exemple 7.2 $\mathrm{TL}[\exp(-at)] = 1/(s+a)$

Exemple 7.3 $\mathrm{TL}[t] = 1/s^2$. Pour le démontrer, il suffit d'effectuer une intégration par

partie :

$$\begin{aligned}\int_0^{+\infty} te^{-\omega t}\quad dt = &\quad 0 + \frac{1}{s}\int_0^{+\infty} e^{-st}dt \\ = &\quad \frac{1}{s^2}\end{aligned}$$

Exemple 7.4 $\mathrm{TL}[t^k] = k!/s^{k+1}$ (démontrez cette relation par récurrence.)

Le désavantage des TL est que nous perdons le concept de bases orthogonales. Nous avons vu qu'en étendant un peu notre espace de fonctions à l'espace des distributions, nous pouvions considérer les fonctions $\exp(i\omega t)$ comme une base orthogonale. Rien de tel n'existe pour les TL et les fonctions $\exp(-st)$, quoi qu'on fasse, *ne sont pas* orthogonales les unes aux autres. Avec la perte d'orthogonalité, nous perdons également la possibilité d'inverser (facilement) une transformée de Laplace et la belle symétrie entre une fonction et sa transformée.

7.2 Opérations sur les TL.

Les opération sur les TL sont très similaire, à un facteur i près, aux opérations sur les TF. Par contre, il faut vraiment bien les maîtriser, puisque prendre la TL inverse est souvent une opération complexe (au sens propre) et qu'on préfère toujours se ramener à des expressions connues.

Changement d'échelle. $\mathrm{TL}[f(t/a)] = a\tilde{f}(as)$ La démonstration est triviale.

Translation. $\mathrm{TL}[\exp(-at)f(t)] = \tilde{f}(s+a)$ Multiplier l'originale par une exponentielle revient à translater l'image. Par exemple, $\mathrm{TL}[1] = 1/s$, donc $\mathrm{TL}[\exp(-at)] = 1/(s+a)$.

Multiplication par t. Si on dérive $\tilde{f}(s)$ par rapport à s, nous avons $d\tilde{f}(s)/ds = -\int tf(t)\exp(-st)dt$. Donc, $\mathrm{TL}[tf(t)] = -d\tilde{f}(s)/ds$ [1] . Par exemple, comme $\mathrm{TL}[1] = 1/s$, alors $\mathrm{TL}[t] = 1/s^2$

Dérivation. Elle contient un élément supplémentaire, et c'est cela le grand intérêt des TL.

$$\int_0^{\infty} f'(t)\exp(-st)dt = f(t)\exp(-st)|_{t=0}^{t=\infty} + s\int_0^{\infty} f(t)\exp(-st)dt$$

1. Vous remarquerez que nous avons souvent été négligent avec l'orthodoxie des convergences et des dérivations sous le signe somme. Mais vous pouvez démontrer qu'ici au moins, nous n'avons pas enfreint de règles (démontrez le).

Ce qui nous amène à

$$\mathrm{TL}[f'(t)] = -f(0) + s\mathrm{TL}[f(t)]$$

En généralisant cela, nous voyons que $\mathrm{TL}[f"(t)] = s^2\tilde{f}(s) - sf(0) - f'(0)$, et ainsi de suite.

Intégration. Il n'est pas difficile, en utilisant la règle de dérivation ci-dessus, de démontrer que

$$\mathrm{TL}[\int_0^t f(\tau)d\tau] = (1/s)\tilde{f}(s)$$

Exemple 7.5 Résolvons l'équation différentielle

$$x'(t) + \nu x(t) = \alpha t \tag{7.1}$$

avec la condition initiale $x(t=0) = x_0$. Par la méthode classique, on résout d'abord l'équation homogène pour obtenir $x = C\exp(-\nu t)$, ensuite nous supposons que $C = C(t)$ et nous obtenons une autre équation différentielle pour $C(t)$; la résolution de cette dernière et finalement l'utilisation de la condition initiale nous donne la solution finale.

Prenons plutôt la TL des deux cotés de l'éq.(7.1) :

$$-x_0 + (s+\nu)\tilde{x}(s) = \frac{\alpha}{s^2}$$

Nous avions déjà, à l'exemple 3 ci-dessus, calculé la TL$[t]$, et nous avons juste utilisé ce résultat. En général, les TL des fonctions les plus connues sont entreposées dans des tables et on ne fait souvent que les consulter au lieu de recalculer la TL (comme pour les tables de logarithme). En décomposant en fraction simple, nous avons

$$\frac{\nu}{s^2(s+\nu)} = \frac{1}{s^2} - \frac{1}{\nu}\frac{1}{s} + \frac{1}{\nu}\frac{1}{s+\nu}$$

et la solution de notre équation s'écrit :

$$\tilde{x}(s) = \frac{\alpha}{\nu}\left[\frac{1}{s^2} - \frac{1}{\nu}\frac{1}{s} + \frac{1}{\nu}\frac{1}{s+\nu}\right] + \frac{x_0}{s+\nu} \tag{7.2}$$

Bon, nous connaissons la TL de la solution, et il faut inverser le processus pour calculer $x(t)$. Or, nous savons que l'originale de $1/s^2$ est t, l'originale de $1/s$ est 1 , l'originale de $1/(s+\nu)$ est $\exp(-\nu t)$ (souvenez vous de la règle de translation). Nous avons donc

$$x(t) = \frac{\alpha}{\nu}t - \frac{\alpha}{\nu^2}(1 - e^{-\nu t}) + x_0 e^{-\nu t} \tag{7.3}$$

On peut vérifier, en l'injectant directement dans l'équation (7.1) que ceci est bien la solution. Notez avec qu'elle facilité la condition initiale a été prise en compte dans la solution.

$f(t)$	$\tilde{f}(s)$
$f(t/a)$	$af(as)$
$\exp(-at)f(t)$	$\tilde{f}(s+a)$
$tf(t)$	$-\frac{d}{ds}\tilde{f}(s)$
$f'(t)$	$s\tilde{f}(s)-f(0)$
$f"(t)$	$s^2\tilde{f}(s)-sf(0)-f'(0)$
$f^{(n)}(t)$	$s^n\tilde{f}(s)-\sum_{k=1}^{n}s^{n-k}f^{(k-1)}(0)$
$\int_0^t f(\tau)d\tau$	$\frac{1}{s}\tilde{f}(s)$
1	$1/s$
t	$1/s^2$
$\exp(-at)$	$1/(s+a)$
$\sin(at)$ ou $\cos(at)$	$a/(s^2+a^2)$ ou $s/(s^2+a^2)$
$\sinh(at)$ ou $\cosh(at)$	$a/(s^2-a^2)$ ou $s/(s^2-a^2)$
$-t\cos(at)+(1/a)\sin(at)$	$2a^2/(s^2+a^2)^2$
$1/\sqrt{t}$	$\sqrt{\pi}/\sqrt{s}$
$\sqrt{t}$	$(\sqrt{\pi}/2)s^{-3/2}$
$1/(t+1)$	$\exp(s)\Gamma(0,s)$

TABLE 7.1 – Résumé des règles de manipulation des TL et un petit dictionnaire des TL élémentaires.

Exemple 7.6 Résoudre $x^{(3)}+3\ddot{x}+3\dot{x}+x=1$ avec les conditions initiales nulles ($x^{(n)}$ désigne la dérivée $n-$ième de x).

La TL nous donne $\tilde{x}(s)=1/s(s+1)^3=(1/s)-1/(s+1)^3-1/(s+1)^2-1/(s+1)$. En se reportant à la table (7.1), on trouve immédiatement $x(t)=1-(t^2/2+t+1)\exp(-t)$.

7.3 Décomposition en fraction simple.

Comme nous avons à utiliser souvent les décompositions en fraction simple, nous allons faire un petit détour pour rappeler les grands principes.

Cas des racines simples. Soit $\tilde{f}(s)=p(s)/q(s)$, où $p(s)$ et $q(s)$ sont des polynômes et qu'en plus, $q(s)$ n'a que des racines simples, c'est à dire $q(s)=(s-a_1)(s-a_2)...(s-a_n)$. Nous voulons écrire $f(s)$ comme

$$\tilde{f}(s)=\frac{A_1}{s-a_1}+\frac{A_2}{s-a_2}+...+\frac{A_n}{s-a_n}$$

Soit $q_i(s) = q(s)/(s-a_i)$. Nous voyons que $q_i(s)$ n'a pas de zéro en $s = a_i$. Quand $s \to a_i$, le terme dominant dans $f(s)$ est

$$\tilde{f}(s) = \frac{p(s)}{q_i(s)}.\frac{1}{s-a_i} = \frac{p(a)}{q_i(a)}.\frac{1}{s-a_i} + \mathcal{O}(1)$$

d'où on déduit que $A_i = p(a_i)/q_i(a_i)$. En plus, comme $q(a_i) = 0$,

$$\lim_{s\to a_i} q_i(s) = \lim_{s\to a_i} \frac{q(s)-q(a_i)}{s-a_i} = q'(a_i)$$

quand $s \to a_i$. Nous pouvons donc écrire l'original de $\tilde{f}(s)$directement comme

$$f(t) = \sum_n \frac{p(a_n)}{q'(a_n)} \exp(a_n t)$$

où la sommation est sur les zéros de $q(s)$.

Exemple 7.7 $\tilde{f}(s) = (3s^2-3s+1)/(2s^3+3s^2-3s-2)$. Nous avons $p(s) = 3s^2-3s+1$, $q(s) = 2s^3+3s^2-3s-2$ et $q'(s) = 6s^2+6s-3$. Les zéro du dénominateur sont aux $s = 1, -2, -1/2$. Comme $p(1)/q'(1) = 1/9$, $p(-2)/q'(-2) = 19/9$ et que $p(-1/2)/q'(-1/2) = -13/18$, nous avons

$$\tilde{f}(s) = \frac{1}{9}\frac{1}{s-1} + \frac{19}{9}\frac{1}{s+2} - \frac{1}{6}\frac{1}{s+1/2}$$

Cas des racines multiples. Soit maintenant $\tilde{f}(s) = R(s)/(s-a)^n$où $R(s)$ est un quotient de polynôme qui n'a pas de pôles en a. Nous voulons l'écrire sous forme de

$$\tilde{f}(s) = \frac{A_0}{(s-a)^n} + \frac{A_1}{(s-a)^{n-1}} + ... + \frac{A_{n-1}}{(s-a)} + T(s)$$

où $T(s)$ contient le développement en fractions simples autour des autres pôles. Pour déterminer les coefficients A_i nous avons à nouveau à calculer le comportement de $\tilde{f(s)}$ pour $s \to a$. Comme $R(s)$ est tout ce qui a de plus régulier autour de a, nous pouvons le développer en série de Taylor autour de ce point :

$$R(s) = R(a) + R'(a)(s-a) + (1/2)R''(a)(s-a)^2 + ...$$

Ce qui nous donne immédiatement

$$\begin{aligned} A_0 &= R(a) \\ A_1 &= R'(a) \\ &... \end{aligned}$$

Exemple 7.8 Trouvons l'originale de $\tilde{f}(s) = 1/(s^2+a^2)^2$. Nous avons

$$\tilde{f}(s) = \frac{A_0}{(s-ia)^2} + \frac{A_1}{(s-ia)} + \frac{B_0}{(s+ia)^2} + \frac{B_1}{(s+ia)}$$

Nous pouvons bien sûr tout calculer, mais remarquons simplement que dans l'expression de $\tilde{f}(s)$, le changement de s en $-s$ laisse ce dernier invariant. Pour avoir cette même invariance dans l'expression de $\tilde{f}(s)$ une fois décomposée en fraction simple, nous devons avoir $B_0 = A_0$ et $B_1 = -A_1$. Or, d'après ce qu'on vient de dire, autour de la racine $s = ia$, $R(s) = 1/(s+ia)^2$ et

$$A_0 = \frac{1}{(s+ia)^2}\Big|_{s=ia} = -\frac{1}{4a^2}$$

De même,

$$A_1 = \frac{-2}{(s+ia)^3}\Big|_{s=ia} = \frac{1}{4ia^3}$$

Comme l'originale de $1/(s \mp ia)^2$ est $t\exp(\pm iat)$ et que l'originale de $1/(s \mp ia)$ est $\exp(\pm t)$, en regroupant correctement les termes, on trouve que

$$f(t) = -\frac{1}{2a^2} t\cos(at) + \frac{1}{2a^3}\sin(at).$$

7.4 Comportement asymptotique.

7.4.1 Comportement pour $t \to +\infty$.

Si nous connaissons la transformée de Laplace d'une fonction $f(t)$, nous pouvons parfois trouver des approximations de cette fonction quand $t \to \infty$. Par exemple, les fonctions de Bessel $I_n(t)$ sont définies par

$$I_n(t) = \frac{1}{\pi}\int_0^\pi e^{t\cos\theta}\cos(n\theta)d\theta$$

et il n'est pas difficile de démontrer [2] que la transformée de Laplace de $I_0(t)$ est $\hat{I}_0(s) = 1/\sqrt{s^2-1}$. Cette transformée nous permet facilement d'approximer, pour $t \gg 1$, la fonction de Bessel par $\exp(t)/\sqrt{2\pi t}$ (figure 7.1). Voyons voir le comment du pourquoi.

Nous nous sommes peu intéressés jusque là au domaine d'existence de la Transformée de Laplace. Il est évident que pour que la TL ait un sens, il faut que $\int_0^\infty f(t)\exp(-st)dt$ existe. Pour certaines fonctions comme $\exp(-t^2)$, cette condition est toujours réalisée. Pour d'autres, comme $\exp(t^2)$, elle ne l'est jamais. Enfin, pour la plupart de fonctions

2. en changeant l'ordre d'intégration sur θ et t, cf exercice 7.8.

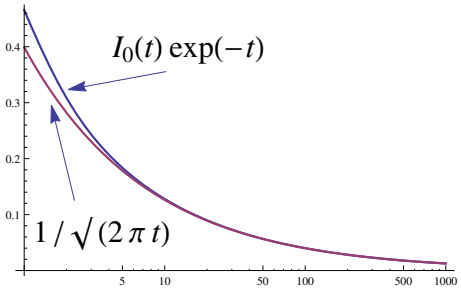


FIGURE 7.1 – Comparaison des la fonction de Bessel $I_0(t)$ et son approximation assymptotique (l'axe x est logarithmique)

usuelles [3] , la condition est réalisée si $Re(s) > s_0$, où s_0 est un réel. Par exemple, pour toutes les fonctions polynomiales ou toute puissance positive de t, $s_0 = 0$. Pour la fonction $\cosh(t)$, $s_0 = 1$ [4].

Souvent, nous nous intéressons surtout au comportement de $f(t)$ pour t grand : nous voulons savoir rapidement si notre particule revient à une position donnée ou si au contraire, elle part à l'infini, et si elle part à l'infini, à quelle vitesse elle le fait. Nous allons voir dans la suite que le comportement de $\tilde{f}(s)$ autour de son pôle le plus à droite s_0 nous renseigne directement sur le comportement asymptotique de l'originale. Sans perte de généralité, nous allons supposer par la suite que $Re(s_0) = 0$, puisque si la TL de la fonction $f(t)$ a un pôle en $s = a$, la fonction $\exp(-at)f(t)$ a un pôle en $s = 0$. Le comportement asymptotique de la fonction $f(t)$ s'en déduit donc immédiatement.

Revenons maintenant à notre fonction $f(t)$. Si $I = \int_0^\infty f(t)dt < +\infty$, c'est que $f(t) \to 0$ quand $t \to +\infty$ et nous n'avons pas trop de questions à nous poser pour son comportement asymptotique. Supposons donc que I n'existe pas, mais que la TL de $f(t)$ est bien définie pour $Re(s) > 0$. Nous pouvons toujours écrire $f(t) = g(t) + h(t)$, où $g(t)$ contient le terme dominant de $f(t)$ quand $t \to \infty$ et $h(t)$ tous les autres. Par exemple, le terme dominant de $1/\sqrt{t} + \exp(-5t) + 1/(1+t^2)$ est $1/\sqrt{t}$. Nous pouvons formellement écrire que $h(t) = o(g(t))$ [5]. Il est évident que pour $s \to 0$, la transformée de Laplace est dominée par la TL de $g(t)$, c'est à dire $\tilde{h}(s) = o(\tilde{g}(s))$ quand $s \to 0$ (exercice : le démontrer). Un simple développement autour du pôle le plus à droite de la TL nous donne donc directement le comportement asymptotique de l'originale.

Exemple 7.9 $\tilde{f}(s) = 1/s(s+a)$ pour $a > 0$ a son pôle le plus à droite à $s = 0$. Autour de ce point, $\tilde{f}(s) = 1/as + \mathcal{O}(1)$. Donc, $f(t) \approx 1/a$ quand $t \to \infty$ (l'original de

3. lire "qui ne croissent pas plus vite qu'une exponentielle"
4. Puisque la fonction $\cosh(t)$ comporte un terme en e^t.
5. C'est à dire que $\lim_{t\to\infty} h(t)/g(t) = 0$. Les notations $\mathcal{O}$ et o sont dues à Edmund Landau, mathématicien du premier tiers du vingtième siècle.

$1/s$ est bien sûr 1). Dans cet exemple, et ceux qui suivent, le lecteur est encouragé à calculer l'originale exacte et vérifier le développement asymptotique.

Exemple 7.10 $\tilde{f}(s) = 1/s(s-a)^2$ pour $a > 0$. Le pôle le plus à droite est en $s = a$. $\tilde{f}(s) \approx (1/a)(s-a)^{-2}$ quand $s \to a$ et donc $f(t) \approx (t/a)\exp(-at)$ quand $t \to \infty$. Remarquer que nous aurions pû pousser l'approximation un peu plus loin : $\tilde{f}(s) \approx (1/a)(s-a)^{-2} - (1/a^2)(s-a)^{-1}$ et donc $f(t) \approx (t/a - 1/a^2)\exp(-at)$.

Exemple 7.11 $\tilde{f}(s) = 1/(s^2+a^2)^2$. Là, nous avons deux pôles de même partie réelle $s = \pm ia$, et nous devons tenir compte des deux. Nous laissons le soin au lecteur de démontrer que le terme dominant doit être $-t\cos(at)/2a^2$.

Nous avons en fait souvent recours au développement asymptotique parce que nous ne savons pas calculer exactement l'originale. Prenons l'équation $\ddot{x} + \dot{x} = \sqrt{t}$ avec des conditions initiales nulles. C'est l'équation du mouvement d'un corps soumis à un frottement visqueux et à une force qui grandit comme la racine du temps. La solution est facilement trouvée en terme de TL : $\tilde{x}(s) = (\sqrt{\pi}/2)s^{-5/2}(s+1)^{-1}$. Nous ne savons pas calculer [6] l'originale de cette fonction. Par contre, comme il existe un pôle à zéro, le développement asymptotique s'écrit $x(t) \approx (4/3\sqrt{\pi})t^{3/2}$(Le démontrer ; pouvez vous calculer les deux prochaines corrections à ce développement ?).

7.4.2 Comportement pour $t \to 0$.

Nous disposons d'un théorème analogue pour trouver le comportement de $f(t)$ autour de $t = 0^+$ si nous disposons de sa transformée de Laplace. Il n'est pas difficile de voir que

$$f(0) = \lim_{s\to\infty} s\hat{f}(s) \tag{7.4}$$

Ceci découle simplement des règles de TL :

$$\mathrm{TL}[f'(t)] = \int_0^\infty e^{-st} f'(t)dt = s\hat{f}(s) - f(0)$$

Or, quand $s \to \infty$, l'intégrale tend vers zéro, d'où l'égalité (7.4). Nous pouvons bien sûr aller plus loin. Le développement de Taylor de $f(t)$ proche de $t = 0$ s'écrit

$$f(t) = f(0) + f'(0)t + (1/2)f''(0)t^2 + \ldots$$

et résulte de la TL inverse du développement asymptotique de $sf(s)$ pour $s \to \infty$. (voir l'exercice 7.9).

6. Pas avec notre dictionnaire actuel.

7.5 Produit de Convolution.

Le produit de convolution de deux fonctions est donné par

$$h(t) = \int_0^t f(\tau)g(t-\tau)d\tau \tag{7.5}$$

On note cela par $h(t) = (f \star g)(t)$. Il est facile de démontrer, en échangeant l'ordre d'intégration, que

$$\tilde{h}(s) = \tilde{f}(s).\tilde{g}(s)$$

La solution de beaucoup d'équation différentielle se met naturellement sous la forme (7.5).

Exemple 7.12 **la méthode de la variation des constantes.** Nous voulons résoudre l'équation ordinaire à coefficient constant *avec* second membre

$$\dot{x}(t) + ax(t) = f(t)$$

En prenant la TL, nous trouvons que

$$\tilde{x}(s) = \frac{1}{s+a}\tilde{f}(s) + \frac{x_0}{s+a}$$

Comme l'originale de $1/s + a$ est $\exp(-at)$, en utilisant le résultat sur les produits de convolution, nous trouvons

$$x(t) = e^{-at}\left(x_0 + \int_0^t e^{a\tau}f(\tau)d\tau\right)$$

Ce résultat est connu sous le nom de la méthode de la variation des constantes et se généralise (à l'aide des décompositions en fraction simple) aux équations de degrés quelconques.

7.6 Aperçu des équations intégrales.

Il existe une classe d'équations intégrales (qu'on appelle de Voltera) qui s'écrivent sous la forme :

$$f(t) = \int_0^t f(\tau)K(\tau,t)d\tau + \lambda$$

Dans le cas où le noyau K est symétrique, c'est à dire qu'il s'écrit sous la forme $K(t-\tau)$, ces équations admettent une solution simple en terme de transformées de Laplace. En prenant la TL des deux cotés, on trouve :

$$\tilde{f}(s) = \tilde{f}(s)\tilde{K}(s) + \frac{\lambda}{s}$$

c'est à dire que $\tilde{f}(s) = \lambda/s(1-\tilde{K}(s))$. C'est ensuite un exercice de trouver l'originale ou en tout cas son développement asymptotique.

7.7 Aperçu des systèmes de contrôle asservis (*feedback systems*).

Quand on conduit une voiture et que l'on tourne le volant ou que l'on appuie sur la pédale de frein, on n'exerce pas directement une action sur les roues, mais on actionne des circuits hydrauliques qui s'en chargent. Ces circuits sont munis d'automatismes qui règlent la pression sur les roues exactement comme demandée, quelque soit les conditions extérieures. Pour pouvoir effectuer cela, il faut qu'ils soit munis des mécanismes correcteurs qui constamment comparent la direction ou la pression des roues actuelles à la consigne demandée et réduisent l'erreur. Si l'on regarde autour de nous, des objets les plus simples comme un réfrigérateur qui maintient sa température quand on ouvre ou ferme sa porte aux objets les plus complexes, comme l'ABS ou le pilotage d'un avion, nous sommes entouré d'automates. En cela bien sûr nous ne sommes qu'entrain d'imiter le monde vivant qui a implanté ces mécanismes à tous les niveaux, de la reproduction de l'ADN aux mouvements d'une bactérie ou à la marche d'un bipède.

Il se trouvent que la très grande majorité des automates est constitué d'automate dont l'action est gouvernée par des équations différentielles linéaires [7]. L'outil primordial pour étudier et concevoir les automates est la transformée de Laplace. On peut dire sans exagérer que les "*automateurs*" passent la majorité de leur temps dans l'espace de Laplace [8]. Nous allons étudier l'exemple fondamental des régulateurs PID.

Le régulateur PID est apparu dans les années 1920 ; les opérations d'intégration et de dérivation que nous allons voir étaient effectuées par des éléments mécaniques (ressort, masse,..) ou électroniques (circuits RLC).

Supposons que nous voulons maintenir un bain thermique à une température de consigne T_c dans une chambre à température T_R. Nous avons de nombreuses sources de perturbation, comme des courants d'air ou une température fluctuante dans la chambre [9]. Notre automate doit maintenir le bain à $T_c > T_R$ malgré ces perturbations (Figure 7.2).

En l'absence de source de chaleur, un bain à température $T > T_R$ perd de la chaleur et se refroidit :

$$\frac{\partial T}{\partial t} = r(T_R - T)$$

r est un coefficient qui reflète l'échange de la chaleur entre le bain et la chambre et dépend de l'isolation du bain. Nous ne connaissons pas la valeur exacte de r.

Nous pouvons injecter de la puissance $P(t)$ dans le bain à l'aide d'un générateur électrique, auquel cas l'évolution de la température dans le bain s'écrit

$$\frac{\partial T}{\partial t} = r(T_R - T) + P(t) \tag{7.6}$$

7. Depuis les années 1990 et la disponibilité des microcontrôleur, le paysage a pas mal changé.
8. Comme les cristallographes passent la majorité de leur temps dans l'espace de Fourier.
9. La porte !

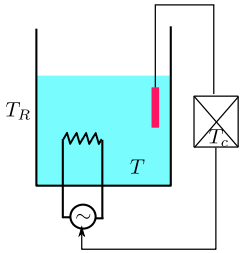


FIGURE 7.2 – Schéma général du dispositif. Le bain doit être amené et maintenu à température de consigne T_c, tandis que la température de la chambre est à T_R. La température du bain peut être augmenté en faisant passer un courant électrique dans une résistance à l'intérieur du bain. Un automate mesure constamment la température du bain T à l'aide d'une sonde, la compare à la consigne T_c et régule la tension du générateur électrique.

Le contrôleur doit décider à chaque instant t de la puissance à injecter $P(t)$ pour atteindre la consigne. Nous supposons qu'à $t = 0$, $T = T_R$.

La première idée serait de programmer le contrôleur de façon *proportionnelle* (d'où le P du PID) :

$$P(t) = \alpha(T_c - T)$$

plus on est loin de la consigne, plus on injecte de la puissance. Notre équation (7.6) s'écrit alors $\partial_t T = r(T_R - T) + \alpha(T_c - T)$ et sa transformée de Laplace nous donne

$$(s + \alpha + r)\hat{T}(s) = \frac{\alpha T_c + rT_R}{s} + T_R$$

Le développement asymptotique pour $s \to 0$ finalement nous montre que pour $t \to \infty$, le bain atteint la température

$$T_{eq} = \frac{\alpha T_c + rT_R}{\alpha + r} < T_c$$

Notre dispositif n'est pas très bon, puisque le bain ne peut pas atteindre T_c. Le problème est que si le bain atteint T_c, le générateur cesse d'y injecter de la puissance et le bain se met à refroidir. Il nous faut quelque chose qui continue d'injecter de la puissance même quand on est à T_c.

L'idée extrêmement élégante était d'apprendre à l'automate *ses erreurs passées*, en y ajoutant un terme qui somme l'historique des écarts à la consigne :

$$P(t) = \alpha(T_c - T) + \beta \int_0^t (T_c - T)d\tau$$

Cette fois, la TL de l'équation (7.6) nous donne

$$\left(s+\alpha+r+\frac{\beta}{s}\right)\hat{T}(s)=\frac{\alpha T_c+rT_R}{s}+\frac{\beta T_c}{s^2}+T_R$$

et le développement asymptotique pour $s\to 0$ nous montre que pour $t\to\infty$, le bain atteint la température

$$T_{eq}=T_c$$

et ceci quelque soit T_R, r, α et β. Notre simple automate se débrouille fort bien. Évidemment, en réalité, nous devons atteindre la température de consigne rapidement et nous y maintenir de façon stable ; c'est pourquoi dans les régulateur PID, il existe aussi un terme différentielle et que les coefficients α et β sont ajustable, mais le principe général est là.

7.8 La physique statistique.

En physique statistique, la fonction de partition Z est une sorte de moyenne (pondérée par $\beta=1/KT$) des énergies que peut atteindre un système. Si l'indice i dénombre les états possibles du système, chacun avec l'énergie E_i, alors

$$e^{-\beta F}=Z(\beta)=\sum_i e^{-\beta E_i}$$

La quantité F est appelé l'énergie libre du système.

Il arrive souvent que beaucoup d'état ont la même énergie et dans ce cas, ont peut les regrouper dans la somme :

$$Z(\beta)=\sum_E e^{-\beta E}n(E)$$

où cette fois, nous sommons sur les énergies disponibles au système ; $n(E)$ désigne le nombre d'états ayant l'énergie E. Si la différence entre les niveaux d'énergie est faible par rapport à notre mesure, nous pouvons récrire la somme ci-dessus sous forme d'une intégrale

$$Z(\beta)=\int_0^\infty e^{-\beta E}f(E)dE \tag{7.7}$$

où $f(E)dE$ est le nombre d'état avec une énergie entre E et $E+dE$; toutes les énergies sont mesurées par rapport à l'énergie minimum du système E_0 que nous choisissons comme référence : $E_0=0$.

Ce que nous voyons là est très simple : la fonction de partition est la transformée de Laplace de la densité d'énergie.

Exemple 7.13 **l'oscillateur harmonique.**

Soit une particule dans un champ quadratique, son énergie étant fonction de sa distance au centre : $E(x) = kx^2$. x ici est la variable qui dénombre les états. Le nombre d'état entre E et $E + dE$ est $1/\sqrt{kE}$. En nous reportant à la table des T.L., nous trouvons la fonction de partition

$$Z(\beta) = \sqrt{\frac{\pi}{k}}\beta^{-1/2}$$

Exemple 7.14 **Problème : Transition de phase.**

En utilisant la définition (7.7), démontrez que la fonction $Z(\beta)$ est forcément continue.

Par définition, une transition de phase (comme la transformation de l'eau en glace) est une discontinuité de la fonction de partition (ou de l'une de ses dérivées). Vous venez de démontrer que les transitions de phases ne peuvent pas exister. Où est l'erreur ? Ceci était un problème majeur de la physique statistique jusque dans les années 1920 et l'invention du modèle d'Ising par le scientifique du même nom. Ce modèle n'a reçu une solution qu'en 1944 par Onsanger. Les années 1970 ont vu apparaître les théories mathématiquement "sales" (dites groupes de renormalisation) pour traiter les transitions de phases de second ordre. Nous ne disposons à ce jour pas de théories mathématiques générales satisfaisantes pour les transitions de phase.

7.9 TL inverse.

Pour pouvoir effectuer les TL inverses, il faut connaître un minimum de la théorie d'intégration dans le plan complexe. Pour les lecteurs qui en sont familier, mentionnons la procédure qui est juste une adaptation des TF. Considérons la fonction $f(t)$ telle que $f(t < 0) = 0$. Nous pouvons écrire la fonction $e^{-ct}f(t)$ comme la TF inverse de sa TF :

$$e^{-ct}f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_0^{\infty} e^{-ct}f(t)e^{-i\omega t}dt\right)e^{i\omega t}d\omega$$

La borne inférieure de la deuxième intégrale commence à zéro puisque la fonction est nulle pour $t < 0$. (i) En multipliant les deux côtés par e^{ct} ; (ii) en posant $s = c + i\omega$; (iii) en prenant soin dans la deuxième intégrale du changement de variable $d\omega = ds/i$; (iv) et enfin en reconnaissant une TL classique dans l'intégrale intérieure, nous aboutissons à

$$f(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\hat{f}(s)e^{st}ds$$

l'intervalle $]c - i\infty, c + i\infty[$ désigne une droite parallèle à l'axe imaginaire et de coordonnée réelle c dans le plan complexe. Il faut choisir $c > Re(s_0)$, où s_0 est le pôle le plus à droite de la fonction $\hat{f}(s_0)$. Nous déférons une discussion plus détaillée de cette procédure jusqu'au chapitre sur les fonctions complexes.

Exercices.

§ 7.1 Trouver la TL des fonctions suivantes : $\sin(at)$; $\cos(at)$; $\sinh(at)$; $\cosh(at)$;

§ 7.2 Trouver, par la méthode de votre choix, l'original de $a^2/(s^2+a^2)^2$. Help : Remarquez que vous pouvez écrire cette fonction comme $(s^2+a^2)/(s^2+a^2)^2 - s^2/(s^2+a^2)^2$, et que le dernier terme vaut $(1/2)s(-d/ds)(1/(s^2+a^2))$. Il suffit ensuite d'utiliser les règles de manipulation des TL pour remonter à l'originale.

§ 7.3 Démontrez que $\mathrm{TL}[t^k] = k!/s^{k+1}$.

§ 7.4 Trouver la TL du delta de Dirac.

§ 7.5 La TL de la fonction escalier $\sum_{n=0} H(t-n)$ est $1/s(1-e^{-s})$.

§ 7.6 La TL d'une fonction $a-$periodique $f(t)$ ($f(t+a) = f(t)$) est $\hat{f}(s)/(1-e^{-as})$, où $\hat{f}(s) = \int_0^a f(t)\exp(-st)dt$.

§ 7.7 Représentez graphiquement, et trouvez la TL des fonctions suivantes : $(t-a)H(t-a)$; $H(t)-H(t-a)$; $\sum_{n=0}^{\infty}(-1)^n H(t-na)$; $\sum_{n=0}^{\infty}(-1)^n (t-na)H(t-na)$

§ 7.8 Les fonctions de Bessel jouent un rôle très important en physique mathématique. Elles jouent un rôle important pour l'équation de Laplace en coordonnées cylindriques, analogue à celui des fonctions circulaires à une dimension. L'une d'elle est définie par

$$I_0(z) = (1/2\pi)\int_0^{2\pi} e^{z\cos\theta}d\theta$$

Démontrer que sa T.L. est

$$\tilde{I}_0(s) = \frac{1}{\sqrt{s^2-1}}$$

Démontrez que $I_0(t) \approx (1/\sqrt{2\pi})\sqrt{t}\exp(t)$ quand $t \to +\infty$. Help : Pour Calculer des $\int R(\cos\theta, \sin\theta)d\theta$, on a intérêt à effectuer le changement de variable $u = \tan(\theta/2)$

§ 7.9 Les fonctions de Bessel $I_n(t)$ obéissent à l'équation différentielle

$$t^2u''(t) + tu'(t) - (t^2+n^2)u(t) = 0$$

Démontrer alors que la TL de la fonction $u(t)$ obéit à l'équation

$$(s^2-1)\hat{u}''(s) + 3s\hat{u}'(s) + (1-n^2)\hat{u}(s) = 0$$

Résoudre cette équation pour $n = 1$ (l'équation se ramène alors à une équation de premier ordre) et démontrer que sa solution est de la forme

$$\hat{u}(s) = C_0 + C_1/\sqrt{s^2-1}$$

Sachant que $I_1(0) = 0$, $I'(0) = 1/2$ et en utilisant le comportement asymptotique de $s\hat{u}(s)$, démontrer que $C_1 = -C_0 = 1$. Sachant que $I_0'(t) = I_1(t)$, déduire également la TL de la fonction $I_0(t)$. Enfin, en utilisant la relation de récurrence

$$2I_n'(t) = I_{n-1}(t) + I_{n+1}(t)$$

Obtenez la forme générale des TL des fonction de Bessel I.

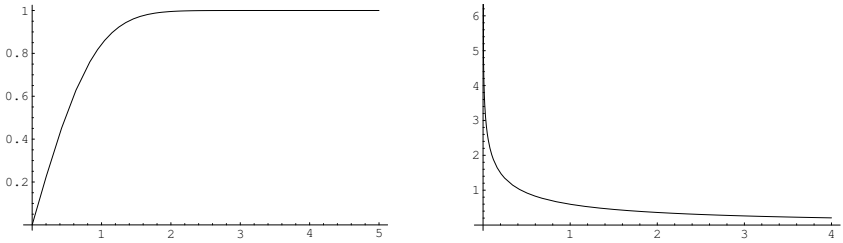


FIGURE 7.3 – Les fonctions $\mathrm{erf}(t)$ et $\exp(t)\Gamma(0,t)$

§ 7.10 La fonction d'erreur est définie par $\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-u^2} du$ (voir figure 7.3). Elle joue un rôle fondamentale en probabilité. Démontrer que la TL de $\exp(-t^2)$ est $(\sqrt{\pi}/2)\exp(s^2/4)(1 - \mathrm{erf}(s/2)$. En déduire également la TL de la fonction $\mathrm{erf}(t)$.

§ 7.11 La fonction Gamma d'Euler est définie par $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}\exp(-t)dt$. Il n'est pas difficile de démontrer que $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ (le faire) et donc que cette fonction est la généralisation de la fonction factorielle $n! = \Gamma(n+1)$. La fonction d'Euler incomplète est définie par $\Gamma(\alpha, z) = \int_z^\infty t^{\alpha-1}\exp(-t)dt$ (voir figure 7.3). Son développement asymptotique est donné (pour $z \to \infty$) par $z^{\alpha-1}\exp(-z)$. Tout ça pour vous demander de démontrer que $\mathrm{TL}[1/(1+t)] = \exp(s)\Gamma(0,s)$. Généraliser se résultat aux puissance négative de $(1+t)$.

§ 7.12 Trouver le comportement asymptotique de l'originale de la fonction $1/(s^2+a^2)^2$. Attention : les deux pôles sont imaginaires pures et contribuent également.

§ 7.13 La fonction de Bessel J d'ordre 0 est définie par $J_0(z) = I_0(iz)$,et sa TL est $(s^2+1)^{-1/2}$ (Pouvez vous le démontrer ?). Démontrer que son comportement asymptotique est donnée par $J_0(z) \approx \sqrt{2/\pi z}\cos(z - \pi/4)$.

§ 7.14 Résoudre $\ddot{x} + \omega^2 x = b\sin(\omega t)$ avec des conditions initiales $x(0) = x_0$ et $\dot{x}(0) = v_0$. Notez que c'est l'équation d'un oscillateur harmonique forcé à sa fréquence de résonance.

§ 7.15 Résoudre $x^{(3)} + 3\ddot{x} + 3\dot{x} + x = 1$ avec les conditions initiales nulles.

§ 7.16 Résoudre $x^{(4)} + 2\ddot{x} + x = \sin t$ avec les conditions initiales nulles.

§ 7.17 Le mouvement d'une particule dans un champs magnétique peut être ramené à la résolution du système suivant :

$$\dot{x} = \alpha y \ ; \ \dot{y} = -\alpha x$$

où x, y sont les composantes du vecteur vitesse et α une constante proportionnelle au champs magnétique et à la charge de la particule. Les conditions initiales sont à $t = 0$, $x = x_0; y = y_0$. Résoudre ce système à l'aide des transformées de Laplace.

§ 7.18 Résoudre

$$\begin{aligned} \ddot{x} - x + y + z &= 0 \\ x + \ddot{y} - y + z &= 0 \\ x + y + \ddot{z} - z &= 0 \end{aligned}$$

avec les conditions initiales $x(0) = 1$ et $y(0) = z(0) = \dot{y}(0) = \dot{z}(0) = 0$.
Sol. : $x(t) = (2/3)\cosh(t\sqrt{2}) + (1/3)\cos t$; $y(t) = z(t) = -(1/3)\cosh(t\sqrt{2}) + (1/3)\cos t$.

§ 7.19 Donner la solution générale, sous forme de produit de convolution, de l'équation

$$y'' + 2ay + b = f(t)$$

avec les conditions initiales $y(0) = y_0$; $y'(0) = v_0$. Considérer les deux cas où $a^2 - b \neq 0$ et $a^2 - b = 0$.

Problèmes avancés.

Problème 7.1 la propagation des ondes.

Nous nous intéressons à nouveau à la résolution de l'équation des cordes vibrantes. Nous l'avons déjà rencontré dans le chapitre sur les distributions, et nous le verrons encore dans la section consacrée aux fonctions de Green. Nous allons ici utiliser les TL pour résoudre ces équations, *avec des conditions initiales données*. Nous avons vu que les TL sont très bon quand il s'agit d'avoir un début des temps.

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \tag{7.8}$$
$$u(x,0) = f(x) \tag{7.9}$$
$$\partial_t u(x,0) = g(x) \tag{7.10}$$

Nous savons que sa solution est donnée par

$$u(x,t) = f(x-ct) + f(x+ct) + \frac{1}{2c}\int_{x-ct}^{x+ct} g(\xi)d\xi \tag{7.11}$$

Nous allons établir la même chose, mais en utilisant de façon combiné les TF *et* les TL, ces dernières ayant l'avantage de gérer automatiquement les conditions initiales. Le schéma de la résolution que nous allons mener est le suivant : $u(x,t) \overset{\text{TL}}{\to} \hat{u}(x,s) \overset{\text{TF}}{\to} \tilde{\hat{u}}(q,s) \overset{\text{TL}^{-1}}{\to} \tilde{u}(q,t) \overset{\text{TF}^{-1}}{\to} u(x,t)$. Noter que $t \in [0, +\infty[$, donc nous allons effectuer des TL par rapport à cette variable. Par contre, $x \in]-\infty, +\infty[$, donc nous allons procéder à des TF pour cette dernière.

1. En prenant la TL de l'équation (7.8) par rapport à la variable t, démontrer que

$$-c^2 \frac{d^2\hat{u}(x,s)}{dx^2} + s^2\hat{u}(x,s) = sf(x) + g(x).$$

2. En prenant la TF de cette dernière, démontrer que

$$\tilde{\hat{u}}(q,s) = \frac{s}{c^2q^2 + s^2}\tilde{f}(q) + \frac{1}{c^2q^2 + s^2}\tilde{g}(q)$$

où $\tilde{f}$ et $\tilde{g}$ sont les TF des fonctions f et g .

3. En prenant la TL inverse, démontrer qu'on obtient

$$\tilde{u}(q,t) = \tilde{f}(q)\cos(ctq) + \frac{\sin(ctq)}{cq}\tilde{g}(q) \tag{7.12}$$

4. Résultat intermédiaire. Démontrer que si la TF de $g(x)$ et $\tilde{g}(q)$, alors

$$\int_{x-a}^{x+a} g(\xi)d\xi \xrightarrow{\text{TF}} \frac{2\sin(aq)}{q}\tilde{g}(q)$$

5. En utilisant le résultat intermédiaire ci-dessus, et les règles de translations pour les TF, prendre la TF inverse de (7.12) pour obtenir l'équation dans l'espace direct.

Problème 7.2 Diffusion sur réseau.

Nous souhaitons résoudre un système infini d'équations différentielles couplées :

$$\frac{du_n}{dt} = -2u_n + u_{n+1} + u_{n-1} \qquad n = \ldots -2, -1, 0, 1, 2, \ldots \tag{7.13}$$

Nous cherchons les fonctions $u_n(t)$ avec les conditions initiales

$$\begin{aligned} u_0(0) &= 1 \\ u_n(0) &= 0 \quad n \neq 0 \end{aligned}$$

Évidement, ce n'est pas par hasard que nous sommes intéressés par ces équations, elles constituent la version "discrète", ou "sur réseaux", de l'équation de diffusion que nous avons abondement étudié en cours.

1. Prenez la Transformée de Laplace des équations (7.13), en prenant bien soin de distinguer le cas $n = 0$ du cas $n \neq 0$.

2. Comme vous le voyez, la TL a transformé notre système d'équations différentielles en un système d'équations algébriques linéaires sur les $\tilde{u}_n(s)$. Nous chercherons la solution sous la forme de

$$\tilde{u}_n(s) = f(s)^{|n|}g(s) \tag{7.14}$$

où les fonctions $f(s)$ et $g(s)$ sont à déterminer [10]. Quelle est la relation entre $\tilde{u}_n(s)$ et $\tilde{u}_{-n}(s)$? En utilisant l'expression (7.14) dans les équations que vous avez obtenu pour le cas $n > 0$, démontrez que $f(s)$ doit obéir à l'équation

$$f(s) + \frac{1}{f(s)} = s + 2 \tag{7.15}$$

dont une des solutions est

$$f(s) = \frac{s + 2 + \sqrt{s^2 + 4s}}{2} \tag{7.16}$$

Que vaut $-f(s) + 1/f(s)$?

3. Connaissant $f(s)$, il vous reste maintenant à utiliser l'équation que vous aviez obtenu pour le cas $n = 0$ pour déterminer $g(s)$. Donnez alors l'expression complète de la fonction $\tilde{u}_n(s)$. Si nous avions plus de temps, nous aurions tenté de d'inverser la TL et de montrer la relation qui existe entre la solution (7.16) et (i) les fonctions de Bessel $I_n(t)$; (ii) la solution de l'équation de diffusion continue $\partial_t u - \partial_x^2 u = \delta(x)\delta(t)$. Mais ce sera de la gourmandise. C'est déjà pas mal de connaître les solutions sous forme de leurs TL.

4. Démontrez que le pôle le plus à droite des fonctions $\tilde{u}_n(s)$ est $s = 0$. Démontrer alors que le terme dominant du développement de $\tilde{u}_n(s)$ au voisinage de $s = 0$ est $1/\sqrt{s}$. En déduire le développement asymptotique de $u_n(t)$ quand $t \to +\infty$.

10. Ceci est une technique classique de résolution dont le nom savant est "matrice de transfert".

Problème 7.3 le théorème H.

Introduction. Nous souhaitons déterminer la fonction $c(x,t)$ obéissant à l'équation

$$\int_{p=0}^{1}\int_{x_1=0}^{\infty}\int_{x_2=0}^{\infty} c(x_1)c(x_2)\delta\left(p(x_1+x_2)-x\right)dx_2dx_1dp$$
$$-c(x)=0 \tag{7.17}$$

δ désigne bien-sûr la distribution δ de Dirac. Cette équation est à la base du théorème H établit par Boltzmann vers 1870 et forme le cœur de la théorie cinétique des gaz et de la physique statistique. Nous allons voir ici qu'aussi intimidant qu'elle paraisse à priori, cette équation se traite en faite facilement par les outils que nous avons vus dans notre cours. Du point de vue de la physique, la variable x dénote l'énergie cinétique ; $c(x)$ est le nombre de molécules ayant l'énergie x ; comme les molécules ont une énergie cinétique positive,

$$c(x)=0 \quad \text{si } x<0 \tag{7.18}$$

p désigne le taux de distribution d'énergie après un choc entre deux particules : après un choc, deux molécules d'énergie initialement x_1 et x_2 deviennent d'énergie $p(x_1+x_2)$ et $(1-p)(x_1+x_2)$, où p est un nombre aléatoire entre 0 et 1. L'équation ci-dessus est une équation de bilan, mesurant la variation du nombre de molécule à énergie x après chaque choque. A l'équilibre, $c(x)$ doit rester stable ; c'est cette distribution d'équilibre que nous souhaitons trouver.

1. Nous pouvons effectuer l'intégrale triple ci-dessus dans l'ordre que nous voulons. Nous commencerons par intégrer sur x_2. Démontrer alors que l'intégrale triple se transforme en une intégrale double

$$\int_{p=0}^{1}(1/p)\int_{x_1=0}^{\infty} c(x_1)c(x/p-x_1)dx_1dp \tag{7.19}$$

En utilisant la condition (7.18), montrer que l'intégrale double ci-dessus se ramène à

$$\int_{p=0}^{1}(1/p)\int_{x_1=0}^{x/p} c(x_1)c(x/p-x_1)dx_1dp$$

L'intégrale sur x_1 commence à prendre la tête d'un produit de convolution, nous avons intérêt à passer en TL.

2. Concentrons nous sur l'intégrale sur x_1 :

$$I_1(x)=\int_{x_1=0}^{x/p} c(x_1)c(x/p-x_1)dx_1$$

et passons en Transformé de Laplace $x \overset{\text{TL}}{\to} \beta$, $c(x) \overset{\text{TL}}{\to} \hat{c}(\beta)$, $I_1(x) \overset{\text{TL}}{\to} \hat{I}_1(\beta)$. Nous savons, d'après la règle des dilatation en TL, que $\text{TL}[f(x/p)]=p\hat{f}(\beta p)$. En utilisant cette relation, et le résultat sur les produits de convolution, démontrer que

$$\hat{I}_1(\beta)=p\hat{c}(\beta p)^2$$

Et donc que l'équation du bilan (7.17) se met sous la forme

$$\int_{p=0}^{1}\hat{c}(\beta p)^2dp-\hat{c}(\beta)=0$$

Effectuez un dernier changement de variable évident pour mettre le résultat sous la forme de

$$\frac{1}{\beta}\int_0^\beta \hat{c}(u)^2 du - \hat{c}(\beta) = 0 \tag{7.20}$$

3. Vérifier que la fonction

$$\hat{c}(\beta) = \frac{A}{\beta + A}$$

où A est une constante est solution de l'équation (7.20) ci-dessus. [Pour trouver cette solution, il suffit de remarquer que l'équation (7.20) peut se transformer en une équation différentielle de Riccati en dérivant une fois].

4. Soit $T = \int_0^\infty xc(x)dx$; T représente l'énergie totale du gaz. Démontrer que

$$T = -\left.\frac{\partial \hat{c}(\beta)}{\partial \beta}\right|_{\beta=0}$$

En déduire que $A = 1/T$. En inversant la TL, déduire que la distribution des énergies dans le système à l'état stationnaire est

$$c(x) = \frac{1}{T}e^{-x/T}$$

5. En réalité, ce que nous venons de trouver n'est pas vraiment le théorème H. Nous pouvons écrire la version temporelle de cette équation en calculant $c(x,t)$. Dans ce cas, le membre de droite de l'équation (7.17) n'est pas nulle, mais vaut $\partial_t c(x,t)$. On peut alors démontrer que la quantité

$$S(t) = -\int_0^\infty c(x,t)\log\left(c(x,t)\right)dx$$

qu'on appelle *entropie* est une fonction croissante du temps : $\partial_t S \geq 0$, l'égalité ne se réalisant qu'à l'équilibre.

Problème 7.4 le modèle de Glauber.

Les transitions de phase paraissait hors de porté de la physique statistique jusqu'à ce que Ising propose son modèle de magnétisme dans les années 1920. Il existe de nombreuse façon de résoudre ce modèle à une dimension où paradoxalement, il n'existe pas de transition de phase. Plus exactement, la transition de magnétisation arrive seulement quand la température est abaissée à 0K. En 1963, Glauber a proposé un modèle cinétique équivalent au modèle d'Ising à une dimension.

Considérons une chaîne de dipôle magnétique, ou spin. Nous supposons que les spins ne peuvent prendre que deux valeurs, ± 1, correspondant aux dipôles pointant vers le "haut" ou vers le "bas". Quand une majorité de spin pointe dans une direction, le matériaux devient magnétique. Chaque spin n'interagit qu'avec ses deux plus proches voisins, et tente d s'aligner sur eux. Plus exactement, il a une certaine probabilité de s'aligner sur ces voisins ... Plus tard.aboutir au résultat (7.11)

8 Les fonctions de Green.

8.1 Entrée en matière

Les fonctions de Green constituent une méthode assez général de résolution d'équations différentielles, ou de transformation d'équations différentielles en équations intégrales. Elles sont extrêmement utilisées en mécanique quantique, où on les appelle des propagateurs, et en théorie des processus stochastiques. Nous n'aborderons ce sujet que très légèrement ici, juste pour rappeler les grands principes de la méthode.

Supposons que nous voulons résoudre l'équation différentielle

$$a\frac{d^2x}{dt^2} + b\frac{dx}{dt} + cx = f(t) \tag{8.1}$$

avec les conditions initiales $x(0) = x_0$ et $x'(0) = \tilde{x}_0$. Ceci est par exemple l'équation du mouvement d'une particule soumise à une force $f(t)$. a et b et c peuvent être fonction du temps. Pour résoudre cette équation différentielle, il nous faut trouver la solution de l'équation homogène, et lui ajouter une solution particulière. Nous cherchons justement une solution particulière.

Supposons que nous savons calculer la réponse de la particule à une force impulsionnelle (genre δ de Dirac) appliquée au temps t'. Saurions nous calculer la réponse de la particule à une force générale $f(t)$? La réponse est oui : la force $f(t)$ peut être vue comme une superposition d'impulsions appliquées à différent temps. Il suffit donc de superposer les réponses aux divers impulsions pour obtenir la réponse à la force $f(t)$. Plus exactement, on peut écrire

$$f(t) = \int_0^\infty f(t')\delta(t-t')dt' \tag{8.2}$$

ce qui veut dire que la force $f(t)$ est la superposition d'impulsions appliquées au temps t', *avec le poids* $f(t')$. Revenons à notre équation différentielle, et appelons $G_{t'}(t)$ la réponse à l'impulsion appliquée au temps t'. Comme mettre les indices est un peu lourd comme notation, nous noterons cette fonction plutôt $G(t,t')$. De par sa définition, elle doit satisfaire à

$$a\frac{d^2G(t,t')}{dt^2} + b\frac{dG(t,t')}{dt} + cG(t,t') = \delta(t-t')$$

Notez que toutes les dérivations sont faites par rapport à t. Multiplions les deux côtés de l'équation par $f(t')$. Comme $f(t')$ ne dépend pas de t, on peut la rentrer à l'intérieur de l'opérateur différentiel, et écrire :

$$a\frac{d^2[f(t')G(t,t')]}{dt^2} + b\frac{d[f(t')G(t,t')]}{dt} + cf(t')G(t,t') = \delta(t-t')f(t')$$

Intégrons maintenant les deux cotés par rapport à t'. Comme la dérivation est par rapport à t, nous pouvons (jetant par dessus bord la décence et l'exigence à priori de la convergence uniforme) échanger la dérivation et l'intégration.

$$a\frac{d^2}{dt^2}\int_0^\infty f(t')G(t,t')dt' + b\frac{d}{dt}\int_0^\infty f(t')G(t,t')dt' + c\int_0^\infty f(t')G(t,t')dt' = \int_0^\infty \delta(t-t')f(t')dt' \qquad (8.3)$$

Nous remarquons, d'après (8.2), que le côté droit de l'équation ci-dessus est juste $f(t)$. Appelons

$$y(t) = \int_0^\infty f(t')G(t,t')dt' \qquad (8.4)$$

et nous voyons, d'après (8.3), que $y(t)$ est solution de l'équation (8.1) ! Remarquez l'élégance, nous devons calculer une seule fois la fonction de Green $G(t,t')$ pour une équation différentielle. Ensuite, quelque soit le membre de droite, la solution s'obtient par une simple intégration. La solution générale de l'équation différentielle s'écrit maintenant

$$x(t) = C_1x_1(t) + C_2x_2(t) + y(t)$$

où C_1 et C_2 sont choisit pour satisfaire les conditions initiales.

Nous avons occulté pas mal de point important. Voyons quelques exemples.

Exemple 8.1 Equation de première ordre

Soit l'équation

$$dx/dt + \alpha x = f(t) \qquad (8.5)$$

La fonction de Green est la solution de

$$dG(t,t')/dt + \alpha G(t,t') = \delta(t-t')$$

Prenons la TF des deux côtés de l'équation (par rapport à t bien sûr)

$$\tilde{G}(\omega,t') = \frac{\exp(-i\omega t')}{i\omega+\alpha}$$

$H(t)$ étant la fonction de Heaviside, nulle pour $t<0$ et 1 pour $t>0$. Comme vous vous souvenez, la TF de $H(t)\exp(-\alpha t)$ est justement $1/(i\omega+\alpha)$. Donc,

$$G(t,t') = H(t-t')\exp(-\alpha(t-t'))$$

Comme vous le remarquez, $G(t,t') = 0$ si $t' > t$. Cela est normal, puisque $G(t,t')$ est la réponse, au temps t, à une impulsion au temps t'. Si t' est *plus tard* que t, la réponse est nulle.

Prenons maintenant plusieurs formes de $f(t)$ de l'exercice précédent.

Exemple 8.2 $f(t) = H(t)t$

$$\begin{aligned} y(t) &= \int_0^\infty H(t')t'H(t-t')\exp(-\alpha(t-t'))dt' \\ &= \int_0^\infty t'H(t-t')\exp(-\alpha(t-t'))dt' \\ &= \int_0^t t'\exp(-\alpha(t-t'))dt \\ &= (1/\alpha^2)(\exp(-\alpha t)-1) + (1/\alpha)t \end{aligned}$$

Exemple 8.3 $f(t) = H(t)\sin(\beta t)$. En suivant les même étapes,

$$\begin{aligned} y(t) &= \int_0^t \sin(\beta t')\exp(-\alpha(t-t'))dt \\ &= \frac{1}{\alpha^2+\beta^2}\left[\beta e^{-\alpha t} + \beta\cos(\beta t) + \alpha\sin(\beta t)\right] \end{aligned}$$

Vous voyez ici comment on résout une fois l'équation différentielle pour la fonction de Green, et qu' ensuite, il suffit d'appliquer une intégration pour trouver la solution générale.

8.2 Généralisation.

En langage opératoriel, on écrit une équation différentielle comme

$$\mathcal{L}[x(t)] = f(t)$$

où L est un opérateur différentiel (dans l'exemple 8.1 ci-dessus $\mathcal{L} = d/dt + \alpha$), c'est à dire qui transforme une fonction en une autre fonction. La solution de cette équation s'écrira

$$x(t) = \mathcal{L}^{-1}[f(t)] \tag{8.6}$$

Trouver la fonction de Green revient à trouver l'opérateur $\mathcal{L}^{-1}$ et ce n'est pas un hasard donc qu'il comporte une intégration. Si on s'est donné une base, on peut représenter L par une matrice (infinie) et trouver la fonction de Green revient à inverser cette matrice.

Concrètement, nous résolvons l'équation

$$\mathcal{L}\left[G(t,t')\right] = \delta(t-t') \tag{8.7}$$

où la fonction $G(t,t')$ est la réponse du système [1] au temps t à une impulsion au temps t'. L'opérateur $\mathcal{L}$ s'applique à la variable t, et il faut considérer, pour l'instant, t' comme un paramètre, une constante. Une fois ce problème résolu, nous trouvons la solution de l'équation (8.6) par une intégration :

$$x(t) = \int_{-\infty}^{\infty} G(t,t')f(t')dt'$$

Invariance par translation. Un cas particulier important se présente quand l'opérateur $\mathcal{L}$ est invariant par translation. Par exemple, l'opérateur $\mathcal{L} = ad/dt + b$ est invariant par translation dans le temps si a, b ne dépendent pas explicitement de t. Cela veut dire que si l'on change $t \to \tau + t_0$ (où t_0 est une constante) , la forme de l'opérateur ne change pas et s'écrit en fonction de la nouvelle variable $ad/d\tau + b$. Dans ce cas, il est évident que la solution $G(t,t')$ ne doit dépendre [2] que de la distance entre t et t' :$G(t,t') = G(t-t')$ et donc

$$x(t) = \int_{-\infty}^{\infty} G(t-t')f(t')dt'$$

et nous reconnaissons là un produit de convolution : $x = G * f$.

Équation intégrale. La démarche de Green est très générale et va au delà de simple fonction de t au second membre d'équation (8.7). Supposons que notre équation est un peu plus compliquée :

$$\mathcal{L}[x(t)] = f(t,x)$$

Le membre de droite comporte explicitement un terme en x, comme par exemple $t.x^{1/2}$ ce qui rend la résolution de l'équation nettement plus ardue par les techniques classiques. Mais cela ne change rien pour les fonctions de green. La solution s'écrira toujours

$$x(t) = \text{Sol.Homogène} + \int_{-\infty}^{\infty} f(t',x)G(t,t')dt' \tag{8.8}$$

Nous avons transformé une équation différentielle en une équation intégrale. A priori, nous n'avons pas gagné grand chose, ces dernières étant plus compliquées à résoudre que les premières. Mais souvent, et surtout en mécanique quantique, la forme (8.8) se traite bien par la technique des perturbations (objet du prochain chapitre), et c'est un grand avantage que de pouvoir en disposer. Nous en verrons des exemples plus bas.

1. Nous utilisons t comme le temps ici pour illustrer, mais elle peut représenter n'importe quelle quantité bien sûr.
2. Poser $u = t - t'$ dans l'équation (8.7)

8.3 Le potentiel électrostatique.

Nous avons peut être présenté les fonctions de Green comme quelque chose de compliqué, mais le lecteur peut remarquer qu'il utilise les fonctions de Green depuis qu'il a appris l'électrostatique. Le potentiel électrostatique $\phi(\mathbf{r})$ créé par une charge *ponctuelle* unité en $\mathbf{r}'$ est

$$G(\mathbf{r},\mathbf{r}') = \frac{1}{4\pi\epsilon_0}\frac{1}{|\mathbf{r}-\mathbf{r}'|} \tag{8.9}$$

Si maintenant nous avons une distribution de charge $\rho(\mathbf{r}')$ dans l'espace, le potentiel crée par elle au point $\mathbf{r}$ vaut

$$\phi(\mathbf{r}) = \int G(\mathbf{r},\mathbf{r}')\rho(\mathbf{r}')d\mathbf{r}' \tag{8.10}$$

Nous connaissons cette formule depuis la première année de l'université. Nous savons par ailleurs que le potentiel obéit à l'équation de Poisson

$$-\Delta\phi = \rho/\epsilon_0 \tag{8.11}$$

Nous oublierons dorénavant le facteur ϵ_0 pour alléger les notations. Il n'est pas difficile, vue les équations (8.9-8.11) de suspecter que $G(\mathbf{r},\mathbf{r}')$ est la fonction de Green de l'équation de Poisson, c'est à dire qu'elle obéit à

$$-\Delta G(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') \tag{8.12}$$

Démontrons ce résultat. Jusque là, nous n'avions manipulé que des TF et des distributions à une dimension. Leurs généralisation à trois dimensions n'est pas vraiment compliquée. Par exemple, la TF est définie par

$$\tilde{f}(\mathbf{q}) = \int f(\mathbf{r})\exp(-i\mathbf{q}.\mathbf{r})d\mathbf{r}$$

où $\mathbf{q}$ et $\mathbf{r}$ sont des vecteurs à trois dimension et $d\mathbf{r}$ désigne le volume infinitésimal $dxdydz$. En général, les vecteurs sont notés par des caractères gras droits, et leurs normes par le même symbole mais non gras et en italique. Par exemple, $q = |\mathbf{q}|$. Les opérations sur les TF se généralise également assez facilement. Prenons la TF des deux cotés de (8.12) par rapport à $\mathbf{r}$:

$$\tilde{G}(\mathbf{q},\mathbf{r}') = \frac{e^{-i\mathbf{q}.\mathbf{r}'}}{q^2} \tag{8.13}$$

puisque le numérateur est la TF de la fonction δ translatée de $\mathbf{r}'$, et que la TF du Laplacien d'une fonction est $-q^2$ fois la TF de la fonction (pouvez vous démontrez ce résultat ?). Il nous faut maintenant inverser la TF pour retrouver la fonction de Green :

$$\tilde{G}(\mathbf{r},\mathbf{r}') = \frac{1}{(2\pi)^3}\int \frac{e^{i\mathbf{q}(\mathbf{r}-\mathbf{r}')}}{q^2}d\mathbf{q} \tag{8.14}$$

Pour effectuer l'intégration, passons aux coordonnées sphériques, où nous prenons l'axe q_z parallèle à $(\mathbf{r}-\mathbf{r}')$. Dans ce cas, $\mathbf{q}(\mathbf{r}-\mathbf{r}') = q|\mathbf{r}-\mathbf{r}'|\cos\theta$ et $d\mathbf{q} = q^2 \sin\theta dq d\theta d\phi$. L'intégrale (8.14) s'écrit alors

$$\tilde{G}(\mathbf{r},\mathbf{r}') = \frac{1}{(2\pi)^3}\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \int_0^{\infty} dq \sin\theta . e^{iq|\mathbf{r}-\mathbf{r}'|\cos\theta}$$

Une première intégration sur ϕ ne mange pas de pain et nous sort un facteur 2π. Ensuite, en posant $u = \cos\theta$, le reste s'écrit

$$\tilde{G}(\mathbf{r},\mathbf{r}') = \frac{1}{(2\pi)^2}\int_0^{\infty} dq \int_{-1}^{+1} e^{iq|\mathbf{r}-\mathbf{r}'|u} du$$

et en intégrant sur u, nous trouvons

$$\tilde{G}(\mathbf{r},\mathbf{r}') = \frac{1}{2\pi^2}\int_0^{\infty} \frac{\sin q|\mathbf{r}-\mathbf{r}'|}{q|\mathbf{r}-\mathbf{r}'|} dq$$

Un changement de variable $q|\mathbf{r}-\mathbf{r}'| \to q$ nous donne

$$\tilde{G}(\mathbf{r},\mathbf{r}') = \frac{1}{2\pi^2}\frac{1}{|\mathbf{r}-\mathbf{r}'|}\int_0^{\infty} \frac{\sin q}{q} dq$$

Nous avons donc bien mis en évidence la dépendance en $1/|\mathbf{r}-\mathbf{r}'|$. L'intégrale maintenant n'est qu'une constante que nous pourrions calculer à l'aide de la théorie des fonctions analytiques et vaut $\pi/2$. Ce qui nous donne exactement l'expression (8.9).

Problème 8.1 Il n'est pas difficile de généraliser la technique ci-dessus et trouver la fonction de Green de l'opérateur $\Delta - \lambda^2$, où λ est un réel. En langage claire, résolvez

$$\Delta G(\mathbf{r},\mathbf{r}') - \lambda^2 G(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r},\mathbf{r}')$$

Ceci est extrêmement utilisé en mécanique quantique. Le lecteur y reconnaîtra peut être un semblant d'équation au valeur propre.

8.4 La propagation des ondes

Avant d'attaquer le problème de la fonction de Green de la corde vibrante, nous avons besoin de quelques résultats intermédiaires. Quelle est par exemple la TF de la fonction $f(t) = H(t)\sin(\omega_0 t)$? Cette question n'a pas de sens à priori, puisque la fonction sin n'est pas sommable (et ne tend sûrement pas vers zéro quand $t \to \infty$). Mais nous pouvons calculer la TF de $H(t)\exp(-\nu t)\sin(\omega_0 t)$ et une fois la TF calculée, faire $\nu \to 0$. Cela nous donnera, et on laisse au lecteur le soin de le démontrer, que

$$\tilde{f}(\omega) = \frac{\omega_0}{-\omega^2 + \omega_0^2}$$

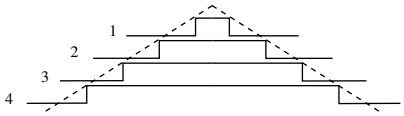


FIGURE 8.1 – La solution $u(x,t)$ en fonction de x pour les temps t_0, $2t_0$,...

Quelle est maintenant la réponse d'un oscillateur (initialement au repos) à une force impulsionnelle ? Nous devons résoudre

$$\frac{d^2y}{dt^2} + \omega_0^2 y = f_0\delta(t) \tag{8.15}$$

En faisant un aller–retour dans l'espace de Fourier, nous voyons que la solution est

$$y(t) = H(t)\frac{f_0}{\omega_0}\sin(\omega_0 t)$$

Nous somme maintenant bien outillé pour calculer la réponse d'une corde vibrante (initialement au repos) à une force impulsionnelle. Nous notons $u(x,t)$ la hauteur de la corde à l'abscisse x et au temps t :

$$\frac{\partial^2 u}{\partial t^2} - c^2\frac{\partial^2 u}{\partial x^2} = f_0\delta(x)\delta(t)$$

En prenant la TF par rapport à la variable x, nous trouvons pour $\tilde{u}(q,t)$

$$\frac{\partial^2 \tilde{u}}{\partial t^2} + c^2q^2\tilde{u} = f_0\delta(t)$$

Mais cela est justement l'équation (8.15) que l'on vient de résoudre, et nous avons donc

$$\tilde{u}(q,t) = f_0 H(t)\frac{\sin(ct\,q)}{cq}$$

Il nous reste maintenant à inverser la TF, ce qui est facile si on se souvient de la TF de la fonction Porte $\Pi(x/a)$ rencontrée au chapitre 3 :

$$u(x,t) = \frac{f_0}{c}H(t)\Pi(\frac{x}{ct})$$

(Exercice : Est-ce tout cela dimensionnellement correct ?) Cette solution est représenté sur la figure 8.1 .

Il est évident que si au lieu d'appliquer la force f_0 en $x = 0$ nous avions appliqué la force $f_{x'}$ en $x = x'$, la solution, qui est la fonction de Green de la propagation, s'écrit

$$G(x,t;x',0) = \frac{f_{x'}}{c} H(t)\Pi(\frac{x-x'}{ct})$$

Si la corde était initialement au repos et on y appliquait la force distribuée $f(x)$, la déformation de la corde est donnée par

$$\begin{aligned} u(x,t) &= c^{-1}H(t)\int_{-\infty}^{+\infty} f(x')\Pi(\frac{x-x'}{ct})dx' \\ &= c^{-1}H(t)\int_{x-ct}^{x+ct} f(x')dx' \end{aligned}$$

L'influence d'un événement en x' (à l'instant $t = 0$) ne peut être ressenti en x à l'instant t que si cet événement était à l'intérieur du cône d'influence de ce dernier, c'est à dire que si $x - ct < x' < x + ct$. En terme moins mystique, une perturbation se propage à vitesse c.

8.5 Disposer d'une base propre.

La lecture de cette section nécessite des connaissances importé du chapitre 10 sur les opérateurs linéaires.

Nous souhaitons résoudre l'équation

$$\mathcal{L}G(t,t') = \delta(t-t') \tag{8.16}$$

où $\mathcal{L}$ est un opérateur linéaire opérant sur la variable t. L'opérateur

$$\mathcal{L} = ad_t^2 + bd_t + c$$

de l'exemple 8.1 est un échantillon.

Supposons que nous disposons d'une base orthonormale [3] dans l'espace des fonctions (voir chapitre 2) $\{\phi_0(t), \phi_1(t), ...\}$. Rappelons que Cela veut dire que n'importe quelle fonction $f(t)$ peut s'écrire sous la forme

$$f(t) = \sum_n f_n \phi_n(t)$$

où les coefficients f_n de projection sur $\phi_n(t)$ sont donnés par [4].

$$f_n = \int_I f(t)\phi_n(t)dt \tag{8.17}$$

3. La base peut être dénombrable comme les séries de Fourier ou continue comme les transformées de Fourier, ceci ne change pas la discussion ici du moment que l'on connaît les distributions.

4. Les notations de Dirac sont plus concises et précises, nous les rencontrerons au prochain chapitre.

Supposons de plus que la base utilisée est une *base propre* de l'opérateur $\mathcal{L}$; cela veut simplement dire que

$$\mathcal{L}\phi_n(t) = \lambda_n \phi_n(t) \tag{8.18}$$

où les nombres λ_n sont connu. Par exemple, la fonction $e^{i\omega t}$ est une fonction propre de l'opérateur d_t avec la valeur propre $i\omega$.

Nous pouvons décomposer la fonction (encore) inconnue $G(t, t')$ sur cette base

$$G(t, t') = \sum_n g_n(t')\phi_n(t) \tag{8.19}$$

et nous ne connaissons pas encore les coefficients $g_n(t')$. Par contre, nous connaissons la décomposition de $\delta(t - t')$ dans cette base :

$$\delta(t - t') = \sum_n \phi_n(t')\phi_n(t) \tag{8.20}$$

qui découle simplement de la relation (8.17). En remplaçant les décomposition (8.19,8.20) dans l'équation (8.16), nous avons alors

$$\sum_n \left(\lambda_n g_n(t') - \phi_n(t')\right) \phi_n(t) = 0 \tag{8.21}$$

et nous savons que, puisque les $\phi_n(t)$ sont une base, cela implique

$$g_n(t') = \frac{1}{\lambda_n}\phi_n(t')$$

et la fonction de Green est simplement donnée par

$$G(t, t') = \sum_n \frac{1}{\lambda_n}\phi_n(t')\phi_n(t) \tag{8.22}$$

8.6 Propagateur pour l'équation de Schrödinger.

9 Calcul des perturbations.

Le calcul des perturbations n'est pas une méthode scientifique utilisée dans les hôpitaux psychiatriques pour évaluer les symptômes d'un patient. Il existe peu de problèmes exactement solubles en physique et il faut souvent recourir aux techniques d'approximations. Une des techniques les plus utilisées est celle qui porte le nom de ce chapitre. L'idée de base est une généralisation du développement de Taylor : si nous connaissons la valeur d'une fonction au point x_0 , nous pouvons calculer, sous certaines conditions, la valeur de la fonction au point $x_0 + \epsilon$:

$$f(x_0 + \epsilon) = f(x_0) + A_1\epsilon + A_2\epsilon^2 + ...$$

Le développement de Taylor nous fournit en plus un moyen de trouver les coefficients A_i : ce sont, si elles existent, les i-ème dérivées de la fonction f au point x_0, divisé par $i!$. De plus, si ϵ est petit et que nous ne sommes pas très exigent quant à la précision, nous pouvons nous contenter du premier ou des deux premiers termes du développement. Le calcul des perturbations généralise cette démarche au calcul des solutions des équations différentielles, des racines des polynômes, des équations intégrales, des vecteurs propres des matrices, ... C'est le premier outil utilisé par le physicien qui tombe sur un problème ardu dont la solution n'est pas connue : si on connaît un problème proche dont la solution est connue, on peut tenter les perturbations. Mentionnons tout de suite que cette technique ne marche pas toujours. On tombe parfois (souvent), sur des perturbations dites singulières et il faut alors sortir l'artillerie lourde. Les perturbations qui se traitent facilement sont dites régulières. Nous nous intéresserons surtout aux perturbations régulières, mais dirons quelques mots sur les perturbations singulières.

9.1 Les perturbations régulières.

La meilleure façon de s'habituer aux calcul des perturbations est à travers des exemples.

Les racines d'un polynômes. Supposons que nous ne connaissons pas la résolutions des équations algébriques de second ordre, mais que nous savons résoudre l'équation $x^2 - x = 0$, dont les racines sont $x_0 = 0, 1$. Nous cherchons la solution de l'équation

$$x^2 - x + \epsilon = 0 \tag{9.1}$$

où nous supposons ϵ petit. Cherchons la solution sous forme de

$$X = x_0 + \epsilon x_1 + \epsilon^2 x_2 + ... \tag{9.2}$$

Nous cherchons la solution sous cette forme puisque nous pensons que comme ϵ est petit, la nouvelle racine ne doit pas être trop loin de l'ancienne, et l'écart doit être justement fonction de ϵ : pour $\epsilon = 0$, nous devons trouver la solution originale. Nous connaissons déjà x_0, et il nous faut trouver une méthode pour calculer $x_1, x_2, ...$ Injectons maintenant (9.2) dans (9.1) et regroupons les en fonction des puissances d'ϵ :

$$(x_0^2 - x_0) + [(2x_0 - 1)x_1 + 1]\,\epsilon + (x_1^2 + 2x_0x_2 - x_2)\epsilon^2 + ... = 0 \tag{9.3}$$

Le membre de droite de l'équation ci-dessus est un polynôme en ϵ et il est uniformément nul. Nous en déduisons donc que tous les coefficients du polynôme doivent être nuls, c'est à dire :

$$\begin{aligned} x_0^2 - x_0 &= 0 && (9.4)\\ (2x_0 - 1)x_1 + 1 &= 0 && (9.5)\\ x_1^2 + 2x_0x_2 - x_2 &= 0 && (9.6)\\ ... &= 0 \end{aligned}$$

L'équation (9.4), donnée par le coefficient de ϵ^0 et appelé le terme d'ordre zéro, est notre équation originale non perturbée que nous savons résoudre. l'équation (9.5) nous donne x_1 :

$$x_1 = 1/(1 - 2x_0)$$

Comme nous connaissons déjà x_0, nous déterminons facilement que $x_1 = 1$ ou -1. L'équation (9.6) nous détermine le coefficient x_2 :

$$x_2 = \frac{x_1^2}{1 - 2x_0}$$

et donc $x_2 = 1$ ou -1. Nous pouvons continuer ainsi (cela dépend de notre patience) et trouver les coefficient x_n. Ce que nous devons remarquer est que : (i) pour déterminer x_k , nous n'avons que besoin des $x_{k-1}, x_{k-2}, ...$ (ii) l'équation qui détermine x_k est *linéaire* en x_k, c'est à dire ne comporte que des puissances unité de x_k. C'est deux points nous permettent assez aisément de calculer la solution aussi précisément que l'on souhaite. Nous avons donc, pour les deux racines de l'équation (9.1),

$$\begin{aligned} X_1 &= 0 + \epsilon + \epsilon^2 + ... && (9.7)\\ X_2 &= 1 - \epsilon - \epsilon^2 + ... && (9.8) \end{aligned}$$

Dans ce cas précis, nous connaissons la solution exacte

$$X = \frac{1 \pm \sqrt{1 - 4\epsilon}}{2}$$

Un développement de Taylor de cette dernière nous rassure sur l'exactitude des résultats (9.7,9.8).

Généralisation. Nous pouvons généraliser l'approche ci-dessus, sans la formaliser plus : on cherche la solution d'un problème avec perturbation comme un développement en puissance de la perturbation. Il faut alors que les coefficients de chaque puissance de la perturbation soit nulle. Le mieux reste encore d'illustrer cela à travers des exemples.

Recherche des valeurs propres d'une matrice symétrique. Supposons que nous connaissons une valeur et un vecteur propre d'une matrices symétrique, c'est à dire que nous connaissons un scalaire λ_0 et un vecteur ϕ_0 tel que $A\phi_0 = \lambda\phi_0$. Une matrice symétrique par définition est égale à sa transposée $A^T = A$. Nous cherchons la valeur propre proche de λ_0 de la matrice $A + \epsilon B$. Appelons cette valeur propre μ. On cherche donc à résoudre

$$(A + \epsilon B)\psi = \mu\psi \tag{9.9}$$

Procédons comme nous l'avons mentionné plus haut. Nous chercherons la solution sous la forme

$$\mu = \lambda_0 + \epsilon\lambda_1 + ... \tag{9.10}$$
$$\psi = \phi_0 + \epsilon\phi_1 + ... \tag{9.11}$$

et nous cherchons à déterminer $\lambda_1, ...$ et $\phi_1, ...$ En injectant (9.10-9.11) dans (9.9), nous avons :

$$(A + \epsilon B)(\phi_0 + \epsilon\phi_1 + ...) = (\lambda_0 + \epsilon\lambda_1 + ...)(\phi_0 + \epsilon\phi_1 + ...)$$

et en regroupant les termes en puissance de ϵ, nous trouvons les équations suivantes :

$$A\phi_0 = \lambda_0\phi_0 \tag{9.12}$$
$$A\phi_1 + B\phi_0 = \lambda_0\phi_1 + \lambda_1\phi_0 \tag{9.13}$$
$$... = ...$$

La première équation, c'est à dire les terme d'ordre 0 en ϵ, ne nous rapporte bien sûr rien que ne l'on connaisse déjà. Dans l'équation (9.13), nous avons deux inconnus, le vecteur ϕ_1 et le scalaire λ_1 à déterminer. Prenons le produit scalaire des deux côtés par le vecteur ϕ_0 :

$$(\phi_0, A\phi_1) + (\phi_0, B\phi_0) = \lambda_0(\phi_0, \phi_1) + \lambda_1(\phi_0, \phi_0) \tag{9.14}$$

Nous avons supposé que A est symétrique. Donc,

$$(\phi_0, A\phi_1) = (A^T\phi_0, \phi_1) = (A\phi_0, \phi_1) = \lambda_0(\phi_0, \phi_1)$$

et en injectant ce résultat dans (9.14), nous aboutissons finalement à

$$\lambda_1 = (\phi_0, B\phi_0)/(\phi_0, \phi_0)$$

Nous connaissons donc la correction d'ordre 1 à la valeur propre.

Si tout cela paraît un peu abstrait, cherchons la valeur propre de la matrice

$$\begin{pmatrix} 1+\epsilon & 2\epsilon \\ 2\epsilon & 2+3\epsilon \end{pmatrix}$$

Cette matrice est la somme d'une matrice diagonale $A = \text{diag}(1,2)$ et de la matrice B tel que $B_{ij} = (i+j-1)\epsilon$. La matrice B est la perturbation si on suppose $\epsilon \ll 1$. La matrice A possède bien sur les deux valeurs propres 1 et 2 et les vecteurs propres associés $(1,0)$ et $(0,1)$. Cherchons la valeurs propre proche de 1 de la matrice perturbée. D'après ce que nous avons dit, nous devons calculer $(1,0)B(1,0)^T$:

$$(1,0).\begin{pmatrix} \epsilon & 2\epsilon \\ 2\epsilon & 3\epsilon \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \epsilon$$

Comme $(1,0)(1,0)^T = 1$, la valeur propre proche de 1, à l'ordre 1 en ϵ, s'écrit

$$\mu = 1 + \epsilon$$

Le lecteur peut chercher directement la valeur propre de la matrice perturbée et vérifier le résultat ci-dessus. Bien sûr, quand la matrice est plus grande que 2×2, la recherche directe peut être extrêmement fastidieuse.

La stabilité d'un système différentiel. Nous connaissons parfois un point fixe d'une équation, et nous souhaitons savoir si ce point est stable ou non. Par exemple, le point le plus bas pour un pendule est un point fixe stable, tandis que le point le plus haut est un point instable. Intuitivement, pour connaître la stabilité, nous écartons un peu le système de son point fixe. Si le système revient à son point de départ ou reste proche, l'équilibre est stable. Si au contraire, la perturbation tende à grandir, le point est alors instable.

Considérons l'équation du mouvement d'un pendule amorti

$$\ddot{\theta} + \rho\dot{\theta} + \omega^2 \sin\theta = 0 \tag{9.15}$$

où θ est l'angle avec l'axe vertical et $\theta = 0$ correspond au point le plus bas. Il est évident que les points $\theta = 0$ et $\theta = \pi$ sont les points fixes de cet équation. Supposons maintenant que nous partons très proche du point $\theta = 0$, c'est à dire avec la condition initiale $\theta(t=0) = \epsilon$. Cherchons comme d'habitude la solution sous forme de $\theta = 0 + \epsilon\theta_1 + \ldots$ En réalité, comme nous nous intéressons qu'à la stabilité, le terme d'ordre 1 en ϵ nous suffit. En injectant cette solution dans (9.15) et en développant la fonction sin, nous trouvons :

$$\ddot{\theta}_1 + \rho\dot{\theta}_1 + \omega^2\theta_1 = 0$$

la solution générale de cette dernière est de la forme $\theta_1 = \exp(-\rho t)\exp(\pm i\omega' t)$ et elle tend effectivement vers 0 quand $t \to \infty$. Le point $\theta = 0$ est donc un point stable. En répétant les même opérations pour le point $\theta = \pi$, et en n'oubliant pas que $\sin(\pi + x) = -\sin x$, nous aboutissons à

$$\ddot{\theta}_1 + \rho \dot{\theta}_1 - \omega^2 \theta_1 = 0$$

et cette fois, il est claire que quand $t \to \infty$, $\theta_1 \to \infty$. Le point $\theta = \pi$ est donc instable.

9.2 Les perturbations singulières.

Nous avons supposé, lors de l'étude des perturbations singulières, que la solution perturbée est proche de celle non perturbée. Cela n'est pas toujours le cas et l'ajout d'un petit terme peut radicalement changer la solution. Prenons le cas de l'équation algébrique

$$\epsilon x^2 + x - 1 = 0 \tag{9.16}$$

La solution non perturbée, *i.e.* pour $\epsilon = 0$ vaut $x = 1$. La solution exacte pour $\epsilon \neq 0$ s'écrit

$$x = \frac{-1 \pm \sqrt{1 + 4\epsilon}}{2\epsilon}$$

et un petit développement nous montre que les racines sont, pour $\epsilon \ll 1$, de la forme

$$\begin{aligned} x_1 &= 1 - \epsilon \\ x_2 &= -\frac{1}{\epsilon} \end{aligned}$$

Nous avons donc l'apparition d'une nouvelle racine qui est d'autant plus grande que la perturbation est petite. Cela est un phénomène générale : à chaque fois que la perturbation est sur un terme d'ordre supérieur, la perturbation est singulière. Il existe parfois des changements de variable qui rendent la perturbation régulière. Par exemple, dans l'équation (9.16), en posant $x = 1/y$, nous avons

$$y^2 - y - \epsilon = 0$$

qui peut se traiter par la méthode habituelle. D'après notre traitement de (9.1), ses solutions sont

$$\begin{aligned} y_1 &= -\epsilon \\ y_2 &= 1 + \epsilon \end{aligned}$$

qui nous redonne bien les racines en $x = -1/\epsilon$ et $x = 1 - \epsilon$.

La même remarque s'applique aux équations différentielles. L'équation

$$\epsilon \ddot{x} + \dot{x} + 1 = 0$$

est celle d'un oscillateur harmonique amortie. Si la masse ϵ est nulle, la solution est de la forme $x = A\exp(-t)$. Si la masse est non nulle, la solution, à l'ordre le plus important en ϵ, est de la forme $A\exp(-t) + B\exp(-t/\epsilon)$ et le lecteur peut vérifier que les deux solutions sont radicalement différentes. Remarquons à nouveau que nous pouvons chercher un changement de variable de la forme $t = \epsilon^p t'$ et $x = \epsilon^q y$ qui rendrait la perturbation régulière. Nous en laissons le soin au lecteur intéressé.

L'ennui avec les équations différentielles est que les termes les plus inoffensifs peuvent rendre les perturbations singulières. Considérons l'exemple de l'oscillateur suivant :

$$\ddot{x} + \omega^2 x + \epsilon x^3 = 0 \tag{9.17}$$

ceci est l'équation d'un mobile dans un potentiel en $kx^2 + k'x^4$. La solution général de l'équation non perturbée est $a\cos(\omega t + \phi)$. Sans perte de généralité, on supposera $\phi = 0$. Nous cherchons alors la solution de l'équation perturbée sous forme de $x(t) = a\cos(\omega t) + \epsilon x_1(t) + \ldots$ En injectant dans l'équation et en collectant les termes d'ordre 1 en ϵ , nous trouvons que

$$\ddot{x}_1 + \omega^2 x_1 = -a^3\cos^3(\omega t) \tag{9.18}$$

Or, $\cos^3(u) = (1/4)\cos(3u) + (3/4)\cos(u)$ et donc la solution de (9.20) est donnée par la somme de la solution des deux équations suivante :

$$\ddot{x}_1 + \omega^2 x_1 = (-a^3/4)\cos(3\omega t) \tag{9.19}$$
$$\ddot{x}_1 + \omega^2 x_1 = (-3a^3/4)\cos(\omega t) \tag{9.20}$$

La première équation ne présente pas de danger : c'est l'équation d'un oscillateur harmonique de fréquence propre ω forcée à la fréquence 3ω et possède une solution du genre $\cos(3\omega t + \alpha)$.

Par contre, la deuxième équation (9.20) est celle d'un oscillateur harmonique de fréquence propre ω forcée justement à ω. Il y a donc résonance et la solution, qui est de la forme $t\cos(\omega t + \beta)$, va devenir très large. Cela viole notre hypothèse de départ sur les développements en puissance d'ϵ. Nous avions écrit

$$x(t) = a\cos(\omega t) + \epsilon x_1(t) + \ldots$$

en supposant $x_1(t)$ bornée, de façon à ce que $\epsilon x_1(t)$ reste toujours petit par rapport à la solution non perturbée $a\cos(\omega t)$. Or, nous voyons qu'au bout d'un temps $t > 1/\epsilon$, le soit disant petit terme devient en faite le terme dominant.

Il existe de nombreux types différents de perturbations singulières et au moins autant de façons de les traiter. L'objet de ce livre n'étant pas un cours détaillé sur le calculs de perturbations, nous nous en tiendrons presque là. Nous trouvons cependant utile de montrer comment traiter les cas similaires à l'exemple ci-dessus. La technique est a appelée "élimination des termes séculaires ou résonnants".

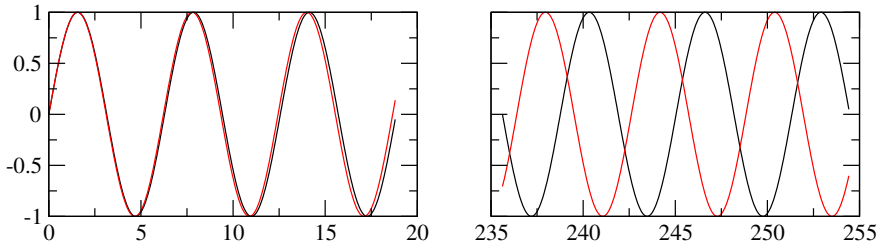


FIGURE 9.1 – $\sin(t)$ (en noir) et $\sin(1.01t)$ (en rouge) pour les temps courts et longs.

Le traitement des termes séculaires.

La perturbation que nous avons considérée plus haut intervient dans pratiquement tous les problèmes d'oscillations, et c'est pourquoi il nous semble important de la traiter. En regardant de plus près l'équation (9.17), nous voyons qu'il n'y a rien d'anormal ou de divergent. Elle décrit simplement des oscillations au fond d'un puits peut être un peu plus raide qu'un puits harmonique et il n'y a aucune raison que quelque chose diverge. L'erreur ne peut venir que de notre traitement. En supposant que la solution s'écrit sous la forme $a\cos(\omega t) + \epsilon x_1(t)$, nous avons fait l'erreur de penser que le terme d'ordre 0 continue à présenter une oscillation à fréquence ω. Il n'y a aucune raison pour cela, et la fréquence peut également être de la forme $\omega + \epsilon\omega_1 + ...$ La figure (9.1) montre la différence entre $\sin(t)$ et $\sin(1.01t)$ pour les temps inférieurs à vingt et pour les temps autour de 250. Nous voyons que la différence est à peine perceptible pour les temps courts, tandis que les deux fonctions n'ont plus rien à voir aux temps longs. Ce problème avait été observé d'abord en astronomie, où les calculs perturbatifs des siècles précédents commençaient à s'écarter des observations (d'où le terme séculaire). Lindstedt (vers 1880) a remédié à ces carences par sa technique de renormalisation qui est de chercher la solution sous la forme

$$x(t) = a\cos[(\omega + \epsilon\omega_1)t] + \epsilon x_1(t) \tag{9.21}$$

En injectant la forme (9.21) dans (9.17) et en collectant les termes d'ordre 1 en ϵ, nous trouvons cette fois

$$\ddot{x}_1 + \omega^2 x_1 - 2a\omega\omega_1 \cos[(\omega+\epsilon\omega_1)t] = (-a^3/4)\cos[3(\omega+\epsilon\omega_1)t] - (3a^3/4)\cos[(\omega+\epsilon\omega_1)t]$$

Il nous suffit maintenant de choisir

$$\omega_1 = \frac{3a^2}{8\omega}$$

pour éliminer le terme résonnant. La solution perturbative à l'ordre 1 s'écrit alors :

$$x(t) = a\cos[(\omega + \epsilon\frac{3a^2}{8\omega})t] + \epsilon A\cos[3(\omega + \epsilon\frac{3a^2}{8\omega})t + \alpha]$$

où les coefficients a, A, α sont déterminés à partir des conditions initiales.

Exercices.

Racine des polynômes.

§ 9.1 Soit $P(x) = \sum a_n x^n$ un polynôme dont une des racines, x_0 est connue, c'est à dire $P(x_0) = 0$. Soit le polynôme $P'(x) = P(x) + \epsilon x^p$. Soit x' la racine proche de x_0 de ce dernier. Montrer qu'à l'ordre 1,

$$x' = x_0 - \frac{x_0^p}{\sum n a_n x_0^{n-1}} \epsilon$$

§ 9.2 Pouvez vous exhiber la correction à l'ordre 2 ?

§ 9.3 Soit l'équation

$$x^6 - 4x^5 + 3x^4 - 6x^2 + 6x + 1 = 0 \tag{9.22}$$

Nous remarquons qu'en écrivant par exemple le coefficient de x^4 comme $2 + \epsilon$ (où $\epsilon = 1$), la somme des coefficients de l'équation non perturbée vaut 0. $x = 1$ est donc une solution de l'équation non perturbée (*i.e.* pour $\epsilon = 0$). Calculez la correction à cette racine à l'ordre 1 en ϵ et comparez à la solution exacte $x = 1.10565$. Et si au lieu du coefficient du x^4, nous avions choisit un autre terme, qu'aurait on obtenu ?

§ 9.4 Calculer la correction à l'ordre 2. Pouvez vous alors trouver un critère pour le choix du coefficient pour que la correction à l'ordre 1 soit la meilleure ?

Équation transcendante. Une équation transcendante est une équation qui fait intervenir des fonctions non-algébriques, comme par exemple $x \sin x = \cos x$. Il est fréquent de rencontrer ces équations en physique.

§ 9.5 $x \log x = 0$ admet une solution pour $x = 1$. Calculez, à l'ordre 3, la racine proche de 1 de l'équation $x \log x = \epsilon$. Comparez à la solution exacte 1.09557 pour $\epsilon = 0.1$. Pouvez-vous utiliser la même approche pour trouver la solution proche de 0 ?

§ 9.6 Trouver la racine, proche de π, de $x \sin x = \epsilon$. Comparez, pour $\epsilon = 0.1$, à la solution exacte $x = 3.10943$.

Équation intégrale. Nous avons rencontré les équations intégrales lors de notre discussion des fonctions de Green. Nous allons étudier ci-dessous un schéma itératif de leurs résolution. Cependant, ces schémas sont en général extrêmement fragiles, et il faut toujours s'assurer de leur convergence.

§ 9.7 Une équation intégrale de Fredholm de deuxième espèce est de la forme

$$f(x) = g(x) + \mu \int_a^b K(x, x') f(x') dx'$$

Proposez un schéma de résolution par le calcul des perturbations. Ne vous contentez pas de l'ordre 1. Trouvez la perturbation d'ordre n en général.

§ 9.8 Profitez pour exhiber la solution exacte de

$$f(x) = 1 + \lambda \int_0^\infty e^{-(x+y)} f(y) dy$$

En calculant les perturbations à tout ordre et en sommant la série ainsi obtenue.

§ 9.9 Faites la même chose pour

$$f(x) = 1 + \lambda \int_0^x f(y) dy$$

Attention, les bornes de l'intégrale dépendent de x. On appelle ces équations Voltera de deuxième espèce.

Stabilité linéaire : Croissance des bactéries. L'équation de croissance de bactéries, connue comme l'équation logistique, est la suivante :

$$\frac{dc}{dt} = ac - bc^2 \tag{9.23}$$

où $c(t)$ est la concentration de bactérie au temps t, a le taux de croissance et b un coefficient qu'on appel de saturation. a et b sont des constantes > 0.

§ 9.10 Trouver les deux solutions stationnaires de l'équation (9.23).

§ 9.11 En utilisant le calcul des perturbations autour de ces deux solutions, démontrer qu'une est instable, tandis que l'autre est stable.

§ 9.12 Sans faire aucun calcul et en vous servant des résultats ci-dessus : tracer l'allure générale de la solution si on utilise la condition initial $c(t = 0) = \epsilon \ll a/b$. Expliquer votre interprétation.

Stabilité linéaire : traitement général. Soit l'équation différentielle

$$\frac{dy}{dt} = -\alpha y + f(y) \tag{9.24}$$

où f est une fonction quelconque. Nous supposons que cette équation possède un point stationnaire y_s tel que

$$\alpha y_s = f(y_s)$$

§ 9.13 En considérant les perturbations à l'ordre 1, discuter la stabilité de ce point en fonction de la valeur de $f'(y_s)$.

§ 9.14 Soit les fonction $f(y)$ et αy telles que montrées dans la figure (9.2). Comme nous pouvons le voir, l'équation différentielle (9.24) possède dans ce cas trois solutions stationnaires. D'après votre analyse précédente, lesquelles sont stables et lesquels instables ?

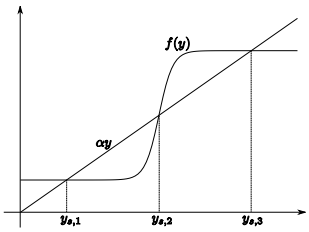


FIGURE 9.2 – Exemple de la fonction f, utilisée pour réaliser des mémoires biochimiques.

Oscillateur de Van der Pol. Van der Pol a proposé l'équation suivante dans les années 1920 pour modéliser les oscillateurs auto-entretenus comme le battement de cœur

$$\ddot{x} + \epsilon(x^2 - 1)\dot{x} + \omega_0 x = 0$$

Le coefficient du terme $\dot{x}$ est équivalent à un frottement. Nous voyons qu'il est *négatif* si l'amplitude x est petite (< 1), c'est à dire que le système reçoit de l'énergie de l'extérieur, ce qui va l'amener à augmenter son amplitude. Par contre si l'amplitude devient trop grande (> 1) le frottement devient positif et le système perd de l'énergie vers l'extérieur, ce qui va diminuer son amplitude. Nous voyons que le système maintient une oscillation stable quelque soit les conditions de départ.

§ 9.15 Montrer que le point fixe $x = 0$ est instable.

§ 9.16 En partant de la solution non perturbée $x = a\cos\omega_0 t$, montrez que les perturbations régulières génèrent des termes résonnants.

§ 9.17 Utilisez la renormalisation de Lindstedt pour éliminer les termes résonnants. Pour cela, chercher la solution sous forme de

$$x(t) = a\cos\Omega t + \epsilon x_1(t)$$

où $\Omega = \omega_0 + \epsilon\omega_1$. Vous pouvez apercevoir que l'élimination des termes résonnants impose une condition sur l'*amplitude* de l'oscillation, ce que l'on appelle un cycle limite.

Écosystème de prédateurs–proies. Une des interactions fondamentales en écologie est celle des prédateurs et des proies. Le premier modèle pour la dynamique de ces deux populations a été proposé par Lotka et Voltera au début des années 1930. Soit P le nombre des prédateurs et N le nombre des proies dans l'écosystème. Lotka et Voltera ont proposé

$$dN/dt = \alpha N - \beta NP \tag{9.25}$$
$$dP/dt = \gamma NP - \delta P \tag{9.26}$$

α est le taux de croissance naturel des proies en l'absence des prédateurs. La présence des prédateurs cause également la disparition des proies, proportionnellement au nombre de prédateur et de proie, d'où le terme en $-\beta NP$ dans la première équation, β étant l'efficacité de la chasse. Dans l'équation qui régit la dynamique des prédateurs, nous voyons que la croissance est fonction du nombre de proie disponible, et le terme δ est le taux de mort naturel des prédateurs.

§ 9.18 Montrez que ce système possède un point fixe, c'est à dire des valeurs N_0, P_0 pour lesquels $dN/dt = dP/dt = 0$.

§ 9.19 Étudiez la solution de ce système pour les faibles écarts au point fixe. Cela veut dire que nous prenons des conditions initiales du genre $N(t=0) = N_0 + \epsilon$ et $P(0) = P_0$. Cherchez la solution sous la forme $N(t) = N_0 + \epsilon N_1(t)$ et $P(t) = P_0 + \epsilon P_1(t)$, et en collectant les termes d'ordre 1 en ϵ, obtenez un système linéaire pour N_1 et P_1. Résolvez ce système et déduisez également la forme du cycle limite, c'est à dire N_1 en fonction de P_1.

§ 9.20 Poussez les calculs à l'ordre 2 en ϵ et étudiez l'apparition d'harmonique supérieurs.

§ 9.21 Vous pouvez également remarquer que le cycle limite peut s'obtenir en divisant directement (9.25) par (9.26) et en résolvant l'équation différentielle du premier ordre. Comparez le résultat de ce calcul au résultat de la question §9.19.

Équation de Riccati. Nous souhaitons résoudre, par le calcul des perturbation au premier ordre de ϵ, l'équation différentielle de Riccati,

$$\frac{dv}{dt} + v^2 + \alpha^2 + \epsilon b(t) = 0 \tag{9.27}$$

Cette équation se rencontre quand nous essayons de résoudre les équations différentielles linéaire de second ordre à coefficients variables comme par exemple l'équation de Mathieu

$$\frac{d^2u}{dt^2} + \left(\omega^2 + \epsilon\, b(t)\right) u = 0 \tag{9.28}$$

que nous avons déjà étudié en cours pour le cas particulier de l'équation de l'oscillateur paramétrique. On transforme en général l'équation (9.28) en l'équation (9.27) en effectuant le changement de fonction $u(t) = \exp\left(V(t)\right)$ où $V'(t) = v(t)$.

§ 9.22 Vérifier que la solution de l'équation différentielle

$$\frac{du}{dt} + a(t)u = b(t) \;\;;\;\; u(0) = u_0$$

est

$$u(t) = e^{-A(t)}\left(u_0 + \int_0^t e^{A(\tau)} b(\tau) d\tau\right) \tag{9.29}$$

où $A'(t) = a(t)$ et $A(0) = 0$.

§ 9.23 Vérifier que la l'équation différentielle

$$\frac{dv}{dt} + v^2 + \alpha^2 = 0 \ ; \ v(0) = 0 \tag{9.30}$$

possède la solution

$$v(t) = -\alpha \tan(\alpha t)$$

§ 9.24 Soit l'équation

$$\frac{dv}{dt} + v^2 + \alpha^2 + \epsilon b(t) = 0$$

En écrivant la solution à l'ordre 1 sous la forme de

$$v(t) = v_0(t) + \epsilon v_1(t)$$

où v_0 est la solution de l'équation (9.30), obtenir l'équation différentielle qui gouverne $v_1(t)$. Résoudre l'équation sur v_1, avec la condition initiale $v_1(0) = 0$, en utilisant le résultat (9.29).

§ 9.25 Trouver la solution explicite à l'ordre 1 dans le cas $b(t) = 2\alpha \tan(\alpha t)$.

§ 9.26 Quelle est la limitation du calcul et la limite de validité du calcul à l'ordre 1 dans le cas général ?

Stabilité d'interface. Soit une interface $u(x,t)$ (par exemple entre solide et liquide lors de la coulée continue en métallurgie) décrite par l'équation

$$\frac{\partial u}{\partial t} = -au - bu^3 + c\frac{\partial^2 u}{\partial x^2} - d\frac{\partial^4 u}{\partial x^4}$$

où nous supposons les coefficients $a, b, d > 0$.

§ 9.27 Discuter la stabilité linéaire de la solution $u(x,t) = 0$ selon que c est positif ou négatif et chercher les seuils d'instabilité.

10 Les opérateurs linéaires.

10.1 Introduction

Une des tâches que l'on rencontre régulièrement en physique est de résoudre des équations différentielles linéaires. De façon générale, nous pouvons représenter ces équations par $\mathcal{L}y = f$, où y est la fonction inconnue à rechercher, f une fonction connu, et $\mathcal{L}$ un opérateur différentiel. Par exemple, l'équation de la chaleur peut s'écrire $\mathcal{L}u = q(x,t)$, où $\mathcal{L} = \left(\partial_t - D\partial_x^2\right)$ et $q(x,t)$ est le terme de source. A priori, la recherche des solutions est du domaine de l'analyse. Nous allons voir cependant que nous pouvons ramener la résolution de ces équations dans le domaine de l'algèbre matricielle des systèmes à n équations et n inconnus [1] du genre $AX = B$. Le très grand avantage est que pour faire de l'algèbre, nous n'avons, en gros, que besoin d'addition et de multiplication [2]. Les transformées de Fourier et de Laplace que nous avons rencontrés dans ce cours étaient des exemples particuliers d'outils bien adaptés à une certaine classe d'équations qui nous permettaient de ramener l'analyse à l'algèbre. Nous allons généraliser cette approche et voir toute la puissance de feu que cela nous procure.

Depuis le début de ce cours, nous insistons fortement sur le concept de vecteur. Au premier chapitre, nous avons vu que dans un espace vectoriel, nous pouvons définir des bases : cela nous permet de manipuler les vecteurs à l'aide de colonnes (ou de lignes) de nombres. Nous avons également vu que si nous disposons d'un produit scalaire, cela facilite grandement la tache de trouver les coefficient d'un vecteur dans une base orthogonale.

Un vecteur peut être un objet aussi simple qu'un vecteur usuel du plan euclidien, ou un objet beaucoup plus complexe tel qu'une fonction. Prenons le cas d'une fonctions f . Une fonction est une *machine* qui prend un nombre en entrée et produit un nombre en sortie. Nous pouvons choisir plusieurs représentations pour une même fonction. Par exemple, si nous choisissons la base de Fourier avec les vecteurs de la base $\exp(iqx)$, f s'écrira comme une superposition de ces fonctions, chacun avec un poids $\tilde{f}(q)$ (nous avons absorbé ici le facteur $1/2\pi$ dans la définition de $\tilde{f}$) :

$$f(x) = \int_{-\infty}^{+\infty} \tilde{f}(q) \exp(iqx) dq$$

L'intégral ici n'effectue rien d'autre que la superposition des vecteurs de base avec leurs

1. n étant infini en l'occurrence, mais ceci n'induit pas de difficultés particulières
2. Voir également le chapitre 21 sur la signification d'une équation différentielle

poids correspondants. Il est également usuel de représenter la fonction f par un tableau qui à chaque entrée numérique, associe un nombre, et on note cela $f(x)$ (la notation est un peu confuse). En réalité, cela revient à représenter une fonction f sur la base des δ de Dirac, où chaque valeur $f(x)$ est le poids associé à un Dirac centré sur x :

$$f(x) = \int_{-\infty}^{+\infty} f(y)\delta(x-y)dy$$

A nouveau, l'intégral ne fait rien d'autre que de superposer des vecteurs de la base.

Revenons maintenant au concept général de vecteur. Soit l'espace vectoriel $\mathcal{E}$. Nous pouvons définir des opérations qui transforment un vecteur en un autre, et cela de façon linéaire. Dans l'espace des vecteurs du plan euclidien, la rotation ou la projection sur un axe sont de telles opérations. Par exemple, la rotation de la somme de deux vecteurs égale la somme de leurs rotations : $R(e_1 + e_2) = Re_1 + Re_2$. Dans l'espace des fonctions infiniment dérivables, l'opération dérivation D est une opération linéaire : elle transforme un vecteur (une fonction) un un autre, et cela de façon linéaire. De même, l'opération intégration $I[f] = \int_0^x f(y)dy$. De façon général, nous appelons, dans l'espace des fonctions, un opérateur linéaire comme une machine qui prend en entrée une fonction et produit en sortie une autre fonction, et fait cela de façon linéaire :

$$O[\lambda f_1 + \mu f_2] = \lambda O[f_1] + \mu O[f_2]$$

où λ et μ sont des scalaires et f_1 et f_2 des fonctions.

Exercices.

§ 10.1 Soit l'opérateur X qui prend une fonction f en entrée et produit la fonction multipliée par x en sortie : $X[f(x)] = xf(x)$ [exemple : $X[\sin(x)] = x\sin(x)$]. Démontrer que c'est un opérateur linéaire.

§ 10.2 Qu'en est il de l'opérateur X^2 : $X^2[f(x)] = x^2 f(x)$? Soit $V(X)$ un opérateur tel que $V(X)[f(x)] = V(x)f(x)$. Est ce que ce dernier est linéaire ?

§ 10.3 Même question pour l'opérateur 1 : $1[f(x)] = f(x)$. L'opérateur 0 est l'opérateur qui associe à n'importe qu'elle fonction la fonction 0. Est-il linéaire ? De façon général, l'opérateur λ associe à une fonction f la fonction λf (il faut avouer que la notation est vraiment confuse entre l'opérateur λ et le scalaire λ ; on s'y habitue vite cependant).

§ 10.4 Soit l'opérateur de translation $T_\epsilon[f(x)] = f(x+\epsilon)$; démontrer qu'il est linéaire.

Note sur les notations. Pour manipuler les opérateurs linéaires, la coutume est de laisser tomber les signes du genre $()$ et $[]$. Ainsi, nous écrivons Of où même pire, $Of(x)$ à la place de $O[f(x)]$. La confusion est gênante quand on écrit par exemple, $Xf(x) =$

$xf(x)$. X ici est un opérateur, $f(x)$ et $xf(x)$ sont des fonctions [3] ; $Xf(x)$ est la fonction qui résulte de l'application de X à f. Pour éviter un peu ces confusions, la convention que nous suivrons dans ce cours et de toujours noter les opérateurs par des lettres majuscules.

10.2 L'algèbre des opérateurs.

Se donner une algèbre est se donner un ensemble $\mathcal{E}$ ou l'on définit les deux opérations $+$ et . (produit) entre ses membres, avec toutes les propriétés d'associativités usuelles que vous connaissez : soit $a, b, c \in \mathcal{E}$, alors

$$\begin{aligned} a+b\ ,\ a.b &\in \mathcal{E} \\ a+b &= b+a \\ a(b+c) &= ab+ac\,;\ (a+b)c = ac+bc \\ (ab)c &= a(bc) \end{aligned}$$

Nous devons avoir quelques propriétés de plus pour mériter le nom d'algèbre. Il faut qu'il existe des éléments neutre vis à vis des deux opérations, qu'on appelle 0 et 1 : $a+0=0$ et $a1=1a=a$. De plus, les inverses des éléments vis à vis de $+$ et de . doivent exister : pour chaque élément a, il doit exister un élément unique, qu'on appelle $-a$, tel que $a+(-a)=0$; de même, il doit exister un élément unique qu'on note $1/a$ ou a^{-1} tel que $a.(1/a)=(1/a)a=1$ (pour $a \neq 0$).

L'ensemble des nombres (rationnels ou réel), équipé de + et de . usuel, constitue une algèbre. C'est un cas un peu particulier, puisqu'en plus, la multiplication y est commutative ($ab=ba$). Mais tous les théorèmes qui ont été démontrés pour l'algèbre des nombres sans invoquer la commutativité du produit sont valable pour n'importe quelle autre algèbre.

Nous pouvons définir une algèbre pour les opérateurs linéaires. Nous devons d'abord préciser le sens de l'égalité entre opérateurs. Nous dirons que $O_1 = O_2$ si le résultat de l'application de ces deux opérateurs à une fonction est le même, quelque soit la fonction [4] : $\forall f, O_1 f = O_2 f$.

3. A vrai dire, c'est encore pire : f est une fonction, $f(x)$ est un nombre. A nouveau, on a l'habitude de ne pas toujours distinguer explicitement les deux choses et laisser le boulot au cerveau. Dans ce cas, le cerveau agit comme un vulgaire compilateur C, testant constamment le *type* des notations qu'on utilise.

4. Il est évident que $\forall f$ est une condition trop exigeante et n'a pas de sens en général. Si par exemple, l'opérateur contient des dérivées d'ordre n, nous devons comprendre $\forall f$ comme : quelque soit f "$n-$fois dérivable". A chaque fois, nous supposerons l'ensemble des fonctions comme compatible avec la définition de l'opérateur. Nous n'entrerons pas plus dans le détail lors de ce cours, pour ne pas alourdir chaque assertion par un train de précautions et de conditions d'applicabilité. Dans la majorité des cas, nous supposons que nous travaillons avec l'ensemble des fonctions $\mathcal{L}^2[-\infty,+\infty]$ au moins deux fois continuement dérivable. Ceci surtout impose à nos fonctions qu'elles et leurs dérivées $\to 0$ quand leur argument $\to \infty$.

L'opération $+$ entre opérateurs hérite directement sa définition de l'opération $+$ entre les fonctions [5] : L'opérateur $O_1 + O_2$ est l'opérateur qui, appliqué à une fonction, produit la fonction $O_1 f + O_2 f$ (bien noter que là, l'addition est entre deux fonctions). L'opération . est la combinaison d'opérateur : $O_1 O_2$ est l'opérateur qui à la fonction f, associe la fonction $O_1[O_2[f]]$. On peut affirmer maintenant que l'ensemble des opérateurs linéaires muni de $+$ et de . constitue une algèbre. Les opérateurs 0 et 1 sont les éléments neutres de l'addition et de la multiplication. Dans la suite de ce cours, l'ensemble des fonctions auxquels ces opérateurs s'appliquent est l'ensemble des fonctions $\mathcal{L}_2$ au moins deux fois dérivable.

Toutes les notations que nous utilisons dans l'algèbre classique peuvent être utilisées pour l'algèbre des opérateurs. Par exemple, X^2 est effectivement $X.X$. Si l'on désigne par ∂_x et ∂_y les opérateurs de dérivation par rapport à x et y, alors $\partial_x^2 - \partial_y^2 = (\partial_x - \partial_y)(\partial_x + \partial_y)$. Il faut juste faire attention à la commutativité : en général, deux opérateurs ne commutent pas : $O_1 O_2 \neq O_2 O_1$. Cela ne veut pas dire que l'on ne peut pas trouver deux opérateurs qui commutent, l'exemple précédent de ∂_x et ∂_y le montre. Simplement, il ne faut pas le présumer à l'avance.

Exemple 10.1 Démontrons que $DX - XD = 1$ (bien noter que ceci est une égalité entre opérateurs). Nous avons $DX[f(x)] = (d/dx)\,(xf(x)) = xf'(x) + f(x)$ et $XD[f(x)] = xf'(x)$. Donc $(DX - XD)[f(x)] = f(x)$: Le résultat d'application de $DX - XD$ à une fonction est le même que l'application de l'opérateur 1, d'où l'égalité entre les opérateurs.

Définition 1 On appelle commutateur de deux opérateurs O_1 et O_2, l'opérateur $O_1 O_2 - O_2 O_1$. Il est usuel de noter ce dernier $[O_1, O_2]$.

Le commutateur joue un très grand rôle en mécanique quantique. Depuis le livre de Paul Dirac en 1930, la mécanique quantique est formulée à travers les relations de commutations.

Fonctions d'opérateurs. Beaucoup de fonctions usuelles sont définies à l'aide de séries algébriques, comme par exemple, $\exp(x) = \sum_n x^n/n!$. Puisque l'on dispose d'une algèbre pour les opérateurs linéaires, nous pouvons faire de même et définir des fonctions d'opérateurs, comme par exemple $\exp(O)$ ou $\log(1+O)$, qui sont elles mêmes des opérateurs linéaire. Par exemple, pour l'opérateur D et le scalaire ϵ,

$$\exp(\epsilon D) = \sum_{n=0}^{\infty} (1/n!)\epsilon^n D^n$$

5. De même que dans l'espace des fonctions, l'opération + est héritée de l'addition entre les nombres : la fonction $f + g$ est la fonction qui associe au nombre x le nombre $f(x) + g(x)$.

et le résultat de son application à une fonction $f(x)$ produit la fonction

$$\sum_{n=0}^{\infty}(1/n!)\epsilon^n f^{(n)}(x) = f(x+\epsilon)$$

L'opérateur $\exp(\epsilon D)$ n'est donc rien d'autre que l'opérateur de translation T_ϵ vu plus haut. Bien sûr, dès que l'on parle de suites et de séries infinies, nous avons besoin de la notion de *convergence*. La convergence dans l'espace des opérateurs hérite sa définition de la convergence dans l'espace des fonctions : nous dirons que la suite O_n converge vers O si la suite des *fonctions* $O_n f$ converge vers la fonction Of quelque soit f (cf la note 4).

Disposer des fonctions d'opérateurs nous permet de résoudre symboliquement nombre d'équations à dérivées partielles (EDP). Prenons d'abord l'équation différentielle ordinaire

$$dy/dt = ay \tag{10.1}$$

avec la condition initiale $y(t=0) = y_0$ où a et y_0 sont des scalaires ne dépendant pas du temps. La solution bien connue est $y(t) = \exp(ta)y_0$. La valeurs de la fonction à un temps ultérieur t s'obtient en appliquant (multipliant) le scalaire $\exp(ta)$ à la condition initiale.

Soit maintenant l'EDP de premier ordre

$$\partial_t\psi - \partial_x\psi = 0 \tag{10.2}$$

où $\psi(x,t)$ est une fonction des deux variables x, t ; la condition initiale étant $\psi(x, t = 0) = \psi_0(x)$, $\psi_0(x)$ étant une fonction connue de la variable x. Nous pouvons écrire cette équation sous forme opératorielle

$$\partial\psi/\partial t = D\psi$$

où $D = \partial/\partial_x$ est un opérateur qui ne dépend pas de t. En s'inspirant de l'exemple de l'équation (10.1), (voir le problème 10.2 pour un traitement rigoureux) nous pouvons donner la solution comme

$$\psi(x,t) = \exp(t.D)\psi_0(x)$$

Nous savons par ailleurs que l'opérateur $\exp(tD)$ n'est rien d'autre que l'opérateur "translation d'une quantité t". La solution s'écrit donc

$$\psi(x,t) = \psi_0(x+t)$$

Ceci est la solution exacte de l'EDP (10.2) que nous pouvons obtenir soit par des transformées de Fourier-Laplace , soit par la méthode des caractéristiques du chapitre 15. Ceci n'est pas une analogie. Les même règles d'algèbre (et d'analyse) que nous appliquons aux fonctions ont été définies pour les opérateurs et nous donnent le droit de

les manipuler symboliquement de la même façon. Voyons cela d'un peu plus près. La fonction $f(t) = \exp(ta)$ est donnée comme la série

$$\exp(ta) = \sum_{n=0} t^n a^n / n!$$

sa dérivée (qui coïncide avec la dérivée terme à terme de la série) s'écrit (en jouant sur l'indice de sommation)

$$\begin{aligned} f'(t) &= \sum_{n=0} t^n a^{n+1} / n! \\ &= a \sum_{n=0} t^n a^n / n! \\ &= a f(t) \end{aligned}$$

et c'est pour cela que cette fonction est la solution générale de (10.1). En utilisant les mêmes règles de manipulation pour l'algèbre d'opérateurs, nous voyons que l'opérateur $\exp(tD)$ possède comme dérivée par rapport au temps l'opérateur $D\exp(tD)$. La *fonction* $\psi(x,t) = \exp(tD)\psi_0(x)$ a donc pour dérivée par rapport au temps la fonction $D\exp(tD)\psi_0(x)$, c'est à dire $D\psi(x,t)$.

De façon générale, nous pouvons avoir des EDP du genre

$$\partial\psi/\partial t = H\psi$$

où H est un opérateur spatial (ne dépendant que de la variable x) plus compliqué que le simple opérateur de dérivation spatiale D, mais la discussion ci-dessus reste valide et nous pouvons donner la solution comme

$$\psi(x,t) = \exp(tH)\psi(x, t=0)$$

L'opérateur $\exp(tH)$ n'est plus alors une simple translation, mais le principe reste le même : la fonction solution à un temps ultérieur t est donnée par l'application de l'opérateur $\exp(tH)$ à la fonction "condition initiale". L'opérateur $\exp(tH)$ est appelé, à juste titre, l'opérateur de l'évolution temporelle. Cette façon de présenter les choses est appelée, en mécanique quantique l'interprétation d'Heisenberg (voir plus bas pour une digression historique). Évidemment, cette façon de résoudre l'équation ne nous sert à rien si nous ne savons pas calculer $\exp(tH)$. Nous verrons plus bas les outils que les mathématiciens ont développé pour calculer efficacement ce genre de fonctions d'opérateurs.

Exercices.

§ 10.5 Identité de Jacobi.

Démontrer que pour trois opérateurs A, B, C, nous avons

$$[A,[B,C]] + [B,[C,A]] + [C,[A,B]] = 0$$

§ 10.6 Démontrer que $D^2+X^2=(D+X)(D-X)+1=(D-X)(D+X)-1$. En déduire $[D-X,D+X]$ et $[D^2+X^2,D\pm X]$. [indication : Utiliser la relation $[D,X]=1$].

§ 10.7 Commutateur des opérateurs. Démontrer que les opérateurs O et $f(O)$ commutent. $f(x)$ est une fonction analytique dans le voisinage de $x=0$. Même chose pour $g(O)$ et $f(O)$. Démontrer que si A et B commutent, alors $f(A)$ et $g(B)$ commutent.

§ 10.8 Dans l'espace des opérateurs linéaires sur les fonctions à trois variables, nous définissons $L_z = X\partial_y - Y\partial_x$; L_x et L_y sont définis cycliquement à partir de cette dernière définition. Calculer $[L_\alpha, L_\beta]$($\alpha,\beta = x,y,z$). Donner la définition de L_α^2 et de $L^2=\sum_\alpha L_\alpha^2$. Calculer $[L^2,L_\alpha]$.

§ 10.9 Opérateur rotation. En utilisant les règles de dérivation en chaîne, démontrer que

$$\frac{\partial}{\partial\phi} = L_z$$

que représente alors l'opérateur $\exp(\alpha L_z)$? [(r,θ,ϕ) sont les coordonnées du point en coordonnées sphérique]

§ 10.10 Exponentiel d'un opérateur. démontrer que

$$\frac{d}{dt}\exp(tA) = A\exp(tA)$$

où A est un opérateur linéaire. **Help :** Utiliser le développement de l'exponentiel et les règles habituelles de la dérivation.

§ 10.11 Exponentiel d'un opérateur (encore). Démontrer que $\exp(P^{-1}AP) = P^{-1}\exp(A)P$ où A,P sont deux opérateurs linéaires.

§ 10.12 Commutation et exponentiel. Soit la matrice

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

Calculer A^2. En déduire une expression générale pour A^n selon que n est pair ou impair. Démontrer alors que

$$\exp(A) = \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}$$

Help : Décomposer la somme en termes pairs et impairs, et utiliser le développement en série des fonctions $\sin$ et $\cos$. Soit maintenant les deux matrices

$$C = \begin{pmatrix} 0 & -x \\ 0 & 0 \end{pmatrix}\ ,\ D = \begin{pmatrix} 0 & 0 \\ x & 0 \end{pmatrix}$$

Démontrer que $C^2=D^2=0$ et en déduire $e^C.e^D$. Que peut ont dire de e^Ce^D et e^{C+D} ? Est ce que C et D commutent ?

§ 10.13 $\exp^{A+B}$ Pour deux opérateurs A et B qui ne commutent pas à priori, démontrer que

$$e^{(A+B)t} = e^{At} + \int_0^t e^{A(t-s)} B e^{(A+B)s} ds$$

Cette relation, appelé en mécanique quantique relation de Dyson, est très utile pour évaluer l'exponentiel d'un opérateur, si $\exp(At)$ est connu et que la matrice de l'opérateur B est très creuse. [indication : Quel est la solution de l'équation $y' - ay = f(x)$? Obtenir une équation analogue pour l'opérateur $\exp(A+B)t$]. En profiter pour donner l'expression de $\exp(A+B)$.

§ 10.14 **Équation d'onde.** Résoudre par la méthode des opérateur l'équation

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0$$

avec les conditions initiales $u(x,0) = f(x)$ et $\partial_t u(x,t)|_{t=0} = g(x)$ en l'écrivant sous la forme symbolique $\partial^2 u/\partial t^2 - c^2 D^2 u = 0$ et en vous inspirant de la solution de l'équation ordinaire $u'' - a^2 u = 0$.

10.3 Représentation matricielle des opérateurs.

Le grand avantage de disposer d'une base est de pouvoir manipuler les vecteurs à l'aide des nombres. Le même avantage est obtenu pour les opérateurs. Par exemple, étant donné un vecteur du plan euclidien, nous n'avons pas à nous munir de compas et de rapporteur pour effectuer une rotation, mais seulement à effectuer des additions et des multiplications sur des colonnes de chiffres. Voyons comment cela fonctionne.

Soit un opérateur linéaire R dans un espace $\mathcal{E}$ que nous avons équipé d'une base $\{e_1, ...e_n\}$. Soit un vecteur v quelconque. Notre but est de pouvoir calculer le résultat de l'application de R à v, qui est un autre vecteur de $\mathcal{E}$. Nous pouvons décomposer v dans la base donnée :

$$v = \sum_j a_j e_j$$

et comme R est linéaire,

$$R.v = \sum_j a_j (R.e_j)$$

Pour entièrement caractériser l'opérateur R, nous avons simplement besoin de connaître l'action qu'il effectue sur chaque élément de la base. Par ailleurs, chaque $R.e_i$ (pour $i = 1, ..., n$) est un vecteur de l'espace $\mathcal{E}$, et donc décomposable sur la base $\{e_1, ...e_n\}$:

$$R.e_j = \sum_i r_{ij} e_i \tag{10.3}$$

Donc, pour entièrement connaître R, il suffit de connaître les $n \times n$ nombres r_{ij}. Nous pouvons alors connaître le résultat de l'application de R à n'importe quel vecteurs de l'espace.

$$R.v = \sum_{i,j} r_{ij} a_j e_i \tag{10.4}$$

Il est habituel de représenter la décomposition de $R.e_j$ comme la $j-$ième colonne d'un tableau de n lignes et appeler le tableau une matrice.

Exemple 10.2 Rotation dans le plan euclidien.

Caractérisons la rotation de $\pi/2$ dans la base orthonormée habituelle du plan euclidien (e_x, e_y). Nous savons que $R.e_x = e_y = 0e_x + 1e_y$. La première colonne de la représentation matricielle de R dans cette base sera donc $(0, 1)$. De même, $R.e_y = -1e_x + 0e_y$. La représentation de R est donc

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Il est très important de faire la distinction entre l'opérateur R et sa représentation matricielle. La représentation matricielle est comme une photo de l'opérateur : elle dépend de la base choisie, comme la photo dépend du point de vue du photographe et de l'humeur de la personne qui se fait photographier. Mais la photo *n'est pas* la personne.

Exemple 10.3 Dérivation dans l'espace des fonctions.

Choisissons la base de Fourier $\{e_q = \exp(iqx)\}$. La base est bien sûr infinie, (et même très infinie !), mais cela n'a pas d'importance. Nous avons

$$D.e_q = iq\, e_q$$

La représentation matricielle de D dans la base de Fourier ne comporte que des éléments diagonaux. Les éléments $d_{qq'}$ de la matrice de D sont donc donnés par $d_{qq'} = 0$ si $q \neq q'$ et par $d_{qq} = iq$. Le fait que D soit diagonale dans la base de Fourier est la raison de son utilisation pour la résolution de l'équation d'onde ou de la chaleur. L'application de D à une fonction quelconque donne

$$Df(x) = \int_{-\infty}^{+\infty} iq\tilde{f}(q) \exp(iqx) dq$$

ce qui n'est rien d'autre qu'une réécriture de l'expression (10.4).

Nous voyons à travers ce dernier exemple un fait important. Peu importe que notre espace soit fini ou infini, à partir du moment où nous disposons de bases (finies, infinies discrètes ou infinies continues) et que les convergences sont assurées, nous pouvons manipuler les opérateurs linéaires comme des matrices.

Base orthonormale. Nous avons vu au premier chapitre que disposer d'un produit scalaire $(.,.)$ et d'une base orthonormale facilite énormément les choses. Soit $\{e_1, ...e_n\}$ une base orthonormale, c'est à dire $(e_i, e_j) = \delta_{i,j}$ et soit la décomposition d'un vecteur quelconque v : $v = \sum a_i e_i$. Alors

$$\begin{aligned}(v, e_k) &= (\sum_i a_i e_i, e_k) \\ &= \sum_i a_i (e_i, e_k) \\ &= a_k\end{aligned}$$

Le coefficient a_k de la décomposition du vecteur v sur la base $\{e_i\}$ n'est rien d'autre que le produit scalaire entre v et e_k. Le même principe nous donne les coefficients r_{ij} d'un opérateur dans une base orthonormée. En partant de l'expression (10.3), nous voyons que

$$r_{ij} = (e_i, R.e_j)$$

Pour trouver r_{ij}, il suffit d'appliquer R à e_j et ensuite prendre le produit scalaire de ce dernier avec e_i.

Exercices.

§ 10.15 Soit la base de Fourier $\{1, \cos(2\pi nx/L), \sin(2\pi nx/L)\}$ pour les fonctions définies sur $[0, L]$. En ordonnant correctement les éléments de la base, donner l'expression de l'opérateur D et D^2 dans cette base.

§ 10.16 Soit les fonctions d'Hermite

$$h_n(x) = C_n(-1)^n \exp(x^2/2) \frac{d^n}{dx^n} \exp(-x^2) \tag{10.5}$$

Les premières fonctions sont (en multipliant par $\exp(-x^2/2)$) $1, 2x, 4x^2 - 2$,... Les coefficients C_n, que nous n'explicitons pas, assurent que les fonctions sont normées. On peut démontrer, avec un peu d'effort, que les fonctions h_n sont deux à deux orthogonales et forment une base.

Démontrer que pour l'opérateur

$$H = -D^2 + X^2$$

nous avons $Hh_n(x) = (2n + 1)h_n(x)$. Donner alors la représentation de H dans la base des h_n.

§ 10.17 **Opérateur D dans la base des Bessel.** Soit les fonctions de Bessel I_n définies par

$$I_n(x) = \frac{1}{\pi} \int_0^\pi e^{x \cos\theta} \cos(n\theta) d\theta$$

Démontrer que

$$I'_n(x) = (1/2) \left(I_{n-1}(x) + I_{n+1}(x)\right)$$

Les fonctions $\exp(-x)I_n(x)$ $(n = 0, 1, ...)$ forment une base. Donner l'expression de la matrice de l'opérateur D dans cette base. Cette matrice joue un rôle très important dans les problèmes de matrices tridiagonales.

§ 10.18 Montrer que si l'opérateur M commute avec tous les autres opérateurs, alors $M = \alpha I$. [Help : donnez vous une base quelconque $(e_1, ..., e_n)$ et considérer l'opérateur projection sur e_1 : $P_1 e_j = \delta_{i,j} e_1$. En utilisant la commutation de M avec cet opérateur et d'autres opérateurs de projection, démontrer que la matrice de M dans cette base est diagonale. Ensuite, considérer un opérateur de permutation cyclique $\Pi e_i = e_{i+1}$ et considérer sa permutation avec M : en déduire que tous les éléments diagonaux de M sont alors égaux.]

§ 10.19 **Laguerre et autres.**

10.4 Valeurs et vecteurs propres.

Quand on travail avec un opérateur linéaire et que l'on souhaite résoudre une EDP qui lui est associé, certaines fonctions (ou vecteur) ont un rôle privilégié. On appelle vecteur propre (*eigenvector* en anglais) d'un opérateur O un vecteur $v_n \neq 0$ tel que

$$Ov_n = \lambda_n v_n$$

où λ_n est un scalaire qui est appelé valeur propre du vecteur v_n.

En géométrie plane, pour l'opérateur de projection sur l'axe x, les deux vecteurs $\mathbf{e}_x$ et $\mathbf{e}_y$ sont des vecteurs propres, avec les valeurs propres 1 et 0. En géométrie à trois dimensions, les trois vecteurs $\mathbf{e}_x$ et $\mathbf{e}_y$ et $\mathbf{e}_z$ sont des vecteurs propres de la projection sur l'axe x, le premier associé à $\lambda = 1$, les deux autres à $\lambda = 0$. Dans le plan, Pour l'opération de rotation de π, tous les vecteurs du plan sont des vecteurs propres associés à la valeur $\lambda = -1$ (remarquer cependant que $R_\pi = -I$). Enfin, la rotation de $\pi/2$ n'a pas de vecteurs propres.

Exemple 10.4 Dans l'espace des fonctions, $\exp(iqx)$ est vecteur propre de l'opérateur $D = \partial_x$ avec la valeur propre iq. Les fonctions $\sin(qx)$ et $cos(qx)$ sont les vecteurs propres de l'opérateur D^2 avec $\lambda = -q^2$. Les fonctions d'Hermite $h_n(x)$ que nous avons rencontré plus haut sont fonctions propres de l'opérateur $-D^2 + X^2$ avec la valeur propre $\lambda_n = (2n + 1)$.

Si les vecteurs propres sont suffisamment nombreux pour former une base, on les appelle une base propre. Évidemment, la représentation d'un opérateur dans sa base propre est diagonale.

Disposer d'une base propre accélère (ou rend possible) la solution des problèmes. Prenons le système à deux équations et deux inconnus

$$\begin{aligned} 2x + 3y &= 4 \\ x - 2y &= 1 \end{aligned}$$

formellement, nous pouvons le représenter comme $Au = B$ où A est la matrice

$$\begin{pmatrix} 2 & 3 \\ 1 & -2 \end{pmatrix}$$

$u = (x, y)^T$ et $B = (4, 1)^T$. A est la représentation matricielle d'une application linéaire dans la base $e_1 = (1, 0)^T$ et $e_2 = (0, 1)^T$. La résolution de ce système nécessite quelques opérations d'additions et de combinaisons. Par contre, le système

$$\begin{aligned} 2x &= 4 \\ 3y &= 1 \end{aligned}$$

se résout immédiatement comme deux équations indépendantes. Formellement, on pouvait l'écrire $Au = B$ où cette fois A est la matrice

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

dans la base e_1, e_2, la matrice de l'opérateur A est diagonale, ce qui permet de ramener la résolution d'un système de deux équations à deux inconnus à la résolution de deux équations à *une* inconnue. Évidemment, un système 2×2 est facile à résoudre, pour un système de 4×4 ou 1000×1000, le gain est déjà plus appréciable. Dans l'espace de dimension *infinie* de Hilbert, trouver une base propre est souvent la seule possibilité de résoudre un problème, comme nous le verrons plus bas.

10.5 Disposer d'une base propre orthogonale.

Le très grand avantage de disposer d'une base propre orthogonale (ou encore mieux, orthonormale) est de pouvoir résoudre des équations aux dérivées partielles. Nous avons vu une application de cette méthode dans le chapitre sur les séries et transformées de Fourier. Nous allons voir cela dans des cas plus généraux.

Supposons que nous voulons résoudre l'équation aux dérivées partielles

$$\partial_t \psi(x, t) = H\psi(x, t) \tag{10.6}$$

où l'on suppose que l'opérateur H n'a que des composantes spatiales. Les conditions initiales et aux limites de cette équation sont les suivantes :

$$\begin{aligned} \psi(x \to \pm\infty, t) &= 0 \\ \psi(x, 0) &= g(x) \end{aligned}$$

la fonction tend vers zéro pour les grand x ; à l'instant $t = 0$, la fonction recherchée prend la forme de $f(x)$. Supposons que nous connaissons une base orthonormale

$\{f_n(x)\}$ dans laquelle l'opérateur H est diagonale [6] :

$$Hf_n(x) = \lambda_n f_n(x)$$

A chaque instant t, nous pouvons décomposer la fonction (inconnu) $\psi(x,t)$ sur la base des $\{f_n\}$:

$$\psi(x,t) = \sum_{i=0}^{\infty} a_n(t) f_n(t)$$

Les $a_n(t)$ sont les coefficients de cette décomposition et varient bien sûr d'un instant à un autre. En utilisant cette décomposition dans l'équation (10.6) nous obtenons [7]

$$\sum_{i=0}^{\infty} a'_n(t) f_n(x) = \sum_{i=0}^{\infty} \lambda_n a_n(t) f_n(x)$$

Comme les f_n constituent une base, les coefficients de la décomposition sont uniques et

$$a_n(t) = a_{n,0} \exp(\lambda_n t)$$

Il nous reste juste à trouver les coefficients $a_{n,0}$; ceci s'obtient en utilisant la condition initiale

$$g(x) = \sum_{i=0}^{\infty} a_{n,0} f_n(x)$$

et donc

$$a_{n,0} = \langle f_n(x), g(x) \rangle$$

Résumons la méthode. On décompose la condition initiale sur la base des fonctions propres. L'amplitude de chaque composante $a_n(t)$ varie exponentiellement avec l'exposant λ_n. Il suffit de recombiner ces composantes à un temps ultérieurs t pour recalculer la fonction ψ en ce temps là :

$$\psi(x,t) = \sum_{i=0}^{\infty} \langle \psi(x,0), f_n(x) \rangle \exp(\lambda_n t) f_n(x) \tag{10.7}$$

Cette façon de calculer la fonction ψ est appelé en mécanique quantique l'interprétation de Schrödinger [8].

6. Nous supposons que les fonctions f_n sont compatibles avec les conditions aux bords : $f_n(x \to \pm\infty) = 0$.

7. Nous supposons que l'on peut intervertir les opérations de sommation et de dérivation par rapport au temps.

8. Heisenberg et Schrödinger ont formulé ces deux approches dans les années 1922-28 quand les méthodes d'analyse fonctionnelle n'étaient pas encore popularisées chez les physiciens. La contribution de Von Neumann était de montrer, en important les concepts développés par Hilbert en mathématiques, que les deux approches étaient équivalentes. Dirac à mis la dernière main à l'édifice en 1932 en formulant l'ensemble de façon extrêmement élégante.

Exemple 10.5 oscillateur harmonique.

En mécanique quantique, l'équation d'une particule dans un potentiel harmonique s'écrit (en choisissant convenablement les unités)

$$i\frac{\partial\psi}{\partial t} = \left(-\frac{\partial^2}{\partial x^2} + x^2\right)\psi$$

D'après l'exercice §10.16 que nous avons vu sur les fonctions de Hermite, il n'est pas difficile d'expliciter la solution sous la forme de

$$\psi(x,t) = \sum a_n e^{-i(2n+1)t} h_n(x)$$

Comparaison Schrödinger-Heisenberg. L'interprétation de Schrödinger et Heisenberg sont bien sûr équivalente. Reprenons l'équation (10.6) . Dans l'interprétation d'Heisenberg, la solution s'écrit

$$\psi(x,t) = e^{tH}\psi(x,0) \tag{10.8}$$

comme nous l'avons indiqué, la fonction ψ au temps t s'obtient en appliquant l'opérateur $\exp(tH)$ à la fonction ψ au temps 0. Dans sa base propre (les $f_n(x)$ ci-dessus), l'opérateur H est diagonal. En utilisant la définition de l'opérateur

$$\exp(tH) = \sum_{n=0} (1/n!) t^n H^n$$

nous voyons que l'opérateur $\exp(tH)$ est également diagonal dans cette base, et ses éléments diagonaux s'écrivent ($(\exp(t\lambda_1), (\exp(t\lambda_2), ...(\exp(t\lambda_n), ...)$. En notation matricielle, nous avons

$$H = \begin{pmatrix} \lambda_1 & 0 & . & . & \\ 0 & \lambda_2 & 0 & & \\ . & 0 & \lambda_3 & . & \\ . & . & . & . & \\ & & & . & . \end{pmatrix}$$

et

$$e^{tH} = \begin{pmatrix} e^{t\lambda_1} & 0 & . & . & \\ 0 & e^{t\lambda_2} & 0 & & \\ . & 0 & e^{t\lambda_3} & . & \\ . & . & . & . & \\ & & & . & . \end{pmatrix}$$

Si maintenant nous exprimons la fonction $\psi(x,0)$ dans la même base

$$\psi(x,0) = \sum_n \langle \psi(x,0), f_n(x) \rangle \, f_n(x)$$

nous voyons que l'interprétation d'Heisenberg (10.8) exprime exactement la même chose que l'interprétation de Schrödinger (10.7).

Fonctions d'opérateurs. Le paragraphe précédent nous indique comment calculer une fonction d'opérateur, même quand nous ne connaissons pas la série de la fonction . Soit l'opérateur A, diagonal dans sa base propre $(e_1, e_2, ...)$ et possédant les valeurs propres $(\lambda_1, \lambda_2, ...)$. L'opérateur $f(A)$ est définie de telle façon que sa représentation dans la même base soit diagonale, avec les éléments $(f(\lambda_1), f(\lambda_2), ...)$.

10.6 Opérateurs hermitiens.

Dans la théorie des opérateurs linéaires, une certaine classe d'opérateurs qu'on appelle hermitiens joue un rôle fondamental, puisqu'on peut démontrer que ces opérateurs sont diagonalisables et qui apparaissent très souvent dans les problèmes de physique.

Nous supposons dorénavant que nous disposons d'un produit scalaire $(.,.)$.On appelle l'adjoint d'un opérateur O l'opérateur $O^\dagger$ tel que quelque soit les deux vecteurs u, v

$$(u, Ov) = (O^\dagger u, v)$$

Supposons que le produit scalaire est défini sur le corps des réels. Alors $(u, v) = (v, u)$. Si dans une base (que nous prenons orthonormale pour plus de simplicité, mais sans nécessité aucune) l'opérateur O est représenté par la matrice $O_{ij} = (e_i, Oe_j)$, on peut démontrer sans difficulté que pour la matrice de l'adjoint, $O^\dagger_{ij} = O_{ji}$: on interverti les lignes et les colonnes d'une matrice pour obtenir celle de son adjoint. La matrice résultante est appelé la transposée. Si le produit scalaire est défini sur le corps des complexes, alors $O^\dagger_{ij} = O^*_{ji}$.

Un opérateur est dit self-adjoint ou hermitien si $O^\dagger = O$.

Exemple 10.6 Dans l'espace des fonctions réelles ou complexes $\mathcal{L}^2$(qui donc tendent vers 0 pour $x \to \pm\infty$), l'opérateur D n'est pas hermitien : une intégration par partie montre que $(u, Dv) = -(Du, v)$. Par contre, deux intégrations par partie montrent que l'opérateur D^2 est hermitien. De même, l'opérateur iD est Hermitien pour les fonctions complexes.

Les opérateurs hermitiens ont quelques propriétés remarquables, parmi lesquels les suivantes que nous demandons au lecteur de démontrer/découvrir à travers les exercices.

Exercices.

§ 10.20 Les valeurs propres d'un opérateurs hermitiens sont réelles : Si $Ax = \lambda x$, $A = A^\dagger$, alors $\lambda \in \mathbb{R}$. [Indication : calculer (x, Ax) et $(x, A^\dagger x)$]

§ 10.21 Les vecteurs propres associés à deux valeurs propres distinctes sont orthogonales : Si $Ax_1 = \lambda_1 x_1$, $Ax_2 = \lambda_2 x_2$, $A = A^\dagger$, $\lambda_1 \neq \lambda_2$, alors $(x_1, x_2) = 0$. [indication : calculer (x_1, Ax_2) et $(x_1, A^\dagger x_2)$.

§ 10.22 Démontrer, dans l'espace des fonctions $\mathcal{L}^2$ que les valeurs propres de l'opérateur $-D^2$ sont positifs. [indication : il suffit de démontrer que $\forall x, (x, -D^2 x) \geq 0$, ce qui peut s'obtenir en effectuant une intégration par partie. De tels opérateurs sont appelés défini positif.].

§ 10.23 Soit la fonction réelle $V(x)$ possédant un minimum absolu $E : \forall x \in \mathbb{R}, V(x) \geq E$. Démontrer d'abord que l'opérateur V est hermitien [définition : $V.f(x) = V(x)f(x)$]; Démontrer ensuite que toutes ses valeurs propres sont $\geq E$. [indication : il suffit de considérer $(u, (V-E)u)$ pour une fonction u quelconque et démontrer que sa valeur est toujours ≥ 0]. En combinant avec le résultat de la question précédente, que pouvez vous déduire pour l'opérateur

$$H = -D^2 + V$$

§ 10.24 Pour un opérateur hermitien, le minimum de (x, Ax) avec la contrainte $(x, x) = 1$ est fournit par le vecteur propre de la matrice associée à la plus petite des valeurs propres; la valeur propre est juste un multiplicateur de Lagrange (voir le chapitre sur le calcul variationnel pour la définition des multiplicateurs de Lagrange).

§ 10.25 Dans l'espace des fonctions $\mathcal{L}^2$ à trois variables définies sur $\mathbb{R}^3$, démontrer que l'adjoint de l'opérateur gradient est l'opérateur ($-$divergence), et vice et versa, tandis que l'adjoint du rotationnel est lui même.

10.7 Méthodes opératorielles, algèbre de Lie.

En algèbre classique, pour résoudre un problème, on essaye de le décomposer en termes plus simples. Par exemple, pour résoudre $x^2 - 3x + 2 = 0$, on peut remarquer que l'on peut la ramener à la forme $(x-1)(x-2) = 0$; on remarque ensuite que pour que ce produit soit 0, il suffit qu'un des termes soit nul, et qu'il suffit donc de résoudre les deux équations $x - 1 = 0$ et $x - 2 = 0$ pour avoir les solutions du problème. Or, ces deux dernières équations sont solubles par des techniques que nous connaissons déjà.

Les opérateurs linéaires possèdent une algèbre. On peut se demander si on ne peut pas entreprendre la même démarche pour décomposer des EDP compliquées en systèmes plus simples. La réponse est oui si on fait attention aux conditions aux limites et si on garde en tête que l'algèbre des opérateurs est non commutative. Nous n'allons pas donner ici un exposé de ces méthodes, mais préférerons l'illustrer à travers deux exemples. Le problème 10.1 généralise cette approche à toute une classe d'opérateur qu'on appelle supersymétrique.

10.7.1 L'oscillateur harmonique en mécanique quantique.

L'opérateur $H = -D^2 + V(X)$ joue un rôle très important en mécanique quantique. Appelant ψ la fonction d'amplitude, l'équation de Schrödinger est de la forme

$$i\frac{\partial \psi}{\partial t} = H\psi$$

Nous avons vu au §10.6 que H est hermitien et que toutes ses valeurs propres sont supérieures à E_0, la valeur minimum de la fonction $V(x)$. L'opérateur $-D^2$ joue le rôle de l'énergie cinétique et l'opérateur $V(x)$ celui de l'énergie potentielle ; l'opérateur H est appelé Hamiltonien du système. Nous travaillons dans l'espace $\mathcal{L}^2]-\infty,+\infty[$ et la gestion des conditions aux limites ne pose pas trop de problème ($\psi(x \to \pm\infty) = 0$). Nous avons vu au §10.5 que résoudre cette équation revient à trouver les valeurs et vecteurs propres de l'opérateur H. Prenons le cas particulier du Hamiltonien

$$H = -D^2 + X^2$$

d'une particule se trouvant dans un potentiel harmonique. Nous savons que toutes les valeurs propres sont > 0. Nous avions donné en exemple plus haut les fonctions et valeurs propres de cet hamiltonien. Nous allons voir que nous pouvons trouver ces fonctions sans résoudre aucune équation différentielle, un peu comme faire du beurre avec de l'eau. Il suffit juste de manipuler l'algèbre des opérateurs et surtout de leurs commutateurs. L'exemple que nous allons suivre est le prototype de ce genre de calcul et nous le donnons donc avec un peu de détail.

Comme $[D, X] = 1$, nous pouvons décomposer l'hamiltonien :

$$H = (-D + X)(D + X) + 1 = (D + X)(-D + X) - 1$$

Notons que l'opérateur X est hermitien $X = X^\dagger$ et que $D^\dagger = -D$. Nous pouvons donc poser

$$\begin{aligned} A &= D + X \\ A^\dagger &= -D + X \end{aligned}$$

et récrire $H = AA^\dagger - 1 = A^\dagger A + 1$. Ce qui implique naturellement que $[A, A^\dagger] = 2$.

Par ailleurs, on peut démontrer facilement que pour trois opérateurs quelconque, $[FG, F] = F[F, G]$ et $[F, FG] = F[F, G]$. Cette petite gymnastique nous permet d'obtenir

$$\begin{aligned} [H, A^\dagger] &= [A^\dagger A + 1, A^\dagger] \\ &= [A^\dagger A, A^\dagger] \\ &= 2A^\dagger \end{aligned} \tag{10.9}$$

L'équation (10.9) est un cas particulier d'un opérateur d'échelle. Avoir à sa disposition de telles relations est d'un très grand avantage comme nous allons le voir plus bas.

Lemme. Si ψ_E est une fonction propre de H associée à la valeur propre E, alors $A^\dagger\psi_E$ est une fonction propre de H associée à la valeur propre $E + 2$.

Vous voyez ici l'avantage : si on connaît une valeur propre et une fonction propre, on peut en trouver beaucoup d'autre par l'application successive de l'opérateur $A^\dagger$. La démonstration est immédiate :

$$\begin{aligned} HA^\dagger\psi_E &= (2A^\dagger + A^\dagger H)\psi_E \\ &= (E+2)A^\dagger\psi_E \end{aligned}$$

Nous pouvons effectuer la même démarche pour A et démontrer que $HA\psi_E = (E-2)A\psi_E$. Comme nous savons que toutes les valeurs propres sont > 0, il existe donc forcément une valeur propre minimum que nous appelons ϵ (noter que $0 < \epsilon < 2$). Toutes les autres valeurs propres sont donc de la forme $(2n+\epsilon)$, avec $n \in \mathbb{N}$.

Il nous reste maintenant à déterminer cette valeur ϵ. Appelons ψ_0 la fonction propre associée à ϵ. Nous devons alors obligatoirement avoir

$$A\psi_0 = 0 \tag{10.10}$$

sinon nous aurions des valeurs propres négatives. De là, nous pouvons naturellement déduire que

$$H\psi_0 = \psi_0$$

Et par conséquent, $\epsilon = 1$. Les valeurs propres de l'opérateur H sont donc de la forme $E_n = 2n+1$. Quand je vous disais qu'on peut fait du beurre avec de l'eau !

Noter que la relation (10.10) est juste une équation différentielle

$$d\psi_0/dx + x\psi_0 = 0$$

avec les conditions $\psi_0 \to 0$ pour $x \to \pm\infty$. Cette équation est simple à résoudre

$$\psi_0(x) = C\exp(-x^2/2)$$

Les autres fonctions s'obtiennent à partir de là (cf l'exercice 10.27).

10.7.2 Le moment cinétique.

A ajouter.

Exercices.

§ 10.26 Donner les matrices des opérateurs A et $A^\dagger$ de l'oscillateur harmonique dans la base des fonctions propres de H.

§ 10.27 En appliquant de façon récursive l'opérateur $A^\dagger$ à ψ_0, démontrer que la fonction propre ψ_n est bien de la forme donnée par l'équation (10.5).

§ 10.28 Guider suffisamment le lecteur pour résoudre l'atome d'hydrogène.

Problèmes.

Problème 10.1 supersymétrie

Nous souhaitons généraliser la méthode développée en sous-section 10.7.1 pour le traitement de l'oscillateur harmonique à toute une classe d'opérateur qu'on appelle supersymétrique[9]. Nous allons illustrer cette méthode à travers le calcul des valeurs et fonctions propres de l'opérateur

$$H_{sec} = -d^2/dx^2 + 2/\cos^2 x$$

Pour cela, nous allons construire un opérateur jumeau dont le calcul des valeurs et fonctions propres sont plus simple, et ensuite revenir à notre objectif initial.

0. Définition. Soit l'opérateur générique

$$H_1 = -d^2/dx^2 + V_1(x)$$

où $V_1(x) \geq V_{min} \in \mathbb{R}$ (c'est à dire que la fonction $V_1(x)$ possède une borne inférieure). Rappelons que cet opérateur transforme la fonction $f \in \mathcal{L}_2$ en

$$H_1.f(x) = -f''(x) + V_1(x)f(x) \tag{10.11}$$

1. Opérateurs d'échange. Démontrer que

$$[d/dx, W(x)] = W'(x)$$

où d/dx est l'opérateur de dérivation, $W(x)$ est l'opérateur de multiplication par la fonction $W(x)$, $[,]$ désigne le commutateur, et l'égalité est dans l'espace des opérateurs. [Help : nous avions vu en cours le cas particulier $W(x) = x$; appliquer le commutateur à une fonction quelconque f et étudier le résultat]

Soit maintenant la fonction $W(x)$ tel que

$$-W'(x) + W^2(x) = V_1(x)$$

Démontrer alors que l'opérateur H_1 se met sous la forme

$$H_1 = A^\dagger A$$

où

$$\begin{aligned} A^\dagger &= -d/dx + W(x) \\ A &= d/dx + W(x) \end{aligned}$$

2. Jumeaux. Soit l'opérateur $H_2 = AA^\dagger$. Montrer que cet opérateur se met sous la forme

$$H_2 = -d^2/dx^2 + V_2(x)$$

où vous préciserez la forme de $V_2(x)$. V_1 et V_2 sont appelés des potentiels jumeaux.

3. Valeurs et fonctions propres. Démontrer que si $\psi_n(x)$ est une fonction propre de H_1 associée à la valeur propre λ_n, c'est à dire $H_1\psi_n(x) = \lambda_n\psi_n(x)$, alors la fonction $\phi_n(x) = A\psi_n(x)$

9. Un traitement détaillé de cette méthode se trouve dans le livre de Cooper, Khare et Sukhatme, *Supersymmetry in quantum mechanics.*

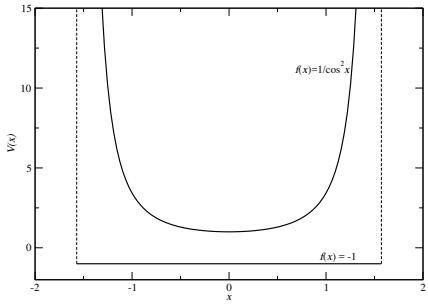


FIGURE 10.1 – Les fonctions $\sec^2 x$ et -1

est une fonction propre de l'opérateur H_2 associée à la même valeur propre. [Help : il suffit juste de se souvenir des définitions de H_1et H_2 en terme de A et $A^\dagger$.]

De même, démontrer que si $\phi_n(x)$ est fonction propre de H_2, c'est à dire $H_2\phi_n(x) = \mu_n\phi_n(x)$, alors $\psi_n(x) = A^\dagger\phi_n(x)$ est fonction propre de H_1 avec la même valeur propre.

Vous voyez ici apparaître un résultat très important : si nous connaissons les valeurs propres et fonctions propre d'un opérateur, nous connaissons immédiatement les valeurs et fonctions propres de son opérateur jumeau.

4. Construction de W.

Soit $\psi_0(x)$ la fonction propre fondamentale de H_1, c'est à dire celle qui est associée à la plus petite valeur propre λ_0 ; sans perte de généralité, nous supposons que $\lambda_0 = 0$. Nous allons considérer cette fonction parce que nous savons (on ne démontrera pas cela ici) que $\psi_0(x) \neq 0\,\forall x$, et cela nous facilite un peu la tâche par la suite de cette question. Demontrer que nous pouvons écrire $W(x)$sous la forme de

$$W(x) = -\psi_0'(x)/\psi_0(x)$$

5. Application : puits de potentiel infini. Soit maintenant l'intervalle $[-\pi/2, \pi/2]$ et le potentiel $V_1(x) = -1$; nous allons nous limiter aux fonctions $\mathcal{L}_2$ sur cet intervalle qui s'annulent sur les bords (en mécanique quantique, on parle d'un puits de potentiel infiniment profond). Vérifier que les fonctions

$$\psi_n(x) = \cos(2n+1)x$$

sont bien des fonctions propres de l'opérateur H_1 ; Quelles sont les valeurs propres λ_n associées à ces fonctions ?

6. Potentiel jumeau. Obtenir explicitement le potentiel $V_2(x)$, jumeau du potentiel $V_1(x) = -1$. En déduire les fonctions et valeurs propres de l'opérateur H_2.

7. Shift et conclusion. Soit $a \in \mathbb{R}$ et H un opérateur quelconque. Démontrer que si $f(x)$ est une fonction propre de H avec la valeur propre λ, alors elle est également une fonction propre de l'opérateur $H + a$, mais avec la valeur propre $\lambda + a$. En déduire les fonctions et valeurs propres de l'opérateur H_{sec}, notre objectif initial.

Problème 10.2 Algèbre des opérateurs et l'opération de dérivation.

Soit un opérateur linéaire $\bar{A}$ défini dans l'espace des fonctions (nous utiliserons un symbole sur-ligné pour désigner un opérateur) . Un exemple typique est l'opérateur $\bar{D} = d/dx$ qui à une fonction $f(x)$ associe la fonction $f'(x)$. Nous avons beaucoup étudié l'algèbre des opérateurs en cours. Nous allons retrouver ici quelques résultats élémentaires. Soit t un scalaire $t \in \mathbb{R}$.

1. Démontrer que $\bar{A}t = t\bar{A}$ [Help : Appliquer les deux opérateurs à une fonction quelconque].
2. Démontrer que $(t\bar{A})^n = t^n \bar{A}^n$. [Help : Rappeler ce que veut dire une puissance d'un opérateur].
3. Soit un opérateur $\bar{H}_t$ dépendant d'un paramètre scalaire (comme par exemple l'opérateur $(t\bar{A})^n$). Nous définissons l'opérateur "dérivée par rapport à t" h_t de la façon suivante :

$$h_t = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(\bar{H}_{t+\epsilon} - \bar{H}_t \right)$$

   Démontrer alors, en utilisant les deux questions précédentes et vos connaissances de la dérivée habituelle, que la dérivée par rapport à t de l'opérateur $(t\bar{A})^n$ est l'opérateur

$$nt^{n-1} \bar{A}^n$$

4. En déduire que la dérivée par rapport à t de l'opérateur $\exp(t\bar{A})$ est l'opérateur

$$\bar{A} \exp(t\bar{A})$$

   [Help : utiliser la définition en terme de série de l'opérateur $\exp(t\bar{A})$]
5. Faire la même chose avec l'opérateur $\left(1 - t\bar{A}\right)^{-1}$ [Help : à nouveau, il faut utiliser le développement en série de cet opérateur].

Problème 10.3 Vibration d'une poutre.

Le but de ce problème est de montrer, à travers un exemple simple, que les outils qu'on utilise en mécanique quantique ont un caractère très général et qu'il n'y a pas à mystifier une équation complexe du genre $i\hbar\partial_t\psi = H\psi$, qui n'est rien d'autre qu'une équation réelle d'ordre supérieur.

Nous souhaitons étudier certains aspects d'une équation d'onde se propageant dans une poutre :

$$\frac{\partial^2 u}{\partial t^2} + \rho^2 \frac{\partial^4 u}{\partial x^4} = 0$$

où ρ est une constante liée à l'élasticité de courbure de la poutre. La fonction $u(x,t)$ est la flèche en un point x à un instant t. La poutre est supposée infiniment longue. Pour faire notre étude, nous utiliserons l'arsenal des opérateurs linéaires sur des fonctions complexes à notre disposition.

1. Passage du réel au complexe. Désignons par $\phi(x,t)$ la fonction complexe

$$\phi(x,t) = \frac{\partial u}{\partial t} + i\rho \frac{\partial^2 u}{\partial x^2}$$

Nous noterons dorénavant H l'opérateur

$$H = \rho\partial^2/\partial x^2 \tag{10.12}$$

La fonction ϕ peut donc être écrite comme $\phi = \partial_t u + iHu$. Démontrer que la fonction ϕ obéit à l'équation

$$\frac{\partial \phi}{\partial t} - iH\phi = 0$$

Notons que la quantité $\phi\phi^* = (\partial u/\partial t)^2 + \rho^2(\partial^2 u/\partial x^2)^2$ est la densité d'énergie en un point x à un instant t.

2. Quelques rappels et définitions. Considérons l'ensemble de fonctions complexes de deux variables t, x muni d'un produit scalaire $(.,.)$. Nous supposerons que le produit scalaire est une intégrale sur la variable x :

$$(\phi, \psi) = \int_{-\infty}^{\infty} \phi^*(x,t)\psi(x,t)dx$$

Étant donné un opérateur quelconque A, sa valeur moyenne sur une fonction ϕ est définie par

$$\langle A \rangle = (\phi, A\phi)$$

Nous supposons dorénavant que les fonctions qui nous intéressent obéissent à une équation d'évolution

$$\frac{\partial \phi}{\partial t} = iH\phi \tag{10.13}$$

où H est un opérateur hermitien qui ne contient que de dépendance en x.

3. Dérivation. Démontrer, en utilisant la linéarité du produit scalaire que

$$\frac{\partial}{\partial t}(\phi, \psi) = \left(\frac{\partial \phi}{\partial t}, \psi\right) + \left(\phi, \frac{\partial \psi}{\partial t}\right)$$

4. Évolution. Démontrer que l'évolution de $\langle A \rangle$, où A est un opérateur quelconque n'ayant que de dépendance en x, est donnée par

$$\frac{\partial}{\partial t}\langle A \rangle = i \langle [A, H] \rangle$$

où $[A, H] = AH - HA$.

5. Densité d'énergie. L'énergie totale dans la poutre est donnée par la quantité $E = \langle 1 \rangle$, où 1 est l'opérateur identité. Démontrer que l'énergie est une quantité conservée, c'est à dire $\partial E/\partial t = 0$.

6. Commutateur de X. Soit l'opérateur $P = \partial/\partial x$. Démontrer que

$$[X, P] = -1$$

où X est l'opérateur de multiplication par x : $X\phi = x\phi$. En déduire que

$$[X, P^2] = -2P$$

Il n'est pas difficile de voir que l'opérateur H défini en (10.12) vaut $H = \rho P^2$.

7. Barycentre de l'énergie. le barycentre de l'énergie est donnée par $\langle X \rangle$. Démontrer que si la fonction ϕ est paire, c'est à dire $\phi(x,t) = \phi(-x,t)$, alors

$$\frac{\partial \langle X \rangle}{\partial t} = 0$$

[Help : Si une fonction est paire, quelle est la nature de sa dérivée ? Que peut ont dire de l'intégrale des fonctions paires et impaires ?]

8. Évolution. Nous savons que la solution de l'équation (10.13) est donnée, *operatoriellement*, par

$$\phi(x,t) = e^{itH}\phi(x,0)$$

En développant l'opérateur exponentiel, démontrer que si la fonction $\phi(x,0)$ (la condition initiale) est paire, alors la fonction aux temps ultérieurs $\phi(x,t)$ reste toujours paire.

Note. Si nous voulions continuer, nous aurions développé une relation entre la vitesse du barycentre est la *chiralité* de la fonction définie par $\int_I \phi^*\phi' dx$; nous aurions également étudié l'élargissement de l'onde, donné par $\langle X^2 \rangle$. Il aurait été utile également de considérer une poutre finie et voire comment on prend en compte les conditions aux bords.

11 Les systèmes de Sturm-Liouville.

11.1 Introduction.

Nous allons traiter dans ce chapitre la théorie générale [1] des équations différentielles *linéaires* de second ordre de la forme :

$$Lu = \lambda u \tag{11.1}$$

où $u(x)$ est une fonction au moins deux fois dérivable, la fonction et ses dérivées étant de carré sommable [2]. λ est un scalaire que nous appelons valeur propre et L est un opérateur différentiel de second ordre

$$L = \alpha(x)\frac{d^2}{dx^2} + \beta(x)\frac{d}{dx} + \gamma(x)$$

Ce genre d'équation se rencontre partout en physique et la solution du système ci-dessus a occupé une grande partie du travail des physiciens-mathématiciens. Les grandes familles de fonctions spéciales, les Bessel, Legendre, Hermite, Laguerre, les hypergéométriques, ... ont été introduites à l'occasion de l'étude de ce genre d'équation. Le but de ce chapitre est en partie de nous habituer à ces fonctions. La théorie générale que nous allons voir dans ce chapitre est une application directe de la théorie des opérateurs linéaires que nous avons rencontré au chapitre précédent ; nous aurons ainsi l'occasion d'approfondir les concepts que nous avons vus.

Illustrons à travers les trois exemples suivants quelques fonctions et équations fondamentales de la physique mathématique.

Exemple 11.1 Nous avons vu que pour trouver $u(x,t)$,solution de l'EDP

$$\frac{\partial^n u}{\partial t^n} = Lu \tag{11.2}$$

où l'opérateur L ne contient pas de dépendance en t, nous pouvons chercher la solution de l'équation aux valeurs propres

$$L\phi_n = \lambda_n \phi_n$$

1. Cette théorie a été formulée par les deux mathématiciens français aux alentour de 1850.
2. Ce genre d'espace est appelé l'espace de Sobolev, mathématicien soviétique du vingtième siècle, précurseur de la théorie des distributions.

où $\phi_n(x)$ est fonction de la seule variable x. Si les $\{\phi_n\}$ constituent une base dans l'espace des fonctions appropriées, alors la solution générale de (11.2) est donnée par

$$u(x,t) = \sum_n f_n a_n(t)\phi_n(x)$$

où les coefficients $a_n(t)$ sont solutions de l'équation différentielle

$$\frac{d^n a_n}{dt^n} - \lambda_n a_n = 0$$

et les scalaires f_n sont donnés par les conditions initiales. Si l'on veut, cette méthode est une "super" méthode de séparation des variables.

Exemple 11.2 **Propagation en coordonnées cylindrique et l'équation de Bessel.**Nous souhaitons résoudre l'équation d'onde à deux dimensions en coordonnées polaires (comme par exemple la vibration d'une membrane) :

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u \tag{11.3}$$

où $u = u(r,\theta,t)$. Si nous savons résoudre

$$\Delta\phi_k(r,\theta) = -k^2\phi_k(r,\theta) \tag{11.4}$$

(Question : pourquoi les valeurs propres de l'opérateur Δ sont réelles *et* négatives ?) alors la solution générale s'écrit [3]

$$u(r,\theta,t) = \sum_k \left(A_k e^{ickt} + A_{-k} e^{-ickt}\right)\phi_k(r,\theta)$$

En coordonnée cylindrique, l'équation (11.4) s'écrit

$$-\frac{\partial^2\phi_k}{\partial\theta^2} = r^2\frac{\partial^2\phi_k}{\partial r^2} + r\frac{\partial\phi_k}{\partial r} + r^2k^2\phi_k \tag{11.5}$$

Si on regarde de plus près cette équation, nous voyons qu'elle a à nouveau exactement la forme (11.2) et nous pouvons appliquer la même méthode en cherchant la solution de l'équation aux valeurs propres

$$r^2\frac{d^2\eta_m}{dr^2} + r\frac{d\eta}{dr} + r^2k^2\eta_m = m^2\eta_m \tag{11.6}$$

3. Nous utilisons le symbol Σ pour signifier "superposition". Si la variable sur laquelle on somme est discrète ($\in \mathbb{Z}$)alors Σ a son sens habituel. Si par contre la variable de sommation est continue ($\in \mathbb{R}$) alors le symbol doit être compris formellement comme une intégrale $\int$.

où $\eta_m = \eta_m(r)$ est la fonction propre et m est la valeur propre. La solution de (11.5) est

$$\phi_k(r,\theta) = \sum_m \left(B_m e^{im\theta} + B_{-m} e^{-im\theta}\right) \eta_m(r)$$

Nous remarquons en plus une symétrie remarquable dans la fonction ϕ_k : $\phi_k(r, \theta + 2\pi) = \phi_k(r,\theta)$. Cette symétrie nous impose que m ne peut pas être quelconque, mais doit être entier : $m = 0, 1, ...$ Nous rencontrerons constamment cette symétrie et sa conséquence dans les équations de la physique écrites en coordonnées cylindriques ou sphériques. Nous pouvons encore quelque peu nettoyer l'équation (11.6) en posant

$$\eta_m(r) = J_m(x)$$

où $x = kr$. Par la règle des dérivations en chaîne, nous avons $d/dr = k(d/dx)$ et nous obtenons

$$x^2 J''_m(x) + x J'_m(x) + (x^2 - m^2) J_m(x) = 0 \tag{11.7}$$

Cette équation s'appelle l'équation de Bessel, les fonctions J_m sont appelées les fonctions de Bessel ; le lecteur notera que cette équation qui joue un rôle primordial dans les équations de propagation en coordonnées cylindriques est de la forme (11.1).

Nous avons beaucoup détaillé les calculs dans cet exemple. Nous pouvons systématiser le travail de la façon suivante : nous cherchons la solution de (11.3) directement sous la forme

$$u(r,\theta,t) = \eta(r) f(\theta) a(t)$$

En remplaçant u dans l'équation (11.3) et en divisant par u, nous trouvons

$$\frac{a''}{a} - c^2 \frac{1}{\eta}\left(\eta'' + \eta'/r\right) - \frac{c^2}{r^2}\frac{f''}{f} = 0 \tag{11.8}$$

Cette équation est valable pour tous (r, θ, t). Supposons que nous gardons constante r et t et que nous varions θ. Les deux premiers termes de cette équation ne dépendent pas de θ et restent inchangés. Pour que l'équation soit satisfaite, nous devons avoir $f''/f = \text{Cte}$. La contrainte $f(\theta + 2\pi) = f(\theta)$ nous impose $\text{Cte} = -m^2$, $m = 0, 1, ...$ De la même manière, si l'on fait varier t, les deux derniers termes restent inchangés et nous devons donc avoir $a''/a = \text{cte}$; la solution devant rester bornée pour tous les temps, nous devons avoir $\text{cte} = -k^2$. L'équation (11.8) se transforme alors en

$$(k/c)^2 + \frac{1}{\eta}\left(\eta'' + \eta'/r\right) - m^2/r^2 = 0$$

en multipliant par η et en nettoyant un peu, cette dernière équation se met sous la forme de l'équation de Bessel. La solution générale est la superposition de ces solutions.

Exemple 11.3 **Équation de Laplace en coordonnées sphériques et équation de Legendre.** Nous souhaitons résoudre, en coordonnées sphérique (r, θ, ϕ), l'équation de Laplace

$$\Delta u = 0 \tag{11.9}$$

Posons $u(r, \theta, \phi) = f(r)g(\theta)h(\phi)$. En suivant notre procédé, nous aboutissons à

$$\frac{1}{f(r)}\frac{d}{dr}\left(r^2\frac{df}{dr}\right) + \frac{1}{g(\theta)\sin(\theta)}\frac{d}{d\theta}\left(\sin(\theta)\frac{dg}{d\theta}\right) + \frac{1}{h(\phi)\sin^2(\theta)}\frac{d^2h}{d\phi^2} = 0$$

Par des arguments analogues à ce que nous avons utilisés précédemment, nous obtenons

$$\frac{h''(\phi)}{h(\phi)} = -m^2$$

et

$$\frac{1}{f(r)}\frac{d}{dr}\left(r^2\frac{df}{dr}\right) = \lambda \tag{11.10}$$

Il n'est pas difficile de vérifier qu'une fonction f de la forme $f(r) = A_n r^n + B_n r^{-(n+1)}$ est fonction propre, avec la valeur propre $\lambda = n(n+1)$. Ceci nous donne pour la dernière équation

$$\frac{1}{\sin(\theta)}\frac{d}{d\theta}\left(\sin(\theta)\frac{dg}{d\theta}\right) + \left(n(n+1) - \frac{m^2}{\sin^2(\theta)}\right)g(\theta) = 0$$

Et nous allons quelques peu nettoyer cette dernière équation. Remarquons d'abord que la fonction $g(\theta)$ est paire ; nous avons donc besoin de la résoudre que sur l'intervalle par exemple $[0, \pi]$. Posons $\cos\theta = t$ ($t \in [-1, 1]$) et $g(\theta) = P(t)$. La dérivation en chaîne nous donne $d/d\theta = -\sin\theta(d/dt)$ et donc

$$\frac{d}{dt}\left((1-t^2)\frac{dP}{dt}\right) + \left(n(n+1) - \frac{m^2}{1-t^2}\right)P(t) = 0 \tag{11.11}$$

Les solutions de cette équation, appelées les fonctions de Legendre associées sont notées $P_n^m(t)$. Pour $m = 0$, L'équation s'écrit

$$(1-t^2)P''(t) - 2tP'(t) + n(n+1)P(t) = 0$$

Notons à nouveau que cette équation, qui joue un rôle primordial dans les équations impliquant le Laplacien en coordonnées sphériques, est de la forme (11.1). Les solutions de cette équation sont appelées les polynômes de Legendre. Nous verrons que les solutions pour $m \neq 0$ s'obtiennent à partir des $m = 0$. La partie angulaire de l'équation de Laplace est donnée par

$$Y_n^m(\theta, \phi) = e^{im\phi}P_n^m(\cos\theta)$$

ces dernières fonctions sont appelées les harmoniques sphériques et jouent le même rôle que les fonctions sinus et cosinus à une dimension. Le lecteur les rencontrera souvent comme les fonctions propres des deux opérateurs L^2 et L_z dans les problèmes à symétries sphériques.

11.2 Reformulation opératorielle.

La méthode que nous avons suivi ci-dessus est quelque peu pédestre et nous oblige à par exemple à séparer les rôles des variables t et x. Ceci est quelque peu artificiel et nous pouvons nous en dispenser. Ce que nous avons distingué dans les exemples que nous avons traités plus haut est que l'équation à dérivées partielles pouvait se mettre sous la forme de $(L_1 + L_2 + ...)u = 0$ où l'opérateur L_i ne contient que des dérivations sur la variable i [4].

Prenons l'espace des fonctions à deux variables (que nous appelons x et y) pour illustrer notre propos. Soit un opérateur H dans cet espace que l'on peut écrire comme la somme de deux opérateurs, chacun ne contenant qu'une seule dépendance : $H = H_x + H_y$ (par exemple, pour le Laplacien en coordonnées cartésiennes nous avons $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2$). Supposons que nous souhaitons résoudre l'équation $Hu(x, y) = 0$ dans le domaine $x \in [a_1, b_1]$ et $y \in [a_2, b_2]$. Supposons de plus que dans l'espace des fonctions sur $[a_2, b_2]$ nous disposons d'une base $\{\eta_i(y)\}$qui de plus est une base propre de l'opérateur H_y : $H_y\eta_i(y) = \lambda_i\eta_i(y)$. Évidemment, nous pouvons, pour chaque x, décomposer la fonction (inconnue pour l'instant) $u(x, y)$ dans cette base :

$$u(x, y) = \sum_i c_i(x)\eta_i(y)$$

connaître les coefficients $c_i(x)$ est équivalent à connaître la fonction $u(x, y)$ [5]. Si nous remplaçons l'expression ci-dessus dans l'EDP que nous tentons de résoudre, nous trouvons, en utilisant les propriétés de linéarité de H,

$$\begin{aligned} Hu &= H\left(\sum_i c_i(x)\eta_i(y)\right) \\ &= \sum_i H\left(c_i(x)\eta_i(y)\right) \\ &= \sum_i \left(H_x.c_i(x) + \lambda_i c_i(x)\right)\eta_i(y) \end{aligned}$$

4. Ceci bien sûr n'est pas un hasard : on choisit (invente) un système de coordonnées dans lequel l'EDP peut s'écrire de cette façon. C'est la raison de la popularité des trois systèmes de coordonnées les plus populaires. Le lecteur rencontrera d'autres systèmes plus exotiques adaptés à des problèmes bien particulier.

5. Souvenez vous, nous appelons cela l'interprétation de Schrodinger.

Puisque les η_i sont linéairement indépendantes (elles forment une base), l'équation à dérivées partielles $Hu = 0$ se transforme en une équation différentielle ordinaire (donc beaucoup plus simple)

$$H_x.c_i(x) + \lambda_i c_i(x) = 0$$

Les exemples que nous avons vu plus haut avaient l'air légèrement plus compliqué que cela. Les EDPs que nous avons vu avaient plutôt la forme

$$(H_x + f(x)H_y)u(x,y) = 0 \tag{11.12}$$

Mais ceci n'est guère différent. En faisant exactement la même démarche de décomposer $u(x,y)$ sur la base des $\{\eta_i(y)\}$, nous obtenons pour les coefficients $c_i(x)$ l'équation différentielle ordinaire

$$H_x.c_i(x) + \lambda_i f(x) c_i(x) = 0$$

Revoyons à nouveau nos exemples.

Dans le cas de la propagation d'onde en coordonnées polaire, nous pouvons écrire l'équation (11.3)sous la forme opératorielle

$$\left(\partial_{tt} - (\partial_{rr} + (1/r)\partial_r) - (1/r^2)\partial_{\theta\theta}\right) u(r,\theta,t) = 0$$

où nous avons posé $c = 1$. Nous voyons que cette équation possède la même structure que (11.12). Pour la variable $\theta \in [0, 2\pi]$, les fonctions $\exp(im\theta)$ sont des fonctions propres de l'opérateur $\partial_{\theta\theta}$ associées à la valeur propre $-m^2$ (de plus, $m \in \mathbb{Z}$) . Pour la variable $t \in]-\infty, +\infty[$, les fonctions $\exp(ikt)$sont des fonctions propres de l'opérateur ∂_{tt} associées à la valeur propre $-k^2$. Si l'on cherche la solution générale sous la forme de

$$u(r,\theta,t) = \sum_{k,m} R_{k,m}(r) e^{im\theta}.e^{ikt}$$

Alors, d'après ce que nous avons dit, la fonction $R_{k,m}(r)$ doit obéir à l'équation

$$\left(-k^2 - (\partial_{rr} + (1/r)\partial_r) + m^2/r^2\right) R(r) = 0$$

Qui n'est rien d'autre que l'équation de Bessel déjà rencontrée.

Prenons maintenant le cas de l'équation de Laplace en coordonnées sphériques $\Delta u = 0$ qui sous forme opératorielle s'écrit

$$\left(\partial_{rr} + \frac{2}{r}\partial_r + \frac{1}{r^2}L^2\right) u = 0$$

où L^2 est l'opérateur qui contient toutes les dérivées par rapport aux variables angulaires θ, ϕ :

$$L^2 = \frac{1}{\sin\theta}\partial_\theta \sin\theta \partial_\theta + \frac{1}{\sin^2\theta}\partial_{\phi\phi}$$

Nous connaissons une base propre [6] de l'opérateur L^2 formée des sphériques harmoniques $Y_n^m(\theta,\phi)$, associées à la valeur propre $-n(n+1)$. Si on cherche la solution sous forme de

$$u(r,\theta,\phi) = \sum_{m,n} R(r) Y_n^m(\theta,\phi)$$

alors la fonction $R(r)$ doit obéir à l'équation

$$R''(r) + \frac{2}{r}R'(r) - \frac{n(n+1)}{r^2}R(r) = 0$$

Exercices.

§ 11.1 Démontrer que la fonction $J_m(x) = (1/\pi)\int_0^\pi \cos(x\sin\theta - m\theta)d\theta$ est solution de l'équation de Bessel si $m \in \mathbb{Z}$. Cette fonction est appelée la fonction de Bessel (d'ordre m) de première sorte.

§ 11.2 Transformée de Bessel-Fourier. Soit une fonction de deux variables $f(x,y)$ dont la transformée de Fourier (bidimensionnelle) est

$$\tilde{f}(u,v) = \int\int_{-\infty}^{\infty} f(x,y)\exp\left(-i\left(ux+vy\right)\right)dxdy$$

Supposons que f est de symétrie cylindrique, c'est à dire qu'on peut l'écrire comme $f(r)$ où $x+iy = r\exp(i\theta)$. Dans ce cas, on peut également choisir les coordonnées polaires dans l'espace réciproque $u+iv = q\exp(i\phi)$. Démontrer alors que $\tilde{f}$ est également de symétrie cylindrique :

$$\tilde{f}(q) = 2\pi\int_0^\infty f(r)J_0(qr)rdr$$

Ceci s'appelle une transformation de Bessel-Fourier ou Hankel et qui est très similaire à la TF, puisqu'on peut démontrer l'inversion :

$$f(r) = 2\pi\int_0^\infty \tilde{f}(q)J_0(qr)qdq$$

[Help : pour démontrer la formule inverse, nous admettons la relation d'orthogonalité des fonctions de Bessel : $\int_0^\infty J_\mu(kr)J_\mu(k'r)dr = \delta(k-k')/k$.]

§ 11.3 Démontrer que pour m entier, $J_{-m}(x) = (-1)^m J_m(x)$. Cela montre que pour m entier, les deux fonctions J_m et J_{-m} ne sont pas linéairement indépendantes. L'équation de Bessel étant de second ordre, elle a besoin de *deux* solutions indépendantes. L'autre solution notée souvent $Y_m(x)$ est appelé la fonction de Bessel de seconde sorte. [Help : effectuer le changement de variable $\theta \to \pi - \theta$].

6. A vrai dire, nous n'avons jamais démontré que ces fonctions constituent une base. La démonstration découle du fameux théorème d'approximation de Weierstrass (1885). Nous admettons le résultat. Nous avions simplement démontré que les sphériques harmoniques sont les fonctions propres de l'opérateur L^2.

§ 11.4 En ramenant l'équation suivante

$$x^2u'' + \nu xu' + \left(x^2 - m^2\right)u = 0$$

à une équation de Bessel, donner sa solution générale. [Help : poser $u = x^\alpha w$ et choisir α en conséquence. Solution : $u(x) = J_{\pm\mu}(x)/\sqrt{x}$ où $\mu^2 = m^2 + (\nu-1)^2/4$.

§ 11.5 Résoudre l'équation

$$x^2u'' + xu' + \alpha^2\left(x^{2\alpha} - m^2\right)u = 0$$

[Help : poser $x = y^\nu$, et choisir ν pour se ramener à une équation de Bessel.Solution : $u(x) = J_m(x^\alpha)$]

§ 11.6 Combiner les deux questions précédentes pour résoudre

$$xu'' + (p+1)u' + u/4 = 0$$

[Solution : $u(x) = x^{-p/2}J_p(\sqrt{x})$].

§ 11.7 Polynômes de Legendre associés. Soit la fonction $u(x)$ solution de l'équation de Legendre

$$(1-x^2)u'' - 2xu' + n(n+1)u = 0$$

Démontrer alors que la fonction $v(x) = (1-x^2)^{m/2}(d^m/dx^m)u(x)$ est solution de l'équation de Legendre associée

$$(1-x^2)v'' - 2xv' + \left(n(n+1) - \frac{m^2}{(1-x^2)}\right)v = 0$$

Si on sait que le polynôme de Legendre P_n est d'ordre n, que peut on déduire pour P_n^m pour $m > n$? Construire explicitement les trois premiers polynômes de Legendre. [Help : Commencer par démontrer par récurrence que

$$d^m\left[(1-x^2)u'' - 2xu' + n(n+1)u\right] =$$
$$(1-x^2)d^{m+2}u - 2(m+1)xd^{m+1}u + \left[n(n+1) - m(m+1)\right]d^m u$$

où d^m est l'opérateur de dérivation d'ordre m.

§ 11.8 **Moment cinétique.** Soit l'opérateur $L^2 = L_x^2 + L_y^2 + L_z^2$. Démontrer que l'opérateur laplacien, en coordonnées sphérique, s'écrit comme

$$\Delta = r^{-2}\left\{\partial_r r^2 \partial_r + L^2\right\}$$

Rappel : $L_z = x\partial_y - y\partial_x$ et les autres s'en déduisent par permutation circulaire des indices (x, y, z). Nous avions déjà vu que $L_z = \partial_\phi$. Démontrer que les harmoniques sphériques sont fonctions propres de L_z, donner les valeurs propres associées.

§ 11.9 Donner la solution générale de l'équation d'onde en coordonnées sphériques

$$\partial_{tt} u = \Delta u$$

Help : en notation opératorielle, l'équation s'écrit

$$\left(\partial_{tt} - r^{-2}\left\{\partial_r r^2 \partial_r + L^2\right\}\right) u = 0$$

Comme nous connaissons les fonctions propres de ∂_{tt} ($= \exp(\pm kt)$) et de L^2 ($= Y_n^m(\theta, \phi)$), il suffit de chercher la solution générique sous la forme de

$$u(r, \theta, \phi, t) = R(r) Y_n^m(\theta, \phi) e^{\sigma i k t}$$

où $\sigma = \pm 1$ et $|m| < n$, et obtenir une équation pour $R(r)$. On devrait obtenir, pour R :

$$R(r) = \frac{1}{\sqrt{kr}} J_{\sigma'(n+1/2)}(kr)$$

où $\sigma' = \pm 1$. Ces fonctions sont parfois appelées des Bessel sphériques.

§ 11.10 **Atome d'Hydrogène.** Résoudre en coordonnées sphériques l'équation de l'atome d'Hydrogène :

$$i\partial_t \psi = (-\Delta - \alpha/r)\psi$$

On suppose $\alpha > 0$ (interaction attractive). Help : En suivant le même procédé que l'exercice précédent, et en renormalisant convenablement, démontrer que la partie radiale de ψ obéit à l'équation

$$x^2 R'' + 2xR' + (-x^2 + \beta x - n(n+1))R = 0$$

où :$-E$ est la valeur propre associée à la fonction propre $\exp(iEt)$ de $i\partial_t$; $\beta = \alpha/\sqrt{E}$ (on suppose $E > 0$) ; $n(n+1)$ est la valeur propre associée à la fonction propre $Y_n^m(\theta, \phi)$. En posant $R(x) = x^s e^{-x} u(x)$ et en choisissant s convenablement, on peut obtenir l'équation

$$xu'' + 2(1 + n - x)u' + [\beta - 2(1 + n)]u = 0$$

Cette équation est connu sous le nom de l'équation de Laguerre.

§ 11.11 **Legendre.** Quelque chose autour de l'expansion

$$\frac{1}{|\mathbf{x} - \mathbf{x}'|} = \frac{1}{\sqrt{r^2 + r'^2 - 2rr'\cos\theta}} = \sum_{n=0}^{\infty} \frac{r'^n}{r^{n+1}} P_n(\cos\theta)$$

11.3 Détour : la mécanique quantique ou pourquoi les valeurs propres ont pris tant d'importance.

Pouvoir développer les solutions d'une EDP comme une somme de fonctions propres était une technique répandue et maîtrisée à la fin du XIXème siècle. Les valeurs propres ont cependant pris une importance extrême avec l'avènement de la mécanique quantique (MQ). La MQ ajoutait quelque chose qui paraissait très étrange aux scientifiques

habitués à la propagation des ondes[7]. Pour comprendre cette *crise* psychologique[8], prenons l'exemple d'une corde vibrante à une dimension sur l'intervalle $[0, L]$ dont l'équation est donnée par

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

La solution sur la base propre de Fourier, d'après ce que nous avons vu plus haut, est donnée par

$$u(x,t) = \sum_n \left(A_n e^{i\omega_n t} + B_n e^{-i\omega_n t} \right) e^{ik_n x} \tag{11.13}$$

où $k_n = (2\pi/L)n$, les $\exp(ik_n x)$ sont une base propre de l'opérateur $\partial^2/\partial x^2$ de valeur propre $k_n^2 = (2\pi n/L)^2$, et $\omega_n = ck_n$. Notons que la fonction $u(x,t)$ représentant un déplacement réel, cela nous impose en plus d'avoir $B_n = A_n^*$.

L'énergie de la corde à un instant donné est la somme de son énergie cinétique et potentielle[9] et s'écrit

$$\mathcal{E}(u) = \int_0^L \left\{ \frac{\rho}{2} u_t u_t^* + \frac{K}{2} u_x u_x^* \right\} dx \tag{11.14}$$

où u_α désigne $\partial u/\partial \alpha$. Par ailleurs, ρ représente la densité de la corde, K sa constante élastique et $c^2 = K/\rho$. En utilisant la solutions (11.13), on trouve facilement

$$\mathcal{E}(u) = (2LK) \sum_n A_n A_n^* k_n^2 \tag{11.15}$$

A_n est l'amplitude du mode n, et nous voyons que le résultat est la somme des amplitudes des modes au carré, pondéré par leur valeurs propres.

Nous pouvons construire un appareil pour mesurer l'énergie de la corde vibrante de la façon suivante : une caméra prend deux photographies successives de la corde. Par

7. De la même manière que la mécanique Newtonienne ajoutait quelque chose de très étrange à la science du mouvement répandu en 1680. Notre intuition nous disait que la vitesse d'une charrette que l'on pousse dépend de la force qu'on exerce dessus et qu'elle s'arrête quand nous ne poussons plus. La physique Aristotélicienne postulait donc, en langage moderne, $v = C.F$. La mécanique Newtonienne a modifié ce postulat en imposant la proportionnalité non pas entre la vitesse et la force, mais entre l'accélération et la force : $a = (1/m)F$. Pour interpréter les résultats expérimentaux dans ce cadre, il a fallu *supposer* l'existence de forces de réaction, de frottement, ... On peut facilement imaginer combien cette mécanique nouvelle paraissait contre-intuitif et étrange aux scientifique de l'époque. Pour nous qui avons assimilé cette mécanique depuis notre enfance, cette mécanique nous paraît aller de soi. De la même manière, la MQ paraissait étrange au début du XXème, puisque les scientifiques s'étaient forgée une intuition qui allait à l'encontre des postulats de la nouvelle mécanique.

8. Parfois appelé dualité onde-corpuscule

9. voir les chapitres consacrés, d'une part aux équations de la physique, d'autre part au calcul variationnel.

l'analyse de la forme de la corde sur une des photos, nous pouvons explicitement mesurer point par point u_x. Par l'analyse de la hauteur de la corde entre les deux photos successives, nous pouvons mesurer point par point u_t. Ces deux mesures nous permettent ensuite d'effectuer une intégrale point par point et remonter à l'énergie (11.14) [10] .

Supposez maintenant que vous avez préparé par exemple $M = 10^6$ cordes vibrantes toutes dans le même état (identiques). En les mesurant les uns après les autres avec l'appareil, on doit toujours trouver la même valeur $\mathcal{E}$. Les premiers chercheurs investiguant les phénomènes atomiques ont trouvé un résultat différent : A chaque mesure, ils trouvaient une valeur différente, correspondant exactement à une des valeurs propres k_n. La *mesure* de l'énergie ne donnait pas la même valeur, ni même des valeurs dispersées de façon continue, mais un ensemble discret de valeur. Il s'avère que quand ont fait cette mesure un grand nombre de fois, la *proportion* de fois où l'on tombe sur la même valeur k_n et proportionnelle au coefficient $A_n A_n^*$. La *moyenne* sur les M mesures donne en effet le résultat attendu $\mathcal{E}$ d'un système classique [11].

Cette observation a été intégrée comme un postulat à la mécanique quantique [12], en sus des équations d'évolution de la mécanique quantique (comme par exemple l'équation de Schrödinger). Ce n'est que récemment (les années 1990) que ce postulat a été compris comme une conséquence du phénomène d'interaction entre un système macroscopique et microscopique, appelé *décohérence*.

Dans beaucoup de branche de physique atomique et moléculaire, la quantité pertinente étant l'énergie, les scientifiques ont élaborer des règles de manipulations impliquant uniquement la connaissance des valeurs propres, sans même passer par la case "résoudre l'équation d'évolution temporelle". Voilà pourquoi les valeurs propres sont devenues une partie si fondamentale de la culture des physiciens.

11.4 Les systèmes de Sturm-Liouville.

Revenons à notre propos original sur l'équation $Lu = \lambda u$ où L est un opérateur différentiel de second degrés. Nous nous plaçons dorénavant dans l'ensemble des fonctions à valeurs complexes. Nous équipons notre espace vectoriel des fonctions d'un produit

10. Pour l'investigation des phénomènes à l'échelle atomique, les scientifiques du début de XXème siècle avaient à leur disposition un appareil de mesure basée sur la spectroscopie. Ce sont le caractère discret des raies qui ont commencé à poser problème. Notez que le problème dont on parle ici n'est pas du tout celui de la "divergence ultra-violette" du corps noir. Ce dernier est dû a un postulat de la physique statistique qui exige, dans le cadre de la physique classique, qu'à l'équilibre thermodynamique, on ait $2LKA_nA_n^*k_n^2 = T/2$, ce qui rend la somme (11.15) divergente.

11. La présentation que je donne est bien sur anachronique. Le concept de décomposition sur base propre à travers l'équation de Schrodinger post-date de quelques 20 années la découverte de la nature discrète des raies atomiques.

12. Cela s'appelle l'interprétation de Copenhague (1928) : Une mesure d'un observable $\mathcal{O}$ projette un état quantique mixte sur un des états propres de cet observable, avec une probabilité donnée par la norme au carré de l'amplitude de ce mode propre.

scalaire, mais nous allons quelque peu élargir notre définition du produit scalaire :

$$(f,g) = \int_{-\infty}^{\infty} f^*(x)g(x)w(x)dx$$

Le lecteur peut vérifier que cette définition possède toutes les bonnes propriétés que l'on exige d'un produit scalaire, *si* $w(x) \geq 0$ [13] . La fonction $w(x)$ est dite le poids (weight en anglais).

Dans la très grande majorité des cas que l'on rencontre en physique-mathématique, l'opérateur L est hermitien [14] : $(f, Lg) = (Lf, g)$C'est d'ailleurs précisément ces systèmes où l'opérateur L est hermitien que nous appelons Sturm-Liouville. Les opérateurs hermitiens ont des propriétés remarquables : ils sont diagonalisables et leurs valeurs propres sont réelles. Ces faits ne sont pas étranger au fait qu'ils apparaissent aussi souvent en physique. Nous avons vu que la forme générale de l'opérateur L est

$$L = \alpha(x)d_{xx} + \beta(x)d_x + \gamma(x)$$

où $d_x$et d_{xx} dénotent les dérivées première et seconde et α, β, γ sont des fonctions à valeurs réelles. Comme on exige de L d'être hermitien, cela limite le choix des fonctions α, β, γ. Si on regarde d'un peu plus près, aucune contrainte ne pèse sur $\gamma(x)$. Si L est hermitien, il est trivial de démontrer que $L+\eta(x)$ est également hermitien. Pour étudier les contraintes qui pèsent sur α et β, nous choisissons donc pour l'instant $\gamma(x) = 0$, cela nous simplifie l'écriture. Formons la différence $I = (Lf, g)-(f, Lg)$. En effectuant une intégration par partie, nous parvenons facilement à

$$\begin{aligned} I &= [(f'g - fg')(\alpha w)]_{-\infty}^{\infty} \\ &+ \int_{-\infty}^{\infty} (fg' - f'g)\left[(\alpha w)' - \beta w\right] dx \end{aligned}$$

L'hermicité de L nous exige donc que

$$(\alpha w)' = \beta w \tag{11.16}$$

$$[(f'g - fg')(\alpha w)]_{-\infty}^{\infty} = 0 \tag{11.17}$$

L'équation (11.16) définit en fait le poids w que l'on doit choisir en fonction de α et β. La condition (11.17) nous indique que selon l'espace des fonctions que nous avons choisi, nous devons exiger de $w(x)$ de tendre suffisamment rapidement vers 0 quand $x \to \infty$. Remarquez que $w = 0$ est solution de l'équation (11.16), donc nous pouvons éventuellement connecter une région où $w > 0$ à une région où $w = 0$ et satisfaire ainsi la condition (11.17). Remarquer qu'une fois w choisie convenablement, nous pouvons

13. Il faut que $w(x) > 0$ sur un ensemble de mesure non-nulle, par exemple sur un intervalle
14. Nous avons le choix du poids $w(x)$

mettre l'équation différentielle $Lu = \lambda u$ sous la forme de

$$\frac{d}{dx}\left[(\alpha w)\frac{du}{dx}\right] + (\gamma - \lambda)wu = 0$$

C'est sous cette forme que les systèmes Sturm-Liouville ont été formulés. Notez que l'équation (11.16) est une équation différentielle simple dont la solution est donnée par

$$w(x) = \frac{C}{\alpha(x)} \exp\left(\int \frac{\beta}{\alpha} dx\right) \tag{11.18}$$

où par $\int$ nous dénotons une primitive.

§ 11.12 Démontrer qu'utiliser un poids $w(x)$ pour le produit vectoriel revient à faire un changement de variable $x = \phi(s)$ dans un produit scalaire avec un poids 1.

11.5 Les solutions polynomiales de Sturm-Liouville.

Nous allons à partir de maintenant nous restreindre aux solutions polynomiales des systèmes Sturm-Liouville

$$LP_n = \lambda_n P_n \tag{11.19}$$

où $P_n(x)$ est un polynôme de degré n. Si nous arrivons à trouver de telles solutions pour un système donné, nous aurons la garantie que ces fonctions $P_n$constituent bien une base de notre espace des fonctions, pourvu que tous les $n \in \mathbb{N}$ soient représentés [15]. La recherche des solutions polynomiales impose une forte contrainte pour satisfaire la condition d'hermicité (11.17) : $(w\alpha)$ doit tendre très rapidement vers 0 quand $x \to \infty$, ou doit être $= 0$ en dehors d'un certain intervalle (à charge pour nous de connecter les régions $w = 0$ et les régions $w > 0$). Nous allons voir plusieurs exemples qui éclaireront cela.

Comme L ne contient au plus que des dérivées d'ordre 2, cela nous impose que les coefficients α, β, γ soient eux même des polynômes, de degrés 2 pour α, 1 pour β et 0 pour γ. (Exercice : le démontrer en considérant les trois premiers polynômes $n = 0, 1, 2$). Nous devons donc considérer tous les systèmes Sturm-Liouville où [16]

$$\begin{aligned} \alpha(x) &= \alpha_2 x^2 + \alpha_1 x^1 + \alpha_0 \\ \beta(x) &= \beta_1 x + \beta_0 \\ \gamma(x) &= \gamma_0 \end{aligned}$$

Nous devons donc explorer un espace à six dimensions pour épuiser toutes les combinaisons des coefficients. Pas tout à fait en réalité, cela peut être beaucoup plus simple.

15. Cela s'appelle la complétude dont nous avions parlé au chapitre sur l'analyse fonctionnelle et qui découle d'un théorème démontré par Weierstrass que nous n'avons pas donné ici.

16. Vous remarquerez donc que les fonctions de Bessel ne sont pas polynomiales.

Prenons l'équation algébrique élémentaire $ax^2 + bx + c = 0$. Au lieu de résoudre directement cette équation, nous pouvons remplacer x par $x + x_0$ et résoudre

$$ax^2 + (2ax_0 + b)x + (ax_0^2 + bx_0 + c) = 0$$

pour une valeur arbitraire de x_0. Une fois cette équation résolue, nous pouvons toujours remplacer x par $x - x_0$ pour retrouver la solution de l'équation originale. Si nous choisissons judicieusement $x_0 = -b/2a$, l'équation se transforme en $ax^2 + c' = 0$. On peut encore simplifier la forme de cette équation (en remplaçant par exemple x par $x/\sqrt{a}$) pour obtenir l'équation

$$x^2 + c' = 0$$

Si nous savons résoudre cette dernière nous savons résoudre l'équation général de second degrés : au lieu d'explorer une famille à trois paramètres, nous n'avons qu'à explorer une famille à un seul paramètre.

Nous avons pu effectuer cette simplification formelle et éliminer deux paramètres parce que nous avions 2 degrés de libertés à notre disposition : le choix de l'origine et de l'échelle de l'axe x. Dans une équation de type (11.19), nous avons 4 degrés de libertés. Par exemple, au lieu de considérer l'opérateur L, nous pouvons étudier l'opérateur aL, où a est un scalaire. Les fonctions propres de cette nouvelle équation restent inchangées, les valeurs propres deviennent $a\lambda$. Cela nous donnera par exemple la possibilité de toujours choisir $\alpha_2 = -1$ (quand $\neq 0$). Nous pouvons également étudier l'opérateur $L + a$: a nouveau, les fonctions propres restent inchangées, les valeurs propres seront déplacées de a. Nous utiliserons cela pour toujours choisir $\gamma_0 = 0$. Nous avons également toujours la liberté de choisir l'origine et l'échelle de l'axe x. Nous aurons donc au plus deux paramètres libres.

Exercices.

§ 11.13 Comment se transforme l'équation de Jacobi

$$(1 - x^2)u'' + [p - q + (p + q + 2)x]u' = \lambda u$$

avec le changement de variable $x \to 1 - 2x$?

§ 11.14 Démontrer que nous ne pouvons pas satisfaire les conditions d'hermicité si $\beta = 0$.

Nous allons maintenant considérer tous les cas possibles.

11.5.1 $\alpha(x)$ quadratique.

Comme nous l'avons dit, nous utilisons un degrés de liberté pour choisir $\alpha_2 = -1$. Nous pouvons donc écrire $\alpha(x) = (-x_0+x)(x_1-x)$ où x_0 et x_1 sont les deux racines de $\alpha(x)$. Ces deux racines sont soit complexes (et conjuguées l'une de l'autre) soit réelles. Le lecteur peut démontrer qu'avec des racines complexes, il n'est pas possible de remplir les conditions d'hermicité. Nous utilisons maintenant deux degrés de liberté dans le chois de l'origine de l'axe x et son échelle pour choisir $x_0 = -1$ et $x_1 = 1$.

§ 11.15 Trouver le changement de variable $y = a(x - b)$ qui transforme effectivement l'opérateur $(-x_0 + x)(x_1 - x)(d^2/dx^2)$ en opérateur $A(1 - y^2)(d^2/dy^2)$

Il nous reste deux paramètres et il est de coutume de les noter $\beta_1 = -(p + q + 2)$ et $\beta_0 = q - p$. La décomposition en fraction simple de β/α s'écrit alors

$$\frac{\beta(x)}{\alpha(x)} = \frac{q+1}{x+1} - \frac{p+1}{1-x}$$

ce qui nous amène finalement, d'après l'éq.(11.18), à

$$w(x) = (1 + x)^q(1 - x)^p$$

N'oublions pas que nous exigeons $w\alpha \to 0$ "très rapidement" quand $x \to \infty$. On peut prendre p, q aussi négative que l'on veut, on ne peut pas satisfaire cette exigence telle quelle. Par contre, si $p, q > -1$, nous voyons que $w\alpha$ s'annule sur les bords. Nous choisissons donc

$$\begin{aligned} w(x) &= (1 + x)^q(1 - x)^p \quad \text{si } |x|<1 \\ &= 0 \quad \text{sinon} \end{aligned}$$

Les polynômes associés à ce choix des paramètres sont appelés des polynômes de Jacobi $J_n^{(p,q)}(x)$. Certains choix de p, q ont obtenus des noms propres. Par exemple, $p = q = -1/2$ nous donne une fonction $w(x) = (1 - x^2)^{1/2}$ et les solutions s'appellent les polynômes de Tchebychev [17]. Pour $p = q = 0$, nous obtenons les polynômes de Legendre, avec le poids $w(x) = 1$.

11.5.2 $\alpha(x)$ linéaire.

Nous considérons le cas $\alpha_2 = 0$. Cette fois, nous n'avons qu'un seul paramètre libre. Nous choisissons $\alpha(x) = x$; $\beta(x) = -x + s + 1$.Cela détermine

$$w(x) = x^s e^{-x}$$

Cela nous convient parfaitement pour $x \to +\infty$, mais pas pour $x \to -\infty$. Cependant, ont peut remarquer que si $s > -1$, nous pouvons connecter la fonction ci-dessus pour $x > 0$ à la fonction $w(x) = 0$ pour $x < 0$. Les solutions de ce système sont appelés les polynômes de Laguerre associés $L_n^{(s)}(x)$. Pour $s = 0$ les solutions s'appellent simplement les polynômes de Laguerre et nous les rencontrerons pour la résolution de l'atome d'hydrogène.

17. Les polynômes de Tchebychev ont été trouvé par l'auteur du même nom (professeur à l'université de Saint-Pétersbourg dans les années 1850-80) à l'occasion de l'étude de la théorie d'interpolation : Soit une fonction donnée sur l'intervalle $[-1, 1]$. On cherche un polynôme de degrés n $T_n(x)$ qui coïncide avec $f(x)$ en n points et qui minimise l'écart ($max_{x\in[-1,1]}|f(x) - T_n(x)|$) avec ce dernier. Les solutions sont nos polynômes.

11.5.3 $\alpha(x)$ constant.

Nous n'avons plus de paramètres libres et tous les cas se ramène au choix $\alpha(x) = 1$ $\beta(x) = -2x$ (le coefficient de x négatif ne pourra pas satisfaire l'hermicité) et nous obtenons

$$w(x) = e^{-x^2}$$

Les solutions sont appelées les polynômes d'Hermite [18]. Nous les rencontrerons très fréquemment dans les problèmes d'oscillateurs harmoniques.

11.6 Valeurs et fonctions propres.

Les valeurs propres des systèmes Sturm-Liouville prennent une forme particulièrement simple. Pour les déterminer, il suffit de regrouper, dans un polynôme d'ordre n, le terme d'ordre n. Nous trouvons alors que

$$\lambda_n = \alpha_2 n(n-1) + \beta_1 n = n\left(\alpha_2 n + \beta_1 - \alpha_2\right)$$

Pour les Jacobi, l'équation différentielle est

$$(1-x^2)u'' - \left((p+q+2)x - (p-q)\right)u' + n(n+p+q+1)u = 0$$

Pour les Tchebychev, les valeurs propres sont de la forme $\lambda_n = -n^2$ et pour les les Legendre de la forme $\lambda_n = -n(n+1)$.

L'équation de Laguerre s'écrit

$$xu'' - \left(x - (s+1)\right)u' + nu = 0$$

avec les valeurs propres $\lambda_n = -n$.

L'équation d'Hermite est

$$u'' - 2xu' + 2nu = 0$$

et les valeurs propres sont $\lambda_n = -2n$.

Notez que les valeurs propres sont espacées linéairement dans les deux derniers cas et quadratiquement dans le premier.

On peut calculer de différentes façons les polynômes orthogonaux. On peut poser $f_n(x) = \sum_{i=0}^{n} a_i x^i$, remplacer l'expression dans l'équation différentielle du polynôme et déduire les coefficients explicitement. Par exemple,

$$\begin{aligned} P_n(x) &= \sum_{i=0}^{[n/2]} \left(\begin{array}{c} n \\ m \end{array} \right) \left(\begin{array}{c} 2n-2m \\ n \end{array} \right) x^{n-2m} \\ H_n(x) &= \sum_{i=0}^{[n/2]} \frac{1}{m!(n-2m)!} (2x)^{n-2m} \end{aligned}$$

18. Professeur à l'Ecole Polytechnique et à l'ENS dans les années $1850 - 1880$.

où P_n et H_n sont les polynômes de Legendre et d'Hermite. De façon plus générale, la formule de Rodrigues donne les solutions polynomiales des systèmes Sturm-Liouville :

$$f_n(x) = C_n \frac{1}{w} \frac{d^n}{dx^n} (\alpha^n w)$$

Nous omettons la démonstration de cette formule [19], mais un des sous produit de cette démonstrations nous fournit la norme des polynômes orthogonaux :

$$\int_{-\infty}^{+\infty} w(x) f_n^2(x) dx = (-1)^n n! a_n \int_{-\infty}^{+\infty} \alpha^n(x) w(x) dx$$

Comme nous l'avons souvent indiqué, les polynômes orthogonaux sont définis à un coefficient multiplicatif près, et on peut utiliser l'expression ci-dessus pour les normaliser.

11.7 La seconde solution : Le Wronskien.

Nous savons qu'une équation différentielle de second ordre homogène possède *deux* solutions indépendante. Or, dans ce que nous avons discuté plus haut, nous n'avons pris en compte qu'une seule de ces solutions. Cependant, trouver la deuxième solution n'est pas trop compliqué.

Soit l'équation différentielle

$$\alpha(x)u'' + \beta(x)u' + \gamma(x)u = 0 \tag{11.20}$$

et notons ses deux solutions indépendante $u_1(x)$ et $u_2(x)$. Supposons que nous connaissons $u_1(x)$, nous allons voir que nous pouvons trouver $u_2(x)$. Considérons la fonction $w(x)$ qu'on appelle le Wronskien :

$$w(x) = u_1(x)u_2'(x) - u_2(x)u_1'(x) \tag{11.21}$$

En dérivant une fois la relation ci-dessus, en la multipliant par $\alpha(x)$ et en utilisant la relation (11.20), nous trouvons

$$\alpha w' = \beta w \tag{11.22}$$

Ceci est une équation de premier ordre que nous savons (en théorie) résoudre exactement. Une fois connue $w(x)$, reconsidérons la relation (11.21). Nous voyons que c'est une équation différentielle de premier ordre en $u_2(x)$, puisque nous connaissons $u_1(x)$ et $w(x)$. Nous pouvons donc trouver $u_2(x)$.

19. Il suffit d'abord de démontrer que f_n est un polynôme de degrés n et de démontrer ensuite que f_n et f_m sont orthogonale avec le poids w. Comme nous savons que les systèmes de polynômes orthogonaux avec un poids w sont unique (à un coefficient multiplicatif près, voir un des exercices plus bas) nous tenons notre démonstration.

11.8 Les solutions non-polynomiales.

La fonction hypergéométrique, les fonctions de Bessel.

11.9 Exercices.

§ 11.16 Calculer la fonction poids pour les équations suivantes ; mettre ces équations sous forme SL :

1. Équation de Legendre $(1-x^2)y'' - 2xy' + n(n+1)y = 0$
2. Équation de Mathieu $(1-x^2)y'' - xy' + n(n+1)y = 0$
3. Équation de Bessel $x^2y'' + xy' + (k^2x^2 - m^2)y = 0$
4. Équation de Bessel $y'' + (1/x)y' + (k^2 - m^2/x^2)y = 0$
5. Équation de Bessel modifiée : $x^2y'' + \nu xy' + (x^2 - m^2) = 0$
6. Équation de Laguerre $xy'' + (1-x)y' + ny = 0$
7. Équation d'Hermite $y'' - 2xy' + 2ny = 0$
8. $x^3y'' + xy' + 2y = 0$

§ 11.17 Démontrer que le système suivant, où tous les coefficients sont réels,

$$(x^2 + \alpha_1 x + \alpha_0)y'' + (\beta_1 x + \beta_0)y' - \lambda y = 0$$

n'a pas de solution polynomiale si $\alpha(x) = x^2 + \alpha_1 x + \alpha_0$ a deux racines complexes [Help : démontrer d'abord que $w\alpha$ n'est pas suffisamment rapidement décroissant et ensuite qu'on ne peut pas trouver $w\alpha$ continue s'annulant en deux points, pour pouvoir ensuite le connecter à la solution $w\alpha = 0$. Pour démontrer ce dernier point, remarquer que $\alpha(x)$ peut se mettre sous la forme de $(x-a)^2 + b^2$ et que $\int dx/\alpha(x) = (1/b)Arctg\,(x-a)\,/b$].

§ 11.18 Démontrer que le système SL

$$\alpha(x)y'' + (\beta_1 x + \beta_0)y' - \lambda y = 0$$

n'a pas de solution polynomiale si $\beta_1 x + \beta_0 \equiv 0$.

§ 11.19 Soit $\{P_n\}$ et $\{Q_n\}$ deux ensembles de polynômes orthogonaux, avec le même poids $w(x)$. Démontrer alors que les deux polynômes sont proportionnel : $P_n = a_n Q_n$. Autrement dit, à un coefficient multiplicatif près, les polynômes orthogonaux associés à un poids sont uniques.

§ 11.20 Démontrer que si les polynômes $\{P_n(x)\}$ sont les fonctions propres d'un système SL avec le poids $w(x)$, alors les polynômes $\{P'_n(x)\}$ sont orthogonaux avec le poids $w(x)\alpha(x)$.

§ 11.21 Déduire de la question précédente que les polynômes ultra sphériques $G_n^m(x)$ s'obtiennent à partir des polynômes de Legendre :

$$G^m_{n-m}(x) = \frac{d^m}{dx^m} P_n(x)$$

Quelle est la relation entre les ultra sphériques et les fonctions de Legendre associés ?

§ 11.22 Les polynômes de Tchebychev $T_n(x)$ sont des polynômes de Jacobi pour $p = q = -1/2$. En écrivant l'orthogonalité des ces fonctions (et en supposant $T_n(1) = 1$) déduire que $T_n(\cos(\theta)) = \cos n\theta$.

§ 11.23 A quelle équation obéit la fonction $f_n(x) = (1/\sqrt{x})H_{2n+1}(\sqrt{x})$? Pouvez vous établir une relation entre cette fonction est un polynôme de Laguerre associé ?

§ 11.24 Soit l'équation aux valeurs propres

$$v'' - x^2 v = -\lambda v$$

D'après ce que nous avons dit, cette équation n'a pas de solution polynomiale. On peut cependant la mettre sous une meilleure forme. Poser

$$v(x) = e^{f(x)}\phi(x)$$

Et déduire une équation pour la fonction $\phi(x)$. Comment faut-il choisir la fonction $f(x)$ pour éliminer le terme en $-x^2\phi(x)$ (attention au signe choisi) ? Que devient alors l'équation au valeurs propres pour ϕ ? Comment les valeurs propres sont distribuées ? Utiliser ces résultats pour donner la solution complète de l'équation de Schrödinger de l'oscillateur harmonique

$$i\frac{\partial\psi}{\partial t} = -\frac{\partial^2\psi}{\partial x^2} + x^2\psi$$

sur l'intervalle $x \in]-\infty, +\infty[$

§ 11.25 Quelle est la relation entre la fonction hypergéométrique $F(a, b, c; x)$ solution de

$$x(1-x)u'' + [c - (a+b)x]u' - abu = 0$$

et les polynômes de Jacobi ?

§ 11.26 Démontrer que les polynômes orthogonaux obéissent à la relation de récurrence

$$f_{n+1}(x) = (A_n + B_n x)f_n(x) + C_n f_{n-1}(x)$$

où vous obtiendrez les coefficients A_n, B_n, C_n. [Help : Choisir A_n, B_n pour que $\Delta = f_{n+1} - (A_n + B_n x)f_n$ soit un polynôme de degrés $n-1$ et projeter cette différence sur la base des f_i pour $i \leq n-1$. Il faut utiliser le fait que $(f_k, x^l)_w = 0$ si $l < k$]

§ 11.27 Soit $u = \sqrt{w\alpha}y$, où y est la solution d'un système Sturm-Liouville $\alpha y'' + \beta y' = \lambda y$. Démontrer que u obéit à une équation du genre

$$u'' + Ru = 0$$

où il faut expliciter la fonction $R(x)$. Cette forme est le point de départ d'une approximation célèbre, appelée WKB.

12 Le calcul variationnel

12.1 Introduction.

Dès que vous avez vu les bases de l'analyse, vous avez appris à répondre à la question suivante : comment trouver le point x pour lequel la fonction $f(x)$ est maximum (ou minimum) ? f est une machine qui prend un nombre en entrée et produit un nombre en sortie. La question ci-dessus en réalité est celle de trouver un extremum *local* : un point qui produit la sortie la plus grande (ou la plus petite) que tous ses voisins immédiats. Nous savons que pour un tel point $x*$,

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

Donnons nous maintenant une fonctionnelle S. Ceci est une machine qui prend *une fonction* en entrée et produit un nombre en sortie. Par exemple

$$\mathcal{S}(f) = \int_a^b \left[f(x)^2 + f'^2(x)\right] dx \tag{12.1}$$

est une fonctionnelle qui prend une fonction, ajoute son carré et le carré de sa dérivée et les intègre entre deux bornes pour produire un nombre. Si on entre la fonction $\sin x$ dans cette machine, elle produit le nombre $b - a$. Si on y entre la fonction $\exp x$, elle produit le nombre $2\exp(2b) - 2\exp(2a)$.

Le calcul variationnel consiste à répondre à la question suivante : quelle est la fonction f qui produit la plus grande sortie $\mathcal{S}(f)$? La réponse que nous allons voir par la suite est que f doit satisfaire une équation différentielle qui est reliée à la forme de la fonctionnelle $\mathcal{S}$.

Donnons deux exemples avant d'aller plus loin.

Exemple 12.1 **Le Brachistochrone.** L'exemple le plus important historiquement est celui du brachistochrone. Soit un point $A(0, 0)$ situé dans le plan vertical, relié à un point $B(x_1, y_1)$ par un toboggan dont la forme est donnée par la fonction $y = f(x)$. On laisse un objet glisser sans frottement du point A le long du toboggan. Comment choisir la forme du toboggan, (la fonction f), pour que le temps d'arrivé au point B soit minimum ? Vous voyez qu'une fois qu'on se donne un toboggan, c'est à dire une fonction, en utilisant quelques notions de mécanique et de conservation d'énergie, on peut calculer le temps de parcours, c'est à dire un scalaire. Essayons de mettre cela en forme. La vitesse de l'objet à l'ordonnée y vaut $\sqrt{2gy}$. L'élément d'arc $ds =$

$\sqrt{dx^2+dy^2} = \sqrt{(1+f'(x)^2}dx$ est parcouru en un temps $dt = ds/v$. Le temps total du parcours est donc

$$T = \int_0^{x_1} \sqrt{\frac{1+f'(x)^2}{2gf(x)}}dx$$

Et il faut trouver la fonction f qui minimise cette intégrale [1]. Ce problème avait été lancé comme un défi par un des frères Bernoulli vers 1690 (à peine dix ans après l'invention du calcul différentiel) à la communauté scientifiques. Tous les grands (Newton, Leibnitz, l'autre Bernoulli, Hospital, ...) y répondirent. Euler (~1740) a trouvé une solution générale de ce genre de problème et Lagrange (~1780) les a généralisé à la mécanique.

Exemple 12.2 **La Mécanique analytique.** Prenons une particule de masse m qui quitte le point $x = 0$ au temps $t = 0$ et arrive à un autre point $x = x_1$ au temps $t = t_1$. Cette particule est soumise à un potentiel $V(x)$. Le mouvement de la particule est donné par la fonction $x(t)$. La fonction $x(t)$ qui minimise

$$S = \int_0^{t_1} \left[(m/2)\dot{x}^2(t) - V(x(t))\right] dt \tag{12.2}$$

est la trajectoire suivie par la particule. Ceci est une nouvelle formulation de la mécanique. Classiquement, nous résolvons l'équation différentielle $F = ma$ où la fonction $F(x) = -dV/dx$ et $a = d^2x/dt^2$pour remonter à la trajectoire $x(t)$. Ici, la démarche est différente : de toute les trajectoires possibles qui relie $(0,0)$ à (t_1, x_1), la particule choisit justement celle qui minimise l'intégrale (12.2). Comme si un dieu calculait le coût (qu'on appelle l'action S) de chaque trajectoire et choisissait la meilleure. Bien sûr, cette formulation de la mécanique et la formulation Newtonienne sont équivalente, bien que la formulation lagrangienne soit beaucoup plus profonde et pratique. Notez que la quantité dans l'intégrale n'est pas l'énergie totale, mais l'énergie cinétique *moins* l'énergie potentielle. On appelle cette quantité le lagrangien.

12.2 Calcul des variations.

Formulons correctement notre problème (nous n'allons pas attaquer le cas le plus général pour l'instant). Soit la fonctionnelle

$$S[f] = \int_a^b \mathcal{L}[f(t), f'(t), t]dt \tag{12.3}$$

1. A première vue, il semble qu'il manque quelque chose à cette formulation : l'intégrale ne contient pas de référence à y_1 et nous n'exigeons apparemment pas que la particule finisse sa trajectoire à l'ordonné y_1. Nous y reviendrons plus tard, quand nous aborderons les contraintes.

Trouver la fonction $f(t)$, *avec* les conditions $f(a) = y_0$ et $f(b) = y_1$ pour laquelle l'intégrale est un extremum.

Traditionnellement, la fonction $\mathcal{L}$ qui se trouve sous l'intégrale est appelée le lagrangien. C'est une fonction tout ce qu'il y a de plus normal. Par exemple, $\mathcal{L}(x, y) = x^2+y^2$. Comme à un instant donné, $f(t)$ et $f'(t)$ sont des nombres, il est tout à fait légitime de calculer $\mathcal{L}[f(t), f'(t)]$ qui dans ce cas, vaut $f(t)^2 + f'(t)^2$ comme l'expression que nous avions écrit dans (12.1). En plus, nous avons le droit de prendre les dérivées partielles de $\mathcal{L}$: par exemple, dans ce cas, $\partial\mathcal{L}/\partial x = 2x$ et $\partial\mathcal{L}/\partial y = 2y$. Il est usuel, si nous avions noté $\mathcal{L}[f(t), f'(t)]$, de noter ces dérivées partielles par $\partial\mathcal{L}/\partial f$ et $\partial\mathcal{L}/\partial f'$: cela veut juste dire "dérivée partielle par rapport au premier ou deuxième argument". Dans ce cas, nous aurions eu par exemple $\partial\mathcal{L}/\partial f' = 2f'(t)$. D'ailleurs, $\partial\mathcal{L}/\partial f'$ est ici une fonction de t et on peut par exemple prendre sa dérivée par rapport au temps : $d[\partial\mathcal{L}/\partial f']/dt = 2f''(t)$. Si au début, vous trouvez cette notation abrupte, remplacez f et f' avant les dérivations par x et y, et remettez les à leur place une fois les opérations terminées. Mais on prend vite l'habitude de ces notations. Notez également que l'on ne cherche pas n'importe quelle fonction, mais les fonctions qui prennent des valeurs bien déterminées (y_0 et y_1) aux bords a et b.

Avant d'aller plus loin, revenons un instant au cas d'une fonction $f(x)$ dont ont veut trouver l'extremum. Si nous connaissons la valeur de la fonction au point x, alors son accroissement quand on se déplace au point $x + \epsilon$ est donnée par

$$df = f(x + \epsilon) - f(x) = A(x)\epsilon + \text{termes d'ordres } \epsilon^2 + ...$$

La première partie de l'accroissement (celle qui est importante quand ϵ est petit) est linéaire en ϵ : si nous avions pris un ϵ deux fois plus grand, nous aurions eu un accroissement deux fois plus grand également. La fonction $A(x)$ est le coefficient de proportionnalité entre df et ϵ au point x, et nous avons plus l'habitude de la noter par $f'(x)$. Le point x est un extremum si le coefficient de proportionnalité $f'(x) = 0$, c'est à dire qu'en se déplaçant autour du point x, l'accroissement de f (à l'ordre 1 en ϵ) est nulle.

Nous n'avons qu'à suivre cette méthodologie pour trouver l'extremum de notre fonctionnelle $S[f]$: Nous allons ajouter la fonction $\epsilon g(t)$ à la fonction $f(t)$, et calculer l'accroissement de la fonctionnelle $dS = S[f + \epsilon g] - S[f]$. Avec un peu de chance, cette accroissement comporte un terme linéaire en ϵ :

$$dS = S[f + \epsilon g] - S[f] = A[f, g].\epsilon + \text{termes d'ordre } \epsilon^2 + ...$$

où $A[f, g]$ est un coefficient de proportionnalité qui dépend des fonctions f et g. Nous disons que f est un extremum si $A[f, g] = 0$ quelque soit g ! Cela est analogue à trouver l'extremum d'une fonction de plusieurs variables : le point $(x^*, y^*, z^*, ...)$ est un extremum de la fonction $f(x, y, z, ...)$ si en ce point, *quelque soit* le déplacement (par exemple $(dx, 0, 0, ...)$ ou $(0, dy, dz, ...)$) la partie linéaire de l'accroissement est nulle. Une fonction, comme nous l'avons vu au chapitre 1, n'est finalement qu'un point dans un espace à dimension infini ; une fonctionnelle est comme une fonction d'une

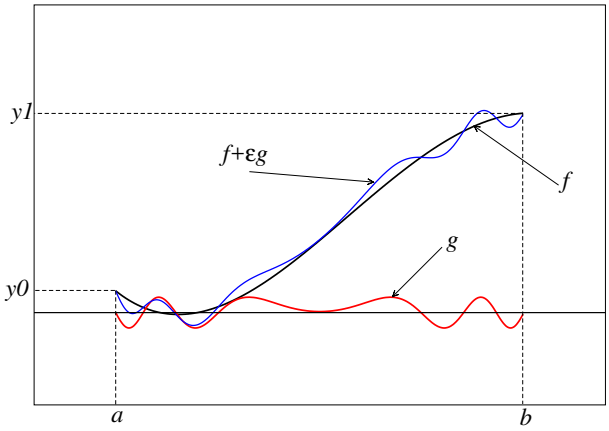


FIGURE 12.1 – Une fonction f et une variation g autour de cette fonction.

infinité de variables. Quelque soit g dans l'expression précédente veut simplement dire quelque soit le déplacement dans cet espace.

Il faut prendre une précaution : nous ne cherchons que des fonctions pour lesquelles $f(a) = y_0$ et $f(b) = y_1$. Comme f satisfait déjà à cette condition, pour que $f + \epsilon g$ la fasse également, nous ne devons considérer que des fonctions g telle que $g(a) = g(b) = 0$. Les fonctions g ne sont donc pas tout à fait quelconque. Nous obtenons :

$$\begin{aligned} S[f+\epsilon g] &= \int_a^b \mathcal{L}[f(t)+\epsilon g(t)\,,\, f'(t)+\epsilon g'(t)]dt \\ &= S[f] + \epsilon \int_a^b \frac{\partial \mathcal{L}}{\partial f} g(t)dt + \epsilon \int_a^b \frac{\partial \mathcal{L}}{\partial f'} g'(t)dt + \ldots \end{aligned} \tag{12.4}$$

où nous avons simplement utilisé le fait que $\mathcal{L}(x+h, y+k) = L(x,y) + (\partial L/\partial x)h + (\partial L/\partial y)k + \ldots$ On peut déjà voir la partie linéaire apparaître. Nous pouvons mettre la deuxième intégrale un peu plus en forme en faisant une intégration par partie :

$$\int_a^b \frac{\partial \mathcal{L}}{\partial f'} g'(t)dt = \left[\frac{\partial \mathcal{L}}{\partial f'} g(t)\right]_a^b - \int_a^b \frac{d}{dt}\left[\frac{\partial \mathcal{L}}{\partial f'}\right] g(t)dt$$

La première partie est nulle, puisque la fonction g vaut justement 0 sur les bords. En remettant ce qui reste dans l'expression (12.4), nous avons :

$$dS = \epsilon \int_a^b \left\{ \frac{\partial \mathcal{L}}{\partial f} - \frac{d}{dt}\left[\frac{\partial \mathcal{L}}{\partial f'}\right] \right\} g(t)dt + \ldots$$

L'intégrale est notre facteur de proportionnalité entre dS et ϵ. Si la fonction f est un extremum, alors l'intégrale doit être nulle *quelque soit* la fonction g. La seule possibilité est donc que f satisfasse à l'équation différentielle

$$\frac{\partial \mathcal{L}}{\partial f} - \frac{d}{dt}\left[\frac{\partial \mathcal{L}}{\partial f'}\right] = 0 \tag{12.5}$$

qui est appelé l'équation d'Euler-Lagrange. Notez que cette équation est homogène dimentionnellement. Faisons quelques exercices pour nous fixer les idées.

Identité de Beltrami. Évaluons l'expression

$$\begin{aligned}\frac{d}{dt}\left\{f'\frac{\partial \mathcal{L}}{\partial f'} - \mathcal{L}\right\} &= f''\frac{\partial \mathcal{L}}{\partial f'} + f'\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial f'}\right) - \frac{\partial \mathcal{L}}{\partial f}f' - \frac{\partial \mathcal{L}}{\partial f'}f'' - \frac{\partial \mathcal{L}}{\partial t} \\ &= f'\left(\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial f'}\right) - \frac{\partial \mathcal{L}}{\partial f}\right) - \frac{\partial \mathcal{L}}{\partial t}\end{aligned}$$

Si $\mathcal{L}$ ne dépend pas explicitement de t, alors $\partial\mathcal{L}/\partial t = 0$; par ailleurs, l'expression entre parenthèse n'est que l'équation d'Euler-Lagrange et vaut zéro par construction. Nous en déduisons que si le lagrangien ne dépend pas explicitement du temps, alors

$$f'\frac{\partial \mathcal{L}}{\partial f'} - \mathcal{L} = \text{Cte}$$

Ceci est appelé l'identité de Beltrami. En mécanique, ceci n'est rien d'autre que la conservation d'énergie (exercice : le démontrer) ; elle est cependant de portée plus générale et facilite grandement les calculs dans les problèmes où la variable indépendante n'intervient pas explicitement, comme dans le problème du brachistochrone.

Exemple 12.3 Mécanique et loi de Newton. En mécanique analytique d'un point matériel, le lagrangien est $\mathcal{L} = T - V$, où T est l'énergie cinétique est V l'énergie potentiel. Si on se restreint au cas unidimensionnel où une particule est soumis à un potentiel $V(x)$, alors pour la particule de trajectoire $x(t)$,

$$\mathcal{L}(x, \dot{x}) = (1/2)m\dot{x}^2 - V(x).$$

Le premier terme de l'équation d'Euler-Lagrange est

$$\frac{\partial \mathcal{L}}{\partial x} = -\frac{dV}{dx}$$

Comme nous l'avons mentionné ci-dessus, le seul terme qui dépend de la première variable x est $V(x)$. Remplacer mentalement la deuxième variable $\dot{x}$par y dans l'expression du lagrangien avant la dérivation si cela vous dérange. La dérivation par rapport à la deuxième variable ($\dot{x}$) donne

$$\frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x}$$

et la dérivation par rapport au temps de cette dernière nous donne

$$\frac{d}{dt}\left[\frac{\partial \mathcal{L}}{\partial \dot{x}}\right] = m\ddot{x}$$

et l'équation d'E.L. s'écrit finalement

$$m\ddot{x} + dV/dx = 0$$

Ce qui est bien sûr la même chose que l'équation de Newton $F = ma$.

Allons un peu plus loin dans le traitement de l'exemple ci-dessus. Supposons que le potentiel est constant $V(x) = Cte$, s'est à dire que la particule se meut dans une région de l'espace où il n'est pas soumis à une force. On peut également dire que cette région de l'espace possède une symétrie d'invariance par translation : deux particules identiques placées à deux endroits différents de l'espace réagiraient exactement de même ; dit encore autrement, nous n'avons aucune méthode pour déterminer où l'on se trouve dans l'espace. Dans ce cas, $\partial \mathcal{L}/\partial x = 0$, et l'équation d'E.L. s'écrit

$$\frac{d}{dt}\left[\frac{\partial \mathcal{L}}{\partial \dot{x}}\right] = 0$$

ou encore la quantité $p = \partial \mathcal{L}/\partial \dot{x} = Cte$. Or, p n'est autre chose que la quantité de mouvement $p = m\dot{x}$. Donc, la symétrie d'invariance par translation dans l'espace nous impose la conservation de la quantité du mouvement. C'est un concept extrêmement profond [2] : à chaque symétrie du lagrangien correspond une loi de conservation. La conservation de l'énergie est liée à la symétrie de translation dans le temps, la conservation du moment cinétique est associé à l'invariance par rotation autour d'un axe et ainsi de suite. En physique, une théorie n'est bien formée que si on peut la formuler sous forme variationnelle et chercher, par la symétrie sous-jacente du lagrangien, ses lois de conservation. Plus tard dans ce chapitre, nous verrons les formulations lagrangienne de l'électromagnétisme et de la mécanique quantique quand nous verrons comment traiter les champs.

12.3 Plusieurs degrés de libertés.

Parfois, souvent, le lagrangien dépend de plusieurs fonctions. Par exemple, une particule dans l'espace habituel possède trois degrés de libertés x_1, x_2, x_3 et $\mathcal{L} = \mathcal{L}(x_1, x_2, x_3, \dot{x_1}, \dot{x_2},$ Si les x_i sont les coordonnées cartésiennes, nous avons $\mathcal{L} = (m/2)\sum_i x_i^2 - V(\mathbf{r})$. La

2. Ceci s'appelle le théorème de Noether, du nom de la mathématicienne allemande qui l'a formulé vers 1912.

démarche ci-dessus peut-être répétée mot à mot pour démontrer que nous aurons une équation d'E.L. pour *chaque* degrés de liberté :

$$\frac{\partial \mathcal{L}}{\partial x_i} - \frac{d}{dt}\left[\frac{\partial \mathcal{L}}{\partial \dot{x}_i}\right] = 0 \tag{12.6}$$

L'approche lagrangienne ne dépend pas du choix de coordonnées, bien évidemment. Il suffit donc de pouvoir écrire l'énergie cinétique et potentielle pour déduire les équations du mouvement de façon automatique.

Exemple 12.4 **mouvement dans un champ central.** Il est facile de passer de l'expression de l'énergie cinétique en coordonnées cartésiennes (on pose $m = 1$) $T = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$ aux coordonnées sphériques, $T = \dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta\dot{\phi}^2$. L'énergie potentiel est $V(r)$, ce qui donne le lagrangien

$$\mathcal{L} = \dot{r}^2 + r^2\dot{\theta}^2 + r^2 \sin^2\theta\dot{\phi}^2 - V(r)$$

En écrivant E.L. pour θ :

$$4r\dot{r}\dot{\theta} + 2r^2\ddot{\theta} - 2r^2 \sin\theta\cos\theta\dot{\phi}^2 = 0 \tag{12.7}$$

Nous remarquons que nous avons une symétrie : si nous avions choisi nos axes pour que à $t = 0$, $\theta = \pi/2$ et $\dot{\theta} = 0$ (ce qui revient à choisir le plan xy défini par le rayon vecteur de la particule et sa vitesse), alors $\theta(t) = \pi/2$ vérifie trivialement l'équation (12.7) : la particule reste dans le plan xy. Le lagrangien s'écrit plus simplement donc

$$\mathcal{L} = \dot{r}^2 + r^2\dot{\phi}^2 - V(r)$$

Nous remarquons que ϕ n'intervient pas dans le lagrangien, le moment associé à cette variable se conserve donc :

$$p_\phi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = 2r^2\dot{\phi} = \text{cte}$$

Ceci n'est rien d'autre que la conservation du moment cinétique, comme nous l'avions annoncé ci-dessus. Finalement, il faut écrire l'E.L. pour r et résoudre les équations différentielles résultantes. Le lecteur trouvera la solution de ses équations dans les livres de mécanique.

Exemple 12.5 **Pendule paramétrique.** Soit un pendule de masse $m = 1$ restreint au plan xy (on suppose la gravité selon l'axe y) dont la longueur varie de façon périodique dans le temps : $\ell = \ell(t)$. Quelle est son équation de mouvement ?

Le système possède un seul degré de liberté. Nous choisissons l'angle θ que fait le pendule avec l'axe y comme ce degré. L'énergie cinétique est donnée par $T = (\dot{x}^2 + \dot{y}^2)/2$. Comme $x = \ell\sin\theta$ et $y = \ell\cos\theta$, $T = (\dot{\ell}^2 + \ell^2\dot{\theta}^2)/2$. L'énergie

potentielle est $U = gy = -g\ell\cos\theta$. Pour le premier terme de l'équation E.L., nous avons

$$\begin{aligned}\frac{d}{dt}\left[\frac{\partial\mathcal{L}}{\partial\dot{\theta}}\right] &= \frac{d}{dt}\left[\ell^2\dot{\theta}\right] \\ &= \ell^2\ddot{\theta} + 2\ell\dot{\ell}\dot{\theta}\end{aligned}$$

Comme par ailleurs, $\partial\mathcal{L}/\partial\theta = -\partial U/\partial\theta = -g\ell\sin\theta$, nous trouvons l'équation du mouvement :

$$\ddot{\theta} + 2\left(\dot{\ell}/\ell\right)\dot{\theta} + (g/\ell)\sin\theta = 0$$

Si ℓ est constante, nous retrouvons l'équation classique d'un pendule. Si maintenant ℓ oscille faiblement autour d'une position moyenne $\ell = \ell_0 + \ell_1\sin\omega t$, $\ell_0 \gg \ell_1$, on peut démontrer qu'il y a résonance si la fréquence d'excitation ω est le *double* de la fréquence propre $\Omega = \sqrt{g/\ell_0}$; la phase de l'oscillateur est alors bloquée (à 0 ou π) sur la phase de l'excitation. Ce procédé est largement utilisé dans les circuits électroniques.

Vous avez sans doute remarqué avec quelle facilité l'on obtient les équations du mouvement. En aucun cas on ne doit chercher les forces de réaction des contraintes, comme dans le cas de la formulation Newtonienne. Il existe une interprétation géométrique extrêmement profonde de cette approche qui peut-être trouvée dans les livres de mécanique analytique. Nous ne continuerons pas plus le sujet ici.

Dérivation vectorielle.

Nous avons vu ci-dessus comment dériver les équations d'Euler-Lagrange quand le lagrangien dépend de plusieurs fonctions $x_1, x_2, ...x_n$ d'une seule variable t. Très souvent cependant, les fonctions $x_i(t)$ sont les coordonnées d'un *vecteur* et il est vraiment dommage de devoir décomposer le vecteur en ses composantes. Ainsi, au lieu d'écrire $\mathbf{F} = m\mathbf{a}$, nous sommes amenés à écrire n équations du genre $f_x = m\ddot{x}$, $f_y = m\ddot{y}$, ... Nous pouvons *vectoriser* les équations E-L pour nous éviter cette gymnastique inutile et donner un sens géométrique à nos équations.

Pour cela, nous devons généraliser le concept de dérivation. Premièrement, remarquons que le lagrangien est toujours un *scalaire*. Un lagrangien avec un sens géométrique ne doit donc faire intervenir que des opérations sur les vecteurs dont le résultat est un scalaire intrinsèque, c'est à dire un scalaire dont le résultat ne dépend pas du système de coordonnées que nous avons choisi. Le meilleurs exemple d'une telle opération est le *produit scalaire*.

Prenons par exemple, dans l'espace à trois dimensions, la fonction $f(\mathbf{r}) = \mathbf{r}.\mathbf{r}$; si nous nous sommes équipées de coordonnées cartésiennes, ceci est un raccourci pour écrire

$$f(x, y, z) = x^2 + y^2 + z^2$$

Nous pouvons donc donner un sens à des expressions tel que $\partial_x f$. Pouvons nous donner un sens à l'expression $\partial f/\partial \mathbf{r}$? La réponse est évidemment oui si nous nous souvenons de la définition de la dérivée. Faisons un *petit* déplacement $\mathbf{u}$ autour du point $\mathbf{r}$ et mesurons la partie linéaire du changement dans la fonction f

$$\begin{aligned} df &= f(\mathbf{r}+\mathbf{u}) - f(\mathbf{r}) \\ &= (\mathbf{r}+\mathbf{u}).(\mathbf{r}+\mathbf{u}) - \mathbf{r}.\mathbf{r} \\ &= 2\mathbf{r}.\mathbf{u} + O(u^2) \end{aligned}$$

Nous pouvons donc effectivement écrire la partie linéaire comme

$$df = \left(\frac{\partial f}{\partial \mathbf{r}}\right).\mathbf{u}$$

où $(\partial f/\partial \mathbf{r})$ représente le vecteur $2\mathbf{r}$. Ceci est l'exact équivalent (et la généralisation) de la dérivée de la fonction d'une seule variable scalaire $f(x) = x^2$, où $f'(x) = 2x$. En analyse vectorielle, la quantité $(\partial f/\partial \mathbf{r})$ est souvent notée grad$f$ ou ∇f [3].

Nous pouvons généraliser le produit scalaire en utilisant les notations matricielle. Dans ce cas, le produit scalaire ci-dessus s'écrit $\mathbf{r}^T\mathbf{r}$ où $\mathbf{r}^T$ est un vecteur colonne associé à $\mathbf{r}$. De façon encore plus générale, nous pouvons avoir des expressions du genre $f(\mathbf{r}) = \mathbf{r}^T A\mathbf{r}$ où A est une application bilinéaire, qu'on appelle plus communément un tenseur de rang 2. Il n'est pas difficile alors de voir que

$$df = \left(\mathbf{r}^T(A + A^T)\right).\mathbf{u}$$

Si l'on note $A_s = (A + A^T)/2$ la partie symétrique de l'application A, nous avons $\partial f/\partial \mathbf{r} = 2\mathbf{r}^T A_s$. Cette formulation est développée en détail dans le chapitre consacré aux tenseurs. Notons simplement que l'application A peut être utilisée pour gérer simplement les changements de coordonnées. A 2 dimensions, en coordonnées polaires par exemple, nous avons

$$A = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$$

§ 12.1 Soit le lagrangien $\mathcal{L} = (1/2)m\dot{\mathbf{x}}.\dot{\mathbf{x}} + V(\mathbf{x})$. Démontrer que les équations d'Euler-Lagrange s'écrivent

$$m\ddot{\mathbf{x}} = -gradV$$

12.4 Formulation lagrangienne et équation du mouvement d'un champ.

Le domaine où le calcul variationnel montre toute sa puissance est celui de l'obtention d'équations d'évolution d'un champ : la hauteur d'une corde vibrante, le champ

3. Voir les chapitres sur les formes différentielles et le calcul tensoriel.

électromagnétique, le champ des contraintes élastique dans un solide, le champ des amplitudes en mécanique quantique. Peu importe le champ, il suffit de pouvoir formuler un lagrangien, c'est à dire l'énergie cinétique moins l'énergie potentielle, et le tour est joué. La différence par rapport à ce que nous avons développée ci-dessus est minime. Jusque là, nous avions considéré des lagrangiens fonctions de plusieurs x_i, chacune de ces dernières fonction *d'une* variable t. Là, nous allons considérer des lagrangiens fonction de plusieurs ϕ_i, chacune de ces dernières fonction de plusieurs variables.

Exemple fondamental. Pour fixer les idées, considérons le cas de la corde vibrante fixée à ces deux extrémités ($x = 0$ et $x = L$). A chaque instant t, la hauteur de la corde à la position x est donnée par $\phi(x, t)$. Étant donnée la forme de la corde à l'instant $t_0(\phi(x, t_0) = y_0(x)$) et $t_1(\phi(x, t_1) = y_1(x)$), quelle est l'évolution de la corde qui minimise l'action ? La trajectoire de la corde est alors la surface reliant $y_0(x)$ à $y_1(x)$ dans le temps.

A un instant donnée t, l'énergie cinétique de la corde est donnée par la somme de l'énergie cinétique de tout ses points matériels,

$$T = \int_0^L \rho \left(\partial\phi/\partial t\right)^2 dx$$

et son énergie potentielle est l'énergie élastique des déformations

$$V = \int_0^L k \left(\partial\phi/\partial x\right)^2 dx \tag{12.8}$$

où ρ est la densité de la ligne et k sa constante élastique. Notant $c^2 = k/\rho$, l'action (le *coût* d'une trajectoire $\phi(x, t)$) est

$$S[\phi] = \int_{t_0}^{t_1} \int_0^L \left[\left(\partial\phi/\partial t\right)^2 - c^2 \left(\partial\phi/\partial x\right)^2\right] dxdt$$

En considérant une variation $\epsilon g(x, t)$ autour de la trajectoire $\phi(x, t)$, nous trouvons qu'à l'ordre 1 en ϵ, la variation de l'action est

$$S[\phi + \epsilon g] - S[\phi] = \epsilon \int_{t_0}^{t_1} \int_0^L \left[\frac{\partial\phi}{\partial t}\frac{\partial g}{\partial t} - c^2 \frac{\partial\phi}{\partial x}\frac{\partial g}{\partial x}\right] dxdt$$

En faisant des intégrations par parties, nous trouvons

$$\delta S = \int_{t_0}^{t_1} \int_0^L \left[\frac{\partial^2\phi}{\partial t^2} - c^2 \frac{\partial^2\phi}{\partial x^2}\right] g dxdt$$

$\delta S = 0$ quelque soit la variation g, que si le terme entre [ ] est nul, c'est à dire

$$\frac{\partial^2\phi}{\partial t^2} - c^2 \frac{\partial^2\phi}{\partial x^2} = 0$$

qui est bien sûr l'équation d'onde bien connue.

Équation générale du mouvement de champ. Nous pouvons maintenant généraliser cette approche par des calculs aussi élémentaires que ceux ci-dessus, l'étape la plus technique étant une intégration par partie. Soit un domaine $\mathcal{D}$ dans un espace à n dimensions, où nous cherchons à trouver l'extremum de la fonctionnelle

$$S[u] = \int_D \mathcal{L}(u_{,1}, ..., u_{,n}, u, x_1, ...x_n) d^n x$$

où $u_{,i} = \partial u / \partial x_i$. La valeur de la fonction $u(x_1, ..., x_n)$ sur la frontière du domaine $\mathcal{D}$ étant fixé. L'équation d'Euler-Lagrange pour ce champ est donnée par

$$\sum_{i=1}^{n} \frac{\partial}{\partial x_i} \left(\frac{\partial \mathcal{L}}{\partial u_{,i}} \right) - \frac{\partial \mathcal{L}}{\partial u} = 0$$

Remarquez que cette expression est une simple généralisation d'EL à une dimension : l'expression de dérivation du moment $(d/dx)(\partial \mathcal{L}/\partial u')$ est remplacée par la somme de toutes les dérivées partielles.

Tenseur énergie-impulsion. Nous pouvons pousser un peu plus loin. Nous avons vu l'égalité de Beltrami pour une fonction d'une variable : quand le lagrangien ne dépend pas de la variable indépendante

$$\frac{d}{dx} \left\{ y' \frac{\partial \mathcal{L}}{\partial y'} - \mathcal{L} \right\} = 0$$

ou si nous voulions l'écrire avec nos notations sophistiquées

$$\frac{d}{dx} \left\{ y_{,x} \frac{\partial \mathcal{L}}{\partial y_{,x}} - \mathcal{L} \right\} = 0$$

ou nous avons noté $y_{,x} = y' = dy/dx$. Si $\mathcal{L}$ représente le lagrangien de la mécanique analytique, le terme entre $\{\}$ représente l'Hamiltonien, une quantité scalaire. Nous pouvons généraliser ce concept au lagrangien d'un champ $u(x_1, x_2, ...x_n)$ qui ne dépend pas explicitement des variables indépendantes. En notant [4]

$$T_{ij} = u_{,i} \left(\frac{\partial \mathcal{L}}{\partial u_{,j}} \right) - \delta_{ij} \mathcal{L}$$

nous pouvons obtenir l'équivalent de l'identité de Beltrami pour un champ :

$$\sum_{j=1}^{n} \frac{\partial T_{ij}}{\partial x_j} = 0$$

4. Rappelons que $\delta_{ij} = 0$ si $i \neq j$ et 1 sinon. Ceci est appelé le symbole de Kronecker

la quantité T , qui joue le rôle de l'Hamiltonien pour le champ, est souvent appelée en physique le tenseur énergie-impulsion. Le lecteur intéressé pourra consulter des livres de géométrie pour voir la signification générale de ce tenseur [5].

Exercices.

§ 12.2 Pour une corde vibrante dans un champ de gravitation, nous devons ajouter le terme $V_g = \int_I \rho' \phi dx$ à l'énergie potentielle (12.8) où ρ' est le produit de la densité par l'accélération de la gravité. Déduire l'équation du mouvement du champ dans ce cas. Même question si la corde se trouve dans un potentiel harmonique $V_h = \int_I \kappa \phi^2 dx$.

§ 12.3 Calculer le tenseur énergie-impulsion pour la corde vibrante sans potentiel extérieur. Que représente T_{tt} ? Que veut dire l'identité de Beltrami dans ce cas ? Que représente alors T_{tx} ?

§ 12.4 Déduire les équations du mouvements d'un champ dans un espace à $n-$dimensions, où l'expression de l'énergie potentielle (interne) est donnée par $V = \int_I k(\nabla\phi)^2 d\mathbf{r}$. Ceci généralise l'équation de la corde vibrante.

§ 12.5 Déduire l'équation du mouvement d'un champ dans un espace à $n-$dimensions anisotrope, où l'expression de l'énergie potentielle (interne) est donnée par

$$V = \int_I \sum_{i,j} a_{ij} \frac{\partial \phi}{\partial x_i} \frac{\partial \phi}{\partial x_j} d\mathbf{r}$$

où pour les coefficients, nous pouvons supposer $a_{ij} = a_{ji}$ (pourquoi ?). Ce genre d'expression se rencontre fréquemment dans des problèmes comme l'élasticité des cristaux, où les déformations dans les différentes directions ne sont pas équivalentes.

12.5 Optimisation sous contraintes.

12.5.1 Les déplacements compatibles.

Oublions pour l'instant le problème de l'extremum d'une fonctionnelle, et revenons au problème beaucoup plus simple de l'extremum d'une fonction de plusieurs variables. Soit la fonction $f(x, y)$ dont nous cherchons l'extremum. Nous savons qu'au point extremum, les termes linéaires des variations doivent être nulle

$$df = f(x+h, y+k) - f(x,y) = \left(\frac{\partial f}{\partial x}\right) h + \left(\frac{\partial f}{\partial y}\right) k + O(h^2, k^2)$$

et ceci quelque soit h, k : quelque soit la direction que l'on prend au point extremum, le terme linéaire de la variation doit être nulle.

5. Notons quand même que cela ressemble à une expression du genre $\mathbf{div} T = 0$. Classiquement, la divergence est défini pour un vecteur et est fortement associée au flux de ce vecteur à travers une surface. Il n'est pas trop difficile de donner un sens au concept du flux d'un tenseur.

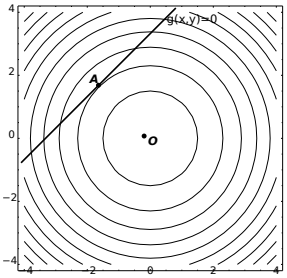


FIGURE 12.2 – La fonction $f(x, y) = x^2 + y^2$ est représentée par ses courbes de niveau. Le minimum absolu de cette fonction est le point $O = (0, 0)$. Le point le "plus bas" de la fonction qui doit également appartenir à la courbe $g(x, y) = 0$ est le point A.

Supposons maintenant que nous ajoutons une contrainte : nous ne cherchons pas l'extremum absolu de f, mais un point x^* qui satisfasse en plus à la contrainte $g(x, y) = 0$. Par exemple, nous savons que le minimum de la fonction $f(x, y) = x^2+y^2$ est le point $(0, 0)$. Mais si nous contraignons notre point à se déplacer sur la courbe $y = ax + b$, $b \neq 0$, le point $(0, 0)$ n'est pas atteignable. On peut cependant chercher un point qui minimise f avec la contrainte donnée [6]. Dans le cas simple que nous somme en train de traiter ici, on peut résoudre une coordonnée par rapport à l'autre et ramener la fonction à une seule variable (degré de liberté) : $f(x) = x^2+(ax+b)^2$ et chercher maintenant le minimum de cette nouvelle fonction (où les contraintes ont été absorbées). En général cependant, la fonction contrainte $g(x, y) = 0$ est suffisamment compliquée pour qu'on ne puisse pas résoudre une des coordonnées par rapport aux autres.

Revenons à la définition générale : la variation linéaire de la fonction autour du point extremum (x^*, y^*) doit être nulle quelque soit le déplacement (h, k) autour de ce point. En présence des contraintes, nous relâchons cette exigence : il faut que la variation linéaire autour du point (x^*, y^*) soit nulle seulement pour les déplacement (h, k) *compatibles* avec les contraintes [7]. Les déplacements compatibles avec la contrainte $g(x, y) = 0$ sont données par

$$(\partial g/\partial x)h + (\partial g/\partial y)k = 0$$

Ceci nous donne une équation supplémentaire à considérer avec l'équation

$$df = (\partial f/\partial x)h + (\partial f/\partial y)k = 0$$

Nous sommes arrivés à un système de deux équations à deux inconnus *linéaire*s. En résolvant k en fonction de h dans la première équation ; en l'injectant dans la deuxième

6. Imaginer que vous marchez en montagne sur un chemin, et vous vous intéressez au point le plus bas *sur ce chemin* et non pas en général.

7. En mécanique, cela est appelé la méthode des travaux virtuels

équation ; et en exigeant maintenant que df soit nulle quelque soit h, nous obtenons [8]

$$(\partial f/\partial x)(\partial g/\partial y) - (\partial f/\partial y)(\partial g/\partial x) = 0 \tag{12.9}$$

Ceci est une condition supplémentaire sur l'extremum de la fonction f.

Exemple 12.6 Trouver l'extremum de la fonction $f(x, y) = x^2 + y^2$ avec la contrainte $g(x, y) = -ax + y - b = 0$. En écrivant la condition (12.9), nous obtenons

$$(2x)1 - (2y)(-a) = 0$$

ou autrement dit, $x + ay = 0$. En combinant cette équation avec la contrainte $y = ax + b$, nous trouvons les coordonnées du point extremum : $x^* = -ab/(1 + a^2)$, $y^* = b/(1 + a^2)$.

12.5.2 Les multiplicateurs de Lagrange.

Lagrange a introduit une méthode équivalente à ce que nous venons de voir ci-dessus. Nous voulons optimiser $f(x, y)$ avec la contrainte $g(x, y) = 0$. Lagrange introduit une variable supplémentaire, λ et considère maintenant le problème de l'optimisation (libre) de la fonction $F(x, y, \lambda) = f(x, y) - \lambda g(x, y)$. Cela nous donne

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0 \tag{12.10}$$

$$\frac{\partial F}{\partial y} = \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0 \tag{12.11}$$

$$\frac{\partial F}{\partial \lambda} = g(x, y) = 0 \tag{12.12}$$

L'équation (12.12) n'est rien d'autre que la contrainte et assure que la solution trouvée est bien conforme. Les deux premières équations (12.10,12.11) ne sont rien d'autre que la condition (12.9) si on y élimine λ.

Exemple 12.7 Minimisons la fonction $f(x, y) = x^2 + y^2$ avec la contrainte $g(x, y) = -ax + y - b = 0$. Nous introduisons la fonction $F(x, y) = x^2 + y^2 - \lambda(-ax + y - b)$. Cela nous donne

$$\begin{aligned} 2x + \lambda a &= 0 \\ 2y - \lambda &= 0 \end{aligned}$$

8. On peut également dire que pour que le système $Ah + Bk = 0, Ch + Dk = 0$ ait une solution non trivial $h = k = 0$, il faut que le déterminant de la matrice

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

soit nulle.

Ce qui bien sûr nous donne $x + ay = 0$. Nous trouvons à nouveau le même point optimum ; de plus, si on le souhaite, on peut trouver également $\lambda^* = 2y^* = 2b/(1 + a^2)$.

Le coefficient λ est appelé un multiplicateur de Lagrange. La méthode a une signification physique profonde : supposons que nous ayons un point matériel dans le potentiel $f(x, y)$ astreint à rester sur la courbe $g(x, y)$. Au point x^*, le point (qui est stable) est soumis à une force non-nulle de la part du potentiel $\mathbf{F} = (\partial_x f, \partial_y f)$. Pour qu'il puisse rester sur la position x^*, il faut que la courbe $g(x, y)$ exerce une force de réaction sur le point qui vaut justement $\lambda^*(\partial_x g, \partial_y g)$. Si on enlevait la contrainte mais qu'on soumettait le point matériel à cette force supplémentaire, il se mettrait exactement à la même position. [9]

12.5.3 Généralisation des multiplicateurs de Lagrange.

Les multiplicateurs de Lagrange se généralisent de façon naturelle aux fonctions de n variables soumis à m contraintes. Pour optimiser la fonction $f(x_1, ...x_n)$ soumise aux contraintes $g_1(x_1, ...x_n) = 0$,...,$g_m(x_1, ...x_n) = 0$, nous considérons la nouvelle fonction F de $n + m$ variables

$$F(x_1, ...x_n; \lambda_1, ...\lambda_m) = f(x_1, ...x_n) - \lambda_1 g_1(x_1, ...x_n) - ... - \lambda_m g_m(x_1, ...x_n)$$

Nous chercherons l'extremum (sans contrainte) de la fonction F.

§ 12.6 Démontrer le cas général, en suivant le chemin déjà tracé.

La méthode des multiplicateurs de Lagrange se généralise très naturellement aux fonctionnelles. Si nous devons trouver le minimum de la fonctionnelle $S[y] = \int_I \mathcal{L}(y', y, x)dx$ avec la contrainte $\int_I g(y', y)dx = A$, il nous suffit d'introduire la fonctionnelle

$$S'[y; \lambda] = \int_I \{\mathcal{L}(y', y, x) - \lambda g(y', y)\}$$

et de chercher le minimum de cette fonctionnelle par les techniques habituelles du calcul des variations (expliquer pourquoi nous pouvons omettre la constante A) .

Exemple 12.8 Le problème isopérimétrique.
Trouver la courbe fermée de longueur donnée qui maximise sa surface.

9. Cette dualité force-contrainte a joué un grand rôle dans le développement de la mécanique, de sa formulation Newtonienne dans les années ~1680 à l'apparition du livre de mécanique analytique par Lagrange en ~1780. Lagrange était fier de n'avoir pas inclus une seule figure dans son livre. En mécanique, la méthode de Lagrange nous permet de calculer une position d'équilibre sans calculer explicitement les forces de réactions.

Supposons [10] que nous pouvons représenter la courbe par une fonction $R(\theta)$. Nous cherchons à minimiser la fonctionnelle

$$S[R] = \int_0^{2\pi} (1/2)R^2(\theta)d\theta \tag{12.13}$$

avec la contrainte

$$\int_0^{2\pi} R(\theta)d\theta = L \tag{12.14}$$

Considérons la nouvelle fonctionnelle

$$S'[R(\theta);\lambda] = \int_0^{2\pi} \left\{(1/2)R^2(\theta) - \lambda R(\theta)\right\} d\theta$$

Le terme entre $\{\}$ (que nous continuons d'appeler le lagrangien) ne contient même pas de dérivée de R ; l'équation d'Euler-Lagrange nous donne directement

$$R - \lambda = 0$$

c'est à dire $R = \lambda = Cte$. Ceci est bien un cercle. Pour trouver λ, nous utilisons la contrainte (12.14), qui nous donne $\lambda = L/2\pi$. La courbe qui minimise la surface est donc un cercle de rayon $R = L/2\pi$.

Exemple 12.9 distribution d'énergie dans un système isolé.

Nous avions déjà rencontré ce problème précédemment sous le nom du théorème H (voir problème 7.3), nous allons le revisiter sous un autre angle. Soit un système isolé ayant N particule ($N \to +\infty$) et soit $c(x)$ le nombre relatif de particules ayant une énergie dans l'intervalle $[x, x + dx[$. Comme le système est isolé, le nombre de particules et l'énergie totale sont conservés :

$$\int_0^{\infty} c(x)dx = 1 \tag{12.15}$$

$$\int_0^{\infty} xc(x)dx = T \tag{12.16}$$

où nous appelons T l'énergie moyenne par particule : $T = E_0/N$ avec E_0 l'énergie totale du système. Le théorème H de Boltzmann nous apprend que la fonction $c(x)$ doit être choisi de façon à ce que la quantité

$$H\,[c] = \int_0^{\infty} c(x).\log\left[c(x)\right] dx \tag{12.17}$$

10. Ceci est une très grosse supposition. Le lecteur a intérêt a tracer quelques courbes qui ne sont pas représentable par une fonction $R(\theta)$ pour se rendre compte de la simplification que nous assumons derrière cette supposition.

soit extremum. La quantité $-H$ est souvent appelé entropie. Avec les deux contraintes (12.15,12.16), nous devons donc chercher l'extremum de la fonctionnelle

$$H'\left[c;\lambda,\mu\right]=\int_0^\infty c(x).\log\left[c(x)\right]dx-\lambda\int_0^\infty c(x)dx-\mu\int_0^\infty xc(x)dx$$

Le calcul est assez simple dans ce cas, puisque nous n'avons pas de dérivé dans la fonctionnelle. En cherchant la valeur de la fonctionnelle pour $H'\left[c+\epsilon g;\lambda,\mu\right]$ et en nous contentant d'ordre 1 en ϵ, nous trouvons

$$H'\left[c+\epsilon g;\lambda,\mu\right]=H'\left[c;\lambda,\mu\right]+\epsilon\int_0^\infty\left(1+\log c-\lambda-\mu x\right)g(x)dx$$

la fonction $c(x)$ est un extremum de la fonctionnelle H' si les variations linéaires sont nulles $\forall g$, c'est à dire si

$$\log c=(\lambda-1)+\mu x$$

ou autrement dit $c(x)=Ae^{-\mu x}$; l'utilisation des deux contraintes nous donne $A=-\mu=1/T$. Autrement dit,

$$c(x)=(1/T)e^{-x/T}$$

12.5.4 Formulation variationnelle des systèmes Sturm-Liouville.

Rappelons qu'un système Sturm-Liouville est une équation aux valeurs propres

$$\alpha(x)y''+\beta(x)y'+\gamma(x)y=\lambda y$$

Nous avons mentionné (et insisté) que si nous trouvons un poids $w(x)$ qui rend l'opérateur hermitien, alors les solutions des systèmes SL minimisent une certaine fonctionnelle. Nous avons vu que si une telle fonction poids existe, alors elle doit obéir à l'équation $(\alpha w)'=\beta w$ et le système peut alors s'écrire sous la forme alternative

$$\frac{d}{dx}\left(w\alpha y'\right)+w\gamma y=\lambda wy$$

Remarquer que cette forme ressemble furieusement à une équation d'Euler-Lagrange. On peut faire le chemin inverse : minimiser la fonctionnelle

$$S[y]=\int_I\left(p(x)y'^2+q(x)y^2\right)dx$$

avec la contrainte

$$\int_I w(x)y^2dx=1$$

Comme on peut le voir, nous pouvons formellement identifier l'équation d'Euler-Lagrange du système ci-dessus à un système SL en posant $p = w\alpha$, $q = -w\gamma$. A regarder de plus près, la formulation variationnelle d'un système SL, si on assimile la variable x au temps, ressemble beaucoup à un oscillateur harmonique avec une masse et une constante de ressort dépendant du temps. Ceci, comme nous l'avons mentionné, est le point de départ de l'approximation WKB.

12.6 Les conditions aux bords "naturelles" 1.

Revenons à notre formulation initiale de problème d'extremum. Nous voulons trouver l'extremum de la fonctionnelle

$$S[y] = \int_a^b \mathcal{L}(y', y, x)dx$$

avec les conditions aux bords fixée

$$y(a) = y_0 \ ; \ y(b) = y_1$$

Pour cela, nous avons écrit la variation de S en fonction d'une petite perturbation $g(x)$ pour obtenir

$$\delta S = \left[\frac{\partial \mathcal{L}}{\partial y'} g\right]_a^b + \int_a^b \left\{\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial y'}\right\} g dx \tag{12.18}$$

et nous avons cherché dans quelles conditions, $\delta S = 0$ quelque soit $g(x)$. Nos conditions aux bords nous ont imposées $g(a) = g(b) = 0$, donc le premier terme est nul ; le deuxième terme nous donne les équations d'E-L.

Ceci dit, nous pouvons relâcher nos contraintes, et ne pas exiger que $y(a) = y_0$, $y(b) = y_0$. Le problème serait alors : parmi toutes les courbes entre a et b, trouver celle qui extrémise la fonctionnelle. Dans ce cas, l'annulation de δS exige toujours l'annulation de l'intégrale, qui nous donnera comme d'habitude les équation d'Euler Lagrange, et l'annulation du terme de bord. Or, cette fois, comme les bords ne sont plus fixe, nous n'avons plus $g(a) = g(b) = 0$, pour annuler les termes de surface, nous devons exiger

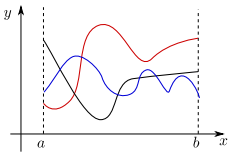


FIGURE 12.3 – Bords libres

$$\left.\frac{\partial \mathcal{L}}{\partial y'}\right|_{x=a,b} = 0$$

ce qui nous fournit deux nouvelles conditions en remplacement des conditions $y(a) = y_0$, $y(b) = y_0$. On peut bien sur mixer les conditions : fixer y en un bord et laisser y varier sur l'autre bord.

Application I : Élasticité des poutres. Pour illustrer ce principe, considérons un lagrangien qui contient des dérivées d'ordre 2 : $\mathcal{L} = \mathcal{L}(y'', y', y, x)$. Dans ce cas, la variation s'écrit

$$\delta S = \left[\frac{\partial \mathcal{L}}{\partial y'} g - \frac{d}{dx} \frac{\partial \mathcal{L}}{\partial y''} g + \frac{\partial \mathcal{L}}{\partial y''} g' \right]_a^b + \int_a^b \left\{ \frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx} \frac{\partial \mathcal{L}}{\partial y'} + \frac{d^2}{dx^2} \frac{\partial \mathcal{L}}{\partial y''} \right\} g dx$$

Pour une poutre élastique soumise à une charge $f(x)$ par exemple, l'énergie s'écrit

$$\mathcal{E} = \int_0^L \left\{ B y''^2 - f(x) y \right\} dx$$

Si la poutre est encastrée (a), nous avons les conditions $y(0) = y(L) = 0$ *et* $y'(0) = y'(L) = 0$. Dans ce cas, le terme de bord est automatiquement nulle. Par contre, si la poutre est seulement posé (b), nous avons seulement $y(0) = y(L) = 0$. Pour trouver les deux autres conditions et annuler le terme de bord, nous devons avoir $\partial \mathcal{L}/\partial y'' = 0$, c'est à dire $y''(0) = y''(L) = 0$. Nous laissons au lecteur le soin de trouver les conditions aux limites pour le cas (c).

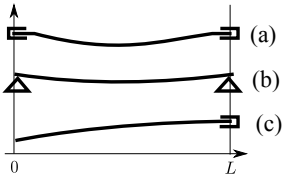


FIGURE 12.4 – poutres fixées.

Application II : angle de contact. Mettons un liquide en contact avec un support solide. Le liquide monte le long de la paroi solide, et par rapport à l'état de référence, l'énergie libre du système s'écrit :

$$F = \gamma_{lg}(\ell - L) + (\gamma_{sl} - \gamma_{sg})h$$

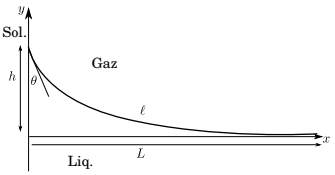
les coefficients γ_{lg},... sont les tensions de surface liquide-gaz, ... ; ℓ est la longueur de l'interface liquide-solide après la montée du liquide, L la longueur de cette même interface avant la montée, et nous supposerons ces deux valeurs très grandes ($\to \infty$). h est la longueur de l'interface solide-liquide.

Repérons la courbe du liquide par la fonction $y(x)$. En mettant le problème sous forme variationnelle, nous avons

$$F[y] = \int_0^\infty \gamma_{lg} \left(\sqrt{1 + \dot{y}^2} - 1 \right) dx + (\gamma_{sl} - \gamma_{sg}) y(0)$$

avec la contrainte de la conservation de la masse qui est

$$\int_0^\infty y dx = 0$$



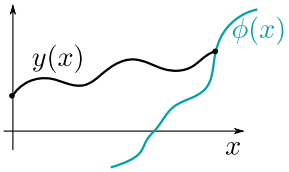


FIGURE 12.5 – Le bord droit de la courbe $y(x)$ est libre de se mouvoir le long d'une courbe $\phi(x)$.

Nous sommes en présence d'une optimisation sous contrainte, en présence de bords libre : la hauteur du liquide à $x = 0$ et $x = \infty$ n'est pas fixée. En ajoutant à $y(x)$ une variation $\epsilon g(x)$ où g ne s'annule pas sur les bords de l'intervalle, nous obtenons d'une part une équation d'Euler Lagrange comme d'habitude, et d'autre part une condition non-trivial sur le bord $x = 0$:

$$\left[\frac{\partial \mathcal{L}}{\partial \dot{y}} + (\gamma_{sl} - \gamma_{sg})\right]_{x=0} = 0$$

Or,

$$\frac{\partial \mathcal{L}}{\partial \dot{y}} = \frac{\gamma_{lg}\dot{y}}{\sqrt{1+\dot{y}^2}} = \gamma_{lg}\cos\theta$$

où θ est l'angle entre la tangente et l'axe y. On en déduit l'angle de contact solide-liquide

$$\cos\theta = \frac{\gamma_{sg} - \gamma_{sl}}{\gamma_{lg}}$$

La relation ci-dessus est connu sous le nom de la relatons d'Young. Comme la mesure de l'angle de contact est facile, on l'utilise en général pour mesurer les tensions de surface.

12.7 Les conditions aux bords naturelles 2.

Dans le paragraphe précédent, l'abscisse x était fixée sur les deux bords, y ayant la liberté de bouger le long d'une droite *verticale*. Nous pouvons généraliser encore plus en relâchant cette contrainte et en permettant aux points sur les bords de se mouvoir le long d'une courbe quelconque. Supposons, pour la simplicité, que seulement le bord droit à $x = b$ peut se mouvoir le long d'une courbe $y = \phi(x)$. L'expression (12.18) que nous avions écrit dépend bien sûr du bord, et on peut écrire la variation causé autour de $y(x)$ par

$$
\begin{aligned}
S[y,b] - S[y+\epsilon g, b+db] &= \left[\frac{\partial \mathcal{L}}{\partial y'}\epsilon g\right]_a^b + \int_a^b \left\{\frac{\partial \mathcal{L}}{\partial y} - \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial y'}\right\}\epsilon g dx \\
&+ \int_b^{b+db} \mathcal{L}(y', y, x) dx \qquad (12.19) \\
&= \left[\frac{\partial \mathcal{L}}{\partial y'}\epsilon g\right]_a^b + \mathcal{L}(y', y, x) db + \{...\} \qquad (12.20)
\end{aligned}
$$

Le terme entre crochet nous donne l'équation d'Euler-Lagrange comme d'habitude sur l'intervalle $[a, b]$. Occupons nous seulement des termes du bord. Nous devons avoir, puisque le point du bord se meut le long de la courbe $y = \phi(x)$,

$$y(b + db) + \epsilon g(b + db) = \phi(x + db)$$

Cela nous donne, à l'ordre 0 en db, $y(b) = \phi(b)$ bien sûr, et de plus, à l'ordre 1 :

$$\epsilon g(b) = (\phi'(b) - y'(b))\, db$$

Finalement, la variation de S due au bord s'écrit

$$\delta S = \left\{\frac{\partial \mathcal{L}}{\partial y'}\left(\phi'(b) - y'(b)\right) + \mathcal{L}(y', y, x)\right\} db$$

et la nullité de ce terme nous impose

$$y'(b)\frac{\partial \mathcal{L}}{\partial y'} - \mathcal{L} = \frac{\partial \mathcal{L}}{\partial y'}\phi'(b)$$

Le lecteur peut noter que le terme de gauche est souvent noté H, tandis que $\partial\mathcal{L}/\partial y'$ est souvent noté p. En mécanique analytique, on les appelles l'Hamiltonien et le moment. L'équation s'écrit donc, sur le bord,

$$H = p\phi'(b) \qquad (12.21)$$

Résumons : Si $db = 0$, nous n'avons rien à faire de plus dans l'expression (12.20), et nous avons notre condition au bord du paragraphe précédent $p = 0$. Si $db \neq 0$, alors nous devons utiliser la condition plus générale (12.21). En particulier, si $\phi'(b) = 0$ (nous n'admettons que des déplacements le long de l'axe x), alors nous devons avoir

$$H = 0$$

Exemple 12.10 **angle de contact d'une goutte posé sur un substrat solide non plane.**

Exemple 12.11 **Brachistochrone générale.**

12.8 Détour : éléments de géométries non-euclidiennes.

Le cinquième axiome d'Euclide est le suivant : "d'un point en dehors d'une droite, on ne peut tracer qu'une et une seule droite parallèle au premier". Pendant deux millénaires, les mathématiciens ont cru que ceci n'est pas un axiome, mais un théorème que l'on peut démontrer à partir des quatre premiers axiomes. A partir de 1830, il est devenu clair que le *cinquième* mérite le titre d'axiome et qu'on peut tout a fait formuler d'autres géométries qui acceptent d'autres axiomes. Nous allons étudier une version simplifiée de la géométrie riemannienne. Pour cela, nous devons donner un sens précis au mot "droite" : nous le définirons dorénavant comme le chemin le plus court entre deux points ; il est plus habituel alors d'appeler cela une géodésique. Supposons que nous avons muni notre espace bidimensionnel d'un système de coordonnées (x_1, x_2). La distance entre deux points infiniment voisin est définie par

$$ds^2 = \sum_{i,j} g_{ij} dx_i dx_j$$

où $g_{ij} = g_{ij}(x_1, x_2)$ est appelé le tenseur métrique. Par exemple, si nous avons muni le plan euclidien de coordonnées polaires, $ds^2 = dr^2 + r^2 d\theta^2$, nous avons (en posant $x_1 = r$, $x_2 = \theta$), $g_{11} = 1$, $g_{22} = x^2$, $g_{12} = g_{21} = 0$.

Nous allons considérer dans la suite le cas très particulier où $g_{11} = 1$, $g_{i\neq j} = 0$ et $g_{22} = g^2(x)$, où g est une fonction quelconque [11]. Le périmètre d'une courbe $y(x)$ reliant deux points est donnée par

$$\ell[y(x)] = \int_1^2 \sqrt{1 + g^2(x) y'^2} dx$$

et il est élémentaire de montrer, d'après les précédents paragraphes, que l'équation d'Euler-Lagrange nous donne immédiatement l'équation de la courbe :

$$y' = a/g\sqrt{g^2 - a^2}$$

où a est une constante d'intégration. Dans le cas des coordonnées polaire par exemple où $g(x) = x$, nous pouvons intégrer l'équation ci-dessus [12] et obtenir l'équation d'une droite $y = \alpha + \arccos(a/x)$ où a et α sont deux constante d'intégration. Pour vous convaincre que cela est effectivement le cas, Il suffit d'interpréter x comme r et y comme θ, faire un petit schéma et quelques manipulations d'angles .

11. Nous noterons par habitude les coordonnées (x, y) sans leur associer l'idée de coordonnées cartésiennes

12. il suffit d'effectuer le changement de variable $u = a/x$

Prenons le cas plus intéressant pour nous de $g(x) = \sin(x)$. A nouveau, l'intégration s'effectue sans difficulté [13] et nous obtenons

$$y = \psi + \arccos(\cos\alpha \cot x)$$

où ψ et α sont deux constantes d'intégration. Nous voyons par exemple que pour $\alpha = \pi/2$, nous avons une famille de droites données par $y = \text{Cte}$. Prenons la géodésique $y = 0$ et considérons le point $P = (\pi/2, \pi/2)$ en dehors de cette droite. Toutes les droites traversant ce point doivent avoir le paramètre $\psi = 0$. Il n'est pas alors difficile de voir que toutes ces droites croisent la droite $y = 0$ au point $\cot x = 1/\cos\alpha$.

Nous venons de démontrer dans ce cas que *toutes* les droites traversant P croisent une droite ne contenant pas P ; cela est très différent du cinquième axiome d'Euclide. Le cas que nous venons de traiter correspond à la géométrie sphérique : sur la sphère unité, la distance entre deux points est donnée par $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$. Mais le point de vue de Riemann est beaucoup plus fondamental que cela : ce qui caractérise l'espace et qui lui donne sa substance est la donnée du tenseur métrique. Les habitants de la surface de la sphère unité ne peuvent pas voir qu'ils sont sur une sphère. Ils peuvent par contre visionner les géodésiques (en suivant les trajets des faisceaux de lumière) et déterminer la nature de leur espace en faisant des mesures par exemple de la somme des angles d'un triangle formé par trois géodésiques. C'est exactement dans ce cadre qu'Einstein a formulé sa théorie de la gravité en 1915, où les masses confèrent de la courbure à l'espace-temps.

A ajouter.

1. Discuter la jauge dans le Lagrangien et éventuellement le théorème de Noether.

Exercices.

§ 12.7 **Surface minimum.**
Soit deux cercles concentriques de rayon à priori différents disposés l'un au dessus de l'autre à une hauteur h. Quelle est la forme de la surface d'aire minimum qui relie les deux cercles (fig. 12.6a) ?

§ 12.8 **Energie de courbure.** Comment faut-il écrire les équations d'Euler-Lagrange si le lagrangien contient des dérivées secondes ? Plus spécifiquement, supposer que le lagrangien est de la forme $\mathcal{L} = y''(x)^2 + V(y)$. Généraliser ensuite au cas $\mathcal{L} = \mathcal{L}(y'', y', y, x)$.

§ 12.9 **Brachistochrone.** Nous voulons minimiser

$$T = \int_0^{x_1} \sqrt{\frac{1+y'^2}{y}}\, dx$$

13. Il suffit de poser $u = \cot x$

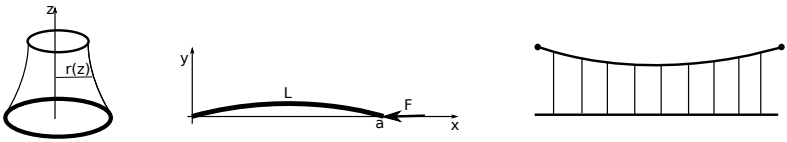


FIGURE 12.6 – (a) : surface minimum entre deux cercle ; (b) flambage d'une barre ; (c) pont suspendu à une chaînette.

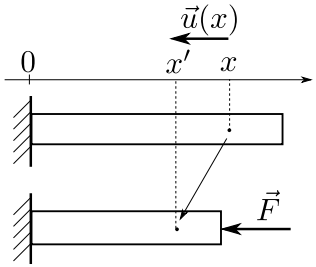


FIGURE 12.7 – Une barre élastique soumise à une force F.

En utilisant l'identité de Beltrami, démontrer que la courbe $y(x)$ doit obéir à l'équation

$$\frac{-1}{y(1+y'^2)} = C$$

où C est une constante. Résoudre cette équation du première ordre en démontrant d'abord que

$$dx = dy\sqrt{\frac{y}{2a-y}}$$

où nous avons posé $1/C^2 = 2a$; intégrer cette dernière équation en posant $y = a(1-\cos\theta)$. [Help : nous devons obtenir l'équation du cycloïde sous forme paramétrique.]

§ 12.10 **Élasticité 1-d.** Soit une barre dont on repère les points (avant déformation) par la coordonnées x (figure 12.7). On appuie sur la barre parallèlement à son axe ; les points de la barre se déplacent aux coordonnées x' (Figure 12.7). Nous appelons *déplacement* $u(x) = x' - x$ la fonction qui traque cette quantité. L'énergie élastique stockée dans la barre est proportionnelle *au carré du gradient* de ce terme :

$$\mathcal{E} = \int_0^L (1/2)ku'(x)^2 dx$$

où k est la constante élastique de la barre (qu'on appelle également le module d'Young). La force a effectué le travail $W = Fu(L)$. L'énergie totale de la barre s'écrit donc

$$\mathcal{E} = \int_0^L \left\{(1/2)ku'(x)^2 - Fu'(x)\right\} dx$$

Démontrer alors que $u(x) = ax$ où a est une constante à déterminer. Déterminer a en utilisant les conditions aux bords naturelles (section 12.6). En déduite la loi d'élasticité de Hook

$$F = Ku(L)$$

Que vaut K ? quelle est sa dépendance en L ?

§ 12.11 **Flambage d'une poutre.** Appuyez sur une règle tenue verticalement sur une table ; au delà d'une certaine force, la règle flambe (fig.12.6b). Ceci est un problème extrêmement important de la résistance des matériaux et conditionne la conception des tours pour qu'elles ne s'écroulent pas (en flambant) sous leurs propres poids. Repérons la barre par son écart à la droite $y(x)$. En supposant faible l'écart de la barre par rapport à la droite, l'énergie de courbure de la barre est donnée par sa courbure locale $B(y''(x))^2$. Nous supposons l'extrémité de la courbe maintenu à $y = 0$, mais pouvant coulisser sur l'axe x et soumise à une force F. Pour trouver la configuration qui minimise l'énergie, nous devons donc trouver l'extremum de la fonctionnelle

$$S[y, a] = \int_0^a B(y''(x))^2 dx - F(L - a)$$

soumis à la contrainte

$$\int_0^a \left\{(1 + (1/2)y'^2\right\} dx = L$$

Nous avons approximé ici l'élément de ligne $ds = \sqrt{1 + y'^2}$ par son développement de Taylor, en supposant les écarts à la ligne (et leurs dérivées) faible. Démontrez alors que pour $F > F_c$, la poutre droite n'est plus la solution optimum ; calculer F_c.

§ 12.12 **Flambage d'une poutre II.** Nous pouvons formuler différemment le problème de flambage, en intégrant directement la contrainte dans le lagrangien. Pour cela il suffit d'utiliser un système de coordonnées plus adaptés que les coordonnées cartésiennes. Repérons un point le long de la poutre par sa longueur d'arc à partir de l'origine s, et par l'angle de la tangente à la courbe en ce point avec l'axe horizontal $\theta(s)$. Ce nouveau système de coordonnées est relié aux coordonnées cartésiennes par la relation

$$\begin{aligned} dx &= \cos\theta ds \\ dy &= \sin\theta ds \end{aligned}$$

La courbure en un point est simplement donnée par $d\theta/ds$. L'énergie s'écrit alors

$$S[\theta] = \int_0^L \left\{B\left(\theta'\right)^2 + F\cos\theta\right\} ds - FL$$

que l'on peut traiter par les équation d'EL sans contraintes : la longueur d'arc s gère automatiquement la constance de la longueur totale de la poutre. Les coordonnées semi-intrinsèque (s, θ) sont très utilisées en géométrie différentielle. Noter de plus la grande similarité de l'action à celle de l'oscillation d'un pendule dans le champs de gravité.

§ 12.13 **Angle de contact.** Trouver l'angle de contact d'une goutte de liquide déposée sur une surface solide.

§ 12.14 isopérimétrique II. Chercher la courbe sous forme paramétrique $x = x(t)$, $y = y(t)$ avec la condition supplémentaire $x(1) = x(0)$ et $y(1) = y(0)$. Écrire au moins les équations d'Euler-Lagrange.

§ 12.15 Isopérimétrique III. Quelle est la courbe de surface donnée qui minimise son périmètre ?

§ 12.16 Équation de la chaînette. Une chaîne $y(x)$ est suspendu entre deux points distant de a. La longueur totale de la chaîne est L. Trouver l'équation de la chaîne. Help : A l'évidence, la chaîne doit minimiser l'énergie potentielle, avec une contrainte sur sa longueur. L'énergie potentielle est de la forme $\int \rho y ds$; on doit donc trouver le minimum de

$$H' = \int_{-a/2}^{a/2} \left\{ \rho y \sqrt{1 + y'(x)^2} - \lambda \sqrt{1 + y'(x)^2} \right\} dx$$

où λ est un multiplicateur de Lagrange. Obtenez également l'équation de la chaînette à laquelle on a suspendu un pont (fig.12.6c)

§ 12.17 Trajectoire complexe. Soit une fonction $y(x)$ complexe ($\mathbb{R} \to \mathbb{C}$) dont l'action est définie par

$$S[y] = \int_I \mathcal{L}(y, y', y^*, y^{*'}; x) dx$$

Comme par exemple $\mathcal{L} = y' y^{*'} + k y y^*$. Obtenir les équations d'Euler-Lagrange de cette fonction. [Help : il faut démontrer que l'on peut considérer y et y^* comme deux composantes indépendantes, et obtenir une équation d'EL pour chacune].

§ 12.18 Équation de Schrödinger. Nous cherchons la fonction complexe $\psi(x, t)$ qui optimise l'action associée au lagrangien suivant

$$\mathcal{L} = i\psi^* \partial_t \psi - a(\partial_x \psi^*)(\partial_x \psi) - V(x)\psi^* \psi$$

Trouver l'équation d'Euler Lagrange à laquelle obéit la fonction ψ. Dans la formulation ci-dessus, ψ et ψ^* ne jouent pas le même rôle. Pouvez vous donner une version plus symétrique de ce lagrangien ?

Problèmes.

Problème 12.1 Champ électromagnétique dans le vide.

Nous définissons un champ $A = (A_0, A_1, A_2, A_3)$ dans un espace à 4 dimensions que l'on repère à l'aide des coordonnées $x = (x_0, x_1, x_2, x_3)$ [14]. Comme nous sommes parfois trop habitué à la séparation en espace (3d) et en temps, nous dissocions parfois les expressions ci-dessus en donnant des noms différents aux différents composants : nous notons par exemple $A = (\phi, -\mathbf{A})$ où nous appelons ϕ le potentiel et $\mathbf{A}$ le potentiel vecteur ; de même, nous notons $x = (t, \mathbf{x})$ où le premier composant est appelé temps et les trois autres l'espace. Les équations d'électromagnétisme ont été formulé dans le cadre de cette séparation étrange et les différentes dérivées du

14. Nous évitons pour l'instant les exposants et notons les coordonnées x_i au lieu de x^i

champ ont reçu des noms différents. Par exemple, on appelle champ électrique le vecteur à 3 dimensions

$$\mathbf{E} = -\partial_t \mathbf{A} - \nabla\phi$$

et champ magnétique

$$\mathbf{H} = \nabla \times \mathbf{A}$$

Revenons à notre formulation générale. Le tenseur électromagnétique est défini par

$$F_{ik} = \frac{\partial A_k}{\partial x_i} - \frac{\partial A_i}{\partial x_k}$$

Ce tenseur est bien sûr anti-symétrique $F_{ik} = -F_{ki}$.

1. Donner l'expression du tenseur F en fonction des champs E_i et H_k.
2. Démontrer que pour trois indices i, j, k, la définition même du tenseur F impose

$$\frac{\partial F_{ij}}{\partial x_k} + \frac{\partial F_{jk}}{\partial x_i} + \frac{\partial F_{ki}}{\partial x_j} = 0$$

   démontrer que les seules équations non-triviales sont celles où $i \neq j \neq l$ et cela nous donne 4 équations que nous pouvons regrouper en

$$\begin{aligned} \nabla \times \mathbf{E} &= -\partial_t \mathbf{H} \\ \nabla.\mathbf{H} &= 0 \end{aligned}$$

   qui ne sont rien d'autre que les deux premières équations de Maxwell.

Dans l'espace-temps relativiste, la "distance" ds entre deux points voisins est donnée par

$$ds^2 = dx_0^2 - (dx_1^2 + dx_2^2 + dx_3^2) = \sum_{i=0}^{3} \epsilon_i dx_i^2$$

où $\epsilon_0 = 1$ et $\epsilon_{i\geq 1} = -1$. Il existe une différence entre une des composantes et les trois autres quant au signe qu'il faut utiliser pour l'élément d'arc. Cette différence apparaît obligatoirement dans toutes les expressions des lois physiques. En particulier, toutes les expressions quadratiques auront une forme similaire à l'expression de l'élément d'arc. Par exemple, l'action du champs électromagnétique est donnée par l'intégrale sur un volume du lagrangien suivant [15] :

$$\mathcal{L} = -\sum_{i,j=0}^{3} \epsilon_i \epsilon_j F_{ij}^2$$

Remarquer la similarité entre cette formulation et la formulation de la théorie de l'élasticité, où le champs A représente le vecteur déplacement et le tenseur F_{ij} le tenseur déformation. Ceci n'est pas un hasard ; Maxwell lui même imaginait l'éther comme un corps élastique et s'inspirait fortement de la théorie de l'élasticité.

1. Déduire que l'on peut mettre le lagrangien sous forme de $\mathcal{L} = \mathbf{E}^2 - \mathbf{H}^2$

15. Le signe moins n'a pas de conséquence pour nos calculs, mais assure que la solution trouvée est un minimum plutôt qu'un maximum de l'action.

2. Déduire les équations du champs

$$\sum_{j=0}^{3} \epsilon_i \epsilon_j \frac{\partial F_{ij}}{\partial x_j} = 0$$

3. Mettre ces équations sous la forme plus usuelle des deux autres équations de Maxwell :

$$\nabla \times \mathbf{H} = \partial_t \mathbf{E} \ ; \ \nabla . \mathbf{E} = 0$$

Problème 12.2 **Elasticité générale.**
Voir mon cours sur l'élasticité.

13 Les opérateurs différentiels.

La plupart des phénomènes physiques sont décrits par des équations différentielles qui impliquent des opérateurs différentielles. On rencontre souvent les gradients, rotationnelles, divergences et laplaciens et le fait que l'espace dans lequel vivent les physiciens ait trois dimensions a peut-être favorisé leurs usages au dépend d'autres formulations plus symétriques [1]. Fondamentalement, ce sont des opérateurs de dérivation et nous allons nous attacher dans ce chapitre à étudier leurs significations et à établir leurs expressions dans divers systèmes de coordonnées.

13.1 Métrique et Système de coordonnées.

Nous repérons les points dans l'espace à l'aide d'un système de coordonnées. Par exemple, pour un espace à trois dimensions, nous associons un triplet (q_1, q_2, q_3) à chaque point P de l'espace. Cet acte fondateur nous permet de ramener la géométrie, science des relations entre points de l'espace, dans le domaine de l'analyse et y appliquer toute la puissance de feu dont on y dispose.

Les points de l'espace ont une existence propre, indépendamment de la représentation en triplet que l'on utilise. Que le triplet qu'on utilise soit les coordonnées cartésiennes ou polaires ne change pas le point P ni (soulignons mentalement deux fois ce *ni*) la distance de ce point à un autre. Si la distance entre deux points est 1 mm, cela ne doit pas dépendre du système de coordonnées cartésienne ou polaire que nous avons choisi. Supposons que nous avons repéré un point P par le triplet (q_1, q_2, q_3) et le point infinitésimalement voisin $P + dP$ par le triplet $(q_1 + dq_1, q_2 + dq_2, q_3 + dq_3)$. Notons ds la *distance* entre P et $P + dP$. La relation entre la distance ds et le triplet (dq_1, dq_2, dq_3) définit la métrique du système de coordonnées. Nous nous contenterons par la suite de systèmes de coordonnées *orthogonales* (cartésien, polaire, cylindrique,...) pour lesquels, de façon général, nous avons :

$$ds^2 = h_1^2 dq_1^2 + h_2^2 dq_2^2 + h_3^2 dq_3^2 \tag{13.1}$$

Les quantités h_1, h_2, h_3 dépendent en général des coordonnées q_i. On les appelle les éléments du tenseur métrique [2]. Il est évident que si pour un certain déplacement, nous

1. Voir le chapitre sur les formes différentielles
2. Dans le cas le plus général, l'élément de distance curviligne s'écrit

$$ds^2 = \sum_{i,j} h_{i,j} dq_i dq_j$$

avons $dq_2 = dq_3 = 0$, alors $ds = h_1 dq_1$ tout simplement.

Coordonnées cartésiennes. On le note souvent par le triplet (x, y, z). C'est le plus simple des systèmes, pour lequel $h_1 = h_2 = h_3 = 1$.

Coordonnées polaires. On le note souvent (r, θ, z). Comme dans ce système,

$$ds^2 = dr^2 + r^2 d\theta^2 + dz^2$$

nous avons $h_1 = 1$, $h_2 = q_1 = r$ et $h_3 = 1$.

Coordonnées sphériques. On le note souvent (r, ϕ, θ). Ici

$$ds^2 = dr^2 + r^2 \sin^2\theta d\phi^2 + r^2 d\theta^2$$

et donc $h_1 = 1$, $h_2 = q_1 \sin q_3 = r \sin\theta$, $h_3 = q_1 = r$.

Coordonnées semi-paraboliques. On le note souvent (σ, τ, z). Ce système est relié au cartésien par les relations

$$x = \sigma\tau \,;\; y = (\tau^2 - \sigma^2)/2 \,;\; z = z$$

Démontrez que les éléments du tenseur métrique sont $h_1 = h_2 = \sqrt{\sigma^2 + \tau^2}$, $h_3 = 1$.

Coordonnées paraboliques. On le note souvent (σ, τ, ϕ), relié au système cartesien par

$$x = \sigma\tau \sin\phi \,;\; y = \sigma\tau \cos\phi \,;\; z = (1/2)(\tau^2 - \sigma^2)$$

Trouvez les éléments du tenseur métrique.

§ 13.1 En coordonnées cartésienne, la surface $q_i = Cte$ est un plan. Donnez des définitions analogues pour les autres systèmes de coordonnées.

En réalité, la démarche est la suivante : une fois que nous avons un système de coordonnées (q_1, q_2, q_3), c'est la donnée du tenseur métrique qui nous indique quel est ce système, où même plus, si l'espace est plat ou courbé (mais ceci est une autre histoire).

et la matrice H dont les éléménts sont les $h_{i,j}$ s'appelle le tenseur métrique. Dans le cas des coordonnées curvilignes orthogonales, les éléments non-diagonaux sont nulles et ds^2 peut s'écrire sous la forme plus simple de 13.1.

13.2 Nabla, div et les autres.

Les opérateurs différentielles ont tous un sens géométrique (disons même plus, physique). Ce ne sont pas juste des règles de dérivation du genre $\partial_1 E_2 - \partial_2 E_1$. Si on connaît ce sens, on peut comprendre le sens profond de l'équation qui les contient. Par ricochet, il devient très facile de déduire leurs expressions dans n'importe quel système de coordonnées. C'est ce à quoi nous allons nous attacher par la suite. Notons quand même que ces opérateurs ne sont pas si dissemblable qu'il n'y paraît et tous relient, d'une façon ou d'une autre, un flux à travers un point, un circuit ou une surface fermée à une intégrale. Dès la fin du dix-neuvième siècle, cette similitude a amené E. Cartan à inventer la notion de formes différentielles qui unifie tous ces opérateurs, qui ne sont alors que l'expression d'une opération de dérivation (qu'on appelle extérieure). Ces formes d'une élégance extraordinaire font l'objet d'un autre chapitre. Ici, nous nous attacherons à une introduction *classique* de ces opérateurs.

13.3 Le gradient.

Soit la fonction $f(P)$ qui a chaque point de l'espace associe une quantité. Le nom savant de cela est un *champ scalaire*. Cela peut être une densité , un potentiel, ... Nous sommes intéressés par savoir de combien cette fonction change si on passe du point P au point voisin $P + d\mathbf{s}$. Le gradient *est* la quantité physique qui nous donne cette information :

$$df = f(P + d\mathbf{s}) - f(P) = \mathbf{grad} f.d\mathbf{s} \tag{13.2}$$

$\mathbf{grad} f$ qu'on note également ∇f est un *vecteur* dont le produit scalaire avec le déplacement $d\mathbf{s}$ donne la variation de f. Ceci est la *définition* du gradient. ∇f à priori dépend du point P. Notez que jusque là, nous avons exprimé la variation indépendamment du système de coordonnées choisi pour repérer les points de l'espace. Une quantité physique ne doit jamais dépendre du système de coordonnées et sa définition doit toujours être donnée de façon intrinsèque, indépendemment des coordonnées. Quand en mécanique, nous écrivons $\mathbf{F} = md^2\mathbf{r}/dt^2$, ceci est une relation qui est valable quelque soit le système de coordonnées. La même chose s'applique aux opérateurs différentiels que nous utilisons en physique.

Évidement, une fois que nous avons exprimé les choses de façon intrinsèque, il faut ensuite faire le boulot et calculer la trajectoire, les lignes du champ, les isopotentiels,...Pour cela, nous devons choisir un système de coordonnées. Donc, nous avons besoin d'exprimer ∇f dans un système de coordonnée, celui qui convient le mieux au problème considéré. Supposons que le point P est repéré par (q_1, q_2, q_3) et le point voisin par $(q_1 + dq_1, q_2, q_3)$. Alors $df = f(q_1 + dq_1, q_2, q_3) - f(q_1, q_2, q_3) = (\partial f/\partial q_1)dq_1$. Le membre de droite de l'équation(13.2) vaut

$$(\nabla f)_1 h_1 dq_1$$

où $(\nabla f)_1$ est la composante du gradient dans la direction 1. Ceci nous donne $(\nabla f)_1 = (1/h_1)(\partial f/\partial q_1)$. En refaisant la même opération pour les trois coordonnées, on obtient :

$$\nabla f = \left(\frac{1}{h_1}\frac{\partial f}{\partial q_1}, \frac{1}{h_2}\frac{\partial f}{\partial q_2}, \frac{1}{h_3}\frac{\partial f}{\partial q_3} \right)$$

Exemple 13.1 En coordonnées cartésiennes, nous avons $\nabla f = (\partial f/\partial x, \partial f/\partial y, \partial f/\partial z)$. En coordonnées polaires, $\nabla f = (\partial f/\partial r, (1/r)\partial f/\partial \theta, \partial f/\partial z)$.

Exercices.

§ 13.2 Donner l'expression du gradient dans les autres systèmes de coordonnées.

§ 13.3 Écrire l'expression général du gradient dans l'espace à n dimensions.

La seule connaissance du tenseur métrique nous permet de donner l'expression du gradient dans le système de coordonnées en question. Nous allons suivre la même démarche pour tous les autres opérateurs différentiels.

Notons également que la définition (13.2) donne la direction selon laquelle le champs f varie le plus rapidement. A ds fixe, c'est la direction donnée par ∇f qui donne la variation la plus importante.

Un corollaire important de cela est que le vecteur gradient est perpendiculaire aux surfaces de niveau. Une surface de niveau est l'ensemble des points sur lesquels f est constante. Si f est une fonction de n variables, $f(x_1, ...x_n) =$ Cte est une (hyper) surface de dimension $n-1$. Un déplacement $d\mathbf{s}$ perpendiculaire au vecteur gradient ne change pas (à l'ordre 1 en $d\mathbf{s}$) la valeur de f ce qui implique que le gradient est perpendiculaire à la surface.

13.4 Champ de vecteurs.

Les quantités que l'on utilise en physique ne sont pas toutes scalaires comme la densité ρ ou le potentiel V. Certaines quantités comme le champs électrique $\mathbf{E}$ ou la vitesse d'un flot $\mathbf{J}$ sont des quantités vectorielles. Nous supposons ici connu la notion de champ de vecteur, le lecteur intéressé par plus de détails et de rigueur pourrait se reporter au cours traitant les variétés différentiables. Pour visualiser un champ de vecteur, il suffit de choisir un ensemble de points représentatifs, souvent régulièrement espacés, et de montrer par une flèche le vecteur associé à ces points.

Les lignes de champs sont faciles à imaginer. En chaque point P, la tangente à la ligne de champs est donnée par le champ en ce point. Les lignes de champs de la figure 13.1a sont des droites passant par l'origine, tandis que les lignes de champs de la figure 13.1b sont des cercles centrés sur l'origine. Le calcul des lignes de champs est élémentaire d'après ce que nous venons de dire. Supposons que nous utilisons les coordonnées (q_1, q_2, q_3) et à chaque point $\mathbf{q} = (q_1, q_2, q_3)$ de l'espace nous avons associé le vecteur

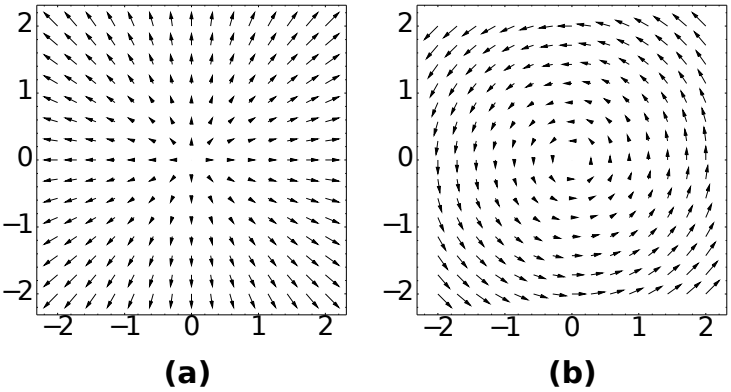


FIGURE 13.1 – Représentation des champs de vecteurs $(f_x = x, f_y = y)$ et $(f_x = -y, f_y = x)$. Une flèche en un point représente la valeur du vecteur $\mathbf{f}$ en ce point.

$(f_1(\mathbf{q}), f_2(\mathbf{q}), f_3(\mathbf{q}))$. Soit une ligne de champ que nous paramétrons par la variable t. Cela veut dire que nous définissons la ligne par trois fonctions $q_1(t), q_2(t), q_3(t)$. D'après ce que nous venons de dire, nous devons avoir

$$dq_i/dt = f_i \quad i = 1, 2, 3$$

ou encore, en regroupant les trois expressions,

$$\frac{dq_1}{f_1} = \frac{dq_2}{f_2} = \frac{dq_3}{f_3}$$

Nous avons écrit ces expressions pour un espace à trois dimensions, mais cela peut s'appliquer à n'importe quelle dimension.

Exemple 13.2 Soit, à deux dimensions, $f_1 = q_1$ et $f_2 = q_2$. Les équations des lignes de champs sont $dq_1/dt = q_1$ et $dq_2/dt = q_2$. La solution est $q_1 = \alpha q_2$, où α est une constante. Si (q_1, q_2) désigne les coordonnées cartésiennes, alors ceci est une famille de droite passant par l'origine, c'est à dire les lignes de champs de la figure 13.1a.

Exemple 13.3 Soit, à deux dimensions, $f_1 = 0$ et $f_2 = -q_1$. Les équations des lignes de champs sont $dq_1/dt = 0$ et $dq_2/dt = -q_1$. La solution est $q_1 = \alpha$, $q_2 = -\alpha t$ où α est une constante. Si (q_1, q_2) désigne les coordonnées polaires , alors ceci est une famille de cercles centrés sur l'origine, c'est à dire les lignes de champs de la figure 13.1b.

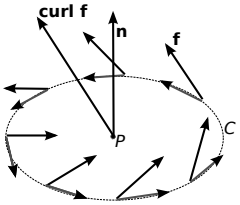


FIGURE 13.2 – Le champ $\mathbf{f}$,sa projection sur le circuit C entourant le point P, la normale à la surface $\mathbf{n}$ et le rotationnel noté $\nabla \times \mathbf{f}$ ou $\mathbf{curl}\,\mathbf{f}$.

13.5 Le rotationnel.

La distinction entre les figure 13.1a et b saute aux yeux : dans le premier, les lignes de champs ne se referment pas sur elles mêmes, dans le deuxième, toutes les lignes se referment sur elle même. Dans le premier, les lignes de champs sont comme provenant d'une source à l'origine, dans le deuxième au contraire, aucune source ne saute au yeux à priori. C'est cela que l'opérateur rotationnel, que l'on note **rot** ou parfois $\nabla\times$ (ou **curl** dans la littérature anglo-saxonne) mesure localement. Précisons les choses.

Soit un champ $\mathbf{f}$. Considérons un point P et un circuit *infinitésimal* C autour de ce point. Si la projection des lignes de champ de $\mathbf{f}$ sur C se referme, alors $I_C = \int_C \mathbf{f}.d\mathbf{s} \neq 0$. Le rotationnel est l'opérateur qui quantifie l'amplitude de I_C. Il y a cependant un petit détail à régler : la direction du circuit C a autant d'importance que sa taille. Soit $A\mathbf{n}$ le vecteur porteur du circuit C, A étant l'aire de la surface enclose par C et $\mathbf{n}$ le vecteur unitaire perpendiculaire à C, alors nous définissons $\nabla \times \mathbf{f}$ telle que

$$A\mathbf{n}.\nabla \times \mathbf{f} = \int_C \mathbf{f}.d\mathbf{s} \tag{13.3}$$

à l'ordre 1 en A (voir note [3]). Si vous n'aimez pas le travail avec les éléments infinitésimaux (quoiqu'ils aient une existence mathématiquement légitime depuis les années 1960), vous pouvez utilisez la définition

$$\mathbf{n}.\nabla \times \mathbf{f} = \lim_{A\to 0} \frac{1}{A} \int_C \mathbf{f}.d\mathbf{s} \tag{13.4}$$

où A est la surface entourée par la courbe C.

Comme vous le savez probablement, le **rot** n'est pas un *vrai* vecteur, mais un pseudo-vecteur. En faite, on ne peut donner un sens vectoriel au rotationnel que dans l'espace à trois dimensions. Nous verrons le sens général du rotationnel dans le chapitre consacré soit aux tenseurs, soit aux formes différentielles. Nous continuerons de les traiter classiquement dans ce qui suit.

Une fois que nous avons défini le rotationnel de façon (eq.13.3), nous pouvons nous en servir pour l'écrire dans n'importe quel système de coordonnées. Soit le système (q_1, q_2, q_3) et le champs $\mathbf{f} = (f_1, f_2, f_3)$ défini dans ces coordonnées : $f_i = f_i(q_1, q_2, q_3)$. Considérons un petit circuit autour du point P, dans le plan (q_2, q_3).

Sur la partie PA, notre intégrale vaut (nous omettons d'écrire la variable q_1 qui ne

3. Précisons quelques notions sur les approximations. Supposons que nous pouvons approximer une fonction autour d'un point x par son développement de Taylor :

$$\Delta = f(x+h) - f(x) = f^{'}(x)h + (1/2)f^{''}(x)h^2 + ...$$

Quand on dit qu'à l'ordre 1 en h, Δ vaut $f^{'}(x)h$, cela veut dire que

$$\frac{1}{h}\lim_{h\to 0} \Delta = f^{'}(x)$$

Concrètement, cela veut dire que nous nous intéressons aux très petits h (infinitésimaux) et le premier terme de l'approximation est amplement suffisant. De façon plus formelle, nous pouvons écrire

$$\Delta = f^{'}(x)h + o(h)$$

où $o(h)$ regroupe tous les termes qui sont négligeable devant h quand $h \to 0$:

$$\frac{1}{h}\lim_{h\to 0} o(h) = 0$$

Si nous avons une idée précise des termes que l'on néglige (comme c'est le cas ici) on peut écrire

$$\Delta = f^{'}(x)h + O(h^2)$$

où $O(h^2)$ veut dire que le plus grand terme que nous avons négligé est au mieux de l'ordre de h^2 :

$$\frac{1}{h^2}\lim_{h\to 0} O(h^2) = \text{Cte} < \infty$$

Pour simplifier, par $o(h)$ il faut entendre "très petit devant h" et par $O(h)$ de l'ordre de h. Les symboles o et O sont appelés les symboles de Landau, du nom du mathématicien allemand Edmund Landau (et non du physicien soviétique Lev Landau). Ils permettent une grande rigueur et concision dans l'écriture des expressions impliquant des limites.

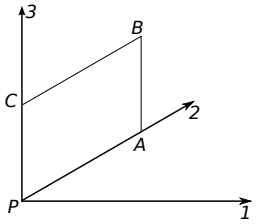


FIGURE 13.3 – un petit circuit autour du point $P = (q_2, q_3)$, où $A = (q_2 + a, q_3)$, $B = (q_2 + a, q_3 + b)$ et $C = (q_2, q_3 + b)$. Noter que le circuit est dans le plan (q_2, q_3) et perpendiculaire à l'axe q_1.

varie pas), à l'ordre 1 en a

$$f_2(q_2, q_3)h_2(q_2, q_3)a \tag{13.5}$$

et sur la partie BC

$$-f_2(q_2, q_3 + b)h_2(q_2, q_3 + b)a$$

et la somme de ces deux parties nous donne

$$-\frac{\partial}{\partial q_3}\left[h_2 f_2\right] ab$$

Par le même mécanisme, l'intégration sur la partie AB et CP nous donne

$$\frac{\partial}{\partial q_2}\left[h_3 f_3\right] ab$$

La partie gauche de l'eq.(13.3) est par ailleurs, à l'ordre le plus bas en a, b : $h_2 h_3 ab \left(\mathbf{rot}\, \mathbf{f}\right)_1$, ce qui nous donne, enfin,

$$\left(\mathbf{rot}\, \mathbf{f}\right)_1 = \frac{1}{h_2 h_3}\left[\frac{\partial(h_3 f_3)}{\partial q_2} - \frac{\partial(h_2 f_2)}{\partial q_3}\right] \tag{13.6}$$

Les autres composantes se trouvent facilement par une permutation circulaire de $(1, 2, 3)$.

Exemple 13.4 En coordonnées polaire (r, θ, z), $h_1 = h_3 = 1$, $h_2 = r$. Nous avons donc

$$\begin{aligned}
\left(\mathbf{rot}\, \mathbf{f}\right)_r &= \frac{1}{r}\left[\frac{\partial f_z}{\partial \theta} - \frac{\partial(r f_\theta)}{\partial z}\right] \\
\left(\mathbf{rot}\, \mathbf{f}\right)_\theta &= \left[\frac{\partial f_r}{\partial z} - \frac{\partial f_z}{\partial r}\right] \\
\left(\mathbf{rot}\, \mathbf{f}\right)_z &= \frac{1}{r}\left[\frac{\partial(r f_\theta)}{\partial r} - \frac{\partial f_r}{\partial \theta}\right]
\end{aligned}$$

§ 13.4 Donnez l'expression du rotationnel en coordonnées sphérique et parabolique.

Généralisation. Quand on se trouve dans l'espace à trois dimensions, nous pouvons caractériser une surface plane par le vecteur orthogonal à celle-ci. Dans le cas général, ceci n'est pas possible et il faut indexer les composants du rotationnel par deux indices désignant le plan qui contient le circuit sur lequel nous avons effectué l'intégral. Ainsi, nous aurions dû noter l'expression (13.6) en réalité

$$(\mathbf{rot}\,\mathbf{f})_{2,3} = \ldots$$

où l'indice $(2,3)$ implique que le circuit C se trouvait dans le plan x_2, x_3. A quatre dimensions par exemple, le rotationnel contient 6 composantes (et $n(n-1)/2$ à n dimensions). Dans l'espace-temps par exemple, le rotationnel d'un vecteur qu'on appelle "potentiel vecteur" possède 6 composantes : les trois où les circuits contenaient une dimension temporelle sont appelés "champ magnétique" et les trois qui ne contiennent que des dimensions spatiales sont appelé "champ électrique". [4]

13.6 La divergence.

Le travail d'un comptable est de faire le bilan des sommes dépensées et gagnées par son entreprise. C'est exactement ce travail qu'effectue l'opérateur divergence. Considérons une surface infinitésimal fermée σ autour du point P : quel est le bilan du flux d'un champs $\mathbf{f}$ à travers cette surface ? C'est ce bilan que la divergence quantifie. Plus exactement,

$$dV\,\mathbf{div}\,\mathbf{f} = \int_\sigma \mathbf{f} d\sigma$$

La démarche pour calculer la divergence est similaire à ce que nous avons fait pour le rotationnel. Considérons le flux (sortant) à travers la surface $ABCD$ (les normales aux surfaces sont par convention orientées *sortant*) :

$$bch_2(q_1+a, q_2, q_3)h_3(q_1+a, q_2, q_3)f_1(q_1+a, q_2, q_3)$$

et le flux (entrant) à travers la surface $PB'C'D'$

$$-bch_2(q_1, q_2, q_3)h_3(q_1, q_2, q_3)f_1(q_1, q_2, q_3)$$

Le bilan de ces deux termes nous donne

$$\frac{\partial(h_2h_3f_1)}{\partial q_1}abc$$

En considérant le flux à travers les quatre autres surfaces, et en notant que $dV = h_1h_2h_3abc$, on obtient finalement

$$\mathbf{div}\,\mathbf{f} = \frac{1}{h_1h_2h_3}\left(\frac{\partial(h_2h_3f_1)}{\partial q_1} + \frac{\partial(h_3h_1f_2)}{\partial q_2} + \frac{\partial(h_1h_2f_3)}{\partial q_3}\right)$$

4. Le rotationnel dans ce cas est appelé "tenseur électromagnétique". Nous référons le lecteur à un livre avancé en électromagnétisme pour voir cela en détail.

Exemple 13.5 En coordonnées polaire (r, θ, z), $h_1 = h_3 = 1$, $h_2 = r$. Nous avons donc

$$\mathbf{div}\,\mathbf{f} = \frac{1}{r}\left(\frac{\partial(rf_r)}{\partial r} + \frac{\partial f_\theta}{\partial \theta} + \frac{\partial(rf_z)}{\partial z}\right)$$

Noter que comme $\partial_z(rf_z) = r\partial_z f_z$, l'expression ci-dessus peut encore se simplifier quelques peu.

Exercices.

§ 13.5 Donnez l'expression de la divergence en coordonnées sphérique et parabolique.

§ 13.6 Comment pourrait on généraliser la divergence pour les espaces à n dimensions ?

Prenons le cas d'un fluide de densité ρ et de champs de vitesse $\mathbf{v}$. La densité de courant (le débit de la masse) vaut en chaque point $\rho\mathbf{v}$.Si le fluide est incompressible (pensez à l'eau), alors $\mathbf{div}\rho\mathbf{v} = 0$. Si au contraire le fluide est compressible, la différence entre le flux entrant et sortant dans un petit volume provoque une accumulation de la masse en ce point, d'où le sens de l'équation

$$\frac{\partial \rho}{\partial t} + \mathbf{div}\rho\mathbf{v} = 0$$

Ceci est également le sens des équations de Maxwell en électromagnétisme, sauf que là, on ne considère pas le flux d'un vecteur mais d'un objet un peu plus complexe qu'on appelle le tenseur électromagnétique.

13.7 Le Laplacien.

Le laplacien d'un champ scalaire est défini en terme des autres opérateurs que nous venons de voir :

$$\Delta f = \mathbf{div}(\mathbf{grad} f)$$

et d'après ce que nous avons dit, s'exprime simplement en coordonnées curviligne comme

$$\Delta f = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial q_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial f}{\partial q_1}\right) + \frac{\partial}{\partial q_2}\left(\frac{h_3 h_1}{h_2}\frac{\partial f}{\partial q_2}\right) + \frac{\partial}{\partial q_3}\left(\frac{h_1 h_2}{h_3}\frac{\partial f}{\partial q_3}\right)\right]$$

Exemple 13.6 En coordonnées polaire (r, θ, z), $h_1 = h_3 = 1$, $h_2 = r$. Nous avons donc

$$\Delta f = \frac{1}{r}\left[\partial_r(r\partial_r f) + (1/r)\partial_\theta^2 f + r\partial_z^2 f\right]$$

qui prend une forme plus simple si l'on fait entrer le facteur $(1/r)$ à l'intérieur du $[\,]$.

La signification du laplacien est l'écart à la moyenne. En un point P, le laplacien mesure de combien le champ scalaire f est différent de la moyenne du champ pris sur les points voisins. Voyons cela de plus près. Prenons d'abord le cas à une dimension et supposons que nous utilisons les coordonnées cartésiennes. Autour du point P d'indice x_0, choisissons deux points distants de h est calculons l'écart à la moyenne d'une fonctions f :

$$\begin{aligned}\Delta &= \frac{1}{2}\left(f(x_0+h)+f(x_0-h)\right)-f(x_0)\\ &= \frac{1}{2}\left(f(x_0+h)-f(x_0)\right)+\frac{1}{2}\left(f(x_0-h)-f(x_0)\right)\\ &= \frac{1}{2}f''(x_0)h^2+O(h^3)\end{aligned}$$

donc, à un facteur $1/2$ près (qui dépend de la dimension de l'espace), l'écart à la moyenne est donnée par la dérivée seconde *multipliée* par la distance des points de voisinage au carré h^2. La généralisation est immédiate. Donnons nous un point P_0 et une sphère de rayon h petit autour de ce point. Calculons la moyenne de l'écart entre la valeur de la fonction au point P de la sphère et le point P_0. Le point P est repéré par le vecteur $h\mathbf{n}$, c'est à dire que $P = P_0 + h\mathbf{n}$

$$\begin{aligned}\Delta &= \frac{1}{\Sigma}\iint_\Sigma \left(f(P)-f(P_0)\right)d\Sigma\\ &= \frac{1}{\Sigma}\iint_\Sigma \left(\mathbf{grad}f\right)h\mathbf{n}d\Sigma\\ &= \frac{h}{\Sigma}\iiint_V \mathrm{div}\left(\mathbf{grad}f\right)dV\\ &= \frac{hV}{\Sigma}\mathrm{div}\left(\mathbf{grad}f\right)+O(h^3)\end{aligned}$$

La quantité $V/\Sigma = Ch$, où C est un facteur géométrique. A trois dimensions [5] par exemple, C=1/3. Nous retrouvons donc bien la signification du laplacien de la moyenne.

Par exemple, l'équation de la vibration d'une membrane élastique, dont la hauteur u est relevée en chaque point est

$$\rho\frac{\partial^2 u}{\partial t^2} = k\Delta u$$

En terme de mécanique du point, l'équation ci-dessus est juste la formule $m\gamma = F$: le terme de gauche est l'accélération verticale ; le terme de droite, l'écart à la moyenne de chaque point par rapport à ses voisins, est la force exercée sur ce point.

Une des équations très importante de la physique (électrostatique sans charge) est celle de Laplace (d'où le nom de laplacien)

$$\Delta V = 0$$

5. De façon générale, $V/S = r/d$, où d est la dimension de l'espace et r le rayon de la sphère.

TABLE 13.1 – Les opérateurs differentiels en coordonnées curviligne. Pour les opérateurs vectoriels, seul la composante selon q_1 est donnée, les autres se déduisent par permutation circulaire $(1, 2, 3)$.

Expression	application
$(\mathbf{grad} f)_1 = \frac{1}{h_1}\frac{\partial f}{\partial q_1}$	$f(b) - f(a) = \int_C \nabla f.d\mathbf{s}$
$(\mathbf{rot}\,\mathbf{f})_1 = \frac{1}{h_2 h_3}\left[\frac{\partial(h_3 f_3)}{\partial q_2} - \frac{\partial(h_2 f_2)}{\partial q_3}\right]$	$\int_C \mathbf{f}.d\mathbf{s} = \int_S \mathbf{rot f}.d\mathbf{n}$
$\mathbf{div}\,\mathbf{f} = \frac{1}{h_1 h_2 h_3}\left(\frac{\partial(h_2 h_3 f_1)}{\partial q_1} + \frac{\partial(h_3 h_1 f_2)}{\partial q_2} + \frac{\partial(h_1 h_2 f_3)}{\partial q_3}\right)$	$\int_S \mathbf{f}d\mathbf{n} = \int_V \mathbf{div f}dV$
$\Delta f = \frac{1}{h_1 h_2 h_3}\left[\frac{\partial}{\partial q_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial f}{\partial q_1}\right) + ...\right]$	

FIGURE 13.4 – Découpage d'un circuit en Nintervalles

Cela veut dire qu'en tout point x, la fonction $V(x)$ est égale à la moyenne de son voisinage. Ceci veut dire que soit la fonction est localement linéaire autour du point P, soit que les variations le long d'une direction sont compensées par des variations en sens inverse dans d'autres dimension. Prenez par exemple l'image d'un col en montagne : dans une direction, on monte, dans l'autre direction, on descend. Cela implique donc qu'une fonction obéissant à cette équation ne peut pas avoir d'extremum local nul part à l'intérieur du domaine où cette équation est valable.

13.8 Résumons.

Il est temps de mettre toutes ces expression côte à côte et voir leur ressemblance. Remarquez, dans la colonne des applications de la table (13.1), la relation entre la partie droite et gauche de chaque égalité. Dans la partie gauche, nous somme entrain de calculer quelque chose comme le *flux* d'un champ sur une courbe de 0, 1, 2 dimensions. Dans la partie droite, nous relions ce flux à l'intégrale d'un opérateur différentiel de ce champ sur une surface de 1,2,3 dimensions qui entoure la courbe précédente. Les trois relations s'appellent formule de Stockes généralisée.

Remarquez que les relations de Stockes s'obtiennent directement à partir des définitions. Prenons l'exemple du gradient, et donnons nous un circuit C commençant par le point A et finissant par le point B. Découpons ce circuit en N intervalles (bien sûr, nous pensons à N très grand, $\to +\infty$). Sur chaque intervalle, nous pouvons écrire, à l'ordre 1 en l'inverse de longueur des intervalles, et en utilisant la définition du gradient :

$$\begin{aligned} f(P_1) - f(A) &= \nabla f|_A . d\mathbf{s}_1 \\ f(P_2) - f(P_1) &= \nabla f|_{P_1} . d\mathbf{s}_2 \\ &\dots \\ f(B) - f(P_{N-1}) &= \nabla f|_{P_{N-1}} . d\mathbf{s}_N \end{aligned}$$

En sommant les deux côtés de ces égalités et en prenant la limite $N \to \infty$, nous obtenons la relation de Stockes pour le gradient. La relation de Stockes pour les deux autres opérateurs s'obtient de façon similaire.

Nous reviendrons beaucoup plus tard sur ces notions en leur donnant le caractère général qui leur sied d'abord à travers le cours sur les formes différentielles et ensuite quand nous aborderons le calcul tensoriel et les variétés différentielles. Les formes différentielles sont plus élégants, mais les physiciens sont plus habitués au calcul tensoriel. Les deux approches sont très complémentaires, des perspectives différentes de la même chose. Notons simplement qu'avec les formes différentielles, les relations de Stockes se notent, de façon très générale,

$$\int_{d\Sigma} \omega = \int_{\Sigma} d\omega$$

où $d\Sigma$ est l'hyper surface qui entoure l'hyper volume Σ, est $d\omega$ est la dérivée extérieure de ω.

Exercices.

§ 13.7 Démontrer les relations suivantes et surtout, donner leur un sens géométrique en vous inspirant des définitions

1. $\mathbf{rot}(\mathbf{grad} f) = 0$ (Help : considérer des circuits infinitésimaux sur des surfaces de niveau entourant un point P).
2. $\mathbf{div}(\mathbf{rot}\, \mathbf{f}) = 0$
3. $\mathbf{div}(\mathbf{A} \times \mathbf{B}) = -(\mathbf{rotA}).\mathbf{B} + \mathbf{A}.(\mathbf{rotB})$

§ 13.8 En utilisant les relations locales, démontrez les formules de Stockes du tableau 13.1.

§ 13.9 Nous avons défini le Laplacien d'un champ scalaire. Le Laplacien d'un champ vectoriel est défini par

$$\Delta \mathbf{f} = \mathbf{grad}(\mathbf{divf}) - \mathbf{rot}(\mathbf{rotf})$$

Exprimez le Laplacien dans les différents systèmes de coordonnées. Pouvez vous en donner un sens géométrique ?

§ 13.10 Démontrer les relations de Stockes pour le rotationnel et la divergence.

13.9 Notes.

Certaines manipulations impliquant les dq_i peuvent paraître approximatives et sans la rigueur nécessaire. Il n'en est rien. Reprenons par exemple le calcul du rotationnel avec autant de précision que souhaitable.

Quelques points à éclaircir d'avance. Si nous connaissons la fonction (très lisse, infiniment dérivable) f et ses dérivées au point $\mathbf{q} = (q_i^0)$, alors nous pouvons connaître sa valeur en un point proche, par exemple $(q_1^0 + dq_1, q_2^0, q_3^0)$:

$$f(q_1^0 + dq_1, q_2, q_3) = f(\mathbf{q}) + dq_1 \partial_1 f(\mathbf{q}) + (1/2) dq_1^2 \partial_1^2 f(\mathbf{q}) + ...$$

C'est ce que nous appelons le développement de Taylor. Maintenant, si nous connaissons la fonction f et ses dérivées au point $\mathbf{q} = (q_i^0)$, alors comment évaluer

$$\int_{q_1^0}^{q_1^0+a} f(q_1, q_2, q_3) dq_1$$

Rien de plus simple à partir du développement de Taylor. Dans l'intervalle d'intégration, nous choisissons le paramétrage $q_1 = q_1^0 + u$

$$\begin{aligned} \int_{q_1^0}^{q_1^0+a} f(q_1, q_2, q_3) dq_1 &= \int_0^a f(q_1^0 + u, q_2, q_3) du \\ &= \int_0^a \left[f(\mathbf{q}^0) + u \partial_1 f(\mathbf{q}^0) + (1/2) u^2 \partial_1^2 f(\mathbf{q}^0) + ... \right] \\ &= a f(\mathbf{q}^0) + (1/2) a^2 \partial_1 f(\mathbf{q}^0) + (1/6) a^3 \partial_1^2 f(\mathbf{q}^0) + ... \end{aligned}$$

La dernière ligne a été possible parce que les quantités $f(\mathbf{q})$, $\partial_1 f(\mathbf{q})$, ... sont juste des constantes pour l'intégrale en question : l'intégration se fait sur u ! Dans l'équation13.5, nous avons simplement écrit le premier ordre, le seul qui est pertinent quand on prend la limite $a, b \to 0$. Pour vous en persuader, il suffit de faire le calcul à l'ordre supérieur.

14 Les tenseurs en physique.

Le mot tenseur peut évoquer des objets avec beaucoup d'indices et d'exposants du genre ξ_{ijk}^{ml} qui multiplient d'autres objets de ce genre et où il faut se souvenir que certains varient de façon contravariant et d'autres de façon covariant. En faite, ce sont des objets très simples qui généralisent les matrices. Le lecteur déjà familier avec ces dernières n'aura aucun mal à manipuler les tenseurs. Comme on va le voir par la suite, les tenseurs sont partout en physique et donnent beaucoup de sens aux diverses formules.

14.1 Les tenseurs de rang 2.

Prenons d'abord le cas d'une force appliquée à une masse m. Comme maître Newton l'avait affirmé, le vecteur accélération de la particule, $\mathbf{a}$, est relié au vecteur force, $\mathbf{F}$, par la relation

$$\mathbf{a} = \frac{1}{m}\mathbf{F} \tag{14.1}$$

Dit en langage de mathématicien, il existe une application linéaire qui, appliquée au vecteur $\mathbf{F}$ produit le vecteur $\mathbf{a}$. Comme nous l'avons mainte fois souligné, linéaire veut dire que (i) si l'on double la force, l'accélération est doublée ; (ii) si la force est considérée comme la somme de deux autres forces, l'accélération produite sera la somme des accélérations produites par chacune de ces forces. Si on note par A cette application, nous avons

$$A(\lambda\mathbf{F_1} + \mu\mathbf{F_2}) = \lambda A(\mathbf{F_1}) + \mu A(\mathbf{F_2}) \qquad \lambda, \mu \in \mathbb{R}$$

C'est la définition d'une application linéaire. Comme c'est très simple *et* très fondamental, ça ne mange pas de pain de le rappeler. En physique, nous avons l'habitude d'appeler cela le principe de superposition.

Notre application linéaire dans ce cas était vraiment triviale : prendre le vecteur $\mathbf{F}$ et le multiplier par un scalaire $1/m$. Prenons le cas maintenant d'une onde électromagnétique qui rentre dans un matériau diélectrique (non métallique). Localement, le matériau devient polarisé, et le vecteur polarisation $\mathbf{P}$ est relié au champs électrique $\mathbf{E}$ par

$$\mathbf{P} = \chi\mathbf{E} \tag{14.2}$$

La polarisation est bien sûr reliée linéairement au champ électrique : si on double le champs, la polarisation est doublé etc. Mais $\mathbf{P}$ n'a aucune raison d'être *parallèle* à $\mathbf{E}$! Le matériau, s'il est cristallin, possède des axes privilégiés. Il est plus polarisable dans

certaines directions et la composante de $\mathbf{E}$ parallèle à ces directions est plus amplifiée. Le vecteur polarisation résultant à priori n'est donc pas parallèle au champ électrique. La susceptibilité électrique (c'est son nom) χ n'est donc pas un scalaire, mais un tenseur (une application linéaire) [1]. Quand on veut manipuler et mesurer ces choses, il faut bien les représenter par des nombres. Comme nous l'avons dit, nous nous donnons alors une base, est nous représentons un vecteur par ses composantes E_i ou P_i. La représentation de χ est une matrice, dont les éléments sont χ_{ij} et numériquement, la relation (14.2) s'écrit

$$P_i = \sum_j \chi_{ij} E_j \tag{14.3}$$

L'expression ci-dessus est juste le produit de la matrice χ par le vecteur $\mathbf{P}$ écrit de façon explicite.

Pour le physicien, la susceptibilité est une propriété fondamentale du matériau, au même titre que sa masse. De même, les vecteurs $\mathbf{E}$ et $\mathbf{P}$ ont une existence propre, indépendemment de comment nous les mesurons. Ceci veut dire que dans différentes bases, les *composantes* des vecteurs, E_i et P_i auront différentes valeurs, mais si on les manipule correctement, on retrouvera les même vecteurs originaux. De même pour les éléments χ_{ij} : quelque soit la base que nous avons choisi [2], en suivant la relation (14.3), on doit toujours retrouver le même vecteurs polarisation. Cela va de soit si nous savons que ces nombres ne sont que des représentations des vecteurs et applications linéaires, et que nous disposons des mécanismes précis pour les calculer, une fois donnée une base. Cela allait un peu moins de soi au début du vingtième siècle quand l'algèbre linéaire n'était pas aussi démocratisé que de nos jours. On présentait alors un tenseur χ_{ij} comme une collection de nombres avec des règles de transformations précises lors des changements de base. C'était un peu comme faire de l'arithmétique avec des chiffres latins (quel est le résultat de MCLLXIV+MLCLXII ?) et malheureusement, cette conception des tenseurs reste encore vivante de nos jours.

Nous reviendrons plus tard à ces règles de manipulation pratique des chiffres. Pour l'instant nous allons nous habituer un peu plus à ces concepts. A propos, le titre de cette section était "les tenseurs de rang 2". C'est un terme savant pour désigner les applications linéaires. Un scalaire est un tenseur de rang 0 (zéro) et une application linéaire est un tenseur de rang 2 (il faut deux indices pour énumérer les éléments). A votre avis, qu'est ce qu'un tenseur de rang 1 ?

1. Dans le temps, on faisait beaucoup de distinction entre un scalaire, un vecteur et un tenseur. Bien sûr, on peut voir un scalaire comme une matrice diagonale avec tous ses éléments égaux : le produit d'une application identité par un nombre.

2. Par exemple, base 1 : x horizontal et selon le rayon laser qui rentre dans le cristal, z vertical et y perpendiculaire aux deux autres ; base 2 : x,y,z parallèle aux axes principaux du cristal ; base 3 : x dans la direction de l'étoile polaire, y pointant vers le soleil et z vers le centre de la galaxie. Les deux premières bases sont couramment utilisées par l'opticien : la première est le référentiel du laboratoire très naturel à utiliser ; Dans la deuxième, la matrice de la susceptibilité est diagonale (au dit autrement, les axes du cristal constituent les vecteurs propres de χ) et donc les calculs y sont très simple. La base 3 n'a jamais été utilisé à ma connaissance pour faire de l'optique.

Prenons maintenant un autre exemple, une fonction vectorielle dans l'espace $\mathbb{R}^3$, $\mathbf{u}(x,y,z) = (u_1(x,y,z), u_2(x,y,z), u_3(x,y,z))$ [3]. Si nous connaissons la valeur de la fonction au point (x_0, y_0, z_0), la valeur de la fonction au point (x_0+dx, y_0+dy, z_0+dz) est $\mathbf{u}(x_0, y_0, z_0) + d\mathbf{u}$, où

$$d\mathbf{u} = D.d\mathbf{r} \tag{14.4}$$

la relation ci-dessus n'est pas autre chose que $du = u'(x_0)dx$ pour les fonctions d'une seule variable et le tenseur D généralise la notion de dérivée. Évidement, nous savons que les composantes de ce tenseur s'écrivent

$$D_{ij} = \frac{\partial u_i}{\partial x_j}$$

où nous avons remplacé les coordonnées x, y, z par x_1, x_2, x_3 (cela est beaucoup plus pratique). Remarquez qu'à priori, nous n'avons rien fait d'autre que d'écrire le développement à l'ordre 1 de chaque composante u_i et de regrouper le tout sous forme d'une matrice. La relation (14.4) est cependant plus profond et met l'accent sur la linéarité de la relation entre $d\mathbf{r}$ et $d\mathbf{u}$ [4].

Ce que vous devez retenir à ce point est que les tenseurs généralisent la notion de multiplication à des objets de dimension supérieurs à 1.

Convention de sommation. Les tenseurs sont devenus populaires à partir des années 1920. Einstein par exemple a formulé sa théorie de relativité générale en termes tensoriels. Il a remarqué que quand on écrivait des sommes du genre (14.3), l'indice sur lequel on sommait (j dans le cas cité) était toujours répété dans deux quantités (χ_{ij} et E_j dans ce cas). Du coup, autant laisser tomber le signe "somme" et accepter que quand un indice est répété deux fois, cela veut dire qu'il faut sommer sur cet indice. Cette convention est tellement pratique qu'elle a été adopté partout. La relation (14.3) s'écrit, avec notre convention, $P_i = \chi_{ij}E_j$. Si nous étions dans un espace à quatre dimensions, la relation $\xi_{ij} = \lambda_{ijkl}\zeta_{kl}$ (peut importe ce que cela veut dire) est une façon plus concise d'écrire

$$\xi_{ij} = \sum_{l=1}^{4}\sum_{k=1}^{4}\lambda_{ijkl}\zeta_{kl}$$

Si les composantes d'un vecteur $\mathbf{x}$ dans la base $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ sont x_1, x_2, x_3, nous pouvons, en suivant cette convention, écrire simplement $\mathbf{x} = x_i\mathbf{e}_i$. Nous suivrons cette convention dans cette section.

3. par exemple la vitesse d'un écoulement en chaque point de l'espace.

4. Si l'on déforme un solide, chacun de ses points est déplacé. Nous pouvons définir une fonction vectorielle $\mathbf{u}(x,y,z)$ qui relate le déplacement de chaque point. Le tenseur D que nous avons défini ci-dessus (ou plutôt sa version symétrisée) est appelé dans ce cas le tenseur des déformations et constitue le socle de la théorie d'élasticité.

Note : Faire plusieurs exercices pour bien habituer à la convention, surtout en préparation de ce qui suit. Surtout les habituer aux permutations d'indice : $p_{ji}x_j$ et $p_{ij}x_i$ c'est la même chose. Donner l'équivalent en langage humain du genre "sommons sur le premier indice". Faire vraiment beaucoup d'exercice sur l'indice muet. Une bonne partie de la pratique des tenseurs c'est seulement cette manipulation d'indice.

Exercices.

Les mettre peut être juste avant les changement de base, ou les distribuer au fur et à mesure.

§ 14.1 Champ électrique créé par un dipôle

§ 14.2 énergie d'interaction entre dipôles

§ 14.3 le tenseur quadripolaire

§ 14.4 élasticité : tenseur de déformation et de contrainte, relation de Hook. Profiter pour habituer aux changements de variables pour écrire ces tenseurs en symétrie sphérique ou cylindrique : conteneur précontraint, les billes d'actine, etc.

§ 14.5 tenseur de déformation crée par une force ponctuelle à la surface ; par un dipôle de force ;

§ 14.6 Énergie d'interaction entre deux marches à la surface d'un cristal. Instabilités élastiques

§ 14.7 tenseur hydrodynamique et la viscosité, à la landau.

§ 14.8 Le tenseur métrique. Les opérateurs différentiels : grad, curl, div.

§ 14.9 Un peu de géométrie différentielle.

§ 14.10 Le tenseur électromagnétique en relativité, le lien avec le chapitre précédent.

14.2 Généralisation des tenseurs.

Les tenseurs de rang 2 que nous avons rencontré plus haut étaient des fonctions linéaires d'un espace vectoriel E_1dans un autre espace vectoriel E_2 : $A : E_1 \to E_2$. Un tenseur de rang 3 est une application bilinéaire qui prend *deux* vecteurs en entrée et produit *un* vecteur en sortie : $A : E_1 \times E_1 \to E_2$. l'application $A(\mathbf{x}, \mathbf{y})$ est bilinéaire, c'est à dire linéaire pour chacun de ses arguments :

$$\begin{aligned} A(\mu\mathbf{x}_1 + \lambda\mathbf{x}_2, \mathbf{y}) &= \mu A(\mathbf{x}_1, \mathbf{y}) + \lambda A(\mathbf{x}_2, \mathbf{y}) \\ A(\mathbf{x}, \mu\mathbf{y}_1 + \lambda\mathbf{y}_2) &= \mu A(\mathbf{x}, \mathbf{y}_1) + \lambda A(\mathbf{x}, \mathbf{y}_2) \end{aligned}$$

où $\lambda, \mu \in \mathbb{R}$. A nouveau, pour faire des calculs, nous nous donnons une base, et cette fois, A doit être donnée par trois indices a_{ijk} (une matrice à trois dimensions si vous

voulez : ligne, colonne, épaisseur). Si $\mathbf{z} = A(\mathbf{x}, \mathbf{y})$, alors la relation entre leurs composantes est juste une généralisation des produits matricielles :

$$z_i = a_{ijk} x_j y_k$$

Reprenons notre exemple de fonction $\mathbf{u}(x_1, x_2, x_3)$ que nous avons rencontré plus haut.

introduire déjà les formes linéaires, mais attendre le produit scalaire pour parler de l'espace dual.

14.3 Les composantes d'un tenseur.

Les tenseurs de rang 2.

Il nous faut maintenant manipuler pratiquement les tenseurs, c'est à dire à calculer leurs composantes. Revenons à ce que nous savons sur les vecteurs. Pour les manipuler, nous les représentons sous formes d'un ensemble de chiffre (souvent sous forme d'une colonne) : nous nous donnons une base $\{\mathbf{e}_1, \mathbf{e}_2, ...\mathbf{e}_n\}$, nous écrivons un vecteur $\mathbf{x}$ sous forme de combinaison linéaire de ces vecteurs, $\mathbf{x} = x_1\mathbf{e}_1 + ...x_n\mathbf{e}_n = x_i\mathbf{e}_i$ et nous représentons le vecteurs par les nombres $(x_1, ...x_n)$.

Pour caractériser une application linéaire A (autrement dit, un tenseur de rang 2), il suffit de connaître l'action de cette application sur chaque vecteur de base [5] :

$$\begin{aligned} A\mathbf{e}_1 &= a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2 + ...a_{n1}\mathbf{e}_n \\ A\mathbf{e}_2 &= a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2 + ...a_{n2}\mathbf{e}_n \\ &... \\ A\mathbf{e}_n &= a_{1n}\mathbf{e}_1 + a_{2n}\mathbf{e}_2 + ...a_{nn}\mathbf{e}_n \end{aligned} \tag{14.5}$$

ou dans notre notation concise tensorielle avec la convention de sommation de l'indice répété,

$$A\mathbf{e}_i = a_{ji}\mathbf{e}_j \qquad i = 1, ..., n \tag{14.6}$$

Notez le sens de la variation des indices : c'est la fameuse convention de présenter les composantes de $A\mathbf{e}_i$ comme des vecteurs *colonnes* et de les aligner côtes à côtes pour former une matrice. Connaissant ces composantes a_{ij}, nous pouvons calculer les composantes de n'importe quel vecteur $\mathbf{y} = A\mathbf{x}$ (on ne le précisera plus, nous suivons

5. Nous considérons seulement le cas où A est une application linéaire d'un espace vectoriel de dimension n dans un espace vectoriel de même dimension. La matrice représentant A est alors carrée. La généralisation au cas les espaces vectoriels de départ et d'arrivé n'ont pas la même dimension est triviale et laissé au soin du lecteur.

notre convention de sommation) :

$$\begin{aligned} A\mathbf{x} &= A(x_i\mathbf{e}_i) \\ &= x_i A\mathbf{e}_i \quad \text{linéraité d'une application linéaire} \\ &= x_i a_{ji}\mathbf{e}_j \quad \text{la relation (14.6) ci dessus} \end{aligned}$$

La i-ème composante du vecteur $\mathbf{y}$ est donc donné par

$$y_i = a_{ij}x_j \tag{14.7}$$

Ceci est bien sûr la règle selon laquelle il faut multiplier la i-ème ligne de la matrice par le vecteur colonne $\mathbf{x}$ élément par élément pour obtenir la i-ème composante de $\mathbf{y}$. Notez à nouveau [6] le sens de la variation de l'indice et comparez à la relation (14.6) : cette fois nous sommes en train de sommer sur le deuxième indice tandis que dans la relation (14.6), nous sommions sur le premier indice. On dit que les composantes d'un vecteur varient de façon *contravariante*, c'est à dire dans le sens contraire à la variation des vecteurs sous l'effet d'une application linéaire. Le mot "contraire" est mal choisi [7], puisque il veut simplement dire qu'il faut sommer sur l'autre indice. C'est malheureusement un mot consacré qu'il faut connaître.

Les tenseurs de rang quelconque.

Ce que nous venons de dire se généralise naturellement aux tenseurs de rang quelconque. Prenons le cas d'un tenseur A de rang 3, c'est à dire une fonction bilinéaire qui prend deux vecteurs en entrée et produit un vecteur en sortie ($A : E_n \times E_n \to E_n$, où E_n est un espace vectoriel de dimension n). On ne peut plus représenter A comme une matrice, il nous faudrait arranger les termes dans un tableau tri-dimensionnel, peu pratique à écrire. Il nous faut nous contenter des coefficients, mais à la longue vous le verrez, ce n'est pas moins pratique. Comme nous l'avons indiqué plus haut, il nous faut connaître l'action de A sur *tous* les *couples* vecteurs de base :

$$A(\mathbf{e}_i, \mathbf{e}_j) = a_{kij}\mathbf{e}_k \quad i,j = 1, ..., n$$

connaissant ces coefficients a_{kij}, nous pouvons connaître l'action de A sur n'importe quels deux vecteurs $\mathbf{x}$,$\mathbf{y}$:

$$\begin{aligned} \mathbf{z} &= A(\mathbf{x}, \mathbf{y}) \\ &= A(x_i\mathbf{e}_i, y_j\mathbf{e}_j) \\ &= x_i y_j A(\mathbf{e}_i, \mathbf{e}_j) \\ &= a_{kij}x_i y_j \mathbf{e}_k \end{aligned}$$

6. Nous avons permuter i et j pour désigner la ligne de la matrice par i et sa colonne par j. Certaines habitudes ont la vie dure.

7. Historiquement, l'application linéaire le plus souvent considérée était un changement de base et le mot contravariant avait plus de sens. Nous verrons cela un peu plus tard.

Autrement dit,

$$z_k = a_{kij} x_i y_j$$

Voilà, c'est ce que nous disions sur la généralisation naturelle des matrices. Notez la concision que la convention de sommation de l'indice répété nous procure.

14.4 Changement de base.

Connaissant les composantes d'un vecteur ou tenseur dans une base, comment en déduire ces composantes dans une autre base ? L'exercice n'est pas futile, et mérite que l'on y passe un peu de temps pour au moins deux raisons : (i) nous mesurons souvent les composantes d'un tenseur dans une base privilégiée où il est diagonal (par ce qu'il est beaucoup moins fastidieux de mesurer n coefficients que n^k), mais nous gagnons notre pain dans la base du laboratoire où nous effectuons nos expériences ; (ii) certaines théories comme la mécanique classique où la relativité sont fondées sur l'invariance de certaines quantités lors de changements de référentiel : il faut donc savoir comment passer d'une base à l'autre si l'on veut comprendre ces théories.

Commençons par le cas d'un vecteur. Nous connaissons les composantes d'un vecteur $\mathbf{x}$ dans la base $\{\mathbf{e}_1, ..., \mathbf{e}_n\}$ et nous souhaitons obtenir ses composantes dans la base $\{\mathbf{f}_1, ..., \mathbf{f}_n\}$. Évidement, il existe une application linéaire P qui transforme les vecteurs $\mathbf{f}_i$ en vecteur $\mathbf{e}_i$: $P\mathbf{f}_1 = \mathbf{e}_1$, $P\mathbf{f}_2 = \mathbf{e}_2$,... Par exemple, si nous avons tourné les axes de la première base de 45 degrés pour obtenir la deuxième base, l'application P est l'application "Rotation de -45 degrés". Notez que c'est l'application qui transforme la deuxième base en première ; pour plus de clarté, on le note parfois $P_{2\to1}$. Pour avoir la représentation matricielle de P dans la base $\{\mathbf{f}_i\}$, il suffit, comme nous l'avons dit plus haut (relation 14.5), d'écrire les vecteurs $\mathbf{e}_i$ comme des combinaisons linéaires des vecteurs $\mathbf{f}_i$:

$$\mathbf{e}_i = p_{ji}\mathbf{f}_j$$

Connaissant ces coefficients p_{ij}, on obtient directement l'expression de $\mathbf{x}$ comme combinaison linéaire des vecteurs $\mathbf{f}_i$:

$$\begin{aligned} \mathbf{x} &= x_i \mathbf{e}_i \\ &= x_i p_{ji} \mathbf{f}_i \end{aligned}$$

Si nous désignons par x'_i les composantes du vecteur dans la nouvelle base, nous avons

$$x'_i = p_{ij} x_j \tag{14.8}$$

Vous avez peut être remarqué la beauté de la relation ci-dessus : les composantes du vecteur $\mathbf{x}$ dans la deuxième base $\{\mathbf{f}\}$ sont égaux aux composantes du vecteur $P\mathbf{x}$ dans la première base $\{\mathbf{e}\}$! Revenons à l'exemple de la base $\{\mathbf{f}\}$ obtenue par une rotation

de +45 degrés de la base $\{\mathbf{e}\}$; les composantes d'un vecteur $\mathbf{x}$ dans cette nouvelle base sont égaux aux composantes du ce vecteur, tourné de *-45 degrés*, dans l'ancienne base. En faite, voilà l'origine du mot contravariant.

Continuons par un tenseur de rang 2. Nous connaissons les éléments a_{ij} d'un tenseur A dans une base $\{\mathbf{e}\}$, et nous souhaitons les exprimer dans la base $\{\mathbf{f}\}$ en nous aidant de l'application de passage $P = P_{2\to 1}$. Cette fois, nous avons également besoin de l'application inverse $Q = P^{-1}$:

$$\mathbf{f}_i = q_{ji}\mathbf{e}_j$$

Connaître l'application A dans la base $\{\mathbf{f}\}$, c'est connaître l'action de A sur chaque vecteur $\mathbf{f}_i$. Allons-y :

$$\begin{aligned} A\mathbf{f}_i &= Aq_{ji}\mathbf{e}_j \\ &= q_{ji}A\mathbf{e}_j \\ &= q_{ji}a_{kj}\mathbf{e}_k \\ &= q_{ji}a_{kj}p_{lk}f_l \end{aligned}$$

Ce qui nous donne :

$$a'_{li} = p_{lk}a_{kj}q_{ji}$$

Cette démarche se généralise immédiatement aux tenseurs de rangs supérieurs. Il ne faut surtout rien retenir pas cœur : si vous avez compris la démarche, la formule de transformation s'obtient en deux coups de cuillère.

Exercice : déduire l'expression des tenseurs a'_{ijk} et a'_{ijkl}

14.5 Le produit scalaire généralisé, covariant et contravariant, les formes linéaires.

Le lecteur qui a rencontré les tenseurs, surtout dans le contexte de la relativité générale ou de la géométrie riemannienne a noté que l'on fait une distinction très nette entre les vecteurs, tenseurs,... covariant et contravariant. Les composantes des uns s'écrivent avec des indices en bas, d'autre avec des indices en haut et nous manipulons constamment des expressions du genre $\xi^{ij}_{kl}y_i x^k$. Nous écrivions jusque là tous les indices en bas, mais essayons d'aller plus loin. Une bonne partie du calcul tensoriel consiste à pratiquer la manipulation des indices. Nous allons nous donner une nouvelle règle de sommation dont le sens sera précisé plus loin.

Règle de sommation. Une expression qui contient des indices répétés représente une somme si un des indices est en haut et l'autre en bas.

C'est une règle grammatical au même titre que les parenthèses : une expression bien formée contient le même nombre de parenthèses ouvertes que fermées. De même pour les indices répétés, l'expression $x^i y_i$ est correctement formé et veut dire $x^1y_1 + x^2y_2 + x^3y_3 + ...$ L'expression $g_{ij}x^iy^j$est bien formée, ainsi que $g^{ij}x_iy_j$ et $g^j_i x_j y^i$. Par contre, $g^i_j x_j y^i$ ou $g_{ij}x_iy^j$ sont grammaticalement incorrect, au même titre que les expressions du genre $(a+b))c+d($ ou $(a+(b+(c)$. Par convention, tout ce qui a un indice en bas est appelé *covariant* ; un indice en haut représente une quantité contravariant.

Donnons nous maintenant une base $(\mathbf{e}_1, \mathbf{e}_2, ...)$ dans un espace vectoriel E. N'importe quel vecteur $\mathbf{x}$ s'écrit comme une combinaison linéaire des éléments de la base, ce qui, avec notre nouvelle convention de sommation, donne :

$$\mathbf{x} = x^i \mathbf{e}_i$$

Si vous vous souvenez, nous avions souligné que les composantes varient dans le sens contraire des vecteurs lors d'un changement de base, cette notation met cela clairement en évidence.

Nous rentrons maintenant dans le vif du sujet. A l'espace E, nous associons un espace dual E^* des formes linéaires. Cet espace représente, si vous voulez, l'autre coté du miroir. Par forme linéaire, il faut entendre l'espace de toutes les fonctions f qui prennent un élément de E, produisent un scalaire, et font cela de façon linaire : $f : E \to \mathbb{R}, f(\lambda \mathbf{e}_1 + \mu \mathbf{e}_2) = \lambda f(\mathbf{e}_1) + \mu f(\mathbf{e}_2)$. E^* ressemble à E : c'est également un espace vectoriel ($af + bg$, où $a, b \in \mathbb{R}$et $f, g \in E^*$a exactement le sens qu'on lui donne) et de plus, E^* a la même dimension que E. (cf exercice plus bas). E^*est l'espace des vecteurs contravariants comme nous allons le voir, mais retenez l'image du miroir : quand vous tournez vers la gauche, votre image dans le miroir tourne vers la droite.

Revenons à notre espace E. Vous avez vu (cf 2.1) que nous pouvons le munir d'un *produit scalaire* $<,>$, c'est à dire une application bilinéaire qui a deux vecteurs de E associe un scalaire. Si nous nous donnons une base $(\mathbf{e}_1, \mathbf{e}_2, ...)$ et que nous connaissons le produit scalaire de chaque couple d'entre eux $< \mathbf{e}_i, \mathbf{e}_j >= g_{ij}$, alors nous pouvons écrire le produit scalaire de n'importe quels deux vecteurs

$$\langle \mathbf{x}, \mathbf{y} \rangle = g_{ij} x^i y^j$$

où x^iet y^j sont les composantes des deux vecteurs. N'oublions pas que le produit scalaire doit être défini positif, c'est à dire $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ si $\mathbf{x} \neq \mathbf{0}$. Cela impose certaines contraintes sur les valeurs g_{ij}que nous verrons plus tard.

Considérons maintenant l'objet $f_v = \langle \mathbf{v}, . \rangle$, où $\mathbf{v} \in E$. Ceci est bien une forme linéaire appartenant à E^*, puisque si $\mathbf{u} \in E$, alors $f_v(\mathbf{u}) = \langle \mathbf{v}, \mathbf{u} \rangle \in \mathbb{R}$.

Note sur le produit tensoriel de deux espaces vectoriel, tenseur d'élasticité qui relie deux tenseurs,...

Indice haut, bas, la réduction, monter ou descendre un indice : essentiellement en relation avec la relativité et le produit scalaire de Minkowski.

espace dual

exercice : démontrer que E^*est un espace vectoriel de même dimension

une base dans l'espace dual sans faire référence au produit scalaire.

produit scalaire généralisé $< x, y >= g_{ij} x^i y^j$

15 Équation à dérivée partielle du premier ordre.

Le monde des équations à dérivées partielles (EDP) est vaste, et des centaines de livres, souvent extrêmement pédagogiques, leur sont consacrées. Les EDP les plus utilisées en physique sont de second ordre pour l'espace, et de premier ou de second ordre pour le temps. Ce sont les équations de la chaleur, de Laplace et de Poisson, et l'équation d'onde. Nous n'allons pas traiter ces équations de façon générale ; les méthodes des chapitres précédents (TF,TL, Green,...) sont les outils de base par lesquels ces EDP linéaires sont abordées.

Les EDP linéaires *de premier ordre* sont par contre exactement soluble par la méthode des caractéristiques, au moins en théorie. Il est utile d'en donner un bref aperçu.

15.1 La méthode des caractéristiques.

Nous souhaitons déterminer la fonction $\phi(s,t)$ obéissant à l'équation :

$$\partial_t \phi + P(s)\partial_s \phi = Q(s)\phi \tag{15.1}$$

Nous cherchons la solution sous forme de $\phi(s,t) = f(s)g(u(s,t))$ où f, g, u sont des fonctions inconnues à déterminer. A priori, nous n'avons rien gagné d'autre que l'augmentation du nombre de fonctions inconnues. Mais nous gagnons la liberté d'imposer des contraintes à ces fonctions qui ramènerons le problème à des choses plus connues.

Nous avons, pour les dérivées partielles de ϕ :

$$\partial_t \phi = \partial_t u f(s) g'(u) \tag{15.2}$$
$$\partial_s \phi = f'(s)g(u) + \partial_s u f(s) g'(u) \tag{15.3}$$

En insérant (15.2,15.3) dans l'équation (15.1), nous avons :

$$(\partial_t u + P(s)\partial_s u) f(s) g'(u) + (P(s)f'(s) - Q(s)f(s))g(u) = 0. \tag{15.4}$$

Une solution évidente de (15.4) est donnée par :

$$P(s)f'(s) - Q(s)f(s) = 0 \tag{15.5}$$
$$\partial_t u + P(s)\partial_s u = 0 \tag{15.6}$$

L'équation (15.5) est une équation différentielle linéaire homogène de premier ordre en f et sa solution est donnée par $f(s) = \exp(A(s))$, où $A'(s) = Q(s)/P(s)$.

L'équation (15.6) , qui est une EDP de premier ordre homogène, a comme solution $u(s,t) = \exp(W(s) - t)$ où $W'(s) = 1/P(s)$. Notons que la fonction g reste indéterminée. Son choix dépend des conditions initiales imposées au système.

Exemple 15.1 Nous souhaitons résoudre une équation de diffusion avec un terme de dérive linéaire, appelé Ornstein-Uhlenbeck :

$$\partial_t p = \partial_x(kxp) + D\partial_x^2 p \tag{15.7}$$

avec la condition initiale (CI)

$$p(x,0) = \delta(x - x_0) \tag{15.8}$$

où δ est la fonction de Dirac. $p(x,t)$ est une densité de probabilité de présence à l'instant t à l'abscisse x, et la condition initiale veut simplement dire que toute la probabilité est condensée en x_0 à l'instant $t = t_0$. Soit $\phi(s,t)$ la transformée de Fourier en x de $p(x,t)$:

$$\phi(s,t) = \int_{-\infty}^{+\infty} e^{isx} p(x,t)dx$$

l'équation (15.7) se transforme alors en

$$\partial_t \phi + ks\partial_s \phi = -Ds^2\phi \tag{15.9}$$

avec la condition initiale

$$\phi(s,0) = e^{isx_0} \tag{15.10}$$

L'équation (15.9) est une EDP de premier ordre similaire à (15.1), avec $P(s) = ks$ et $Q(s) = -Ds^2$. Nous avons alors $A(s) = -Ds^2/2k$ et $W(s) = \log(s)/k$. Les fonctions f,u et ϕ s'écrivent :

$$f(s) = \exp(-Ds^2/2k) \tag{15.11}$$
$$u(s,t) = s\exp(-kt) \tag{15.12}$$
$$\phi(s,t) = \exp(-Ds^2/2k)\, g(s\exp(-kt)) \tag{15.13}$$

Nous pouvons vérifier par dérivation directe que l'expression (15.13) est bien solution de (15.9). Il nous reste à utiliser la CI (15.10) pour trouver g, ce qui donne :

$$g(s) = \exp(Ds^2/2k + iux_0)$$

et la solution complète s'écrit :

$$\phi(s,t) = \exp\left[\frac{-Ds^2}{2k}(1 - e^{-2kt}) + isx_0 e^{-kt}\right]$$

La fonction ϕ est appelé en probabilité la fonction caractéristique associé à la densité de probabilité p. La moyenne et la variance se calcule aisément si on connaît ϕ :

$$\begin{aligned}\langle X\rangle &= \int xp(x)dx = -i\frac{\partial\phi}{\partial s}\mid_{s=0}\\ \langle X^2\rangle &= \int x^2p(x)dx = -\frac{\partial^2\phi}{\partial s^2}\mid_{s=0}\end{aligned}$$

Exercices.

§ 15.1 Calculer, pour le processus d'Ornstein, $\langle X(t)\rangle$ et $Var(X(t)) = \langle X^2(t)\rangle - \langle X(t)\rangle^2$.

§ 15.2 **Croissance exponentielle.** Pour une croissance exponentielle, l'équation maîtresse s'écrit :

$$\partial_t p(n,t) = (n-1)p(n-1,t) - np(n,t)$$

où $p(n,t)$ est la probabilité pour une population d'avoir la taille n à l'instant t. En utilisant plutôt la transformée de Laplace

$$\phi(s,t) = \sum_n p(n,t)e^{-ns}$$

Calculer $\langle n(t)\rangle$et $Var(n(t))$. La condition initiale est $p(n,0) = \delta_{n,n_0}$. Ceci est également connu sous le nom de processus de Poisson.

§ 15.3 **(Plus ardu).**On démarre avec une population initiale de bactérie n_0, dont les taux de mort et naissance sont égaux et valent α. Calculer $\langle n(t)\rangle$ et $Var(n(t))$.

Nous verrons que l'équation maîtresse dans ce cas s'écrit :

$$\partial_t p(n,t) = (n-1)p(n-1,t) + (n+1)p(n+1,t) - 2np(n,t) \tag{15.14}$$

avec la condition initiale $p(n,0) = \delta_{n,n_0}$. $p(n,t)$ est la probabilité pour la population, à l'instant t, d'avoir la taille n.

Pour résoudre l'éq.(15.14), vous avez plus intérêt à utiliser la transformée de Laplace :

$$\phi(s,t) = \sum_n p(n,t)e^{-ns}$$

15.2 Interprétation géométrique.

Nous avons donné la méthode des caractéristiques comme une recette. Mais cette recette découle d'une interprétation géométrique très simple : les dérivées premières sont les pentes de la fonction selon les diverses directions. Prenons d'abord le cas de l'équation

$$P(s,t)\partial_s\phi + R(s,t)\partial_t\phi = 0 \tag{15.15}$$

Nous pouvons représenter la solution $\phi(s,t)$ comme une surface : $\phi(s,t)$ étant la hauteur à la position (s,t). Si nous pouvions connaître les courbes de niveau de cette surface, nous aurions déjà une très bonne connaissance de la solution (voir figure 15.1). Quand on parcours une courbe de niveau, la valeur de la fonction ϕ y reste constante. Supposons maintenant que nous sommes à une position (s,t). Comment se déplacer d'une quantité (ds,dt) pour que la valeur de la fonction ϕ reste constante ? Quelle doit être le rapport entre le déplacement dans la direction ds et le déplacement dans la direction dt pour ne pas changer d'altitude ? Noter que se donner une relation entre ds et dt en tout point défini une courbe dans le plan (s,t). Par exemple, $dy/dx = -x/y$ défini l'équation d'un cercle de centre origine ; le rayon de ce cercle est donné par une condition initiale.

La variation de ϕ en fonction de (ds,dt) est

$$d\phi = (\partial_s\phi)ds + (\partial_t\phi)dt$$

En comparant cette expression à (15.15), nous voyons qu'il suffit de choisir ds proportionnel à $P(s,t)$ et dt proportionnel à $R(s,t)$ pour que $d\phi = 0$. Autrement dit, pour avoir $d\phi = 0$, il suffit de choisir

$$\frac{ds}{P(s,t)} = \frac{dt}{R(s,t)} \tag{15.16}$$

Comme vous le remarquez, l'expression ci-dessus est une équation différentielle ordinaire donnant la forme de la courbe qu'on appelle caractéristique.

Exemple 15.2 L'équation $t\partial_s\phi - s\partial_t\phi = 0$: les courbes de niveaux sont données par $sds+tdt = 0$, autrement dit par $s^2+t^2 = C$. Ce sont des cercles centrés sur l'origine.

Appliquons cela à l'équation plus simple que nous avions traité au début de ce chapitre :

$$\partial_t\phi + P(s)\partial_s\phi = 0 \tag{15.17}$$

Les courbes de niveau sont données par

$$\frac{ds}{P(s)} = \frac{dt}{1}$$

En intégrant les deux côtés, nous avons donc

$$W(s) - t = C$$

où $W'(s) = 1/P(s)$ et C est une constante d'intégration. La fonction $u(s,t) = W(s) - t = C$ nous donne donc les courbes de niveau, et la solution générale de l'équation (15.17) est donc de la forme

$$\phi(s,t) = g\left(u(s,t)\right)$$

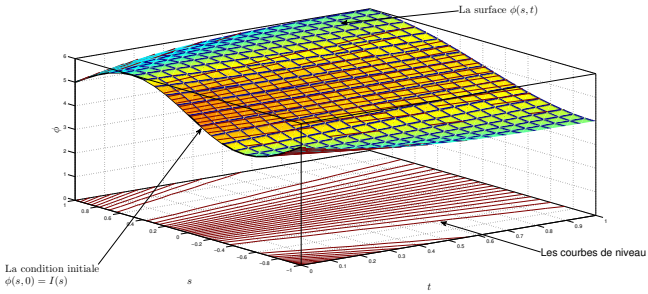


FIGURE 15.1 – Construction d'une solution : nous trouvons d'abord les courbes de niveau dans le plan (t,s). Ensuite, en utilisant la condition initiale $\phi(s,0) = I(s)$, on précise la hauteur de la surface de ϕ sur chacune des courbes et on reconstruit la solution $\phi(s,t)$.

La fonction g doit être trouvée en utilisant les conditions aux bords. Ceci est le sens de la "recette" de résolution que nous avions donné au début de ce chapitre, dans le cas simple où le second membre est nul.

Très bien, nous connaissons les courbes de niveaux. Mais pour vraiment connaître ϕ, il faut connaître la valeur de cette fonction sur chaque courbe. Comment déterminer cela ? Évidement, à l'aide des conditions initiales. Si par exemple, on se donne $\phi(s,0) = I(s)$, nous connaissons alors la valeur de ϕ sur la courbe de niveau qui passe par $(s,0)$ (voir figure 15.1). Autrement dit, en transportant les hauteurs $\phi(s,0)$ le long des courbes de niveau, on reconstruit la surface $\phi(s,t)$. Ceci est le sens de la détermination de la fonction $g(u)$ par les conditions initiales dans la section précédente. Nous avons donc la méthode générale de la résolution d'une EDP de premier ordre.

Exemple 15.3 **Équationd'onde.** Soit l'équation $c\partial_x\phi - \partial_t\phi = 0$, avec la condition initiale $\phi(x,0) = f(x)$. Les courbes caractéristiques sont $dx/c = -dt$, autrement dit $x = -ct + x_0$. En inversant cette relation, nous trouvons $x_0 = x + ct$. Nous trouvons donc que

$$\phi(x,t) = f(x+ct)$$

Nous avons appelé cette équation "équation d'onde" puisque l'équation $c^2\partial^2\phi/\partial x^2 - \partial^2\phi/\partial t^2 = 0$ se factorise en

$$(c\partial_x - \partial_t)(c\partial_x + \partial_t)\phi = 0$$

Si ϕ est solution de $c\partial_x\phi - \partial_t\phi = 0$ ou de $c\partial_x\phi + \partial_t\phi = 0$, alors ϕ est solution de l'équation d'onde.

15.3 Généralisation.

A partir de là, nous pouvons généraliser notre analyse à l'équation

$$P(s,t)\partial_s\phi + R(s,t)\partial_t\phi = Q(s,t,\phi)$$

avec les conditions initiales $\phi(s,0) = I(s)$. Les courbes caractéristiques données par (15.16) ne sont plus des courbes de niveau, mais la variation de ϕ le long des courbes est donnée par une équation différentielle ordinaire. Supposons que la solution de $ds/P = dt/R$ soit donnée par $s = f(t,s_0)$ [1], c'est à dire $ds/dt = f'(t,s_0) = P/R$. Quand on se déplace le long d'une courbe caractéristique $s = f(t,s_0)$, la variation de ϕ est

$$\begin{aligned} d\phi &= ds\partial_s\phi + dt\partial_t\phi \\ &= (P\partial_s\phi + R\partial_t\phi)(dt/R) \\ &= (Q/R)dt \end{aligned}$$

Donc, le long de ces courbes, ϕ est solution de l'équation

$$\frac{d\phi}{dt} = \frac{Q(s,t,\phi)}{R(s,t)} \tag{15.18}$$

La stratégie pour trouver la solution est une modification de ce que nous avons dit précédemment :

1. Trouver la courbe caractéristique $s = f(t,s_0)$, l'inverser pour trouver $s_0 = g(s,t)$.
2. Résoudre l'équation différentielle *ordinaire* (15.18) où $s = f(t,s_0)$ soumise à la condition initiale $\phi(0) = I(s_0)$, *le long* d'une courbe caractéristique.

Exemple 15.4 Nous souhaitons Résoudre l'équation $\partial_s\phi + P(t)\partial_t\phi = Q(t)\phi$ avec la condition initiale $\phi(s,0) = I(s)$. Notez que nous n'avons pas de dépendance explicite en s dans cette équation.

Les courbes caractéristiques sont $ds/1 = dt/P(t)$. Si nous appelons $W(t)$ une primitive de $1/P(t)$, (la fonction W est connue, puisque P l'est) les courbes caractéristiques sont données par $W(t) - s = W(0) - s_0$. Le long de ces courbes, nous choisissons s comme variable indépendante (nous avons le choix du paramétrage) et donc, le long de ces courbes,

$$\frac{d\phi}{dt} = \frac{Q(t)}{P(t)}\phi$$

Si nous appelons $A(t)$ une primitive de Q/P, alors la solution générale de l'équation ci-dessus est

$$\phi(t) = C.\exp(A(t)).$$

1. Pour être plus général, nous aurions du écrire $f(s,t,s_0) = 0$

Toutes les courbes passant par t (quelque soit leur ordonnée s) ont la forme ci-dessus. Vous pouvez voir cela comme le croisement entre la surface ϕet le plan t. La caractéristique passant par le point (s,t) passe par le point $(s_0,0)$ où $s_0 = W(t)-s-W(0)$. La solution complète de l'équation s'écrit donc comme

$$\phi(s,t) = I(W(t) - s - W(0).) \exp(A(t)$$

Nous laissons au lecteur le soin de revoir la première section de ce chapitre à la lumière de ces développements.

16 Les formes différentielles et la dérivation extérieure.

16.1 Introduction.

L'analyse des fonctions d'une seule variable est simple : nous savons prendre la dérivée première, seconde, ... et cela a un sens direct (pente, courbure,...). Quand on aborde les fonctions de plusieurs variables, les choses commencent à se compliquer. Les opérateurs différentielles prennent alors des noms étranges comme gradient, divergence, rotationnel,... Et nous découvrons qu'il existe des relations entre ces êtres : prendre la *circulation* le long d'une courbe fermée revient à calculer le *flux* d'un rotationnel à travers la surface engendrée par cette courbe ; calculer le flux à travers une surface fermée revient à prendre l'intégrale de la divergence *dans* le volume,... Si vous êtes très versé dans la manipulation de ces objets, vous savez peut-être ce que vaut $\mathbf{grad}(\mathbf{div}\, f)$ par cœur !

Tout cela n'est pas très joli. D'abord, nous avons de la peine à distinguer la signification de tous ces opérateurs et des relations qui existent entre eux ; ensuite, cette analyse vectorielle ne marche qu'à *trois* dimensions [1] ; Enfin, les équations mathématiques deviennent confuses : pourquoi la dérivée temporelle du champ magnétique devrait être liée au rotationnel du champ électrique ? On sent bien qu'il y a des arguments de géométrie derrière cela, mais quoi exactement ?

A partir du début du vingtième siècle, des mathématiciens comme Poincaré et Cartan ont réalisé que derrière tout ce chaos, il y avait de l'ordre, exactement comme la découverte de l'existence des atomes a donné un sens à la chimie. Les atomes en question ici s'appellent des formes différentielles, et nous allons les étudier plus en détails. Disons simplement que toutes ces relations entre opérateurs ne sont en faite que des généralisations du Théorème Fondamental de l'analyse :

$$\int_a^b F'(x)dx = F(b) - F(a)$$

1. Et cela par un malentendu qui fait correspondre un vecteur aux produit vectoriel dans le cas des espaces à trois dimensions.

16.2 Les $1-$formes.

Une 1-forme est une application linéaire qui agit sur un vecteur et produit un nombre. Le lecteur est probablement habitué déjà à ce concept : en notation matricielle, les (vrais) vecteurs sont représentés par des colonnes (vecteur colonne) ; ceci dit, nous avons également des vecteurs *lignes*. Ces vecteurs lignes sont ce qu'on appelle des 1-formes. L'application d'une forme $\tilde{u}$ à un vecteur v revient à "multiplier" son "vecteur" ligne par le vecteur colonne de v pour produire un nombre.

Bien. Ceci dit, une colonne de nombre est juste une *représentation* d'un vecteur, qui dépend de la base choisie. Les vecteurs sont des objets géométriques qui ne dépendent évidemment pas de leurs représentations [2]. De la même manière, les 1-formes ne sont pas des vecteurs lignes qui est juste une façon de les représenter par des nombre : Le scalaire $\tilde{u}(v)$ ne dépend pas de la base que nous avons choisie.

Pour développer notre théorie, il est pratique de nous donner une base. Prenons, pour simplifier les choses, un espace *plat* à 3 dimensions, où nous nous donnons trois vecteurs indépendants $\mathbf{e}_x, \mathbf{e}_y$ et $\mathbf{e}_z$. N'importe quel vecteur peut maintenant être représenté par une combinaison du genre $\mathbf{v} = a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z$ où les coefficients a_y sont des nombres.

De la même façon, nous pouvons nous donner une base dans l'espace des formes et écrire une 1-forme ω comme une combinaisons de ces 1-formes de base. C'est ici que les mathématiciens ont introduit une notation qui peut paraître déroutant aux physiciens [3], mais qui s'avère extrêmement féconde. Il existe par exemple un (unique) 1-forme ω_x tel que

$$\omega_x(\mathbf{e}_x) = 1\,;\ \omega_x(\mathbf{e}_y) = 0\,;\ \omega_x(\mathbf{e}_z) = 0$$

Nous appellerons cette 1-forme dx. Ici, dx n'a rien d'infinitésimal, sa représentation est le vecteur ligne $(1, 0, 0)$. De la même façon, nous définissons les 1-formes dy et dz. Ainsi, $dy(a_x\mathbf{e}_x + a_y\mathbf{e}_y + a_z\mathbf{e}_z) = a_y$.

Un exemple de 1-forme est la force f en mécanique [4]. Quand sous l'action de cette force, un point matériel bouge du point P_1 au point P_2, le travail (qui un scalaire) de cette force est $f(\vec{P_1P_2})$, c'est à dire l'application de la 1-forme f au vecteur $\vec{P_1P_2}$. Donnons nous par exemple un champ de force constant $f = 2dx + 3dy$ et supposons que le déplacement [5] $\vec{P_1P_2} = 3\mathbf{e}_x + 2\mathbf{e}_y$. Alors le travail de cette force est

$$W = 2dx(3\mathbf{e}_x + 2\mathbf{e}_y) + 3dy(3\mathbf{e}_x + 2\mathbf{e}_y) = 6 + 6 = 12$$

Généralisons un peu plus. De la même façon que nous définissions un champ de vecteur (pensez le champ de déplacement $\mathbf{u}(x, y, z)$ des points d'un corps matériel sous

2. C'est la différence entre une personne et une photo de cette personne.

3. Utiliser ces notations demande un peu de schizophrénie de la part de l'étudiant physicien, qui doit désapprendre ses propres notations.

4. Oui, la force n'est pas un vecteur, mais un 1-forme

5. Oui, le déplacement est un vecteur.

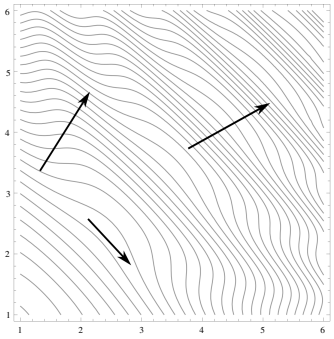


FIGURE 16.1 – Comment représenter les 1-formes ? Dans l'espace à 2 dimensions, une très bonne représentation sont les lignes de flux. Ainsi, l'action d'une 1-forme sur un vecteur est le nombre de ligne que ce vecteur coupe. Ici, nous l'avons ainsi représenté. Le lecteur est déjà familier avec cette représentation comme courbes de niveau pour certaines 1-formes. Pour ces formes là, ceci est une bonne représentation du gradient, comme nous le verrons plus bas. De façon générale, nous pouvons, dans un espace à n dimensions, représenter les 1-formes par des hyper surfaces.

l'action des forces), nous pouvons définir un champ de 1-forme comme par exemple

$$\omega = f(x,y)dx + g(x,y)dy$$

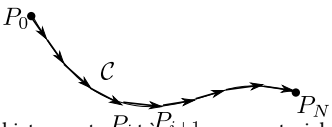
L'application de cette 1-forme à un vecteur $\vec{P_1P_2}$ produit un scalaire qui dépend non seulement du vecteur $\vec{P_1P_2}$, mais également de la localisation de ce vecteur dans l'espace [6]. Si le vecteur est "petit", que le point P_0 qui désigne son milieu est de coordonnées (x_0, y_0), et que nous pouvons le représenter par le vecteur colonne $(h,k)^T$, alors

$$\omega(\vec{P_1P_2}) = f(x_0,y_0)h + g(x_0,y_0)k$$

A vrai dire, le mot "différentielle" dans "forme différentielle" sous-entend bien que nous nous adressons qu'aux "petits" vecteurs ; l'action des formes sur des plus grands objets s'obtient par la sommation de leurs actions sur les petits, ce qu'on désigne par intégration. Nous verrons cela plus bas.

16.3 Intégration des 1-formes.

Nous avons une idée assez claire de ce que veux dire l'intégrale d'une fonction d'une variable sur un intervalle $[a,b]$. Nous devons maintenant définir exactement ce que l'on entend par l'intégration des 1-formes le long d'un chemin $\mathcal{C}$ reliant deux points A et B. Géométriquement, cela est très simple : Donnons nous une 1-forme ω, un chemin $\mathcal{C}$ et $N+1$ points le long de la courbe $(P_0 = A, P_1, ...P_i, ..., P_N = B)$. Nous



6. Il existe une différence entre un vecteur abstrait, c'est à dire un objet appartenant à un espace vectoriel, et un vecteur géométrique reliant deux points P_1 et P_2. Les deux concepts sont fortement connectés, mais différents.

définissons

$$\int_{\mathcal{C}} \omega = \lim_{N\to\infty} \sum_{i=0}^{N} \omega(\overrightarrow{P_i P_{i+1}})$$

En clair, nous appliquons le 1-forme à tous les "petits" vecteurs dont l'union constitue le chemin. Remarquez que nous avons transféré, pour l'intégration ici le poids du "petit" des formes aux vecteurs.

Ceci dit, comment faire l'intégration concrètement ? Soit la forme $\omega = f(x,y)dx + g(x,y)dy$ et le chemin $\mathcal{C}$. Nous pouvons donner l'équation de la courbe sous forme paramétrique $x = x(t)$ et $y = y(t)$ pour $t \in [a,b]$. Nous avons alors, selon les opérations classiques de l'analyse [7], $dx = x'(t)dt$ et $dy = y'(t)dt$. Nous avons alors

$$\int_{\mathcal{C}} f(x,y)dx + g(x,y)dy = \int_a^b \{f(x,y)x'(t) + g(x,y)y'(t)\}\, dt$$

Exemple 16.1 Intégrons la forme $\omega = ydx + xdy$ le long du quart de cercle de rayon 1 parcouru dans le sens positif. Nous pouvons paramétriser le quart de cercle par $x = \cos t$ et $y = \sin t$ pour $t \in [0, \pi/2]$. Nous avons alors

$$\int_{\mathcal{C}} \omega = \int_0^{\pi/2} \left(-\sin^2(t) + \cos^2(t)\right) dt \quad = \quad 0$$

L'intégration de $-ydx + xdy$ le long du même chemin nous donnerai $\pi/2$.

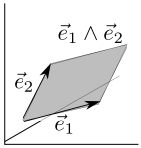
En physique, nous ne faisons souvent pas de distinction entre 1-forme et vecteur. Ainsi, le 1-forme $\omega = adx + bdy$ est souvent remplacé par le vecteur $\mathbf{f} = (a,b)^T$. L'intégrale $\int_{\mathcal{C}} \omega$ est écrit dans ces notations comme

$$\int_{\mathcal{C}} \mathbf{f}.d\mathbf{s}$$

où $d\mathbf{s}$ est un élément infinitésimal et le "point" désigne le produit scalaire.

16.4 les $n-$formes et les $n-$vecteurs.

Dans l'espace, nous avons des points, des vecteurs (reliant deux points proches), et des 1-formes. Nous pouvons maintenant généraliser ces concepts et construire des objets plus complexes. Par exemple, nous pouvons construire des bi-vecteurs. De la même manière qu'un vecteur $\vec{e}$ peut être utiliser pour



7. Les formes différentielles ne font que généraliser les concepts d'analyse.

"porter" un segment de ligne orienté, un bi-vecteur $\vec{e}_1 \wedge \vec{e}_2$ peut être utilisé pour porter un élément de surface *orientée*[8]. Ainsi, nous avons

$$\begin{aligned} \vec{e} \wedge \vec{e} &= 0 \\ \vec{e}_1 \wedge \vec{e}_2 &= -\vec{e}_2 \wedge \vec{e}_1 \end{aligned}$$

un bi-vecteur a toute les propriétés usuelles de distributivité qu'on attend de lui. Par exemple, $\vec{e}_1 \wedge (\vec{e}_2 + \vec{e}_3) = \vec{e}_1 \wedge \vec{e}_2 + \vec{e}_1 \wedge \vec{e}_3$. Si dans l'espace $n = 3$ des vecteurs, nous avons pris $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ comme éléments de la base, Dans l'espace des bi-vecteurs, nous pouvons définir $\vec{e}_1 \wedge \vec{e}_2$, $\vec{e}_2 \wedge \vec{e}_3$, $\vec{e}_3 \wedge \vec{e}_1$ comme les éléments de la base. De façon générale, dans un espace à n dimensions, la dimension de l'espace des bi-vecteurs est $\binom{n}{2}$. Un hasard malheureux (ou heureux, selon les points de vue) fait que si $n = 3$, l'espace des bi-vecteurs est de dimension 3 également. Une certaine habitude s'est alors instaurée de représenter le bi-vecteur par un *vecteur* (normale à la surface), en mettant en garde l'utilisateur que ces vecteurs sont un peu anormaux, qu'il faut les appeler "axial", etc.

Bien, les 2-formes sont de la même manière une généralisation des 1-forme. Une 2-forme ω s'applique à un bi-vecteur pour produire un scalaire de façon bilinéaire. Ainsi, la 2-forme $dxdy$ appliquée[9] à $\vec{e}_1 \wedge \vec{e}_2$ produit le nombre 1. Une 2-forme constante peut-être vue comme un flux ; appliqué à une surface, cela produit le flux à travers cette surface.

16.5 L'intégration des $k-$formes.

Tout ce que nous avons dit sur les 1-forme se généralise aux k-formes. Nous appellerons un $k-$surface un objet que l'on peut paramétrer par k variables. Une courbe par exemple est une 1-surface, une surface au sens habituel est une 2-surface, un volume un 3-surface et ainsi de suite. Dans un espace à n dimensions, nous pouvons donner un sens très précis à l'intégration d'un $k-$forme ω sur un $k-$surface $\mathcal{S}$: nous découpons $\mathcal{S}$ en N "petits" éléments portés par des $k-$vecteurs a_i et nous définissons

$$\int_{\mathcal{D}} \omega = \lim_{N \to \infty} \sum_{i=0}^{N} \omega(a_i)$$

Le calcul effectif se fait par une paramétrisation de $\mathcal{D}$. Par exemple, une 2-formes dans l'espace à 3d peut être représentée par

$$\omega = f(x, y, z)dxdy + g(x, y, z)dydz + h(x, y, z)dzdx$$

8. L'opération $\wedge$ est appelé "produit extérieur".
9. La convention est d'omettre le $\wedge$ entre dx et dy.

Pour calculer concrètement l'intégrale $\int_{\mathcal{S}} \omega$, on paramétrise la surface par deux variables

$$x = a(u,v); y = b(u,v); z = c(u,v)$$

avec $(u,v) \in \mathcal{D} \subset \mathbb{R}^2$. Nous avons alors

$$dx = \frac{\partial a}{\partial u}du + \frac{\partial a}{\partial v}dv$$

et l'élément $dxdy$ par exemple devient

$$dxdy = \left(\frac{\partial a}{\partial u}du + \frac{\partial a}{\partial v}dv\right)\left(\frac{\partial b}{\partial u}du + \frac{\partial b}{\partial v}dv\right)$$

Or, comme $dudu = dvdv = 0$ et $dudv = -dvdu$, nous trouvons simplement

$$dxdy = \left(\frac{\partial a}{\partial u}\frac{\partial b}{\partial v} - \frac{\partial a}{\partial v}\frac{\partial b}{\partial u}\right)dudv$$

et ainsi pour les autres éléments. La parenthèse représente bien sûr ce que nous appelons un Jacobien, c'est à dire le déterminant de la matrice des dérivées. Quand vous passez des vecteurs e_1, e_2 aux vecteurs f_1, f_2 par une transformation linéaire A, le déterminant de A est le scalaire qui relie les surfaces portées par les deux jeux de vecteurs. Ceci est *la définition* du déterminant, indépendamment de la base choisie pour exprimer la matrice de A. Le Jacobien apparaît ici puisque nous sommes passé des dx,dy,dz aux du,dv par une transformation linéaire donnée par la matrice des dérivées.

Connexion avec l'analyse vectorielle.

En analyse vectorielle classique, nous rencontrons souvent l'intégrale d'un champ de vecteur le long d'une courbe, de surface, de volume, ... Par exemple, nous savons que le travail effectué par un champ de force le long d'une courbe est appelé *circulation* du vecteur. Soit le vecteur $\mathbf{f} = (f_x, f_y, f_z)^T$. Sa circulation C le long d'une courbe $\mathcal{C}$ est définie par

$$C = \int_{\mathcal{C}} \mathbf{f}.d\mathbf{t}$$

où $d\mathbf{t}$ est un petit vecteur tangent à la courbe au point P, et $\mathbf{f}.d\mathbf{t}$ représente le produit scalaire entre ces deux vecteurs. En langage de forme différentielles, au vecteur $\mathbf{f}$ est associée la 1-forme $\tilde{f} = f_x dx + f_y dy + f_z dz$ et la circulation est simplement

$$C = \int_{\mathcal{C}} \tilde{f}$$

De même, le flux Φ d'un champ de vecteur $\mathbf{f}$ à travers une surface S est définie par

$$\Phi = \int_{\mathcal{S}} \mathbf{f}.\mathbf{n}.dS \tag{16.1}$$

où dS est un élément infinitésimal de surface, $\mathbf{n}$ est le vecteur "normal" à cet élément, et $\mathbf{f}.\mathbf{n}$ représente le produit scalaire. En langage des formes différentielles, au "vecteur" $\mathbf{f} = (f_x, f_y, f_z)^T$ est associée la 2-forme $\tilde{f} = f_x dydz + f_y dzdx + f_z dxdy$ et le flux est simplement

$$\Phi = \int_S \tilde{f} \tag{16.2}$$

La définition habituelle du flux en analyse vectorielle est problématique à cause de la définition du vecteur $\mathbf{n}$ normal à la surface en un point. A trois dimensions, un petit élément de surface peut éventuellement être représenté par un vecteur (qu'on appelle alors axial) ; à dimension supérieure cependant, un élément de surface ne peut pas être représentée par un vecteur et la définition (16.1) n'est plus valide. La définition (16.2) en terme de forme différentielle reste bien sûr toujours valide.

16.6 La dérivation extérieure.

Jusque là, les $p-$formes ne nous ont apporté rien de nouveau. Leur vrai beauté apparaît avec la dérivation. Nous allons d'abord donner la technique et nous viendrons ensuite sur le sens.

La dérivation extérieur transforme une $p-$forme en une $(p+1)$-forme. Le principe est le suivant : quand on rencontre une expression du genre $A(x, y, z, ...)dxdy...$, on prend le différentiel de A au sens usuel $(\partial A/\partial x)dx + \partial A/\partial y)dy + ...$ et on multiplie par l'élément $dxdy...$ qui était devant. Si on rencontre des $dxdx$, et bien cela vaut zéro et on ne s'en occupe pas ; enfin, on arrange de façon cohérente les divers $dxdz$, ... Cela sera sans doute plus claire à travers des exemples. Dans les exemples ci-dessous, nous prenons un espace à trois dimensions.

Exemple 16.2 Soit la $0-$forme $\omega = f(x, y, z)$. Alors, trivialement,

$$d\omega = (\partial f/\partial x)dx + (\partial f/\partial y)dy + (\partial f/\partial z)dz$$

Nous voyons donc que $d\omega$ représente ce qu'en général nous appelons un gradient et notons ∇f.

Exemple 16.3 Soit la $1-$forme $\omega = A(x, y, z)dx + B(x, y, z)dy + C(x, y, z)dz$. La dérivation de la première partie nous donne

$$\begin{aligned} d[A(x, y, z)dx] &= [(\partial A/\partial x)dx + (\partial A/\partial y)dy + (\partial A/\partial z)dz]\, dx & (16.3)\\ &= -(\partial A/\partial y)dxdy + (\partial A/\partial z)dzdx & (16.4)\end{aligned}$$

Nous somme passé de la première ligne à la seconde en notant que $dxdx = 0$ et nous avons réarrangé le $dydx$ en $-dxdy$. En continuant l'opération sur les deux autres

parties restantes et en regroupant les termes, nous trouvons finalement

$$\begin{aligned} d\omega &= [-(\partial A/\partial y) + (\partial B/\partial x)]\, dxdy \\ &+ [-(\partial B/\partial z) + (\partial C/\partial y)]\, dydz \\ &+ [(\partial A/\partial z) - (\partial C/\partial x)]\, dzdx \end{aligned}$$

Vous avez bien sûr reconnu ce qu'en analyse vectorielle nous noterions par une "rotationnelle". Il suffit, pour reconnaître les notations habituelles, de remplacer A par F_x, B par F_y et C par F_z.

Exemple 16.4 Soit la $2-$forme $\omega = A(x,y,z)dxdy + B(x,y,z)dydz + C(x,y,z)dzdx$. Le seul terme non nul de la dérivation de la première partie est le terme $(\partial A/\partial z)dzdxdy$ que l'on réarrange en $(\partial A/\partial z)dxdydz$ en permutant d'abord dz et dx, ensuite dz et dy ($-1 \times -1 = 1$, d'où le signe positif du réarrangement). En regroupant tout les termes, nous obtenons

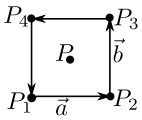
$$d\omega = [(\partial A/\partial z) + (\partial B/\partial x) + (\partial C/\partial y)]\, dxdydz$$

qui, en notations vectorielles, désignerait une "divergence" (Remplacer A par F_z, B par F_x et C par F_y pour les notations habituelles).

Nous voyons donc que les divers opérateurs différentiels de l'espace à trois dimensions ne sont que des formes déguisées de la dérivation (extérieure) des formes différentielles. C'est pour cela que le gradient, associé à une $1-$forme, est toujours intégré le long d'une courbe, tandis que la rotationnelle, associée à une $2-$forme, est intégrée sur une surface (on calcule toujours le flux d'une rotationnelle à travers une surface) ; enfin, la divergence, associée à une $3-$forme est toujours intégrée dans un volume. Noter également que les opérateurs différentielles habituels ne sont bien défini que dans l'espace à trois dimensions, tandis que la dérivation des formes extérieurs se fait indépendamment de la dimension et de façon presque mécanique, sans avoir à apprendre par cœur quoique ce soit. Si par ailleurs, vous êtes incapable de vous souvenir de la forme des opérateurs différentiels dans d'autres systèmes de coordonnées, transformez les formes ci-dessus en coordonnées polaire ou sphérique et dérivez les pour vous convaincre de la simplicité de manipulation des formes (voir les exercices).

Le point le plus fondamental est le corpus géométrique que les formes nous procurent et avec lequel nous allons nous familiariser par la suite.

Que signifie géométriquement la dérivation extérieur ? Commençons par la dérivation d'une 1-forme ω à un point P de l'espace. Donnons nous deux "petits" vecteurs $\vec{a}$ et $\vec{b}$ et faisons un circuits $\mathcal{C}$ autour du point P en suivant alternativement ces deux vecteurs. Appelons I l'intégrale de notre 1-forme le long de ce chemin. Nous définissons alors la 2-formes $d\omega$ comme la $2-$forme qui, appliquée



au bi-vecteur $\vec{a} \wedge \vec{b}$, produirait le même scalaire I :

$$d\omega(\vec{a} \wedge \vec{b}) \stackrel{def}{=} \int_{\mathcal{C}} \omega$$

à la limite quand les deux vecteurs $\to 0$. La technique que nous avons donnée plus haut calcule explicitement la 2-forme. Voyons cela de plus près. Prenons pour simplifier la 1-forme $\omega = f(x,y)dx + g(x,y)dy$ et posons $\vec{a} = (2h, 0)^T$ et $\vec{b} = (0, 2k)^T$. Si (x, y) sont les coordonnées du point P, les coordonnées du milieu du segment P_1P_2 par exemple est $(x, y-k)$. Donc,

$$\begin{aligned} \omega(\overrightarrow{P_1P_2}) + \omega(\overrightarrow{P_3P_4}) &= (f(x, y-k) - f(x, y+k))\,(2h) \\ &= -\frac{\partial f}{\partial y}(4hk) \end{aligned}$$

En calculant l'application de ω aux deux autres vecteurs restants, nous trouvons finalement que

$$\int_{\mathcal{C}} \omega = \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right)(4hk)$$

Ce qui est exactement ce que produit la 2-forme $d\omega$ appliquée à $\vec{a} \wedge \vec{b}$.

Cette construction se généralise aisément à la dérivation des $n-$formes. Pour la dérivation d'une 2-forme ω par exemples, nous nous donnons trois "petits" vecteurs autours d'un point P, et calculons la somme I de l'application de la 2-forme à toutes ces surfaces (correctement orientées). Il existe une 3-forme qui, appliquée au tri-vecteur en question, produit le même nombre et correspond bien sûr à notre $d\omega$.

Lemme de Poincaré.

Ce lemme est quelque chose de tellement évident qui n'a pas mérité le nom de théorème ; ceci dit, nous l'utilisons constamment dans divers contexte en lui donnant des noms différents (par exemple les deux premières équations de Maxwell). Le voici :

$$d(d\omega) = 0$$

Dériver deux fois une $k-$forme produit la $(k+2)-$forme nulle ! Prenons par exemple une 0-forme dans l'espace à 2 dimensions $f(x, y)$ et dérivons là deux fois :

$$\begin{aligned} df &= \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \\ d^2f &= \left(-\frac{\partial^2 f}{\partial y \partial x} + \frac{\partial^2 f}{\partial x \partial y}\right)dxdy = 0 \end{aligned}$$

Ceci a un caractère très général : quelque soit la forme que vous prenez, en la dérivant deux fois, vous tombez sur des expressions où nous avons des dérivées secondes croisées qui apparaissent deux fois avec des signes opposées.

Si par ailleurs, le lecteur a bien compris la signification géométrique de la dérivation extérieure, il n'aura pas de mal à démontrer le lemme de façon générale, sans faire appel aux coordonnées.

§ 16.1 Démontrer pourquoi en calcul vectoriel, nous avons les identités

$$\begin{aligned} \nabla \times (\nabla f) &= 0 \\ \mathrm{div}(\nabla \times \mathbf{u}) &= 0 \end{aligned}$$

Changement de variable.

Une des beautés des formes différentielles est la facilité qu'elles ont à gérer les changements de variables, de façon presque automatique. Prenons par exemple la forme $\omega = dxdy$ en coordonnées cartésiennes à 2 dimensions qui nous sert à mesurer l'aire d'une courbe fermée. En coordonnées polaire nous avons

$$\begin{aligned} x = r\cos\theta \quad &; \quad y = r\sin\theta \\ dx = \cos\theta dr - r\sin\theta d\theta \quad &; \quad dy = \sin\theta dr + r\cos\theta d\theta \end{aligned}$$

et donc $\omega = rdrd\theta$. Obtenir une forme dans un autre système de coordonnées et juste une question de multiplication des formes. Cet exemple nous sert également à illustrer que la dérivation extérieure est indépendante du système de coordonnées choisi. En système cartésien, nous avons par exemple $\omega = d\eta$ où $\eta = (xdy - ydx)/2$. En coordonnées polaires, le même calcul nous amène à

$$\eta = (1/2)r^2 d\theta$$

et nous voyons bien que $d\eta = rdrd\theta$. La "recette" pour effectuer la dérivation extérieure se charge de rendre cette opération libre des coordonnées.

§ 16.2 Démontrer cela de façon générale.

Manipulation des dérivations.

La première chose à maîtriser, comme dans le cas de la dérivation usuelle, est la règle de la dérivation d'un produit, qui s'écrit légèrement différemment :

$$d(\omega_1\omega_2) = (d\omega_1)\omega_2 + (-1)^{\deg(\omega_1)}\omega_1(d\omega_2)$$

où deg mesure la dimensionnalité de notre forme. Le signe est une conséquence de la commutativité des formes :

$$\omega_1\omega_2 = (-1)^{\deg(\omega_1)}\omega_2\omega_1$$

§ 16.3 Démontrer, en analyse vectorielle à 3 dimensions, que

$$\mathbf{div}(\vec{A}\times\vec{B}) = \mathbf{div}(\vec{A})\vec{B} - \vec{A}.\mathbf{div}(\vec{B})$$

Exemple fondamental : le champ électromagnétique.

Soit la 1-forme $A = -A_0dt + A_1dx + A_2dy + A_3dz$ dans l'espace à quatre dimension. Nous avons l'habitude d'appeler $A_0 = V$ le potentiel électrique et $\vec{A} = (A_1, A_2, A_3)^T$ le potentiel vecteur. Le signe $-$ dans la coordonnée associée à dt est dû à la signature de notre espace-temps. Nous pouvons construire la forme dA. La seule complication est d'ordonner correctement les choses à 4d :

$$\begin{aligned} dA &= -\left((\partial_x A_0)dx + (\partial_y A_0)dy + (\partial_z A_0)dz\right)dt \\ &- \left((\partial_t A_1)dx + (\partial_t A_2)dy + (\partial_t A_3)dz\right)dt \\ &+ \left(-\partial_y A_1 + \partial_x A_2\right)dxdy \\ &+ \left(-\partial_z A_2 + \partial_y A_3\right)dydz \\ &+ \left(\partial_z A_1 - \partial_x A_3\right)dzdx \end{aligned}$$

Nous avons pris la peine de séparer les différentes contributions selon les conventions usuelles [10]. Par exemple, la 1-forme qui multiplie dt

$$E = -(\partial_x A_0 + \partial_t A_1)dx + (\partial_y A_0 + \partial_t A_2)dy + (\partial_z A_0 + \partial_t A_3)dz$$

est appelé (tri-) vecteur "champ électrique" et défini (en analyse vectoriel) comme

$$\vec{E} = -\overrightarrow{\nabla V} - \partial_t \vec{A}$$

La 2-forme qui reste (tout ce qui ne contient pas dt)

$$B = \left(-\partial_y A_1 + \partial_x A_2\right)dxdy + \left(-\partial_z A_2 + \partial_y A_3\right)dydz + \left(\partial_z A_1 - \partial_x A_3\right)dzdx$$

est appelé tri-vecteur champ magnétique. Nous voyons ici pourquoi $\vec{B}$ n'est pas un vrai vecteur, puisqu'en réalité, il est associé à une 2-forme :

$$B_x = -\partial_z A_2 + \partial_y A_3; \ldots$$

ou encore, écrit dans les notations pédestres :

$$\vec{B} = \nabla \times \vec{A}$$

10. En donnant les indices 0 à 3 à nos coordonnées, nous avons simplement quelque chose du genre

$$dA = \left(\partial_{x^i} A_j - \partial_{x^j} A_i\right) dx^i dx^j$$

où nous nous sommes données quelques conventions pour ordonner correctement les paires (i, j) et sommer les indices répétés.

Séparer ainsi les coordonnées temporelles et spatiales nous oblige à alourdir inutilement nos notations, et pire, nous aveugler devant des évidences. Par le lemme de Poincaré, nous avons

$$d^2A = 0$$

En séparant péniblement les divers coordonnées, nous obtenons ce que l'on appelle les deux premières équations de Maxwell.

Voyons cela de plus près. Dans la dérivation de dA, nous voyons que seul les termes de la 2-forme B produisent des 3-formes en $dxdydz$, ce que nous écrivons, en notation vectoriel, par

$$\text{div}\vec{B} = 0$$

En regroupant maintenant les autres termes en par exemple $dxdydt$, ..., nous trouvons trois autres identités que l'on écrit, en notation vectorielle, par

$$\nabla \times \vec{E} = -\partial_t \vec{B}$$

Un peu plus de travail sur les formes différentielles nous montrerai que les deux autres équations de Maxwell s'écrivent comme

$$d(*dA) = \mu_0(*J)$$

où $*$ est appelé l'opération de Hodge (nous verrons le sens géométrique plus bas) et J est la 1-forme courant électrique : $J = -\rho dt + j_1 dx + j_2 dy + j_3 dz$.

16.7 théorème de Stockes.

Le théorème fondamental de l'analyse relie l'intégration de la dérivée d'une fonction aux valeurs de cette fonction aux bords :

$$\int_a^b F'(x)dx = F(b) - F(a) \tag{16.5}$$

Ceci est en fait une forme particulière du théorème de Stockes. Donnons nous un domaine [11] S de dimension p dans un espace de dimension n. Une sphère pleine (le point P de coordonnées (x, y, z) appartient à la sphère de rayon R centrée sur l'origine si $x^2 + y^2 + z^2 < R^2$) est par exemple un domaine de dimension trois dans un espace de dimension 3. La boule ($z = 0$, $x^2+y^2 < R^2$) est un domaine de dimension 2 dans un espace de dimension 3. Notons ∂S la frontière du domaine S. La coque $x^2+y^2+z^2 = R^2$ du premier exemple et le cercle $z = 0$, $x^2 + y^2 = R^2$ du deuxième exemple sont les frontières de leurs domaines. Le théorème de Stockes s'écrit

$$\int_{\partial S} \omega = \int_S d\omega \tag{16.6}$$

11. Nous avons à supposer que le domaine est compact.

où ω est une $(p-1)$-formes. C'est pour cela par exemple que l'intégrale d'une fonction (une 1-forme) le long d'une courbe fermée est égale à l'intégrale de la rotationnelle de cette fonction (la $2-$forme dérivée) à travers la surface délimité par cette courbe. En réalité, c'est ce théorème que les physiciens appelle théorème de Stockes

$$\int_S \mathbf{rotf}.d\mathbf{s} = \int_{\partial S} \mathbf{f}.d\mathbf{l} \tag{16.7}$$

Mais nous n'avons pas à nous limiter là. Le flux d'un champ de vecteur à travers une surface S (une $2-$forme) égale à l'intégrale de la divergence de ce champ (la $3-$forme dérivée) dans le domaine D délimité par la surface. En physique, nous écrivons cela comme

$$\int_S \mathbf{f}.d\mathbf{s} = \int_D \mathbf{divf}dV \tag{16.8}$$

Enfin, nous avons appris que si une fonction est le gradient d'une autre, alors son intégrale le long d'un chemin reliant les point A et B ne dépend pas du chemin :

$$\int_C \mathbf{grad}f.d\mathbf{l} = f(B) - f(A) \tag{16.9}$$

Comme vous le constatez, ces théorèmes de grad,rot et div ne sont que des applications du théorème de Stockes aux 1,2 et 3 formes dans un espace de dimension 3.

Aperçu de la démonstration du théorème de Stockes. La démonstration suit de très près la définition de la dérivée extérieure.

16.8 Intégration par partie.

Le théorème de Stockes nous donne la possibilité de généraliser l'intégration par partie

$$\int_I fg'dx = [fg]_I - \int_I f'gdx$$

à n'importe quelle dimension. Nous savons que

$$(d\omega_1)\omega_2 = d(\omega_1\omega_2) - (-1)^{\deg(\omega_1)}\omega_1(d\omega_2)$$

Nous en déduisons que

$$\begin{aligned}\int_S (d\omega_1)\omega_2 &= \int_S d(\omega_1\omega_2) - (-1)^{\deg(\omega_1)}\int_S \omega_1(d\omega_2)\\ &= \int_{\partial S} \omega_1\omega_2 - (-1)^{\deg(\omega_1)}\int_S \omega_1(d\omega_2)\end{aligned}$$

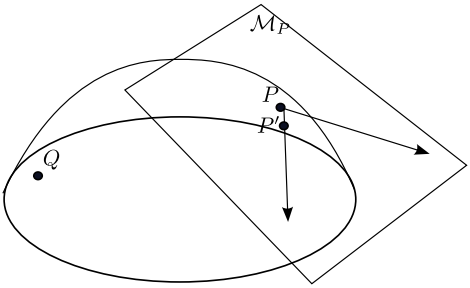


FIGURE 16.2 – L'espace (la variété différentielle) considéré ici est la surface de la sphère. L'espace tangent $\mathcal{M}_P$ au point P où l'on définit les vecteurs est le plan tangent au point P. Tant que le point P' est infiniment proche de P, nous pouvons le considérer appartenant en même temps au sphère et au plan, et utiliser un vecteur de $\mathcal{M}_P$ pour l'atteindre à partir de P par un vecteur. Pour un point plus lointain comme Q, cela n'est plus possible.

16.9 Un peu de géométrie : vecteurs, 1-formes et leurs associations.

16.9.1 La théorie.

Jusque là, nous avons travaillé dans des espaces (que nous avons munis d'un système de coordonnées quelconque), nous y avons défini des vecteurs et des formes et nous avons développer notre théorie. Or, la plupart des espaces qui intéresse les physiciens, comme par exemple notre espace habituel à trois dimensions, possède en plus, une *métrique* : nous savons mesurer la "taille" d'un vecteur et la *distance* entre deux points. Ce concept de métrique était absent dans les sections précédentes ; cependant, toutes les théories physiques que nous élaborons dépendent des *distances* entre points et nous devons maintenant ajouter ce concept si nous voulons que notre mathématiques servent à quelque chose. Cette section nous servira à faire la connexion entre les formes différentielles et d'une part les opérateurs différentiels (gradient, rot, ...) vu au chapitre 13 et d'autre part les tenseurs vu au chapitre 14. Le lecteur remarquera avec quel simplicité les formes automatisent les dérivations fastidieuses du chapitre 13.

Vecteurs. Commençons par les vecteurs. Prenons un espace, c'est à dire un ensemble de points muni du concept de "proche" [12]. Nous pouvons associer de façon bijective cet

12. Un ensemble de points n'a pas de structure, c'est un peu comme un sac de points. En se donnant le concept de "voisinage" et de proche, nous transformons ces points en une sorte de tissu, nous leur donnons une structure. On dit que nous avons muni notre espace d'une topologie. Noter que cela ne contient pas encore le concept de distance, qui est un concept beaucoup plus précis et restrictif que le concept de "proche". En

espace, que nous appelons une *variété différentielle* (*manifold* en anglais) à un ensemble $\mathbb{R}^n$, et nous appelons alors n la dimension de l'espace. Cette association s'appelle "se donner un système de coordonnées" et nous repérons alors un point P par ses coordonnées $(x^1, x^2, ...x^n)$. A chaque points P de l'espace, nous pouvons définir des *vecteurs* $\mathbf{v} = \vec{PP'}$ entre ce point et l'ensemble des points P' (infiniment) proche de lui. L'ensemble $\mathcal{M}_P$ de ces vecteurs $\{\mathbf{v}\}_P$ est appelé l'espace tangente du point P. Nous pouvons utiliser notre système de coordonnées pour définir une base dans cet espace vectoriel. Par exemple, $\mathbf{e}_1 = (1, 0, 0, ..., 0)$, $\mathbf{e}_2 = (0, 1, 0, ..., 0)$, ...Ainsi, pour le point le point $P = (x^1, x^2, ...x^n)$ et $P' = (x^1 + h^1, x^2 + h^2, ...x^n + h^n)$ (où nous supposons les h^i petit), nous pouvons écrire $P' = P + h^i\mathbf{e}_i$ (figure 16.2). Notez que par convention, un indice répété une fois en haut une fois en bas dénote une sommation. Cela nous évite d'alourdir les notations en traînant le symbole $\sum$ partout.

Notez que cette définition de vecteur ne contient pour l'instant aucune notion de la norme d'un vecteur (ceci viendra bientôt) ; elle est basée purement sur le système de coordonnée choisi.

Formes. Passons maintenant aux formes. Une 1-forme est une application linéaire qui s'applique à un vecteur et produit un nombre. L'ensemble $\mathcal{M}_p^*$ des 1-formes s'appliquant aux vecteurs de $\mathcal{M}_p$ forme lui même un espace vectoriel que nous appelons l'espace cotangente. La forme dx^i par exemple est telle que $dx^i(\mathbf{e}_j) = \delta^i_j$; ce sont ces formes (sous le nom de dx ou dy) que nous utilisions aux sections précédentes. Remarquez que comme nous avons un système de coordonnées, nous pouvons associer à chaque vecteur $\mathbf{v} = h^i\mathbf{e}_i$ une 1-forme $\omega = h_i dx^i$ où $h_i = h^i$. Cependant, cette association est quelque peu artificielle est dépend de notre choix de système de coordonnées. Tous les concepts géométriques sont indépendant des systèmes de coordonnées : ces derniers sont juste une aide pour la manipulation, mais ne doivent pas nous cacher les significations géométriques.

Notons que nous pouvons donner une définition un peu plus formelle des 1-formes dérivées des fonctions. Soit une fonction $f(.)$ définie sur notre variété différentielle. La 1-forme df est définie, au point P, par son action sur un vecteur quelconque $\mathbf{v} \in \mathcal{M}_P$ par

$$df(\mathbf{v}) = \frac{1}{\epsilon}\left(f(P + \epsilon\mathbf{v}) - f(P)\right)$$

quand $\epsilon \to 0$. Prenons par exemple la fonction $f(P) = x^i$, *i.e.* la fonction qui associe à un point P sa i-ème coordonnée. Nous voyons que pour la forme dx^i, nous avons $dx^i(\mathbf{e}_j) = \delta^i_j$.

Distance, produit scalaire et le tenseur métrique. Passons maintenant au concept qui transforme un espace topologique en un espace métrique et qui nous permet de

topologie, il n'y a pas de distinction entre cube et sphère, ce sont le même type d'objet.

distinguer une sphère d'une ellipsoïde : le concept de distance et de produit scalaire [13]. Nous avons déjà rencontré ce concept au début de ce cours (chapitre 2) et nous connaissons sa définition axiomatique : le produit scalaire associe à deux vecteurs, de façon linéaire, un *nombre*. En géométrie, se donner un produit scalaire fixe la structure de l'espace. Le produit scalaire $\langle \mathbf{u}, \mathbf{v} \rangle$ entre deux vecteurs $\mathbf{u}$ et $\mathbf{v}$ est indépendant du système de coordonnées. Cependant, si nous nous sommes donnés un système de coordonnées et un produit scalaire, nous pouvons exprimer le produit scalaire entre le vecteur $\mathbf{u} = h^i \mathbf{e}_i$ et le vecteur $\mathbf{v} = k^i \mathbf{e}_i$ par

$$\langle \mathbf{u}, \mathbf{v} \rangle = g_{ij} h^i k^j$$

Les quantités g_{ij} sont appelés les éléments du tenseur métrique g. Si on choisit un autre système de coordonnées, leurs valeurs changent de façon à laisser invariant $\langle \mathbf{u}, \mathbf{v} \rangle$. En général, g dépend du point P de l'espace. La distance entre deux point P et P' infiniment proche est donnée par

$$d(P, P') = \left\| \vec{PP'} \right\| = \sqrt{\left\langle \vec{PP'}, \vec{PP'} \right\rangle}$$

Exemple 16.5 Soit un système de coordonnées (x^1, x^2). Soit $\mathbf{u} = h^i \mathbf{e}_i$ et $\mathbf{v} = k^i \mathbf{e}_i$ deux vecteurs au point $P = (x^1, x^2)$. Le produit scalaire

$$\langle \mathbf{u}, \mathbf{v} \rangle = h^i k^j$$

qui est indépendant du point P définit un espace plat muni de coordonnées cartésiennes. Ici, le tenseur métrique est l'identité $g_{ij} = \delta_{ij}$. De même, si nous pouvons munir un espace de ce produit scalaire, nous disons que nous avons un espace plat muni de coordonnées cartésienne.

Exemple 16.6 Soit un système de coordonnées (x^1, x^2). Soit $\mathbf{u} = h^i \mathbf{e}_i$ et $\mathbf{v} = k^i \mathbf{e}_i$ deux vecteurs au point $P = (x^1, x^2)$. Le produit scalaire

$$\langle \mathbf{u}, \mathbf{v} \rangle = h^1 k^1 + \left(x^1\right)^2 h^2 k^2$$

dépend cette fois explicitement du point de l'espace P. La distance d entre le point P et le point voisin $P' = (x^1 + h^1, x^2 + h^2)$ est donné par

$$d(P, P') = \left(h^1\right)^2 + (x^1)^2 \left(h^2\right)^2$$

Ici, le tenseur métrique est

$$g = \begin{pmatrix} 1 & 0 \\ 0 & (x^1)^2 \end{pmatrix}$$

13. Le concept de distance est lui même contenu dans le concept des invariants. Se donner des *invariants*, comme le fait par exemple que tourner un bâton ne change pas sa taille, c'est se donner une distance. A la fin de XIX-ème siècle, la géométrie a été redéfinie en ces termes par des mathématicien tels que Felix Klein.

Si nous avions utilisé la notation $x^1 = r$ et $x^2 = \theta$, le lecteur aura immédiatement reconnu que nous sommes en train d'utiliser des coordonnées polaires.

Il est important de noter que dans ce système de coordonnées, au point $P = (r, \theta)$, nous avons pour les vecteurs de base $\|\mathbf{e}_r\| = 1$ et $\|\mathbf{e}_\theta\| = r$: un des vecteurs n'est pas normé. Nous pouvons bien sûr choisir d'autres vecteurs de base normés, comme par exemple $\mathbf{e}'_\theta = (1/r)\mathbf{e}_\theta$, mais le choix imposé par le système de coordonnées n'est pas normé.

§ 16.4 Donner le tenseur métrique à trois dimensions pour les coordonnées cylindriques et sphériques.

§ 16.5 Sur la surface de la sphère unité où les points sont repéré par (θ, ϕ), trouver le tenseur métrique hérité de la distance euclidienne.

§ 16.6 Démontrer que même si le tenseur métrique n'est pas symétrique, nous pouvons toujours le symétriser en choisissant

$$g'_{ij} = (g_{ij} + g_{ji})/2$$

et que cette symétrisation ne change pas le produit scalaire. Ceci est la raison pourquoi on choisit toujours un tenseur métrique symétrique.

Association forme-vecteur. Nous sommes maintenant en mesure d'établir une association intrinsèque entre les éléments de $\mathcal{M}_P$ et $\mathcal{M}_P^*$: à un vecteur $\mathbf{u}$, nous associons la forme $\omega = \langle \mathbf{u}, . \rangle$. L'action de cette forme sur un vecteur $\mathbf{v}$ est

$$\omega(\mathbf{v}) = \langle \mathbf{u}, \mathbf{v} \rangle$$

Comme le produit scalaire ne dépend pas du choix du système de coordonnées, cette association l'est également. Cette association est souvent appelée isomorphisme musicale, et on note $\omega = \mathbf{u}^\flat$. En calcul tensoriel, on appelle cette opération "abaissement d'indice". Dans les notations de Dirac en mécanique quantique, un vecteur est noté par un ket, comme par exemple $|u\rangle$, et sa forme associé par un *bra* $\langle u|$. L'opération inverse, l'association d'un vecteur à une forme, est notée $\mathbf{u} = \omega^\sharp$. En calcul tensoriel, on appel cela "élévation d'indice".

Jusque là, nous voyions les formes comme des objets agissant sur des vecteurs, mais la définition ci-dessus nous permet de rendre cette opération complètement symétrique et considérer les vecteurs comme des objets agissant sur des 1-formes. Ainsi, $\mathbf{v}(\omega) = \langle \mathbf{u}, \mathbf{v} \rangle$ où $\omega = \mathbf{u}^\flat$. De même, le produit scalaire dans l'espace cotangente se définit à partir de sa définition dans l'espace tangente : $\langle \omega, \eta \rangle = \langle \omega^\sharp, \eta^\sharp \rangle$.

Nous pouvons expliciter cette association dans un système de coordonnés quelconque. Soit le vecteur $\mathbf{u} = h^i \mathbf{e}_i$ et sa forme associée $\omega = \mathbf{u}^\flat = \alpha_i \tilde{e}^i$. Pour éviter les confusion entre vecteur, nombre et forme, nous mettons pour l'instant un tilde sur les éléments de base des formes : $\tilde{e}^j(\mathbf{e}_i) = \delta_i^j$. Soit maintenant un vecteur $\mathbf{v} = k^i \mathbf{e}_i$. Par définition

des formes et du produit scalaire, nous avons

$$\begin{aligned}\omega(\mathbf{v}) &= \alpha_i k^i \\ \langle \mathbf{u}, \mathbf{v} \rangle &= g_{ij} h^j k^i\end{aligned}$$

En comparant ces deux expressions, nous voyons que nous devons avoir

$$\alpha_i = g_{ij} h^j$$

Si nous connaissons les composants du vecteur, les composants de sa forme associée s'obtiennent par une simple multiplication tensorielle [14]. Il est d'ailleurs d'usage de noter les composant de la forme par le même symbole que les composant de son vecteur associé, mais avec un indice abaissée. Ainsi, à la place de α_i, nous aurions du utiliser h_i et écrire la relation comme $h_i = g_{ij} h^j$.

En inversant cette relation, nous pouvons trouver les composants du vecteur si nous connaissons les composants de la 1-forme :

$$h^i = g^{ij} h_j$$

où les g^{ij} sont les éléments de l'inverse du tenseur métrique g^{-1}. De façon symétrique, le produit scalaire entre deux formes de composantes h_i et k_i est donnée par $g^{ij} h_i k_j$.

Dans la plupart des cas, nous utilisons un système de coordonnées où le tenseur métrique est diagonale : $g_{ij} = 0$ si $i \neq j$. Dans ce cas, le tenseur inverse s'obtient par une simple inversion des nombres : $g^{ii} = 1/g_{ii}$ et le passage des composants de formes au vecteur s'obtient facilement.

Exemple 16.7 Dans l'espace de Minkowski de la relativité restreinte de dimension 4, en coordonnée cartésiennes, le tenseur (pseudo-)métrique est diagonale et donnée par $g_{00} = -1$; $g_{ii} = 1$ pour $i > 0$. Au quadri-vecteur (A_0, A_1, A_2, A_3) est associée la quadri 1-forme $(-A_0, A_1, A_2, A_3)$.

16.9.2 Application au calcul vectoriel.

En physique, nous manipulons constamment des objets que l'on appelle gradient, rotationnel, divergence et laplacien. Au chapitre 13, nous avions vu comment obtenir l'expression de ces objets dans un système de coordonnées quelconque. Nous pouvons maintenant les revisiter à travers le concept des formes différentielles et de façon beaucoup plus générale. Cependant, pour ne pas alourdir trop les expressions, nous nous contentons de cas où le tenseur métrique est symétrique et les coordonnées curvilignes sont orthogonales.

14. En langage courant, on obtient un vecteur ligne en multipliant une matrice et un vecteur colonne

Le gradient. Soit une variété différentielle où nous avons une fonction intrinsèque $f(.)$. Par fonction intrinsèque, nous entendu une fonction qui à chaque point de l'espace associe un nombre, indépendemment du système de coordonnées choisit. Par exemple, une fonction qui mesure la courbure locale de l'espace, ou en physique, une fonction qui mesure la densité ou la température en un point.

Si nous connaissons la valeur de la fonction au point P, sa valeur au point P' proche est donnée par

$$f(P') = f(P) + \nabla f.\vec{PP'}$$

où ∇f est appelé (vecteur) gradient. Ce vecteur est celui associé à la forme df. L'expression des composants de ∇f dans n'importe quel système de coordonnées s'obtient facilement si nous comprenons son association avec les formes.

Donnons nous un système de coordonnées orthogonales x^i et le tenseur métrique g_{ij} diagonal dans ce système. Nous avons

$$df = (\partial f/\partial x^i)dx^i$$

Nous avons donc, pour les composantes du vecteur gradient,

$$(\nabla f)^i = g^{ij}(\partial f/\partial x^j)$$

En coordonnées orthogonales, où $g_{ij} = g^{ij} = 0$ si $i \neq j$, et $g^{ii} = 1/g_{ii}$, nous avons [15]

$$\nabla f = \sum_{i=1}^{n} \frac{1}{g_{ii}} \frac{\partial f}{\partial x^i} \mathbf{e}_i \tag{16.10}$$

Remarquez que la base naturelle est en général non-normée : $\|\mathbf{e}_i\| = \sqrt{g_{ii}}$. Si nous choisissons une base normée $\mathbf{e}'_i = (1/\sqrt{g_{ii}})\mathbf{e}_i$, nous avons alors

$$\nabla f = \sum_{i=1}^{n} \frac{1}{\sqrt{g_{ii}}} \frac{\partial f}{\partial x^i} \mathbf{e}'_i \tag{16.11}$$

Dans la plupart des problèmes de physique, nous utilisons des bases normées et l'expression (16.11) est celle qui est le plus rencontrée.

Exemple 16.8 A deux dimensions et en coordonnées polaires (r, θ), nous avons $g_{rr} = 1$ et $g_{\theta\theta} = r^2$. Le gradient est donc

$$\nabla f = \frac{\partial f}{\partial r}\mathbf{u}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\mathbf{u}_\theta$$

où $\mathbf{u}_r$ et $\mathbf{u}_\theta$ sont les vecteurs normés radial et orthoradial.

§ 16.7 Donner l'expression du gradient en coordonnées cylindriques et sphériques.

15. Nous utilisons explicitement le symbole $\sum$ quand il peut y avoir ambiguïté.

Le rotationnel. Soit un champ de vecteurs intrinsèque $\mathbf{v}(P)$ dans notre espace. Nous parlons par exemple du champ de vitesse dans un fluide ou du champ électromagnétique. Nous souhaitons calculer son rotationnel. Le rotationnel n'est pas un vecteur, mais un bi-vecteur. Dans l'espace à trois dimension, nous pouvons faire une bijection entre les bi-vecteurs et les vecteurs, et c'est pourquoi, à cette dimension, nous parlons du vecteur rotationnel. Pour rester simple, nous nous contentons de dimension trois ici et laissons la généralisation au lecteur à travers un exercice. Le développement consiste à (i) passer du vecteur à la 1-formes ; (ii) dériver la forme ; (iii) revenir de l'espace des deux formes à l'espace des bi-vecteurs.

Au vecteur $\mathbf{v} = v^i \mathbf{e}_i$, nous associons la 1-forme

$$\omega = \mathbf{v}^\flat = \sum g_{ii} v^i \tilde{e}^i$$

où nous avons noté $\tilde{e}^i$ à la place de dx^i pour plus de symétrie. La dérivation extérieure de cette forme nous donne

$$d\omega = \left(\frac{\partial(g_{22}v^2)}{\partial x^1} - \frac{\partial(g_{11}v^1)}{\partial x^2} \right) \tilde{e}^1 \wedge \tilde{e}^2 + \text{permut. circ.}$$

le bi-vecteur associée à cette 2-forme est donc

$$\overrightarrow{\text{rot}}\mathbf{v} = \frac{1}{g_{11}g_{22}} \left(\frac{\partial(g_{22}v^2)}{\partial x^1} - \frac{\partial(g_{11}v^1)}{\partial x^2} \right) \mathbf{e}_1 \wedge \mathbf{e}_2 + \text{permut. circ.} \tag{16.12}$$

Notez à nouveau que le rotationnel est donné ici dans une base non-normée. Si nous utilisons une base normée $\mathbf{e}'_i = (1/\sqrt{g_{ii}})\mathbf{e}_i$, alors le champs de vecteur s'écrit $\mathbf{v} = v'^i \mathbf{e}'_i$ où $v^i = v'^i/\sqrt{g_{ii}}$. En récrivant l'expression (16.12) dans cette nouvelle base, nous avons

$$\overrightarrow{\text{rot}}\mathbf{v} = \frac{1}{\sqrt{g_{11}}\sqrt{g_{22}}} \left(\frac{\partial(\sqrt{g_{22}}v'^2)}{\partial x^1} - \frac{\partial(\sqrt{g_{11}}v'^1)}{\partial x^2} \right) \mathbf{e}'_1 \wedge \mathbf{e}'_2 + \text{permut. circ.} \tag{16.13}$$

Exemple 16.9 En coordonnées cylindrique (r, θ, z), nous avons $g_{rr} = 1$ et $g_{\theta\theta} = r^2$ et $g_{zz} = 1$. Le composant du rotationnel selon $\mathbf{u}_r \wedge \mathbf{u}_\theta$ (que l'on associe à la composante selon le vecteur $\mathbf{u}_z$) est donc

$$\left(\overrightarrow{\text{rot}}\mathbf{v}\right)_z = \frac{1}{r} \left(\frac{\partial(r\, v^\theta)}{\partial r} - \frac{\partial(v^r)}{\partial \theta} \right)$$

§ 16.8 Donner l'expression complète du rotationnel en cylindrique et sphérique.

La divergence. Le calcul de la divergence procède de la même façon que précdément. Il faut seulement noter que la divergence est définie dans un $n-$volume et est le résultat de l'intégration d'un $n-1$ vecteur. Considérons, à trois dimensions, un bi-vecteur

$$\boldsymbol{r} = v^3 \mathbf{e}_1 \wedge \mathbf{e}_2 + v^1 \mathbf{e}_2 \wedge \mathbf{e}_3 + v^2 \mathbf{e}_3 \wedge \mathbf{e}_1$$

auquel on associe un 2-forme

$$\omega = g_{11}g_{22}v^3\tilde{e}^1 \wedge \tilde{e}^2 + \text{permut. circ.}$$

La dérivée de cette forme nous donne la 3-forme

$$d\omega = \left(\frac{\partial(g_{11}g_{22}v^3)}{\partial x^3} + \frac{\partial(g_{22}g_{33}v^1)}{\partial x^1} + \frac{\partial(g_{11}g_{33}v^3)}{\partial x^2} \right) \tilde{e}^1 \wedge \tilde{e}^2 \wedge \tilde{e}^3$$

et un retour dans l'espace des vecteur produit

$$\text{div}\boldsymbol{r} = \frac{1}{g_{11}g_{22}g_{33}} \left(\frac{\partial(g_{11}g_{22}v^3)}{\partial x^3} + \frac{\partial(g_{22}g_{33}v^1)}{\partial x^1} + \frac{\partial(g_{11}g_{33}v^3)}{\partial x^2} \right) \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3$$

Comme précédemment, cette expression est pour des vecteurs de base non-normés. Pour des vecteur normés, on trouve l'expression habituel

$$\text{div}\boldsymbol{r} = \frac{1}{\sqrt{g_{11}g_{22}g_{33}}} \left(\frac{\partial(\sqrt{g_{11}g_{22}}v^3)}{\partial x^3} + \frac{\partial(\sqrt{g_{22}g_{33}}v^1)}{\partial x^1} + \frac{\partial(\sqrt{g_{11}g_{33}}v^3)}{\partial x^2} \right) \mathbf{e}'_1 \wedge \mathbf{e}'_2 \wedge \mathbf{e}'_3$$

Le triple produit vectoriel, au signe près vaut une fois le volume unité. Le signe dépend de l'orientation de l'espace. Bien sûr, un 3-vecteur ne peut pas valoir un scalaire, mais l'espace des 3-vecteurs est de dimensions 1, et nous les associons donc aux scalaires.

16.10 L'opérateur de Hodge.

$n-$forme volume. Un opérateur différentiel fondamental dont nous n'avons pas encore parlé est le laplacien. Nous savons qu'en analyse vectoriel, cet opérateur est relié à une *dérivée seconde*, mais nous ne pouvons évidement pas l'utiliser tel quel, puisque d'après le lemme de Poincaré, pour une $k-$forme ω, $d(d\omega) = 0$. L'opérateur de Hodge nous permet de donner un sens précis au "laplacien" pour les k-forme.

Dans l'espace à n dimension, il existe une $n-$forme fondamentale que nous appelons *volume*. Soit une région de l'espace D et soit une fonction $f()$ tel que $f(P) = 1$si $P \in D$ et $f(P) = 0$ sinon. Alors

$$\int f(P)\phi$$

mesure le *volume* de la région D. Dans un espace plat muni de coordonnées cartésiennes x^i, la $n-$forme volume est simplement

$$\phi = dx^1 ... dx^n$$

Dans un espace muni de coordonnées curviligne et un tenseur métrique g, au $n-$vecteur $\mathbf{e}'_1 \wedge \mathbf{e}'_2 ... \wedge \mathbf{e}'_3$ correspond la $n-$forme volume

$$\phi = \sqrt{|\text{det}g|} dx^1 ... dx^n$$

où det$g$ et le déterminant du tenseur métrique, autrement dit, dans les notations de la sous section précédente où g est supposé diagonal,

$$\text{det}g = g_{11}g_{22}...g_{nn}$$

En coordonnées cylindrique, $\det g = r^2$ et nous déduisons donc que $\phi = rdrd\theta dz$, ce que nous aurions également pu obtenir par changement de coordonnées en partant de $\phi = dxdydz$. De même, en coordonnées sphérique, $\phi = r^2 \sin\theta drd\theta d\phi$.

L'étoile de Hodge. L'opérateur de Hodge $\star$ associe (de façon unique et linéaire) à une $k-$forme $\omega = dx^1...dx^k$ une $(n-k)-$forme $\star\omega$ telle que

$$\eta(*\omega) = \langle \eta, \omega \rangle \, \phi \;\; \forall \eta \tag{16.14}$$

En gros, $*\omega$ complète ω pour former (à un facteur près), le $n-$forme volume. L'association d'un vecteur à un bi-vecteur telle que nous le faisons pour le rotationnel est en faite une telle opération.

Nous avons pas mal utilisé cet opérateur sans lui donner un nom. Prenons l'espace à $n = 3$ muni de coordonnées cartésiens où $g = 1$. Alors, par exemple,

$$*dx = dydz \;;\; *dy = dzdx \;;\; *dz = dxdy$$

Nous voyons alors que $dx^i(*dx^i) = 1dxdydz$ pour $i = 1, 2, 3$. La seule petite complication est de bien gérer le signe, c'est à dire le degrés de permutation que cela nous impose. De même, il est facile de voir que $*1 = dxdydz$ et $*dxdydz = 1$. Listons quelques propriétés de $*$:

1. $*$ est linéaire.
2. $*(f\omega) = f(*\omega)$ où f est une fonction et ω une $k-$forme.
3. $*1 = \phi$ et $*\phi = 1$

La relation (16.14) est en fait une recette explicite pour calculer $*\omega$. Prenons une 1-forme dx^k et calculons $*dx^k$. D'abord, nous voyons que

$$dx^i(*dx^k) = 0 \;\; \text{si}\, i \neq k$$

et par ailleurs,

$$dx^k(*dx^k) = \left\langle dx^k, dx^k \right\rangle \phi = \frac{1}{g_{kk}} \sqrt{|\text{det}g|} dx^1...dx^n$$

Nous en déduisons donc que

$$*dx^k = (-1)^{k+1} \frac{1}{g_{kk}} \sqrt{|\text{det}g|} dx^1...dx^{k-1}dx^{k+1}...dx^n$$

Exemple 16.10 Coordonnées polaires

En coordonnées polaire (r, θ, z), nous avons $g_{rr} = 1$, $g_{\theta\theta} = r^2$ et $g_{zz} = 1$. Nous avons donc

$$*dr = rd\theta dz \tag{16.15}$$
$$*d\theta = -\frac{1}{r} dr dz \tag{16.16}$$
$$*dz = rdrd\theta \tag{16.17}$$

Exemple 16.11 Coordonnées sphériques

En coordonnées sphérique (r, θ, ϕ), nous avons $g_{rr} = 1$, $g_{\theta\theta} = r^2$ et $g_{\phi\phi} = r^2 \sin^2\theta$ et $\sqrt{|\det g|} = r^2 \sin\theta$ Nous avons donc

$$*dr = r^2 \sin\theta d\theta d\phi \tag{16.18}$$
$$*d\theta = -\sin\theta dr d\phi \tag{16.19}$$
$$*d\phi = \frac{1}{\sin\theta} dr d\theta \tag{16.20}$$

Le Laplacien. Prenons maintenant une fonction $f(x, y, z)$, c'est à dire une $0-$forme. Nous avons alors

$$\begin{aligned} df &= \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz \\ *df &= \frac{\partial f}{\partial x}dydz + \frac{\partial f}{\partial y}dzdx + \frac{\partial f}{\partial z}dxdy \\ d(*df) &= \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}\right) dxdydz \end{aligned}$$

Et nous voyons que $d(*df)$ nous donne bien le laplacien sous forme d'une $3-$forme. En d'autres termes,

$$\Delta f = *d * df$$

En notant $\delta = *d$, nous voyons qu'en langage de forme différentielle, $\Delta f = \delta^2 f$. Par ailleurs, il suffit de connaître $*dx^i$ dans un système de coordonnées pour pouvoir dériver aisément l'expression du Laplacien dans ce système.

Exemple 16.12 Coordonnées polaires

Soit une fonction $f(r, \theta, z)$. Nous avons

$$df = \frac{\partial f}{\partial r}dr + \frac{\partial f}{\partial \theta}d\theta + \frac{\partial f}{\partial z}dz$$

et donc, en utilisant les expressions (16.15-16.17), nous avons

$$*df = r\frac{\partial f}{\partial r}d\theta dz - \frac{1}{r}\frac{\partial f}{\partial \theta}dr dz + r\frac{\partial f}{\partial z}dr d\theta$$

et en dérivant une fois de plus

$$\begin{aligned} d(*df) &= \left\{ \frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r}\frac{\partial^2 f}{\partial \theta^2} + r\frac{\partial^2 f}{\partial z^2} \right\} dr d\theta dz \\ &= \left\{ \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial f}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial z^2} \right\} r dr d\theta dz \end{aligned}$$

Pour passer de la première ligne à la seconde, nous avons simplement sorti r de $\{\}$; ainsi, le terme restant en dehors de $\{\}$ a exactement la forme de $n-$volume ϕ. Comme $*\phi = 1$, nous en déduisons, en appliquant un $*$ supplémentaire, que le terme entre $\{\}$ dans la seconde ligne est le Laplacien.

Exemple 16.13 Coordonnées sphériques

Soit une fonction $f(r, \theta, \phi)$. Nous avons

$$df = \frac{\partial f}{\partial r}dr + \frac{\partial f}{\partial \theta}d\theta + \frac{\partial f}{\partial \phi}d\phi$$

En utilisant les expressions (16.18-16.20), nous avons

$$*df = r^2 \sin\theta \frac{\partial f}{\partial r} d\theta d\phi - \sin\theta \frac{\partial f}{\partial \theta} dr d\phi + \frac{1}{\sin\theta}\frac{\partial f}{\partial \phi} dr d\theta$$

En dérivant une fois de plus et en mettant $r^2 \sin\theta$ en facteur, nous avons

$$d*df = \left\{ \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial f}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 f}{\partial \phi^2} \right\} r^2 \sin\theta dr d\theta d\phi$$

Le terme entre $\{\}$ est donc le Laplacien en coordonnées sphérique.

§ 16.9 Donner l'expression complète de Laplacien dans un système de coordonnées curvilignes orthogonales quelconque.

16.11 Quelques applications.

16.11.1 Équation de conservation.

Vous rencontrez cette équation partout en physique, sous des formes diverses et variées. Elle dit simplement que la variation de "quelque chose" dans un volume égale la quantité de ce "quelque chose" qui entre dans ce volume moins la quantité de ce "quelque chose" qui en sort. Le "quelque chose" peut être la concentration d'une substance, l'énergie emmagasinée dans le volume, une probabilité de présence, ... [16] Les

16. Le quelque chose peut être l'argent d'une entreprise, et les comptables sont responsable, sur leur denier personnel, de faire respecter cette loi. L'étudiant en physique perd au plus quelques points à l'examen.

systèmes qui obéissent à cette loi sont dits *conservatif* s. La quantité entrante moins la quantité sortante se dit flux. Notons par ρ le "quelque chose" et par $\mathbf{J}$ son flux. Appliquée à un volume infinitésimal $dxdydz$, l'équation de conservation s'écrit

$$\partial\rho/\partial t + \mathbf{div}\mathbf{J} = 0 \tag{16.21}$$

Si nous parlons d'un fluide en mouvement, alors $\rho(\mathbf{x})$ est la concentration en un point de l'espace, $\mathbf{v}(\mathbf{x})$ la vitesse du fluide et $\mathbf{J} = \rho(\mathbf{x})\mathbf{v}(\mathbf{x})$. Dans le cas d'une substance qui diffuse, le flux $\mathbf{J}$ est proportionnel au gradient de la concentration [17] (la chaleur diffuse du chaud vers le froid, l'encre se dilue dans l'eau,...) $\mathbf{J} = -K\mathbf{grad}\rho$, ce qui, réinjecté dans l'équation de conservation, nous donne l'équation de diffusion $\partial\rho/\partial t = K\nabla^2\rho$.

En langage des formes, cette équation acquiert une interprétation géométrique. Prenons d'abord le cas de l'espace à une dimension spatiale et une dimension temporelle, et considérons un élément infinitésimal dt, dx (les deux cotés d'un carré dans l'espace-temps, si vous voulez) et la forme $\omega = \rho dx - Jdt$ [18]. L'équation de conservation n'est alors rien d'autre que

$$d\omega = 0$$

A 3+1 dimensions (ou plus si affinité), il faut considérer la forme

$$\omega = \rho dxdydz - (J_x dydz + J_y dzdx + J_z dxdy)dt$$

Cette forme est analogue à ce qu'en physique nous appelons un quadri-vecteur. Nous verrons plus bas une application de ce concept au champ électromagnétique.

Exercices.

§ 16.10 **Périmètre et surface.** Soit A l'aire enfermée par une courbe $\mathcal{C}$. Démontrer que nous avons

$$A = (1/2)\int_{\mathcal{C}} xdy - ydx$$

En déduire que l'aire d'une ellipse est πab, où a et b sont les axes majeur et mineur. Comment cette expression est reliée à l'expression habituelle

$$A = \int_a^b f(x)dx$$

de la surface entre une courbe et l'axe x ?

§ 16.11 **Air d'une courbe fermée.** Nous souhaitons calculer l'air enfermée par une courbe $\mathcal{C}$ donnée en coordonnées polaires par l'équation

17. Cela se démontre facilement en physique statistique et s'appelle la réponse linéaire.
18. Le signe "-" vient de notre convention de compter en négatif le flux entrant. Cela paraît aussi arbitraire que la charge de l'électron. Historiquement, cela vient du fait que la surface est orientée pour que la normale pointe vers l'extérieur.

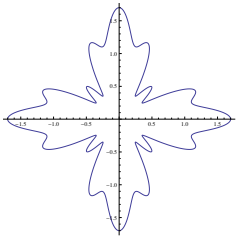


FIGURE 16.3 – Une *feuille* donnée par l'équation $r = 1+(1/2)\cos(4\theta)+(1/5)\cos(16\theta)$

$$r(\theta) = \sum_{n=0}^{N} a_n \cos(n\theta)$$

dont un exemple est donné par la figure 16.3.

1. Démontrer que $\int_0^{2\pi} \cos^2(n\theta)d\theta = \pi$ si $n \neq 0$ et 2π si $n = 0$. De même, démontrer que $\int_0^{2\pi} \cos(n\theta)\cos(m\theta)d\theta = 0$ si $m \neq n$.
2. Démontrer que la 2-forme différentielle $\omega = dxdy$ dérive de la 1-forme $\eta = (1/2)(xdy - ydx)$.
3. Donner l'expression de ω et η en coordonnées polaires.
4. Vérifier qu'en coordonnées polaires, nous avons bien $\omega = d\eta$.
5. En utilisant la forme polaire de ces formes, et en utilisant le théorème de Stockes, démontrer que l'air enfermée par la courbe vaut

$$A = \pi\left(a_0^2 + (1/2)\sum_{n=1}^{N} a_n^2\right)$$

§ 16.12 **Volume.** Démontrer que la 3-forme volume $dxdydz$ dérive de la 2-forme

$$\frac{1}{3}(xdydz + ydzdx + zdxdy)$$

en déduire le volume d'une sphère, d'un ellipsoïde de révolution et d'un ellipsoïde générale.

§ 16.13 **Association bi-vecteur, 2-forme.** Dans un espace à 3 dimensions, donner la relation entre bi-vecteur et 2-formes en en coordonnées généralisées orthogonale. Utiliser ce résultat pour trouver l'association en coordonnés cylindriques et sphériques.

§ 16.14 **Opérateurs différentielles en analyse vectorielle.** Trouver l'expression du laplacien à $n = 3$ en coordonnées généralisées orthogonales. Utiliser ce résultat pour trouver ce résultat en coordonnées cylindriques et sphériques.

§ 16.15 **D'Alembertien.** Dans l'espace-temps ($n = 4$), la 4-forme volume est donnée en coordonnées cartésienne par

$$\phi = -dtdxdydz$$

Étant données la fonction $f(t,x,y,z)$, donner l'expression de $d(*df)$. C'est ce que l'on appelle couramment le D'Alembertien qui gouverne toutes les équations d'ondes.

§ 16.16 **Équations de Maxwell.** En partant du lagrangien (**??**), faire une variation sur A, intégrer par partie les formes résultantes pour déduire les équations du Maxwell en présence du champ.

quelques exemples d'intégration des formes.
est ce que $dw = 0$ implique $w = d\alpha$? exemple pratique sur le rot.
équation de maxwell complète
transformation de Lorentz et les équations de Maxwell, la modification du courant,
relativité
démonstration de stockes pour les deux formes.
l'opérateur *, le laplacien $d(*df)$, en déduire le laplacien en coordonnées curviligne, le lien entre la minimisation de $df \wedge *df$ et $d(*df) = 0$. Éventuellement, si possible, faire le lien avec le chapitre précédent et le calcul des surfaces minimales.
Peut être faire une vague introduction à la géométrie des surfaces et la relation avec les courbures.
Faire le lien avec les tenseurs antisymétriques du prochain chapitre.

17 Théorie des fonctions analytiquess.

17.1 Introduction.

De très bon livres sont consacrés à la théorie des fonctions complexes, nous ne donnerons ici qu'un très bref aperçu pour donner le rudiment nécessaire à la manipulation de ces fonctions. Le lecteur intéressé pourra, (devra) se reporter à un livre consacré à ce thème pour goûter toute la beauté de ces fonctions.

Dans ce chapitre, nous allons étudier les fonctions d'une variable *complexe* $f(z)$, en étendant ce que nous savons de la théorie des fonctions réelles. Le lecteur est déjà quelque peu habitué à cette théorie, puisque l'utilisation par exemple de la fonction $\exp(z)$ ne lui pose pas de problème.

Dans ce que nous allons voire, certains aspects des fonctions analytiques peuvent paraître "miraculeux" comparés à ce que nous savons sur les fonctions réelles. Avant d'aller plus loin, j'aimerai relever le caractère miraculeux.

La première surprise que l'on verra plus bas dans les fonctions analytiques est qu'il suffit de les connaître dans un petit domaine pour les connaître partout ! C'est comme si en retrouvant un morceau d'une dent d'un fossile, on arrivait à reconstituer jusqu'à la moindre détail l'organisme entier.

Ceci n'est pas du au caractère complexe, mais à la dimensionnalité. Prenons une fonction réelle $f(x)$ infiniment dérivable au point x_0 et dont nous connaissons toutes les dérivées. Nous pouvons alors connaître la valeur de la fonction en un point proche $x_1 = x_0 + h$:

$$f(x_1) = \sum_{n=0}^{\infty} f^{(n)}(x_0)h^n/n!$$

L'ennui est qu'en général, nous ne pouvons pas nous éloigner trop du point x_0 : pour des valeurs h un peu trop grand, il se peut très bien que la série diverge. Comme exemple, considérer par exemple la fonction $(1-x)^{-1}$ dont la série de Taylor pour $x_0 = 0$ diverge pour $x \geq 1$. Le point $x = 1$ est une barrière infranchissable sur l'axe réelle. Or, pour les fonctions complexes, le point $z = 1$ est un point que l'on peut *contourner* , puisque nous avons deux dimensions à notre disposition.

La deuxième surprise est la caractère infiniment dérivable des fonctions analytiques. Pour une fonction réelle, le fait d'être une fois dérivable ne donne pas d'obligation

particulière et la fonction peut ne pas posséder de dériver seconde. Par contre, pour les fonctions complexes, le concept de dérivation est plus contraignant : la dérivabilité impose des relations entre les composantes tellement fortes que la fonction devient infiniment dérivable [1]. Cela a également d'autres conséquences. Par exemple, une fonction analytique ne peut pas être bornée dans le plan [2] !

17.2 Les fonctions complexes.

Nous allons dans ce chapitre étudier d'un peu plus près les fonctions complexes $f : \mathbb{C} \to \mathbb{C}$. Ce sont des "machines" qui prennent un nombre complexe en entrée et en produisent un nombre complexe en sortie. A priori, nous pouvons voir les fonctions complexes comme des fonctions de $\mathbb{R}^2 \to \mathbb{R}^2$: un nombre complexe z n'est qu'une collection de deux nombres réelles (x, y) qu'on écrit par convention $x+iy$. Considérons par exemple la fonction complexe

$$f(z) = z^2$$

Si nous écrivons $z = x+iy$, alors $z^2 = (x^2-y^2)+2ixy$ et au lieu d'étudier la fonction ci-dessus, nous pouvons étudier les deux fonctions

$$\begin{aligned} u(x,y) &= x^2 - y^2 \\ v(x,y) &= 2xy \end{aligned}$$

simultanément ou séparément. Mais en faisant cela, nous nous priverions d'un résultat important : nous ne définissons pas à priori l'opération multiplication entre deux *couples* de réelle. Mais pour l'ensemble $\mathbb{C}$, l'opération multiplication (et bien sûr l'addition) a un sens. Nous disons que nous avons muni $\mathbb{C}$ d'une structure algébrique, et cela nous permet de faire la même chose avec les nombres complexes que nous faisons avec les nombres réels.

17.3 Les fonctions analytiques.

Nous allons aller un peu plus loin et nous restreindre aux fonctions analytiques. Une fonction $f : \mathbb{C} \to \mathbb{C}$ est analytique au point z si d'abord f est continue en z et qu'il existe une constante A tel que

$$f(z+h) = f(z) + A.h + O(h^2) \ \ \forall h$$

1. Parfois, des contraintes apparemment bénignes ne le sont pas tant que cela. Nous avons vu un exemple avec l'espace $\mathcal{L}_2$ des fonctions : la simple contrainte d'intégrabilité réduit fortement l'ensemble des fonctions et permet de disposer d'une base dénombrable.

2. à moins d'être une constante. Ceci impliquera qu'un polynôme de degré n doit avoir n racines complexes, comme nous le verrons plus bas.

c'est a dire qu'en gros, les petits déplacements autour du point z provoquent un changement *proportionnel* aux déplacements dans la fonction $f(z)$. Cela paraît une généralisation anodine de la dérivée des fonctions réelles, il n'en est rien. Le mot proportionnel ici utilise les multiplications dans $\mathbb{C}$ et les conséquences sont profondes. Citons quelques conséquences étranges :

1. Une fonction analytique dans un domaine D est infiniment analytiques : toutes ses dérivées sont analytiques.
2. Connaissant la fonction dans une petite région D, on la connaît partout dans le plan complexe [3]
3. Intégrer une fonction analytique sur un contour fermé revient à calculer une quantité relié à la fonction en quelques points seulement.

Voyons donc d'abord ce que veut dire analytique. Notons $z = x+iy$ et $f(z) = u(x,y)+iv(x,y)$ (les parties réelles et imaginaire de la fonction f) et $h = h_x + ih_y$. Alors, par définition,

$$A = \lim_{h\to 0} \frac{u(x+h_x, y+h_y) - u(x,y) + i\left(v(x+h_x, y+h_y) - v(x,y)\right)}{h_x + ih_y}$$

Il suffit maintenant de développer les deux fonctions u et v l'ordre 1 en $h_x$et h_y et considérer les deux cas $h_x = 0$ et $h_y = 0$ pour voir que A est défini de façon unique si et seulement si

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \tag{17.1}$$

$$\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \tag{17.2}$$

Autrement dit, pour qu'une fonction complexe soit analytique, ses parties réelle et imaginaire doivent être intimement liées via les relation (17.1,17.2). Notons que cela a pour conséquences que $\Delta u = \Delta v = 0$, c'est à dire que les deux fonctions doivent être harmonique.

Exemple : démontrer que les fonctions $f(z) = z, z^2, \exp(z)$ sont analytique dans le plan entier, mais que la fonction $f(z) = \bar{z}$ ne l'est nulle part.

17.4 Intégration dans le plan complexe.

La définition de l'intégrale dans le plan complexe est hérité directement de l'intégration dans $\mathbb{R}$: Soit une courbe $\mathcal{C}$ reliant deux point z_A et z_B. Nous découpons la courbe

3. Imaginer reconstituer une voiture, peinture y compris, en observant la trace des pneus.

en N morceaux, chaque morceaux étant une courbe $\mathcal{C}_i$ reliant le point z_i au point z_{i+1} le long de la courbe $\mathcal{C}$. Soit $dz_i = z_{i+1} - z_i$. Alors, l'intégrale est définie par

$$\int_{\mathcal{C}} f(z)dz = \sum_{i=1}^{N} f(z_i)dz_i$$

quand $N \to \infty$ et $\max(|dz_i|) \to 0$.

Bien sûr, il serait extrêmement peut pratique d'utiliser cette définition pour calculer effectivement une intégrale. Mais une courbe n'est qu'un objet unidimensionnel, nous pouvons donc ramener le calcul de l'intégrale complexe à ce que nous savons faire en réel, via une paramétrisation de la courbe : Donnons nous une variable réelle $t \in [a, b]$ et une fonction complexe $z = z(t)$ dérivable telle que quand t parcourt l'intervalle $[a, b]$, z parcourt la courbe $\mathcal{C}$. Alors,

$$\int_{\mathcal{C}} f(z)dz = \int_a^b f(z(t))z'(t)dt \tag{17.3}$$

La théorie d'intégration nous assure bien sûr que l'intégrale ne dépend pas du choix du paramètre t.

Exemple 17.1 Intégrer $1/z$ le long du cercle $|z| = 1$. Nous pouvons paramétrer le cercle par $z = e^{i\theta}$, $\theta \in [0, 2\pi]$, auquel cas $z'(\theta) = iz(\theta)$ et

$$\int_{|z|=1} dz/z = \int_0^{2\pi} id\theta = 2\pi i$$

Le lecteur peut vérifier que l'intégrale sur le même cercle de la fonction $f(z) = z^n$, $n \neq -1$ $n \in \mathbb{Z}$ donnera zéro. Nous aurions aussi pu paramétrer le cercle par $z = e^{i\theta^2}$, $\theta \in [0, \sqrt{2\pi}]$, le lecteur vérifiera que cela ne change rien.

Exemple 17.2 L'intégrale de la fonction $1/z$ le long du demi-cercle supérieur C_1 : $|z| = 1, Im(z) \geq 0$, $z_0 = 1$, $z_1 = -1$ en utilisant la même paramétrisation vaut

$$\int_{C_1} dz/z = \int_0^{\pi} i\theta = \pi i$$

tandis que la même intégrale, mais pris le long du cercle inférieure C_2, nous donne

$$\int_{C_2} dz/z = \int_0^{-\pi} i\theta = -\pi i$$

Exemple 17.3 Intégrer la fonction z^2 le long de la droite C_1 menant de $z_0 = 0$ à $z_1 = 1 + i$. En posant $z = (1 + i)t$, nous avons

$$\begin{aligned} \int_{C_1} z^2 dz &= \int_0^1 (1+i)^2 t^2 (1+i) dt \\ &= (1+i)^3 \left[t^3/3\right]_0^1 = (1+i)^3/3 \end{aligned}$$

Pour intégrer la même fonction le long du quart de cercle $C_2 : |z - i| = 1$ menant au mêmes points finaux nous pouvons utiliser le paramétrage $z = i + \exp(i\theta)$, $dz = i\exp(i\theta)d\theta$. Nous aurons alors

$$\int_{C_2} z^2 dz = \int_{-\pi/2}^{0} (i + \exp(i\theta))^2 i \exp(i\theta) d\theta$$

Ce qui en développant les parenthèses, nous donne à nouveau $(1 + i)^3/3$. Le lecteur pourra vérifier que l'intégration de la fonction z^2 le long de n'importe quel courbe simple ayant les mêmes points finaux donnera toujours le même résultat.

En comparant les exemples 17.2 et 17.3, nous voyons une différence fondamentale : dans le deuxième cas, l'intégrale ne dépend pas de la courbe reliant les deux points z_0 et z_1, mais *seulement* de ces deux points. Dans le premier cas au contraire, deux trajets reliant les deux points finaux donnent des intégrales différentes. Ce qui fait la différence entre ces deux cas est le point fondamental de la théorie des fonctions analytiques : Dans le premier cas, la fonction $1/z$ n'est pas analytique en $z = 0$ et les deux courbes C_1 et C_2 sont de part et d'autre de la singularité. Dans le deuxième cas, z^2 est analytique partout, et l'intégrale ne dépend pas du trajet. Cela nous mène au :

Théorème fondamental de l'intégration (Cauchy-Goursat). Si la fonction $f(z)$ est analytique dans un domaine D et si la courbe fermée C est à l'intérieure de ce domaine, alors

$$\oint_C f(z)dz = 0$$

Pour démontrer la version "light" de ce théorème, nous supposons que $f'(z)$ est continue dans le domaine D. Dans ce cas, en notant $f(z) = u(x, y) + iv(x, y)$ et $dz = dx + idy$, nous avons

$$\oint_C f(z)dz = \oint_C udx - vdy + \oint_C udy + vdx$$

Prenons la première intégrale, et appelons R le domaine enclos par la courbe C. Dans ce cas, nous pouvons appliquer le théorème de Green (ou de Stockes) connu pour l'intégration dans le plan :

$$\oint_C udx - vdy = -\iint_R (\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x})dxdy$$

mais le terme entre parenthèse est nul à cause de l'analycité de la fonction f. Le même raisonnement tient pour le second intégrande, et achève la démonstration.

Cette démonstration n'est pas très bonne, puisqu'elle exige la continuité de $f'(z)$ à priori. Ce pré-requis est de trop, ce qui a été démontré par Goursat vers 1900.

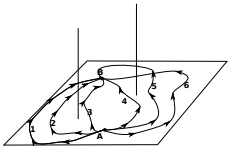


FIGURE 17.1 – Intégration d'une fonction $f(z)$ entre deux points A et B le long d'un chemin i. Les points singuliers de la fonction sont marqués par les "tiges" verticales. Les chemins qui peuvent se déformer les uns dans les autres sans traverser les tiges donnent la même valeur pour l'intégrale de $f(z)$. Ici les chemins 1 et 2 ; 3 et 4 ; 5 et 6 sont équivalent.

La conséquence direct de ce théorème est que dans un domaine D où $f(z)$ est analytique, son intégrale entre deux points A et B ne dépend pas du chemin d'intégration , mais seulement des points A et B :

$$\int_{C_1} f(z)dz = \int_{C_2} f(z)dz$$

C_1 et C_2 étant deux chemins quelconques reliant A et B. En effet, considérons la courbe fermée C formé de l'union de C_1 et de $-C_2$, c'est à dire de C_2 parcouru dans le sens inverse. Dans ce cas,

$$\begin{aligned} \oint_C f(z) &= 0 \\ &= \int_{C_1} f(z)dz + \int_{-C_2} f(z)dz \\ &= \int_{C_1} f(z)dz - \int_{C_2} f(z)dz \end{aligned}$$

L'image qu'il faut retenir est le plan complexe où les points de singularité pour la fonction $f(z)$ constituent des tiges verticaux : les chemins d'intégration peuvent être déformés à souhait sans modifier le résultat d'intégration tant qu'ils ne croisent pas "une tige" (figure 17.1).

17.5 Conséquences du Cauchy-Goursat.

Le théorème de Cauchy-Goursat a des conséquences profondes d'où coule l'essence de la théorie des fonctions analytiques. Nous énumérons les trois plus importantes.

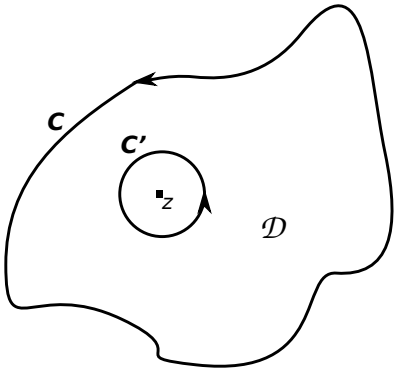


FIGURE 17.2 – Formule intégrale de Cauchy : les intégrales de la fonction $f(\zeta)/\zeta - z$ sur les deux courbes C et C' sont égale.

17.5.1 Connaître localement = Connaître globalement.

Ceci est probablement la propriété la plus frappante des fonctions analytiques : connaître la valeur d'une fonction analytique sur une courbe fermée C revient à connaître la valeur de cette fonction partout à l'intérieur du domaine enclos par cette courbe. De plus, la relation entre les valeurs de la fonction sur la courbe et la valeur de la fonction en un point z à l'intérieur du domaine est très simple : la valeur de f en z est la somme des valeurs que prend f sur la courbe, pondérées par les séparations à ce point (la courbe est parcourue dans le sens positif) :

$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(\zeta)}{\zeta - z} d\zeta$$

Ceci est connu comme la formule intégrale de Cauchy. Pour le démontrer, il suffit de considérer la courbe C et le cercle C' de rayon ϵ centré autour du point z (figure 17.2). Dans le domaine $\mathcal{D}$ entre les deux courbes, la fonction $f(\zeta)/(\zeta - z)$ (considérée comme une fonction de ζ) est analytique ; d'après Cauchy-Goursat donc, son intégrale le long des deux courbes donne le même résultat. Le long de la courbe C', nous pouvons utiliser le paramétrage $\zeta = z + \epsilon \exp(i\theta)$, ce qui nous donne :

$$\begin{aligned} \oint_{C'} \frac{f(\zeta)}{\zeta - z} d\zeta &= \int_0^{2\pi} \frac{f(z + \epsilon e^{i\theta})}{\epsilon e^{i\theta}} i\epsilon e^{i\theta} d\theta \\ &= i \int_0^{2\pi} \left(f(z) + f'(z)\epsilon e^{i\theta} + O(\epsilon^2) \right) d\theta \end{aligned}$$

Nous pouvons choisir le cercle autour de z aussi petite que l'on veut. Quand $\epsilon \to 0$, le seul terme qui contribue [4] à l'intégrale est le premier : $if(z)\int_0^{2\pi} d\theta = 2i\pi f(z)$. Cela démontre donc la formule intégrale de Cauchy. L'essentiel est de retenir que nous avons pu démontrer ce théorème parce que nous avons pu ramener l'intégration le long de C à l'intégration le long de C', une conséquence directe du Cauchy-Goursat.

Réfléchissons un peu à la formule intégrale. En analyse réelle, il ne suffit pas de connaître la valeur d'une fonction en deux points (dimension 0) pour la connaître en *tout* point de l'intervalle (dimension 1). Dans le plan complexe, il suffit de connaître la valeur de la fonction *analytique* le long d'une courbe (dimension 1) pour la connaître partout à l'intérieur de la courbe(dimension 2). Ce résultat ne s'applique évidemment pas à l'ensemble des fonctions complexes, seulement aux analytiques.

17.5.2 Les fonctions analytiques sont infiniment lisse.

En analyse réelle, disposer d'une dérivée première ne garantit en rien de pouvoir disposer des dérivées d'ordre supérieure ; il suffit par exemple de considérer la fonction $|x^3|$ qui possède une dérivée première en $x = 0$, mais pas de dérivée seconde. Les fonctions analytiques par contre ont cette propriété fort sympathique *d'être infiniment dérivable*. Plus exactement, si $f(\zeta)$ est analytique dans un domaine D, y compris sur sa frontière C, alors elle possède des dérivées de toute ordre, qui sont elles même analytiques. Encore plus fort, nous pouvons les calculer exactement par une formule dérivée de la formule intégrale de Cauchy :

$$f^{(n)}(z) = \frac{n!}{2\pi i}\oint_C \frac{f(\zeta)}{(\zeta - z)^{n+1}}d\zeta \tag{17.4}$$

Nous pouvons voir cela comme une généralisation de la formule intégrale de Cauchy qui correspond au cas $n = 0$. La démonstration est assez simple ; nous le montrons pour $f'(z)$ qui se généralise trivialement aux dérivées supérieures. Prenons deux points z et z_1 assez proche et un circuit C dans le domaine D qui les englobes tous les deux. Nous avons :

$$\begin{aligned} f(z) &= \frac{1}{2\pi i}\oint_C \frac{f(\zeta)}{\zeta - z}d\zeta \\ f(z_1) &= \frac{1}{2\pi i}\oint_C \frac{f(\zeta)}{\zeta - z_1}d\zeta \end{aligned}$$

4. Notons pour le lecteur qui n'est pas habitué au symbole O que ce raisonnement est rigoureux. Pour une fonction analytique, par définition nous avons $f(z+h) = f(z) + f'(z)h + R(z,h)h^2$ où $R(z,h) \to Cte$ quand $h \to 0$. Cela veut dire que nous pouvons rendre $|R(z,h) - Cte|$ aussi petite que l'on veut, pourvu que l'on choisissent $|h|$ assez petit. Ce qui veut dire que nous pouvons toujours trouver une constante C tel que pour ϵ suffisamment petit,

$$\left|\int_0^{2\pi} R(z, \epsilon e^{i\theta})\epsilon^2 e^{2i\theta} d\theta\right| < C\epsilon^2(2\pi)$$

Nous laissons au lecteur le soin d'habiller et de finir la démonstration dans toute sa rigueur.

En formant maintenant $(f(z)-f(z_1))/(z-z_1)$ (et en entrant $1/(z-z_1)$ dans l'intégrale sur ζ) et en faisant un peu d'algèbre, nous obtenons :

$$\frac{f(z)-f(z_1)}{z-z_1} = \frac{1}{2\pi i}\oint_C \frac{f(\zeta)}{(\zeta-z)(\zeta-z_1)}d\zeta$$

qui tend bien vers la formule annoncée quand $z_1 \to z$.

17.5.3 Les développements de Laurent.

Les séries de Taylor ont un équivalent autrement puissant pour les fonctions analytiques, appelé développement de Laurent. Considérons deux cercles de rayons R_1 et R_2 ($R_1 < R_2$) centrés sur un nombre z_0 et la fonction $f(z)$ analytique dans la couronne formée par les deux cercles. Soit C un circuit compris entre les deux cercles. Alors nous avons

$$f(z) = \sum_{-\infty}^{+\infty} A_n (z-z_0)^n \tag{17.5}$$

où les coefficients A_n sont données par

$$A_n = \frac{1}{2\pi i}\oint_C \frac{f(\zeta)}{(\zeta-z_0)^{n+1}}d\zeta \tag{17.6}$$

De plus, la convergence de la série (17.5) est uniforme. Comme nous savons par ailleurs que les dérivées (et primitives) de toutes ordres existent, nous pouvons les obtenir en dérivant (ou intégrant) terme à terme la série [5].

Avant de démontrer ce théorème, faisons quelques commentaires. Premièrement, pour développer une fonction en série autour d'un point, nous n'avons pas besoin que cette fonction soit dérivable en ce point ; cela rend les développement de Laurent autrement plus générale que les développement de Taylor.

Ceci dit, si la fonction $f(z)$ est analytique, alors tout les coefficients A_n pour $n < 0$ sont nul et on retrouve les développement de Taylor habituel. Pour voir cela, il suffit de remarquer que les A_n $n < 0$ sont des intégrales sur un circuits fermé du produit de $f(\zeta)$ par un polynôme $(\zeta - z)^m$. Ces deux fonctions étant analytiques, l'intégrale est nulle par Cauchy-Goursat. Pour les $n \geq 0$, nous avons (eq.17.4)

$$A_n = (1/n!)f^{(n)}(z_0)$$

Ce qui nous redonne bien les développement de Taylor habituel.

Pour démontrer la relation (17.5), nous devons d'abord établir que

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

5. Elle est pas belle la vie ?

ce qui se démontrer très facilement en remarquant que

$$\sum_{n=0}^{N} z^n = \frac{1-z^{N+1}}{1-z}$$

qui converge bien vers $1/(1-z)$ pour $|z|<1$.

Considérons maintenant le domaine D délimité par les deux courbes et les deux droites qui les rejoignent (fig.). A l'intérieur de ce domaine,

$$\begin{aligned} f(z) &= \frac{1}{2\pi i}\oint_C \frac{f(\zeta)}{\zeta - z}d\zeta \\ &= \frac{1}{2\pi i}\oint_{C_2} \frac{f(\zeta)}{\zeta - z}d\zeta - \frac{1}{2\pi i}\oint_{C_1} \frac{f(\zeta)}{\zeta - z}d\zeta \end{aligned}$$

Considérons la première intégrale. Nous pouvons écrire

$$\begin{aligned} \frac{f(\zeta)}{\zeta - z} &= \frac{f(\zeta)}{\zeta - z_0 - (z - z_0)} \\ &= \frac{f(\zeta)}{\zeta - z_0}\sum_{n=0}^{\infty}\left(\frac{z-z_0}{\zeta - z_0}\right)^n \end{aligned}$$

puisque $|z - z_0/\zeta - z_0| < 1$. Cela nous donne donc

$$f(z) = \sum_{n=0}^{\infty}\left(\frac{1}{2\pi i}\oint_{C_2}\frac{f(\zeta)}{(\zeta - z_0)^{n+1}}d\zeta\right)(z-z_0)^n$$

En faisant la même chose pour la deuxième intégrale sur C_1, mais en développant en puissance de $(\zeta - z_0)/(z - z_0)$, nous trouvons

$$f(z) = \sum_{n=-\infty}^{-1}\left(\frac{1}{2\pi i}\oint_{C_1}\frac{f(\zeta)}{(\zeta - z_0)^{n+1}}d\zeta\right)(z-z_0)^n$$

Mais la fonction $f(\zeta)/(\zeta - z_0)$ est analytique dans la couronne délimité par les deux cercles C_1 et C_2, on peut donc ramener l'intégration sur ces deux cercles à une intégrale sur n'importe quel chemin fermé contenu entre les deux et retrouver ainsi le développement de Laurent tel qu'annoncé.

Exemple : $\exp(z)$, $\exp(1/z)$, $\exp(t(z-1/z))$ en Bessel, ... [To complete].

17.6 Les résidus et leur application à l'intégration.

Nous avons souvent, en analyse réelle, à calculer des intégrales du genre

$$\int_I f(x)dx$$

Nous connaissons quelques méthodes (changement de variable, intégration par partie) qui avec beaucoup d'entraînement nous permettent, dans certains cas, de nous en sortir.

Nous allons apprendre ici à calculer facilement quelques cas supplémentaires. La méthode est la suivante : au lieu de calculer $\int_I f(x)dx$, nous calculerons $\int_{\mathcal{C}} f(z)dz$ où $\mathcal{C}$ est un circuit fermé dans plan complexe qui contient I. Nous faisons cela quand l'intégrale sur $\mathcal{C}$ est facile à effectuer (voir plus bas) et quand l'intégrale sur la partie du $\mathcal{C}$ qui ne contient pas I est nulle ou très facile à évaluer. Voyons cela de plus près maintenant.

17.6.1 Les résidus.

Dans le développement de Laurent et ses coefficients (17.6), il n'a pas échappé au lecteur que le coefficient A_{-1} joue un rôle particulier, puisque

$$A_{-1} = (1/2\pi i)\int_C f(\zeta)d\zeta$$

On peut dire à l'inverse que si l'on connaissait ce coefficient, on connaîtrai la valeur de l'intégrale. Le rôle particulier joué par A_{-1} est bien sûr lié au fait que $\oint_C z^n dz$ est non nulle seulement pour $n = -1$. Le coefficient A_{-1} est appelé le résidu de la fonction f en z_0 et noté

$$A_{-1} = Res[f(z), z_0]$$

Il est donc évident que connaître les résidus d'une fonction est une donnée précieuse. Voyons quelques exemples.

Exemple 17.4 $Res(1/z(z+1), 0)$. En $z = 0$, la fonction $1/(z+1)$ est analytique et nous pouvons la développer en série de Taylor

$$1/(z+1) = a_0 + a_1 z + ...$$

où $a_0 = 1/(z+1)|_{z=0} = 1$. Nous pouvons donc écrire

$$\frac{1}{z(z+1)} = \frac{1}{z}(a_0 + a_1 z + ...)$$

et donc

$$Res(1/z(z+1), 0) = 1$$

Exemple 17.5 $Res(1/z(z+1), -1)$. En $z = -1$, la fonction $1/z$ est analytique et donc

$$1/z = a_0 + a_1(z+1) + a_2(z+1)^2 + ...$$

où $a_0 = 1/z|_{z=-1} = -1$. Par conséquent, $Res(1/z(z+1), -1) = -1$.

Nous pouvons bien sûr systématiser cet approche. Soit la fonction

$$f(z) = \frac{N(z)}{(z-z_0)^m}$$

où $N(z)$ est analytique. La fonction possède un pôle d'ordre m en z_0. En développant $N(z)$ à l'ordre $m-1$ en z_0, nous voyons que

$$f(z) = ... + \frac{N^{(m-1)}(z_0)}{(m-1)!}\frac{1}{z-z_0} + ...$$

et donc

$$Res(N(z)/(z-z_0)^m, z_0) = N^{(m-1)}(z_0)/(m-1)!$$

Par exemple,

$$Res(\sin z/z^2, 0) = 1$$

On généralise cela trivialement à une fonction rationnelle quelconque

$$f(z) = \frac{N(z)}{D(z)}$$

où $N(z)$ et $D(z)$ sont analytique et $D(z)$ possède un zéro d'ordre m en $z = z_0$: $D(z) = (D^{(m)}(z_0)/m!)(z-z_0)^m + ...$ Nous avons alors

$$Res(N(z)/D(z), z_0) = mN^{(m-1)}(z_0)/D^{(m)}(z_0)$$

En particulier, si le zéro de $D(z)$ est d'ordre 1, nous avons simplement $Res = N(z_0)/D'(z_0)$.

Exemple 17.6 Calculer les résidus de $(\exp kz)/(z^2+a^2)$. D'après ce que nous avons dit, elles valent $\exp(\pm ika)/(\pm 2ia)$.

17.6.2 Application à l'intégration.

Nous allons considérer plusieurs cas.

Les fonctions trigonométriques. Nous allons considérer les intégrale du type

$$I = \int_0^{2\pi} R(\sin\theta, \cos\theta)d\theta$$

Comme par exemple [6]

$$I_1 = \int_0^{2\pi} \frac{1}{a+\cos\theta}d\theta \quad ; \quad a > 1$$

6. Le lecteur se souvient que le changement de variable suggéré dans ce cas était de poser $u = \tan\theta$ ou $u = \tan(\theta/2)$ et utiliser ensuite les relations trigonométriques pour se ramener à des choses connues.

Le passage dans le plan complexe se fait ici assez naturellement, en posant $z = \exp(i\theta)$ et donc $d\theta = -idz/z$. Le contour est le cercle de rayon unité.

Pour l'exemple ci-dessus, nous avons

$$I_1 = -i \oint \frac{2dz}{z^2 + 2az + 1}$$

Or, pour les deux racines du dénominateur, nous avons $z_1 z_2 = 1$. Appelons z_2 la racine telle que $|z_2| > 1$. Seule la racine z_1 est à l'intérieur du cercle $|z| = 1$, et nous devons donc calculer le résidu en ce point.

L'intégrande s'écrit comme

$$f(z) = \frac{2}{(z - z_1)(z - z_2)}$$

et donc

$$Res(f(z), z_1) = \frac{2}{(z_1 - z_2)} = \frac{1}{\sqrt{a^2 - 1}}$$

Finalement,

$$I_1 = (-i)(2\pi i)/\sqrt{a^2 - 1} = 2\pi/\sqrt{a^2 - 1}$$

Les fonctions rationnelles simples. Nous allons considérer des intégrales du type

$$I = \int_{-\infty}^{+\infty} \frac{P(x)}{Q(x)} dx$$

où $P(x)$ et $Q(x)$ sont des polynômes, $\partial Q \geq \partial P + 2$ (pour que l'intégrale converge en infini) et $Q(x) \neq 0$.

Nous allons considérer le circuit fermé $\mathcal{C}$ composé du $\mathcal{C}_1$, le demi-cercle supérieur (ou inférieur) de rayon R et de $\mathcal{C}_2$, l'intervalle $[-R, R]$, et faire tendre $R \to \infty$. Il est facile de démontrer que l'intégrale sur le demi-cercle tend vers 0. Ceci vient du fait qu'en général

$$\int_{\mathcal{D}} f(z)dz < ML$$

où L est la longueur de la courbe $\mathcal{D}$ et M une borne supérieure de $|f(z)|$. Or, sur le demi-cercle, quand $R \to \infty$, $|P(z)/Q(z)| \sim 1/R^p$ où $p \geq 2$, tandis que la longueur du demi cercle $\sim R$. Il n'est pas difficile de donner une démonstration plus rigoureuse en bornant correctement les deux polynômes. Nous avons donc

$$\int_{\mathcal{C}} \frac{P(z)}{Q(z)} dz = \int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} dx$$

et l'intégrale de gauche se calcule par la méthode des résidus.

Exemple 17.7 Calculons

$$I_1 = \int_{-\infty}^{\infty} \frac{x^2}{x^4+1} dx$$

Les pôles de la fonctions $z^4 + 1$ sont les 4 racines simple de -1. En prenant le cercle supérieur comme circuit, nous voyons que $z_1 = \exp(i\pi/4)$ et $z_2 = \exp(3i\pi/4)$ dans à l'intérieur de ce cercle. Par ailleurs, nous savons que

$$Res(f(z), z_1) = \left.\frac{P(z)}{Q'(z)}\right|_{z_1} = \frac{e^{-i\pi/4}}{4}$$

De même,

$$Res(f(z), z_2) = \frac{e^{-3i\pi/4}}{4}$$

Donc,

$$\begin{aligned} I_1 &= 2\pi i \left(Res(f(z), z_1) + Res(f(z), z_2)\right) \\ &= \pi/\sqrt{2} \end{aligned}$$

Produit d'exponentiel et de fonctions rationnelles simples. Considérons maintenant des intégrales du type

$$I(k) = \int_{-\infty}^{+\infty} e^{ikx} \frac{P(x)}{Q(x)} dx$$

où P et Q sont comme dans le cas précédent et $k \in \mathbb{R}$. Supposons que $k > 0$. Sur le cercle supérieur, $z = z_R + iz_I$ où $z_i > 0$. Donc, $\exp(ikz) = \exp(-kz_i)\exp(ikz_R) < 1$. Nous pouvons donc utiliser le même principe que dans le cas précédent. Notez que cela a un fort lien avec les transformées de Fourier.

Nous pouvons même relaxer un peu les contraintes et exiger simplement $\partial Q \geq \partial P + 1$, l'oscillation de l'exponentielle assurant la convergence.

Exemple 17.8 Calculons

$$I_1 = \int_{-\infty}^{\infty} \frac{e^{ikx}}{x^2+1} dx$$

Pour $k > 0$, nous choisissons le cercle supérieur et seul le pôle $+i$ est dans le cercle. Or,

$$Res(f(z), i) = \frac{e^{-k}}{2i}$$

et donc

$$I_1 = \pi e^{-k}$$

Pour $k < 0$, nous choisissons le cercle inférieur. Nous récupérons un signe -1 en plus pour cause de modification de sens de parcours, et donc

$$I_1 = \pi e^k$$

Nous pouvons regrouper les deux résultats en écrivant

$$I_1 = \pi e^{|k|}$$

Exercices.

§ 17.1 Évaluer les intégrales

$$\oint_C \cot z dz \quad C : |z| = 1$$

$$\oint_C \frac{3(z+1)}{z(z-1)^3} dz \quad C : |z - 1/2| = 2$$

$$\oint_C \frac{\exp(-\cosh z)}{z^2+1} dz \quad C : |z| = 2$$

§ 17.2 Calculer les intégrales suivantes :

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

$$\int_0^\infty \frac{\cos kx}{x^4+a^4} dx \; ; \; \int_0^{\pi/2} \sin^4\theta d\theta \; ; \; \int_0^{2\pi} \frac{d\theta}{1+\cos^2\theta}$$

§ 17.3 Quel contour peut-on utiliser pour calculer

$$\int_0^\infty \frac{dx}{x^3+a^3}$$

[Help : Nous pouvons constater que $(xe^{2i\pi/3})^3 = x$; un parcours sur arc ouvert à $2\pi/3$ peut être une bonne idée]. Comment généraliser cette idée pour calculer

$$\int_0^\infty \frac{dx}{x^5+a^5}$$

§ 17.4 Démontrer que

$$\int_0^\infty \frac{\cosh ax}{\cosh \pi x} dx = \frac{1}{2\cos(a/2)}$$

où $|a| < \pi$

§ 17.5 Calculez

$$\int_{-\infty}^\infty \frac{x \cos kx}{x^2+4x+5}$$

où $k \in \mathbb{R}$ et $k > 0$. Vous devez faire un peu attention au parcours que vous choisissez, en n'oubliant pas que votre fonction doit rapidement tendre vers 0 sur l'arc que vous avez choisi.

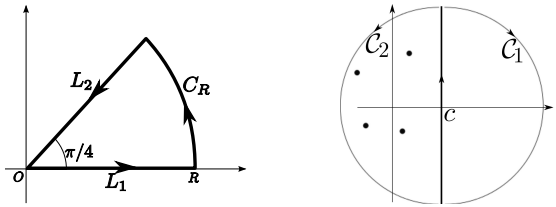


FIGURE 17.3 – (gauche) Contour fermé $C = L_1 + C_R + L_2$ pour le problème 17.1 ; (droite) contour pour le TL inverse du problème 17.3

§ 17.6 Calculez

$$\int_0^{2\pi} \cos^{2n}\theta d\theta$$

où $n \in \mathbb{N}$. Il va de soi que vous connaissez l'expansion binomial $(a+b)^n$.

Problèmes.

Problème 17.1 Intégrale de fonction oscillante.

Nous sommes habitués à l'intégrale gaussienne

$$\int_0^\infty e^{-tx^2} dx = \sqrt{\frac{\pi}{4t}}$$

On rencontre souvent sa version oscillante dans des problèmes d'optique ou de mécanique quantique :

$$\begin{aligned} F_c &= \int_0^\infty \cos(tx^2)dx \\ F_s &= \int_0^\infty \sin(tx^2)dx \end{aligned}$$

ces intégrales sont les limites d'une fonction appelée "Fresnel".

Existence. Indiquez sans trop de démonstration pourquoi ces intégrales existent.

Contour fermé. Que vaut

$$I = \oint_C e^{itz^2} dz$$

où C est l'arc fermé de rayon R et d'angle $\pi/4$ de la figure (17.3)?

Arc. Démontrer que quand $R \to \infty$,

$$\int_{C_R} e^{itz^2} dz \to 0$$

(nous supposons $t > 0$) [Help : Nous avons besoin de borner correctement $\left|\int_{C_R}\right|$; Par ailleurs, sur l'intervalle $[0, \pi/2]$, nous savons que $\sin(u) \geq (2/\pi)u$]

L_2. Que vaut

$$\int_{L_2} e^{itz^2} dz$$

Quand $R \to \infty$?

Synthèse. En synthétisant l'ensemble des résultats précédents, déduire l'intégrale sur L_1 quand $R \to \infty$ et les valeurs de F_c et F_s.

Problème 17.2 Transformée de Laplace des Bessels.

Les fonctions de Bessel I_n et J_n sont définies par

$$\begin{aligned} J_n(z) &= \frac{1}{\pi}\int_0^\pi e^{iz\cos(\theta)}\cos(n\theta)d\theta \\ I_n(z) &= \frac{1}{\pi}\int_0^\pi e^{z\cos(\theta)}\cos(n\theta)d\theta \end{aligned}$$

Démontrer que leurs TL est, pour $n \geq 0$

$$\begin{aligned} \tilde{J}_n(s) &= \frac{\left(\sqrt{s^2+1}-s\right)^n}{\sqrt{s^2+1}} \\ \tilde{I}_n(s) &= \frac{\left(\sqrt{s^2-1}+s\right)^n}{\sqrt{s^2-1}} \end{aligned}$$

Help : Il faut d'abord étendre le domaine d'intégration sur θ, échanger l'ordre d'intégration sur z et θ, et prendre garde à la valeur absolue de n.

Problème 17.3 Transformée de Laplace inverse.

Nous savons que si

$$\hat{f}(s) = \mathsf{TL}[f(t)]$$

alors

$$f(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \hat{f}(s)e^{ts}ds$$

où c est un nombre réel quelconque supérieur à la partie réelle de tous les pôles de la fonction $\hat{f}(s)$ (figure 17.3).

En utilisant vos connaissances de l'intégration dans le plan complexe, et en choisissant sagement vos contours d'intégration, calculer alors la transformée de Laplace inverse de la fonction

$$\hat{f}(s) = \frac{1}{s^2+\omega^2}$$

Help : évidemment, il faut fermer le contour de l'intégration par un arc de cercle de rayon R dont la contribution à l'intégrale est nulle quand $R \to \infty$. Argumentez pourquoi il faut choisir $\mathcal{C}_1$ si $t < 0$ et $\mathcal{C}_2$ si $t > 0$. [Je ne demande pas une majoration rigoureuse, mais un argument convaincant].

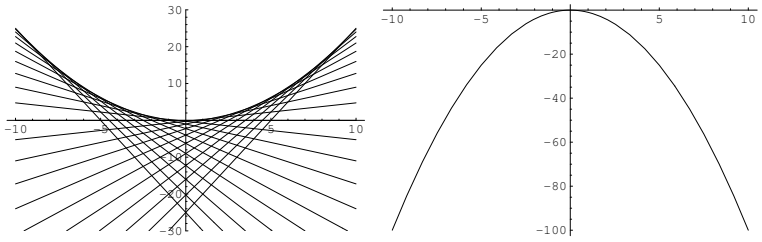
18 Les Transformées de Legendre.

18.1 Définition.

Nous avons vu plusieurs façon de représenter une fonction, dans l'espace direct ou dans l'espace réciproque. Nous avons vu qu'une transformée (de Fourier, de Laplace, ...) nous prenait une fonction et en produisait une autre : dans l'espace de cette nouvelle fonction, certains problèmes peuvent être formulés de façon plus simple. Les représentations que nous avons vu jusque là étaient basées sur le concept d'intégral, nous verrons une nouvelle représentation basée sur la dérivée.

Une question d'enveloppe.

Une des représentations extrêmement utiles a été introduite par Legendre en exploitant l'idée d'enveloppe. Donnons nous une famille de droite $y = px - g(p)$ comme par exemple la figure (18.1). Chaque droite Δ_p est définie par sa pente p et l'ordonnée à l'origine (le croisement avec l'axe y), $-g(p)$. A vrai dire, pour caractériser la famille de droite, il nous suffit seulement de nous donner la fonction $g(p)$. Ce que nous voyons quand on trace une telle famille est l'apparition de l'enveloppe de la famille $f(x)$, c'est à dire une courbe qui est tangente à chacune des membres de la famille. Évidemment, comme nous allons le voir sous peu, n'importe qu'elle dépendance $g(p)$ ne peut pas



(a) (b)

FIGURE 18.1 – (a) Ensemble de droites d'équation $px - g(p)$; (b) la fonction $-g(p) = -p^2$ qui a servi à générer la famille de droite .

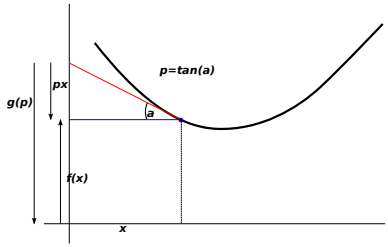


FIGURE 18.2 – Représentation graphique d'une transformation de Legendre.

générer une enveloppe [1]. Mais l'idée à retenir est qu'à la fonction $g(p)$ nous pouvons associer la fonction $f(x)$.

L'opération inverse est encore plus facile : si on se donne une fonction $f(x)$, il est très facile d'obtenir sa famille de droites enveloppes : A chaque point x_0, on calcule la tangente $p = f'(x_0)$. L'équation de la droite est alors donné par $y - f(x_0) = p(x - x_0)$ ou encore

$$y = [f(x_0) - px_0] + px$$

Remarquer que dans le crochet, p et x_0 ne sont pas indépendant : à chaque valeur de p correspond une valeur x_0. Nous posons

$$g(p) = x_0 p - f(x_0),$$

où $g(p)$ est fonction de p seulement. Plus exactement, comme $x_0 = f'^{-1}(p)$, nous avons

$$g(p) = pf'^{-1}(p) - f\left(f'^{-1}(p)\right)$$

Cette notation est plus exacte puisque la dépendance est explicitement soulignée, mais elle est plus lourde. On utilise plus volontiers la première notation, gardant bien en tête que $x_0$est une fonction de p. D'ailleurs, une fois qu'une certaine familiarité a été acquise, l'on oublie même l'indice et on écrit directement [2] $g(p) = px - f(x)$ (Fig.).

Exemple 18.1 Considérons la fonction $f(x) = x^2/2$. En un point x_0, $p = f'(x_0) = x_0$ où encore $x_0 = p$. La famille des courbes enveloppe est donc donnée par $g(p) = p^2 - p^2/2 = p^2/2$.

1. Il faut que la fonction soit convexe ou concave, autrement dit que la dérivée seconde ne change pas de signe ou encore que la dérivée première soit monotone.

2. Voilà pourquoi le signe - a été choisi dans la déifnition de la droite tangente $y = px - g(p)$ plutôt que $y = px + g(p)$. Cela nous permet une symétrie entre la fonction et sa transformée que l'on peut écrire $f(x) + g(p) = px$.

Enveloppe d'enveloppe.

Nous pouvons continuer ce jeux : considérons maintenant la fonction $g(p)$; nous pouvons trouver la famille de droite $\Delta_s : y = sp - h(s)$ dont la fonction $g(p)$ est l'enveloppe. Donnons nous un point p_0, la tangente en ce point est $s = g'(p_0)$ et la droite possède donc l'équation

$$y = [g(p_0) - sp_0] + sp$$

et donc $h(s) = sp_0 - g(p_0)$ où $s = g'(p_0)$.

Exemple 18.2 Continuons avec la fonction $g(p) = p^2/2$. Par définition, $s = g'(p_0) = p_0$ ou encore $p_0 = s$. Nous avons donc $h(s) = s^2 - s^2/2 = s^2/2$. Ce que nous voyons ici est que $h(s) = f(s)$, c'est à dire que l'enveloppe de l'enveloppe est la fonction originale !

Nous allons démontrer cela de façon générale, mais le message à retenir est que les fonctions $g(p)$ et $f(x)$ sont duales vis à vis de cette transformation qu'on appelle transformation de Legendre. La transformée de Legendre revient à définir une fonction non pas en fonction des valeurs qu'elle prend en un point x, mais en fonction de sa pente.

Reprenons, en partant de $f(x)$, nous définissons

$$g(p) = px_0 - f(x_0) \tag{18.1}$$
$$p = f'(x_0) \tag{18.2}$$

Calculons maintenant dg/dp, sachant que les deux variables p et x_0 sont relié par la relation (18.2).

$$dg/dp = x_0 + p.dx_0/dp - df(x_0)/dp$$

Or, $df(x_0)/dp = f'(x_0).dx_0/dp = p.dx_0/dp$. Nous arrivons à cette relation de dualité

$$\frac{dg}{dp} = x_0$$

L'équation de la droite tangente à $g(p)$ en un point p_0 s'écrit donc

$$y = x_0 p - [x_0 p_0 - g(p_0)]$$

L'expression entre crocher n'est rien d'autre que $f(x_0)$d'après (18.1).

Transformée de Legendre et Minimum.

Il existe une interprétation alternative de la transformée de Legendre, basée sur le concept de minimum, et qui est très utilisée par exemple en thermodynamique.

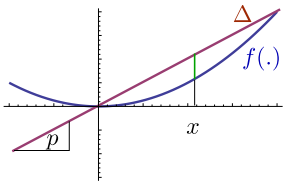


FIGURE 18.3 – Transformée de Legendre vu comme un minimum.

Soit la fonction convexe $y = f(x)$. Pour une pente donnée p, donnons nous la droite $y = px$, et cherchons la valeur de x qui extrémise la distance entre les deux courbes (Fig.18.3)

$$g(x, p) = px - f(x)$$

Il n'est pas difficile de voir que l'extremum est réalisé pour

$$f'(x) = p$$

et la distance extremum entre les deux courbes est fonction de p seul

$$g(p) = px - f(x)$$

A vrai dire, ce que nous venons de faire généralise le concept de minimum d'une fonction (convexe). Son minimum habituel est simplement le point associé à la pente nulle.

Avec cette représentation, il est coutume d'écrire

$$g(p) = \max_x (px - f(x))$$

Recette.

Dorénavant, nous allons laisser tomber les x_0 et les p_0, la signification étant entendu. Nous écrirons

$$g(p) = px - f(x)$$

ayant en tête que x est dépendante de p, avec la dépendance $f'(x) = p$. De même, pour la transformation inverse, nous écrirons

$$f(x) = px - g(p)$$

ayant en tête que p est une fonction de x

§ 18.1 Montrer que la transformée de Legendre de $f(x) = e^x$ est $g(p) = p \log p - p$. En reprenant la Transformée de Legendre, démontrer qu'on retrouve la fonction originale. Trouver la TLg de la fonction $f(x) = x^{2n}$.

Transformée de Legendre à plusieurs dimensions.

La généralisation est triviale. Soit la fonction $f(x_1, x_2, ...)$; nous pouvons définir la fonction $g(p_1, x_2, x_3, ...)$, ou la fonction $g(p_1, p_2, x_3, ...)$ par exactement les mêmes procédures.

18.2 Application à travers la physique.

Les transformées de Legendre apparaissent un peu partout en physique, à chaque fois que les dérivées prennent une vie propre. Par exemple, nous pouvons imposer la valeur d'une fonction en un point comme paramètre de contrôle, où au contraire, la valeur de la tangente comme paramètre de contrôle.

L'action en mécanique.

En mécanique classique, le lagrangien est définie par

$$\mathcal{L}(\dot{x}, x, t) = (m/2)\dot{x}^2 - V(x, t)$$

Noter que $\mathcal{L}$ est considéré comme une fonction de *deux* variables indépendante, x et $\dot{x}$. Nous pouvons prendre la transformée de Legendre du lagrangien par rapport à $\dot{x}$

$$H(p, x, t) = p\dot{x} - \mathcal{L}(\dot{x}, x, t) = (1/2m)p^2 + V(x, t)$$

où

$$p = \frac{\partial \mathcal{L}}{\partial \dot{x}}$$

Le très grand avantage de ce passage est que nous passons des équations de second degrés en $\ddot{x}$ en équation de premier degrés en $\dot{p}$. Formellement, on ne change pas grand chose, une équation de second degrés ou deux équations de premiers degrés à priori nécessitent le même effort ; mais les équations du premier degrés ont une interprétation géométrique immédiate en terme de flux et même si on ne sait pas les résoudre exactement, on peut démontrer beaucoup de choses.

En mécanique relativiste, l'action est donné par

$$\mathcal{S} = \int_a^b -m.ds$$

où $ds^2 = dt^2 - dx^2$ est l'élément d'arc entre deux événements. En notant $\beta = dx/dt$, nous avons

$$\mathcal{L}(\beta) = -m\sqrt{1 - \beta^2}$$

Ce qui nous donne

$$p = d\mathcal{L}/d\beta = m\beta/\sqrt{1 - \beta^2}$$

ou encore $\beta = p/\sqrt{p^2+m^2}$ ou finalement

$$H(p) = \sqrt{p^2+m^2}$$

Relations que nous voyons plus souvent sous la forme

$$E^2 - p^2 = m^2$$

Les potentiels thermodynamiques.

Prenons d'abord un exemple simple. Supposons que l'on connaisse l'énergie libre d'un système $F(V,T,...)$ pour toute température et volume fixée (penser : gaz parfait dans un cylindre à volume fixe). Nous allons mettre en contact notre système avec un réservoir de pression (une force constante par unité de surface) et laisser libre le volume (penser : piston capable de se mouvoir dans le cylindre pour). L'énergie libre de notre système est maintenant $G(V,T,p,...) = F(V,T,...) + pV$.

$F()$est toujours notre énergie libre à volume fixe, pV le travail effectué par le changement de volume. Le *principe du minimum* en thermodynamique nous affirme que le volume V va varier (penser : le gaz se détend ou se compresse) jusqu'à atteindre une valeur d'équilibre V^* telle que l'énergie libre *soit minimum*. Dans ce cas, V^* est une fonction de *p,T,...* est l'énergie libre vaut

$$G(T,p,...) = F(V^*,T,...) + pV^*$$

où V^* est tel que $p = -\partial F/\partial V|_{V=V^*}$.

Nous voyons donc que $G(T,p,...)$ est la transformée de Legendre de l'énergie libre $F(V,T,...)$ par rapport à la variable V. On appelle d'ailleurs G enthalpie libre pour bien marquer cette distinction. Vous avez remarqué que les conventions de signe en thermodynamique ont induit quelques changement de signe par rapport à nos définitions d'avant, mais cela ne change pas les concepts.

De façon générale, considérons l'énergie libre d'un système $F(T,X_1,X_2,...X_n)$ où les X_i sont des paramètres extensifs. La variable conjuguée au paramètre extensif X_i est $x_i = -\partial F/\partial X_i$.

Si maintenant nous mettons notre système en contact avec un réservoir de X_i à x_i fixée, la quantité extensive *X_i* va évoluer pour minimiser la quantité $\Phi = F(T,X_1,X_2,...X_n) + x_iX_i$ où x_iX_i est le travail effectué par le changement de X_i. Le nouveau potentiel thermodynamique Φ dépend du paramètre intensif x_i et non du paramètre extensif X_i. Nous pouvons ainsi passer d'un potentiel thermodynamique à un autre en prenant des transformée de Legendre successif. Le choix du potentiel thermodynamique adéquat pour un problème se fait en fonction des paramètres (intensif ou extensif) que l'expérimentateur peut contrôler. Les délices des définitions de la thermodynamique font que le signe de la variable conjuguée x_i n'est pas toujours - la dérivée partielle , mais parfois + . Il faut prendre cela en compte lors des passages entre les potentiels.

Optique géométrique et caustiques.

Quand on regarde la surface de la mer par une belle journée, on distingue des endroits sur la mer fortement éclairés : la surface n'étant pas plane, certaines directions *concentrent* la lumière. De façon générale, considérons une courbe que l'on présente comme l'interface d'une surface réfléchissante. Quand on éclaire cette courbe par une lumière parallèle, l'enveloppe des rayons réfléchis forme une nouvelle courbe qu'on appelle une *caustique*. Cette courbe n'est rien d'autre que la transformée de Legendre de la courbe réfléchissante.

La forme des cristaux et les évolutes.

Certaines classe d'équations différentielles.

Des équations différentielles où la dérivée apparaît sous une forme plus compliquée que la fonction peuvent avoir une solution simple en transformée de Legendre. Considérons par exemple l'equation pour $y(x)$

$$y = xy' + f(y')$$

Il est évident que si l'on passe de (x, y) à (p, z) où la fonction $z(p)$ est la transformée de Legendre de $y(x)$, nous avons par définition des transformée de legendre $y - xy' = -z$ et donc

$$z(p) = -f(p)$$

Il suffit donc simplement d'inverser la transformée pour trouver la fonction $y(x)$.

Exemple 18.3 Résoudre $y = xy' - y'^2/2$. Nous avons immédiatement $z(p) = p^2/2$,et donc $y(x) = x^2/2$

Les ondes de Fisher.

Fisher est un des pères fondateurs de la théorie de l'évolution moderne. Vers 1930, il a utilisé un modèle simple pour modéliser la propagation des mutants bénéfiques au sein d'une population étendu dans l'espace. Soit $u(x, t)$ la densité locale relative de mutants au temps t à l'abscisse x, et soit s leur avantage sélectif (fitness). L'équation de Fisher est

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + su(1-u) \tag{18.3}$$

nous voyons que l'équation de Fisher est une équation de la chaleur, avec un terme $u(1-u)$ en plus. L'origine de ce terme vient des modèles d'évolution, mais notons que dans les régions de l'espace où soit il n'y a pas de mutant ($u = 0$) soit le mutant est

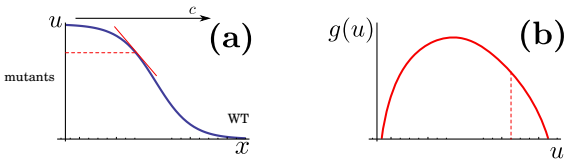


FIGURE 18.4 – La propagation d'onde de Fisher.

dominant ($u = 1$), ce terme vaut zéro. Ce terme supplémentaire ne contribue que dans les régions frontières entre les mutants et les sauvages [3].

Cette équation a été généralisée par Kolmogorov et intervient dans de très nombreux domaine de la physique. Nous nous intéressons au cas où une moitié de l'espace est occupée par le mutant et l'autre moitié par le sauvage (WT), et nous voulons calculer comment la frontière se déplace vers la région des sauvages (figure 18.4a) .

Nous cherchons de plus des solutions propagatives $u(x,t) = u(x-ct)$ où c est la vitesse de propagation du front. En posant $s = x - ct$, l'équation (18.3) se transforme en une équation différentielle ordinaire de second ordre non-linéaire :

$$D\frac{d^2u}{ds^2} + c\frac{du}{ds} + su(1-u) = 0 \tag{18.4}$$

Faisons maintenant un changement de fonction et posons

$$g(u) = -\frac{du}{ds}$$

(figure 18.4b). En dérivant une fois la relation ci-dessus par rapport à s, nous avons

$$\frac{d^2u}{ds^2} = -\frac{du}{ds}\frac{dg}{du} = g\frac{dg}{du}$$

et nous pouvons donc écrire une équation plus simple pour g à partir de (18.4) :

$$Dg\frac{dg}{du} + cg + su(1-u) = 0 \tag{18.5}$$

Remarquons que quand $u \to 0$, $g(u) \to 0$ et donc $dg/du = g/u$. En posant $p = dg/du|_{u=0}$, en divisant par u l'équation (18.5) et en considérant la région $u \to 0$, nous obtenons

$$Dp^2 + cp + s = 0 \tag{18.6}$$

Et nous voyons qu'une solution existe seulement si

$$c \geq 2\sqrt{Ds}$$

3. En évolution, sauvage (Wild type en anglais) veut simplement dire non-mutant

Nous pouvons donc trouver une infinité de solution propagative de vitesses différentes. Cependant, la vitesse qui est sélectionnée et la vitesse *minimum*

$$c = 2\sqrt{Ds}$$

Cela est facile à voir par résolution numérique ; des arguments analytiques peuvent également être donné pour prouver cela, mais cela nous emmènerait dans des détails techniques en dehors de ce cours.

Exemple 18.4 largeur du front

La largeur du front est définie par

$$W = \int_{-\infty}^{\infty} u(1-u)dx$$

Calculer ce largeur pour la solution propagative.

Help : Montrer que vous pouvez définir la vitesse du front par

$$c = \frac{\partial}{\partial t}\int_{\mathbb{R}} [u(x,t) - u(x,0)]\, dx$$

échanger l'ordre de dérivation sur t et intégration sur x, utiliser l'équation de Fisher (18.3) pour démontrer que $W = c/s$.

Un peu plus sur les enveloppes.

Notre façon de calculer les enveloppes était légèrement limité, et par exemple ne nous permet pas de calculer de quelle famille un cercle est l'enveloppe. Cela vient de notre définition d'une droite par l'équation $y = px - g$. Généralisons un peu l'équation d'une famille de droite, paramétré par une variable t :

$$\Delta_t \ : \ u(t)x + v(t)y + w(t) = 0$$

où nous supposons que $u(t)$ et $v(t)$ ne s'annule pas ensemble. L'enveloppe de ces droites peut également être paramétré par la même variable t et nous le représentons donc par $(x(t), y(t))$. L'équation d'enveloppe s'écrit donc

$$\begin{aligned} u(t)x(t) + v(t)y(t) + w(t) &= 0 \qquad (18.7)\\ u(t)x'(t) + v(t)y'(t) &= 0 \qquad (18.8)\end{aligned}$$

En différentiant (18.7) et en y retranchant (18.8), nous obtenons deux nouvelles équations :

$$\begin{aligned} u'(t)x(t) + v'(t)y(t) + w'(t) &= 0\\ u(t)x(t) + v(t)y(t) + w(t) &= 0\end{aligned}$$

Ce qui nous donne directement l'équation de l'enveloppe :

$$\begin{aligned} x(t) &= \frac{vw' - v'w}{uv' - u'v} \\ y(t) &= \frac{vw' - v'w}{uv' - u'v} \end{aligned}$$

Exercices.

§ 18.2 **Parabole.**

Soit une ligne D et un point F en dehors de cette ligne. Soit la famille de ligne bissectrice du segment joignant F à un point de D. Quelle est l'enveloppe de ces lignes ? [Réponse : parabole de foyer F et de directrice D. Pour voir cela, choisir D comme l'axe des x, et $F = (0, a)$ sur l'axe y. L'équation des droites bissectrices est donnée par $y = px - (1/2)ap^2$ où p est la pente de la bissectrice. Comme $g(p) = (1/2)ap^2$, la transformée de Legendre, $f(x) = (1/2a)x^2$ est l'enveloppe].

§ 18.3 **Arc minimum.**Trouver la courbe $y(x)$ entre deux points x_0 et x_1 de longueur d'arc minimum tel que $y'(x_0) = p_0$ et $y'(x_1) = p_1$ et $p_0 p_1 < 0$.

19 Intégrale de Lebesgue.

Note : L'intégrale de Lebesgue est très peu utilisée par les physiciens, puisque le genre de fonctions qu'elle est capable de traiter se rencontre peu en physique [1]. Ce chapitre ne rentre donc pas dans les techniqualités de cette intégrale et il est donné pour la culture générale [2].

19.1 Introduction.

Riemann a donné une définition de l'intégrale d'une fonction dans les années 1840 comme la limite d'une somme finie

$$\mathcal{I} = \int_I f(x)dx = \lim_{N\to\infty} \sum_{i=0}^{N-1} f(x_i)(x_{i+1} - x_i)$$

et a donné les conditions de l'existence de cette intégrale [3]. Très grossièrement, il faut que la fonction soit bornée et continue sauf peut-être en un nombre fini de point. En effet, sur un petit intervalle $[x_i, x_{i+1}[$, la fonction continue $f(.)$ prend des valeurs qui ne sont pas trop loin de $f(P)$ où P est un point dans cet intervalle. Cela nous permet de montrer que la somme des erreurs que l'on fait tend vers 0 quand ces intervalles deviennent de plus en plus petit.

Qu'en est il des fonctions qui sont *très* discontinues, par exemple la fonction $1_Q(.)$ qui vaut 1 si $x \in \mathbb{Q}$ et 0 sinon ? La réponse élémentaire serait de dire que l'intégrale de ces fonctions n'existe pas.

Dire à un mathématicien que quelque chose n'est pas possible est le plus sûr moyen de le lancer sur la piste de « pourquoi pas ? », de lui faire construire des objets qui enfreignent l'impossibilité et de démontrer que ces objets ont une réelle utilité pour des problèmes physiques. $\sqrt{-1}$ et $\log(-1)$ n'existaient pas, ils ont créé un monde de nombres complexes et de surfaces de Riemann. D'un point en dehors d'une ligne, on ne pouvait avoir qu'une seule ligne parallèle à celui-ci ; en violant cette impossibilité, nous avons eu les géométries non-Euclidiennes. On pourrait multiplier les exemples de ce genre.

1. Si l'avions que vous avez construit chute, ce n'est pas parce que vous avez utilisé l'intégrale de Riemann au lieu de l'intégrale de Lebesgue dans vos calculs.

2. Pour goûter la beauté de cette théorie, voir David M. Bressoud : *A Radical Approach to Lebesgue's Theory of Integration.*

3. Le théorème le plus général a été donné par Lebesgue vers 1900.

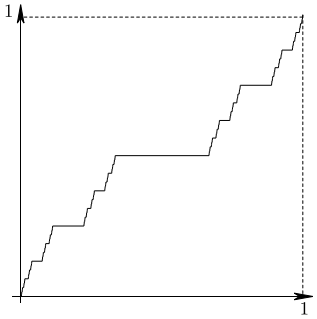


FIGURE 19.1 – La fonction de cantor.

A partir de 1840, les mathématiciens ont construit des fonctions de plus en plus pathologiques pour lesquelles ont pouvait quand même construire une intégrale. Il faut préciser que l'on ne comprenait pas vraiment encore le concept de nombre réel [4] et que ces constructions ont beaucoup contribué à cette compréhension. Voici quelques uns.

Exemple 19.1 L'escalier du diable ou la fonction de cantor

Cantor a construit une fonction $f_c()$ monotone croissante sur l'intervalle $[0,1]$ telle que $f(0) = 0$, $f(1) = 1$ et dont la dérivée est nulle presque partout (figure 19.1).

La fonction se construit de façon itérative. $f_0(x) = x$. Pour $f(1)$,nous prenons la fonction qui est constante sur le tiers central de l'intervalle $[0,1]$: $f_1(x) = 1/2$ si $x \in [1/3, 2/3]$. La fonction est linéaire (affine pour être exacte) sur $[0,1/3]$ et $[2/3,1]$.

On procède de la même façon pour la suite : chaque intervalle où $f_n(x)$ n'est pas constante est découpé en trois morceau où sur la partie centrale $f_n(x)$ est constante et égale à la moyenne de ses deux bords selon f_{n-1}. Sur les deux autres morceaux, la fonction est linéaire et la fonction $f_n(x)$ reste toujours continue.

On peut démontrer que $f_n()$ converge vers $f_c()$ et que

$$|f_c(x) - f_n(x)| \leq 2^{-n} \ \forall x \in [0,1]$$

La fonction de Bolzano est un autre exemple de ce type.

Exemple 19.2 Le monstre de Weierstrass

Ce monstre définie par la série

$$f(x) = \sum_{k=1}^{\infty} a^k \cos(b^k \pi x)$$

4. L'incroyable complexité des nombres réels a été mis à jour par Cantor dans les années 1880 (voir le chapitre 22)

où b est un entier est continue partout et dérivable nulle part. (exercice : à quelle condition doit satisfaire a pour assurer la convergence ?)

Des fonctions aussi pathologiques sont rares ou inexistantes en physique. L'abstraction que l'on appelle le mouvement brownien en est un exemple, de même que la fonction de partition en physique statistique de certains systèmes.

L'exemple le plus utilisé des fonctions pathologiques est celle de Dirichlet $1_Q(x)$ qui vaut 1 si $x \in \mathbb{Q}$ et zéro sinon. Évidemment, nous ne pouvons pas tracer cette fonction, mais nous voyons bien la raison de la non (Riemann) intégrabilité : un intervalle, aussi petit qu'il soit, contient au moins un nombre rationnel et un nombre irrationnel. Nous voyons donc qu'aussi petit soit un intervalle, la fonction 1_Q saute sauvagement d'une valeur à une autre et l'erreur de l'approximer par $f(P)$ où P est un point à l'intérieur de l'intervalle n'est plus petit.

La fonction de Dirichlet nous indique comment construire des fonctions pathologiques en général. Il faut pour cela se donner un ensemble $\mathcal{A} \in I$ où I est un intervalle de $\mathbb{R}$ et exiger que l'ensemble $\mathcal{A}$ soit plus compliqué qu'une simple union finie d'intervalle. Donner alors un certain comportement à la fonction si $x \in \mathcal{A}$ et un autre si $x \in I - \mathcal{A}$.

19.2 Théorie de la mesure.

L'intégration des fonctions pathologiques est/devrait être une généralisation de l'intégrale de Riemann. Après des années de tentative, c'est Lebesgue en 1902 qui a réalisé que la clef de cette généralisation réside dans la *mesure* d'un ensemble $\mathcal{A}$ quelconque.

Nous savons mesurer la longueur d'un intervalle $[a, b]$:

$$\mu([a, b]) = b - a$$

et nous savons mesurer la longueur d'une *union* d'intervalles *disjoints* en sommant chaque mesure. Si

$$\mathcal{A} = \bigcup_{i=1}^{k} [a_i, b_i]$$

alors

$$\mu(\mathcal{A}) = \sum_{i=1}^{k} \mu([a_i, b_i])$$

Nous pouvons généraliser la *mesure* de n'importe quel ensemble en suivant ce chemin.

Définition 2 Une mesure.

Une mesure μ doit avoir les propriétés suivantes :

$$\mu(\mathcal{A}) \geq 0 \tag{19.1}$$
$$\mu(\emptyset) = 0 \tag{19.2}$$
$$\mu\left(\bigcup_{i\in P} \mathcal{A}_i\right) = \sum_{i\in P} \mu(\mathcal{A}_i) \text{ si } \mathcal{A}_i \cap \mathcal{A}_j = \emptyset \tag{19.3}$$

Pour la troisième propriété (19.3), nous exigeons une union *au plus dénombrable*. L'ensemble $\{1, 2, 3\}$ est dénombrable, ainsi que l'ensemble des nombres entiers, ou rationnels ou algébriques. L'ensemble des nombres réels n'est pas dénombrable (voir le chapitre 22).

Pour pouvoir généraliser notre mesure habituelle qui sert à l'intégration, nous choisissons une mesure telle que $\mu[a, b] = b - a$. Cette mesure est appelé la mesure de *Lebesgue*.

Pour mesurer n'importe quel sous ensemble de $\mathbb{R}$ et pas seulement des intervalles, nous utilisons une *couverture* par des intervalles : soit un ensemble $\mathcal{A} \in \mathbb{R}$. Nous pouvons toujours trouver un ensemble $\mathcal{B} \supset \mathcal{A}$ tel qu'il soit formé d'une union *dénombrable* d'intervalles. Par exemple, $[0, 1]$ est une couverture de l'ensemble des nombres rationnels entre 0 et 1. De plus, comme $\mathcal{B}$ est formé d'intervalles, nous savons calculer sa mesure. Notons également que pour un ensemble $\mathcal{A}$ quelconque, nous pouvons trouver une infinité de couverture. La mesure de Lebesgue d'un ensemble $\mathcal{A}$ *est le minimum* des mesures des couvertures $\mathcal{B}$.

Définition 3 La mesure de Lebesgue.

1. La mesure d'un intervalle $[a, b]$ est $b - a$.
2. Soit $\mathcal{B}_\alpha$ des couvertures par intervalle d'un ensemble $\mathcal{A}$. Alors

$$\mu(\mathcal{A}) = \min_\alpha \{\mathcal{B}_\alpha\}$$

Le point essentiel ici est le côté dénombrable, qui peut être plus grand que simplement une union finie, mais pas trop grand pour ne pas tomber dans des complications insurmontable. Si vous avez remarqué, toutes les fonctions pathologiques que nous avons rencontré ont été construit par itération et nous pouvons dénombrer les singularités qu'elles possèdent.

Le point crucial de cette mesure est la suivante : La mesure de Lebesgue d'un *point isolé* est nulle ! (exercice : démontrer en utilisant la définition 3).

$$\text{si } x \in \mathbb{R} \text{ alors } \mu(x) = 0$$

Cela ne choque pas notre intuition qu'un point seul soit de mesure nulle. Par contre, la propriété (19.3) implique maintenant que

$$\mu(\mathbb{Q}) = 0 \tag{19.4}$$

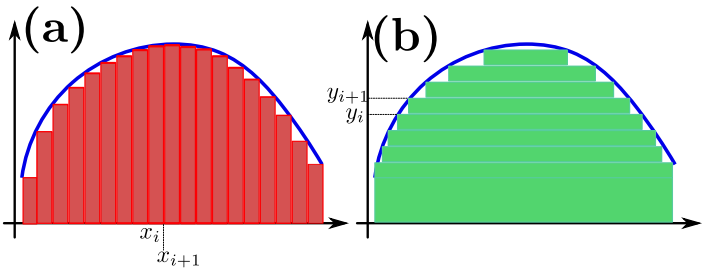


FIGURE 19.2 – L'idée de l'intégrale de Lebesgue.

En effet, l'ensemble des nombres rationnels est dénombrable, donc $\mathbb{Q}$ est une union dénombrable de points isolés, chacun de mesure nulle. L'ensemble $\mathbb{Q}$ ne pèse pas grand chose comparé à n'importe quel intervalle de $\mathbb{R}$! Cette conséquence est compatible avec la classification de Cantor des différents infinis (voir chapitre 22). On appelle un ensemble de mesure nulle un ensemble *négligeable*.

Nous pouvons aussi définir une égalité *presque partout*. Deux ensemble A et B sont égaux (presque partout) si la mesure de leur différence est nulle. De même, deux fonctions $f()$ et $g()$ sont égaux partout si elles diffèrent sur un ensemble négligeable de point.

19.3 L'intégrale de Lebesgue.

Maintenant que nous disposons de la mesure de Lebesgue, nous avons en fait fait l'essentiel du chemin. Toute la complexité a été enfermée dans la mesure et le reste est une petite promenade.

Pour définir son intégrale, Lebesgue, a utilisé notre intuition de l'intégrale de Riemann comme aire sous la courbe, en le couvrant de rectangles *verticales* (figure 19.2a). Bien sûr, nous pouvons utiliser, pour cette couverture, des rectangles horizontales (figure 19.2b).

Soit l'intervalle $[y_i, y_{i+1}]$, et soit l'ensemble $\mathcal{A}_i$ de point x tel que $f(x) \in [y_i, y_{i+1}]$. Nous définissons alors l'intégrale comme la limite de la somme

$$\mathcal{I} = \int f d\mu = \sum_i \mu(\mathcal{A}_i)(y_{i+1} - y_i)$$

quand $\max_i |y_{i+1} - y_i| \to 0$.

On peut également exprimer la même idée en utilisant l'intégrale de Riemann et la fonction adjointe $f^*()$:

$$f^*(t) = \mu\left(\{x | f(x) > t\}\right)$$

et alors

$$\int f d\mu = \int_0^\infty f^*(t)dt$$

Noter que nous avons défini l'intégrale de Lebesgue pour des fonctions bornées positives. La généralisation à des fonctions de signe quelconque est trivial.

§ 19.1 Démontrer que

$$\int 1_Q d\mu = 0$$

Une troisième façon de voir l'intégrale de Lebesgue est d'utiliser la notion de presque partout : Si $g() = f()$ presque partout et que $g()$ possède une intégrale de Riemann, alors

$$\int f d\mu = \int_I g(x)dx$$

20 Les intégrales de chemin.

20.1 Introduction.

Nous connaissons le concept de l'intégrale [1] d'une fonction d'une variable réelle $f(x)$ sur un intervalle $I = [a, b]$

$$\mathcal{I} = \int_I f(x)dx = \lim_{N\to\infty} \sum_{i=0}^{N-1} f(x_i)(x_{i+1} - x_i) \tag{20.1}$$

La recette est simple : l'intégrale $\mathcal{I}$ n'est que la limite d'une somme finie d'éléments : nous décomposons l'intervalle en petits éléments de bord x_i, avec $x_0 = a$, $x_1 = b$. Ensuite, nous multiplions la valeur représentative de la fonction $f(x_i)$ sur chaque sous intervalle par la longueur de ce sous intervalle $x_{i+1} - x_i$ (qui prend le nom de dx dans l'intégrale).

On peut facilement généraliser ce concept et définir l'intégrale des fonctions de plusieurs variables, des fonctions complexes (qui sont, fondamentalement, des fonctions de deux variables), des champs de tenseurs, ... Cela reste toujours le produit d'une valeur par la *mesure* du petit sous domaine.

Peut on généraliser l'intégrale à des objets vraiment différent ? Soit par exemple l'ensemble des fonctions définies sur un intervalle $[a, b]$ et soit une fonctionnelle $\mathcal{F}[.]$ donnée. Rappelons qu'une fonctionnelle est une *fonction de fonction*, qui prend en entrée une *fonction* et produit en sortie un nombre. Par exemple,

$$\mathcal{F}[f] = \int_a^b \left[f'(x)^2 + f(x)^2\right]^2 dx$$

Nous avons vu par exemple que la fonctionnelle *action* joue un rôle fondamental en mécanique ; le chapitre 12 sur le calcul variationnel était dédié à l'étude général de l'extremum des fonctionnelles.

Peut on définir l'intégrale d'une *fonctionnelle* sur un ensemble de fonctions ? De façon analogue à l'expression (20.1), nous souhaitons définir l'intégrale d'une fonctionnelle

1. La description précise de ce concept a été donné par Riemann dans les années 1840 et est appelé l'intégrale de Riemann. Cependant, elle ne s'applique qu'à des fonctions assez simple, en gros des fonctions discontinues en un nombre fini de point. Cela est amplement suffisant pour décrire le monde physique, mais non l'ensemble des fonctions des mathématiciens. Cela a pris environ 80 ans pour qu'une théorie plus générale, appelée l'intégrale de Lebesgue, soit formulée.

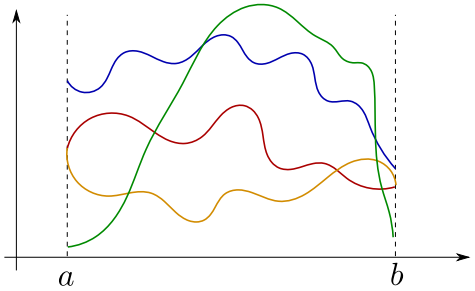


FIGURE 20.1 – Quelques fonctions sur l'intervalle $[a, b]$. Étant donnée une fonctionnelle $\mathcal{F}[.]$, l'intégrale de la fonctionnelle $\int_{\mathcal{E}} \mathcal{F}[f(.)]\mathcal{D}f$ est une somme correctement normalisée de la fonctionnelle pris sur toutes les fonctions de l'ensemble $\mathcal{E}$.

$\mathcal{F}[]$

$$K = \int_{\mathcal{E}} \mathcal{F}[f]\mathcal{D}f$$

en additionnant la valeur de la fonctionnelle sur toutes les fonctions appartenant à un ensemble $\mathcal{E}$ en les normalisant par un facteur $\mathcal{D}f$ qui donnerait une mesure finie à l'intégrale (figure 20.1).

La réponse est oui sous conditions. Il se trouve que les phénomènes physiques qui ont besoin de ces intégrales satisfont ces conditions. Avant de donner plus de sens à cette intégrale, voyons quelques problèmes physiques qui y font appel.

20.2 Exemples fondamentaux.

Mouvement Brownien. Considérons une particule brownien à une dimension, qui démarre au temps t_0 au point x_0. Nous souhaitons de connaître la densité de probabilité

$$P(x_1, t_1|x_0, t_0) \tag{20.2}$$

de trouver la particule au point x_1 au temps t_1. Pour arriver à ce point, la particule peut prendre de nombreux trajectoires $\mathcal{C}$; certains de ces trajectoires sont peu probable (comme par exemple celle qui fait le tour de la terre en 2 ms), d'autres plus. Nous pouvons donner une *densité* de probabilité $\mathcal{F}[\mathcal{C}]$ à chaque trajectoire $\mathcal{C}$. La probabilité (20.2) est alors (figure 20.2)

$$P(x_1, t_1|x_0, t_0) = \frac{\int_{\mathcal{E}_0} \mathcal{F}[\mathcal{C}]\mathcal{D}\mathcal{C}}{\int_{\mathcal{E}_1} \mathcal{F}[\mathcal{C}]\mathcal{D}\mathcal{C}}$$

Nous avons caché quelques difficultés techniques, comme par exemple comment formuler la probabilité d'une trajectoire $\mathcal{F}[\mathcal{C}]$, nous y viendront plus tard. Mais la possibilité

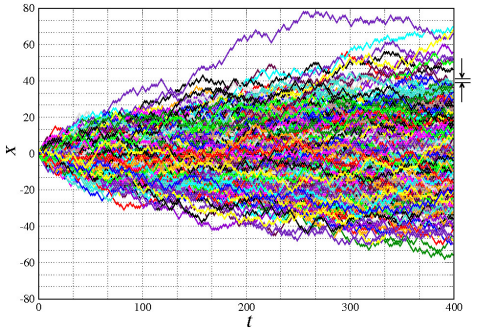


FIGURE 20.2 – Une simulation de $M = 500$ trajectoires brownienne commençant en $x = 0$ au temps $t = 0$. La probabilité $P(x,t)dx$ de trouver la particule brownienne en $[x, x + dx]$ au temps t est le nombre de trajectoire qui arrivent dans cet intervalle divisée par le nombre total de trajectoire, à la limite $M \to \infty$.

de penser au processus stochastic comme un ensemble de trajectoire avec des poids probabilistes est un outil très puissant pour aborder ces phénomènes [2].

Physique statistique des polymères. Soit un système qui peut se trouver dans différents états η d'énergie $E[\eta]$. La loi (l'hypothèse) fondamentale de la physique statistique [3] postule la probabilité de trouver le système en η par

$$P(\eta) = \frac{1}{Z} e^{-\beta E[\eta]}$$

où $\beta = 1/T$ est l'inverse de la température et Z, la constante de normalisation est

$$Z = \sum_{\eta} e^{-\beta E[\eta]} \tag{20.3}$$

la constante de normalisation, appelé la fonction de partition, est bien plus qu'une constante et nous donne l'énergie libre $F = -T \log Z$ du système, d'où nous déduisons toute la thermodynamique.

Si le système que nous étudions est un polymère, η qui désigne sa conformation est un chemin, $E[\eta]$ est la fonctionnelle énergie de cette conformation, et Z est une intégrale de chemin. Il serait très utile donc de pouvoir donner un sens à cet objet.

On peut avoir, sans calcul, une intuition de cette intégrale. soit η_0 la conformation d'énergie minimum ; quand $T \to 0$, l'exponentielle est dominée par $E[\eta_0]$ et les petits

2. Voir mon cour sur les processus stochastiques.
3. Voir mon cours de physique statistique.

écarts par rapport à cette conformation sont très pénalisés : le polymère se trouve dans son état fondamental. Au fur et à mesure que T augmente, le polymère peut explorer des conformations plus lointaines de l'état fondamental.

Comme nous le verrons sous peu, une conformation de polymère est équivalent à une trajectoire de la particule brownienne et les deux problèmes ci-dessus sont en réalité le même.

La mécanique quantique. En mécanique classique, Pour chaque trajectoire $x(t)$ reliant deux point (t_1, x_1) et (t_2, x_2), nous avons un coût appelé *action*

$$S[x(t)] = \int_I \mathcal{L}(\dot{x}, x, t)dt$$

et nous savons que la trajectoire choisie par la particule est celle qui optimise l'action $\delta S = 0$.

Cependant, nous savons que quand nous descendons dans les échelles et que l'action devient comparable à la constante de Planck $\hbar$, les phénomènes quantiques et interférences commencent à apparaître. L'interprétation de Feynman [4] de ce phénomène est la suivante : la particule prend *toutes* les trajectoires reliant les deux points. Une fonctionnelle associe à chaque trajectoire un coût de la forme $\exp\{i\beta S[x(t)]\}$ où $i^2 = -1$ et $\beta = 1/\hbar$ et S est l'action *classique* associée à la trajectoire. . Le propagateur K est donnée par une sommation sur toutes les trajectoires

$$K(x_2, t_2|x_1, t_1) = \sum_{x[t]} e^{i\beta S[x]} \tag{20.4}$$

Le propagateur est un concept similaire à la probabilité $P(x_1, t_1|x_0, t_0)$ de l'expression (20.2). Le calcul effectif du propagateur nous amène naturellement à l'équation de Schrödinger comme nous le verrons sous peu.

Notons l'analogie entre le mouvement quantique (eq. 20.4) et le mouvement Brownien ou le polymère (eq. 20.3) : dans les deux cas, la fonctionnelle est sous forme exponentielle, mais dans le cas quantique, c'est une exponentielle complexe. A priori donc, toutes les trajectoires contribuent le même poids au propagateur. Ceci dit, pour les trajectoires dont $S \gg \hbar$, un petit écart δx de la trajectoire provoque de grande oscillations de l'exponentielle et les contributions se neutralisent. Si la particules est macroscopique, nous avons toujours $S \gg \hbar$ et la seule trajectoire qui contribue effectivement à la somme est celle qui minimise S et nous retrouvons ainsi la mécanique classique.

20.3 Calcul des intégrales de chemin (I).

Nous pouvons calculer une intégrale de Riemann (eq. 20.1) de deux façon : (i) la méthode dure consiste à effectuer directement la somme discrète ; (ii) nous pouvons

4. Feynman & Hibbs, *Quantum Mechanics and Path Integrals*.

aussi utiliser le théorème fondamental de l'Analyse est de faire une sorte d'opération d'antidérivation (trouver la primitive F) et ensuite calculer simplement $F(b) - F(a)$.Évidemment, il faut avoir fait la méthode (i) au moins quelque fois pour comprendre vraiment ce qu'est une intégrale. Ensuite, nous pouvons utiliser la deuxième méthode dans les opérations de tous les jours.

Le calcul des intégrales de chemins suit la même logique. Dans cette section, nous allons utiliser la méthode dure pour comprendre les intégrales de chemins ; la prochaine section est dédiée à une sorte de méthode facile, qui fera intervenir les fonctions de Green.

D'abord, qu'est ce qu'une fonction d'une variable *continue* sur un intervalle $[a, b]$? L'ensemble $\mathbb{R}$ est vraiment trop grand pour la compréhension humaine (voir le chapitre 22). Tout ce que nous savons faire est d'imaginer des ensembles finis et ensuite imaginer une sorte de limite. Pour manipuler et prendre en main une fonction, nous échantillonnons l'intervalle $[a, b]$ par $N + 1$ nombres x_i, et notons la valeur de la fonction pour chacun de ces nombres

#1	$x_0 = a$	x_1	x_2	...	$x_N = b$
#2	y_0	y_1	y_2	...	$y_N = b$

Si notre échantillonnage est suffisamment fin, nous avons une très bonne approximation de la fonction. Dans la limite $N \to \infty$, nous pouvons dire [5] que la fonction et le tableau représentent le même objet. Pour un échantillonnage fin et fixé une fois pour toute, nous pouvons dire que chaque fonction est donnée par un ensemble $\{y_0, y_1, ..., y_N\}$.

Dans ce cas, une *fonctionnelle* est simplement une fonction de $N+1$ variable $\mathcal{F}[f] = \mathcal{F}(y_0, y_1, ...y_N)$ dans la limite où nous faisons $N \to \infty$ (figure 20.3).

Prenons par exemple la fonctionnelle

$$\mathcal{F}[f] = \int_a^b f'(x)^2 dx \tag{20.5}$$

Si nous discrétisons l'intervalle par un pas d'échantillonnage de $\ell = (b - a)/N$, nous avons

$$f'(x_i) \approx (y_{i+1} - y_i)/\ell$$

où $y_i = f(x_i)$ et alors

$$\mathcal{F}[f] = \frac{1}{\ell} \sum_{i=0}^{N-1} (y_{i+1} - y_i)^2$$

Nous voyons que ce processus de discrétisation nous donne maintenant un moyen de calculer l'intégrale de chemins de la fonctionnelle comme l'intégrale d'une fonction de $N + 1$ variables :

$$\int_{\mathcal{E}} \mathcal{F}[f] \mathcal{D}f = A(N) \int_{I^{N+1}} \mathcal{F}(y_0, y_1, ...y_N) dy_0 dy_1 ... dy_N \tag{20.6}$$

5. Plus exactement quand la première ligne liste tous les nombres réelles $\in [a, b]$

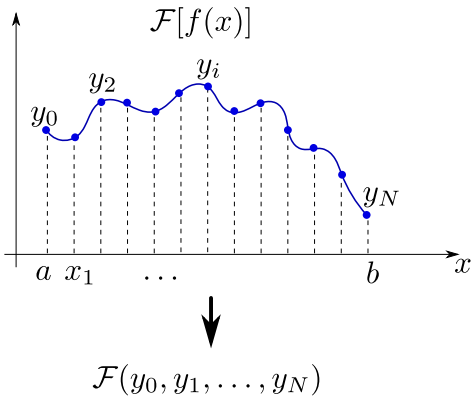


FIGURE 20.3 – Discrétisation d'une fonctionnelle : on échantillonne la fonction en $N+1$ points x_i, et on calcule la fonctionnelle comme une fonction des $N+1$ $y_i = f(x_i)$.

En faisant varier les y_i, nous balayons l'ensemble des fonctions sur l'intervalle $[a, b]$. Nous devons également normaliser la somme par le facteur $A(N)$: dans les intégrales de Riemann, le facteur dx sert à cela ; ici, nous avons également le même problème quand $N \to \infty$ et nous devons trouver l'équivalent de dx pour les intégrales de chemin. L'expression (20.6), correctement normalisée, est la définition de l'intégrale de chemin, quand elle existe.

Exemple 20.1 gradient carré

Soit la fonctionnelle

$$\mathcal{F}[f] = \exp\left(-\gamma \int_a^b f'(x)^2 dx\right)$$

Calculer son intégrale de chemin sur l'ensemble $\mathcal{E}$ des fonctions dérivables sur $[a, b]$ telle que $f(a) = f(b) = 0$.

Avant de calculer l'intégrale dans son entier, calculons une intégrale simple du genre

$$I = \int_{-\infty}^{\infty} \exp\left[-(y_{i+1} - y_i)^2 - (y_i - y_{i-1})^2\right] dy_i$$

où nous sommons sur la variable y_i en tenant constante y_{i+1} et y_{i-1}.

20.4 Digression sur le mouvement Brownien.

Le mouvement Brownien joue un rôle fondamental dans beaucoup de branche de la physique et sous tend l'ensemble des problèmes que nous abordons ici. Les intégrales

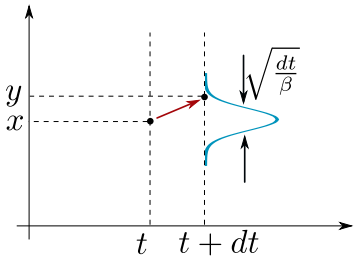


FIGURE 20.4 – Le déplacement d'une particule Brownienne, de la position x à la position y pendant un temps infinitésimal dt est gouvernée par la densité de probabilité (20.7).

de chemin et les équations différentielles stochastiques ont été inventé pour aborder ce problème. Il est utile de s'attarder un peu sur son formalisme.

La façon la plus simple d'aborder ce problème est de considérer un problème spatialement discret, avec la particule qui a une certaine densité de probabilité par unité de temps de sauter instantanément d'un site à un site voisin. Par exemple, nous pouvons avoir

$$W(i \to i \pm 1) = \alpha$$

c'est à dire que pendant un temps dt, une proportion αdt des particules sur le site i sauteront sur le site $i \pm 1$. Nous avons résolu ce problème par exemple dans le problème 3.9. Notons cependant que α doit être un coefficient microscopique ; si le pas de discrétisation de l'espace est ℓ, le coefficient de Diffusion D qui est une quantité mesurable expérimentalement est donnée par $D = \alpha\ell^2$.

On peut également souhaiter aborder ce problème dans un espace continue. Pour cela, nous devons nous donner une probabilité de saut du point x au point y dans l'intervalle de temps dt. Une façon [6] de faire cela correctement est de postuler

$$p(x \to y, dt) = A \exp\left(-(x-y)^2/2Ddt\right) \tag{20.7}$$

Cela veut dire que pendant un temps dt, une proportion $p(x \to y, dt)dy$ sautera de la position x à la position $[y, y + dy[$. La constante de normalisation A est trouvée par la condition $\int_I p(x \to y, dt)dy = 1$ et vaut donc $A = \sqrt{1/(2\pi Ddt)}$. Le déplacement carré moyen est alors

$$\ell^2 = Ddt \tag{20.8}$$

Nous voyons donc que $\ell^2 \sim dt$ où D est un coefficient de diffusion macroscopique.

6. Ceci est le résultat des travaux de Wiener et de Ito dans les années 1930-1950.

20.5 Calcul des intégrales de chemin (II) et les fonctions de Green.

Considérons un chemin $R(x)$ reliant un point $0, R_0$ à un point x, R ; Ce chemin peut par exemple représenter un polymère à deux dimension où x représente l'abscisse et $R(x)$ l'écart du polymère par rapport à un mur ou un autre polymère à l'abscisse x, comme dans le cas des doubles brins de l'ADN. Supposons que l'énergie du polymère est donnée par la fonctionnelle

$$E[R(x)] = \int_0^x \left\{ \kappa \dot{R}^2 + V(R) \right\} dx \tag{20.9}$$

le premier terme désigne le faite que courber le polymère coûte de l'énergie ; le deuxième terme reflète l'énergie d'interaction du polymère avec le mur. Nous souhaitons calculer l'intégrale fonctionnelle

$$Z(R, x|R_0, 0) = \int_{\mathcal{E}} e^{-\beta E[R]} \mathcal{D}R$$

qui est la fonction de partition de l'ensemble des conformations qui relient $(0, R_0)$ à (x, R). Rappelons que $\beta = 1/T$ et que $F = -T \log Z$ est l'énergie libre de ce système. Dans de nombreux domaine, Z est appelé *le propagateur*, pour des raisons que l'on verra sous peu.

Considérons maintenant la quantité $Z(R, x+dx|R_0, 0)$. D'après notre définition des intégrales de chemin (20.6), nous avons

$$Z(R, x + dx, R_0, 0) = \int_{R'} Z(R', x|R_0, 0) Z(R, x + dx|R', x) dx \tag{20.10}$$

où l'intégrale est une intégrale classique de Riemann. L'expression (20.10) veut « en gros » dire que le nombre de chemin pour aller de A à B égale au nombre de chemin pour aller de A à un point intermédiaire C, multiplié par le nombre de chemin allant de C à B, intégré sur tous les C

$$N_{A\to B} = \sum_C N_{A\to C} N_{C\to B}$$

Considérons le terme $Z(R, x + dx|R', x)$ del l'intégrale qui est un propagateur infinitésimal. Sur cette portion [7],

$$E(R'|R) = \left\{ \kappa \left(\frac{R' - R}{dx} \right)^2 + V(R) \right\} dx$$

7. Rappelons qu'au premier ordre,

$$\int_{a-dx}^{a} f(x)dx = f(a)dx$$

D'après notre schéma de discrétisation (20.6), le propagateur infinitésimal se réduit à un simple scalaire (il n'y a qu'une droite reliant R' à R) :

$$Z(R, x + dx|R', x) = Ae^{-\beta E(R'|R)}$$

et vous remarquerez que le propagateur infinitésimale ci-dessus, pour $V = 0$, n'est rien d'autre que la probabilité de saut d'une particule Brownienne (équation 20.7) libre.

Nous avons donc

$$Z(R, x + dx|R_0, 0) = Ae^{-\beta V(R)dx} \int_{R'} Z(R', x|R_0, 0) \exp\left(-\frac{\beta\kappa}{dx}(R' - R)^2\right) dR' \tag{20.11}$$

La quantité $Z(R', x|R_0, 0)$ est une fonction de R' que nous pouvons développer au second ordre

$$Z(R', x|R_0, 0) = Z(R, x|R_0, 0) + (R' - R)\frac{\partial Z}{\partial R} + \frac{1}{2}\frac{\partial^2 Z}{\partial R^2}$$

Nous pouvons maintenant utiliser l'expression ci-dessus pour effectuer l'intégrale (20.11) ; la partie droite de l'équation (20.11) devient

$$\sqrt{\pi}Ae^{-\beta V(R)dx}\left\{\left(\frac{dx}{\beta\kappa}\right)^{1/2} Z + \frac{1}{4}\left(\frac{dx}{\beta\kappa}\right)^{3/2}\frac{\partial^2 Z}{\partial R^2}\right\} \tag{20.12}$$

Pour la partie gauche de l'équation (20.11), nous développons $Z(R, x + dx, R_0, 0)$ à l'ordre 1 en dx :

$$Z(R, x + dx|R_0, 0) = Z(R, x|R_0, 0) + \frac{\partial Z}{\partial x}dx \tag{20.13}$$

Nous pouvons maintenant comparer les expressions (20.12) et (20.13) et exiger leur égalité. A l'ordre 0 en dx, nous avons $\exp(-\beta V dx) = 1$ et nous devons donc avoir

$$\sqrt{\pi}A\left(\frac{dx}{\beta\kappa}\right)^{1/2} = 1$$

ce qui nous donne directement le coefficient de normalisation A. En collectant les termes d'ordre 1 en dx, nous trouvons l'équation auquel la fonction $Z(R, x|R_0, 0)$ doit obéir.

$$T\frac{\partial Z}{\partial x} = \frac{T^2}{4\kappa}\frac{\partial^2 Z}{\partial R^2} - V(R)Z \tag{20.14}$$

Nous voyons ici que Z obéit à une équation de type Schrödinger, où T joue le rôle de $\hbar$, 2κ joue le rôle de la masse, x le rôle de temps et R le rôle de x ; de plus, le temps

est *imaginaire*, puisque cette équation ne contient pas de facteur i. Nous avons omis d'inclure la condition initiale :

$$Z(R,0|R_0,0)=\delta(R-R_0)$$

puisque le polymère est fixée à l'origine à R_0. Nous pouvons maintenant inclure la condition initiale dans l'équation (20.14) (cf la section 5.4) ; en d'autre terme, la fonction $Z(R,x+dx|R_0,0)$ est la fonction de Green de l'équation (20.14) et obéit à l'équation

$$T\frac{\partial Z}{\partial x}-\frac{T^2}{4\kappa}\frac{\partial^2 Z}{\partial R^2}+V(R)Z=\delta(x)\delta(R-R_0) \tag{20.15}$$

Si nous connaissons une base propre de l'opérateur ci-dessus, la fonction de Green se calcule immédiatement (cf 8.5).

La méthode que nous avons développé ci-dessus est une méthode générale pour l'intégrale de chemin du type $\exp(E(y))$: au lieu de calculer l'intégrale directement par la définition (20.6), on déduit une équation à dérivée partielle et on calcule sa fonction de Green.

La similarité entre l'expression (20.14) et l'équation de Schrödinger a emmené Feynman à formuler une autre façon d'interpréter la mécanique quantique. Soit l'action classique d'une trajectoire $x(t)$

$$S[x(t)]=\int_{t_0}^{t}\left\{\frac{m}{2}\dot{x}^2-V(x)\right\}dt$$

Une particule quantique ne prend pas le chemin de moindre action, mais *tous* les chemins en même temps, chaque chemin étant pondérer par un facteur de phase $\exp(iS/\hbar)$. Le propagateur est donné par la somme (fonctionnelle) sur les phases de toutes les trajectoires

$$K(x,t|x_0,t_0)=\sum_{x(t)}e^{i\frac{S}{\hbar}}$$

Nous laissons au lecteur le soin de montrer que le propagateur obéit à l'équation de Schrödinger. Le lecteur remarquera que la convergence de certains calculs est moins assuré quand l'integrand est une exponentielle complexe.

20.6 Problèmes.

Problème 20.1 La fusion de l'ADN

Problème 20.2 Le propagateur du photon

Problème 20.3 L'équation de Kolmogorov

Problème 20.4 Limite $T\to 0$ et le mimimum de l'action

21 Les équations de la physique.

Nous avons, à de nombreuses occasions, rencontré les diverses équations de la physique. Nous voudrions donner ici une dérivation simple et intuitive de ces équations. Comme nous le verrons, pour établir ces équations, nous *discrétisons* le système et faisons ensuite un passage à la limite. La méthode plus rigoureuse (et riche et élégante et ...) d'aborder ces sujets est le calcul variationnel (voir le chapitre correspondant) . La méthode utilisée dans ce chapitre est celle qui avait été utilisée par Euler lui-même dans les années 1740 pour fonder le calcul des variations. Cette vue consiste à regarder les équations différentielles comme des équations aux différences, avec un pas Δx qui peut être rendu aussi petit que l'on souhaite. Cette vue a quelque peu disparu des mathématiques au début du XIXème siècle quand Cauchy &Co ont donné de la rigueur aux mathématiques, mais a donné très naturellement lieu au développement des calculs matriciels et la formalisation des espaces vectoriels un siècle plus tard par Hilbert &Co. Avec l'arrivée des ordinateurs et la résolution numérique des équations, cette approche redevient tout à fait naturelle. Regardons quelques cas particuliers.

21.1 Qu'est ce qu'une équation différentielle ?

Prenons la plus simple des équations différentielles

$$y' = f(x) \tag{21.1}$$

sur l'intervalle $[0, 1]$. Découpons cet intervalle en N morceau de largeur Δx dont les bords sont $x_0 = 0, x_1 = 1/N, x_2 = 2/N, ...x_N = 1$. Nous cherchons à déterminer les $N + 1$ valeurs $y_0 = y(x_0)$, $y_1 = y(x_1)$,...$y_N = y(x_N)$. Nous pouvons approximer le terme y' au point x_i par

$$y'(x_i) = \big(y(x_{i+1}) - y(x_i)\big)/\Delta x$$

L'équation (21.1) se transforme alors en un système d'équations *algébriques* :

$$\begin{aligned} y_1 - y_0 &= \Delta x.f(x_0) \\ y_2 - y_1 &= \Delta x.f(x_1) \\ ... &\quad ... \\ y_N - y_{N-1} &= \Delta x.f(x_{N-1}) \end{aligned}$$

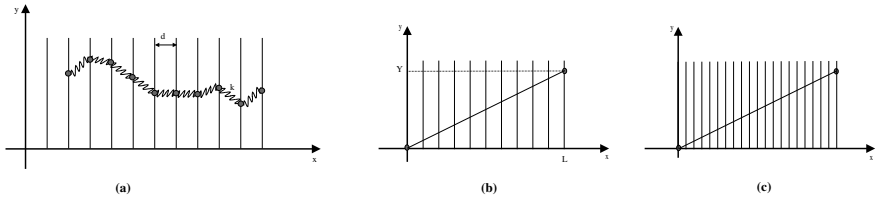


FIGURE 21.1 – Vue discrète d'une corde vibrante.

Si vous regardez bien, nous avons $N+1$ inconnus, mais seulement N équations ! Le système est sous déterminé et n'a pas de solution unique. Pour que le système ait le même nombre d'équations que d'inconnus, il faut ajouter une équation supplémentaire, par exemple $y_0 = a$. C'est cela que nous appelons la condition initiale. Nous avons l'habitude de penser aux équations différentielles comme la donnée de deux choses différentes : une équation de la forme (21.1) et des conditions initiales. En réalité, ces deux choses sont indissociables. Nous pouvons maintenant représenter cela sous forme matricielle

$$\begin{pmatrix} 1 & 0 & ... & & 0 \\ -1 & 1 & 0 & .. & 0 \\ 0 & -1 & 1 & .. & 0 \\ & & ... & & \\ 0 & & ... & -1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_N \end{pmatrix} = \begin{pmatrix} a + \Delta x.f(x_0) \\ f(x_1) \\ . \\ . \\ f(x_N) \end{pmatrix}$$

ou de façon plus succincte $Ay = f$. Nous avons maintenant un système de N équations et N inconnus équilibré. A partir de là, nous pouvons utiliser toute la puissance des techniques d'opérateurs linéaires pour résoudre le problème différentielle. Vous voyez également les différentes généralisations possibles. Si par exemple, nous avons une équation de seconde ordre, nous n'obtiendrons alors que $N-1$ équations pour $N+1$ inconnus et nous devons la supplémenter par *deux* équations aux bords, et ainsi de suite pour des équations d'ordre plus élevés. De la même manière, l'approche se généralise à l'étude des équations aux dérivées partielles. Les techniques de résolution numérique d'équations différentielles ne font que reprendre ces schémas.

21.2 Équation de Laplace.

Considérons un ensemble de N boules de masse m reliées par des ressorts de raideur k. Chaque boule est assujettie à se mouvoir sur une ligne verticale, et les lignes sont espacées de d (figure 21.1.a). L'énergie potentielle totale du système est par conséquent :

$$U(...,y_{n-1},y_n,y_{n+1},...)=\sum_n \frac{1}{2}k(y_{n+1}-y_n)^2 \tag{21.2}$$

où y_n est l'ordonnée de la n-ième boule (et $x_n = nd$ son abscisse). Pour quelles valeurs des y_k le potentiel est minimum ? Comme U est une fonction de N variables, pour être extremum, il faut que sa dérivée par rapport à chaque variable soit nulle. Considérons la n-ième boule. Dans l'expression de l'énergie sous la somme, il y a seulement deux termes qui contienne la coordonnée de y_n qui sont $(y_n - y_{n-1})^2$ et $(y_{n+1} - y_n)^2$. La minimisation de U par rapport à y_n donne donc :

$$\frac{\partial U}{\partial y_n} = k(2y_n - y_{n-1} - y_{n+1}) = 0 \tag{21.3}$$

Cette dernière relation indique simplement que la force exercée sur la n-ième boule doit être nulle : en effet, la force n'est que le gradient (à un signe près) du potentiel. L'extremum du potentiel correspond à une position d'équilibre où les forces exercées s'annulent. Faisons maintenant $d \to 0$ et $N \to +\infty$ pour retrouver le continuum. La variable $x = nd$ devient continue, de même que la fonction $y(x)$. Comme $y_n = y(x_n) = y(nd)$, nous avons, par un simple développement de Taylor,

$$\begin{aligned} y_{n+1} &= y_n + \left.\frac{dy}{dx}\right|_{x=x_n} d + \frac{1}{2}\left.\frac{d^2y}{dx^2}\right|_{x=x_n} d^2 \\ y_{n-1} &= y_n - \left.\frac{dy}{dx}\right|_{x=x_n} d + \frac{1}{2}\left.\frac{d^2y}{dx^2}\right|_{x=x_n} d^2 \end{aligned}$$

L'équation (21.3) se transforme donc en une équation différentielle $d^2y/dx^2 = 0$. A plusieurs dimensions, en appliquant la même démarche, on aboutit à l'équation

$$\Delta y = 0 \tag{21.4}$$

où l'opérateur Δ désigne le laplacien. L'équation (21.4) est appelée justement l'équation de Laplace [1] et comme nous le voyons, est le résultat de la minimisation d'une certaine énergie. C'est exactement cette approche qu'Euler a utilisé pour développer le calcul variationnel et qui a donné lieu aux équations d'Euler-Lagrange.

Que vaut la constante de raideur k ? Elle doit probablement dépendre de notre découpage discret, *i.e.* de l'espacement d entre les éléments discrets que nous avons utilisé pour modéliser le continuum. Mais comment ? La règle fondamentale est que les valeurs que l'on peut mesurer (physiquement) ne doivent pas dépendre de notre découpage. Prenons maintenant une ligne de longueur L que l'on découpe en N morceau

1. Grand mathématicien français de la fin dix-huitième et début dix-neuvième siècle. Très célèbre pour son livre de mécanique céleste, les fondements de la théorie des probabilités (qui l'ont amené à inventer les transformées de Laplace), la théorie moléculaire de la capilarité (quand les molécules n'existaient pas !), ... Ses collègues et contemporins sont Lagrange, Fourier, Poisson et Cauchy. Que du beau monde.

espacés de $d = L/N$. Nous maintenons un coté (disons $x = 0$) à $y = 0$, et l'autre coté ($x = L$) à $y = Y$ (fig. 21.1.b). Comme $(y_{n+1} - y_n) = (Y/L)d$, Selon l'expression (21.2), l'énergie potentielle totale est donnée par

$$U = \sum_n k(Y/L)^2 d^2 = (Y/L)^2 Nkd^2 = (Y/L)^2 Lkd = (Y^2/L)kd$$

Si maintenant nous avions fait un autre découpage en prenant N' boules reliées par des ressorts de constante k' et un espacement d'(fig. 21.1.c), nous aurions trouvé pour l'énergie $U = (Y^2/L)k'd'$. Comme U ne doit pas dépendre de notre découpage, kd doit être constante :

$$k = \frac{K}{d} \tag{21.5}$$

La constante de ressort *microscopique* (résultat de notre découpage discret) k est relié à une constante physique du système K, (qui dénote l'amplitude de la rigidité du système par rapport à un phénomène physique), par la relation (21.5). Comme $(y_{n+1} - y_n) = y'(x).d + \mathcal{O}(d^2)$, l'expression de l'énergie potentielle devient

$$\begin{aligned} U &= (1/2)\sum \frac{K}{d} y'(nd)^2 d^2 \\ &= (1/2)\int K y'(x)^2 dx \qquad \text{quand } d \to 0 \end{aligned}$$

Pour un champ électrique par exemple, l'énergie est donnée par $\int(\epsilon/2)|\mathbf{E}|^2 d\tau = \int(\epsilon/2)|\nabla V|^2 d$ V est le potentiel électrostatique et l'opérateur gradient (∇) généralise la dérivée à plusieurs dimensions. Ici, le rôle de la constante de rigidité du système (vis à vis du champs électrique) est joué par la constante de perméabilité électrique ϵ. En élasticité, la variable du champ est appelé *déplacement*, et la rigidité du système est donnée par le module d'Young [2]. Dans le cas des gaz, nous somme en présence d'un champ de *pression* et K est l'inverse du coefficient de compressibilité.

21.3 Équation d'onde et de chaleur.

Nous nous sommes préoccupé dans la précédente section de phénomènes statiques. Essayons maintenant de formuler la dynamique. Revenons à notre exemple du figure (21.1.a) et supposons que chaque boule a une masse m. Nous pouvons maintenant écrire la relation fondamentale de la dynamique $F = ma$ pour chaque boule. L'accélération de la n-ième boule est donnée par d^2y_n/dt^2. La force sur la n-ième boule étant égale au

2. Maxwell, le fondateur de la théorie électromagnétique dans les années 1860, considérait les phénomènes électromagnétiques comme des déformations élastiques d'une substance hypothétique appelée éther et s'est beaucoup inspiré des travaux sur l'élasticité pour formuler sa théorie.

gradient du potentiel, *i.e.* $F_n = -\partial U/\partial y_n$, nous avons, en suivant ce que nous avons dit plus haut,

$$m\frac{d^2y_n}{dt^2} = -k(2y_n - y_{n+1} - y_{n-1}) \tag{21.6}$$

Comment m dépend de notre découpage ? La réponse est plus simple cette fois. Si nous désignons par ρ la densité (linéaire à une dimension), nous avons $m = \rho d$. Nous avons également, comme indiqué plus haut, $k = K/d$ et $(2y_n - y_{n+1} - y_{n-1}) \approx -(\partial^2 y/\partial x^2)d^2$. Quand $d \to 0$, l'équation (21.6) devient

$$\begin{aligned}\frac{\partial^2 y}{\partial t^2} &= \frac{(K/d)}{\rho d}\frac{\partial^2 y}{\partial x^2}d^2 \\ &= \frac{K}{\rho}\frac{\partial^2 y}{\partial x^2}\end{aligned} \tag{21.7}$$

C'est ce qu'on appelle l'équation d'onde. Elle se généralise de la même manière à plusieurs dimensions (l'opérateur Δ généralise la dérivée seconde). La constante K/ρ possède la dimension d'une vitesse au carré (pourquoi ?) et désigne, comme nous l'avons vu, la vitesse de propagation des ondes.

Question : qu'est ce qui joue le rôle de la densité pour les phénomènes électriques ?

Nous pouvons maintenant aborder plusieurs généralisations. Si les boules "baignent" en plus dans un "liquide" , il faut tenir compte de la force de dissipation visqueuse qui est proportionnelle (et opposée) à la vitesse, et donc à $-dy_n/dt$. L'équation d'onde devient alors, lors de la passage à la limite $d \to 0$,

$$\rho\frac{\partial^2 y}{\partial t^2} + \lambda\frac{\partial y}{\partial t} = K\frac{\partial^2 y}{\partial x^2}$$

En électromagnétisme, λ dénote le coefficient d'absorption d'un matériau (l'inverse de sa transparence). Nous voyons que si la masse des boules (la force inertielle) peut-être négligée par rapport aux autres forces de frottement et appliquée par les voisins (penser aux boules baignant dans du miel), nous pouvons négliger la dérivée d'ordre 2 par rapport au temps et écrire

$$\frac{\partial y}{\partial t} = D\frac{\partial^2 y}{\partial x^2}$$

qui n'est rien d'autre que l'équation de la chaleur. Il est peut-être difficile pour le lecteur de penser au champ de température comme des boules qui se meuvent dans du miel [3]. Nous le référons à la théorie de la réponse linéaire en physique statistique pour une dérivation de l'équation de la chaleur qui ait une plus grande réalité physique.

Revenons encore une fois à notre image de boules de la figure (21.1.a) . Et imaginons qu'en plus d'être reliées par un ressort de raideur k les uns aux autres, elles sont en

3. Même si la conception de la chaleur comme un fluide de "calorique" était populaire jusqu'au début du XIXème siècle.

plus reliées à l'axe x par un ressort de constante Vd (nous normalisons tout de suite la raideur par l'espacement, en laissant le soin au lecteur de démontrer que cela effectivement est la bonne forme). Nous n'avons aucune obligation à penser que V doit être une constante. A certain endroit le long de l'axe x, elle peut être forte, à d'autres endroit, faible. Nous notons donc V_n la constante du ressort qui relie la n-ième boule à l'axe x. L'expression de l'énergie potentielle totale est donc

$$U = \sum_n \frac{1}{2}\frac{K}{d}(y_{n+1} - y_n)^2 + d.V_n y_n^2$$

Il ne sera alors pas difficile pour le lecteur de démontrer que l'équation d'onde s'écrit

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2} - V(x).y$$

et l'expression de l'énergie (potentielle) est de la forme

$$\int (K/2)|\nabla y|^2 + (1/2)V(x)y^2 dx$$

forme couramment utilisée dans la théorie du ferromagnétisme.

le cas de l'équation de Schrödinger comme deux équations couplées, et le rapprochement avec particules dans champs magnétique ou les équations de second degrés ;

Traitement du processus dissipatif en développant le BABA de la réponse linéaire en phys stat. le cas de l'équation de la chaleur, indication (par potentiel chimique) pourquoi ça se généralise à la diffusion de concentration etc.

22 Qu'est ce qu'un nombre ?

Nous avons vu tout au long de ce cours divers outils de mathématiques très utilisés en physique. Ces outils concernaient la manipulation des fonctions dans le but très alimentaire de résoudre des équations issue de la physique. Les fonctions elles mêmes étaient définies comme des boîtes noires transformant un *nombre* en un autre. Nous nous sommes jamais demandé ce qu'est un nombre, nous avons pris cela comme une donnée dont la signification est à priori connue.

Nous allons dans ce chapitre revenir un peu sur ce concept et voir la construction des nombres réels. Nous verrons également que ce n'est pas la seule façon de construire un ensemble complet de nombre, et d'autres ensembles qui défient notre intuition de "proche" et de "loin" sont également constructible. Ce chapitre n'a pas d'autre but que d'éveiller la curiosité du lecteur.

Le plan général que l'on va suivre est de d'abord construire les nombres entiers, ensuite les nombres rationnels. Nous munirons alors notre ensemble d'une *topologie* et construirons soit l'ensemble des nombres réels, soit celui des nombres p-adiques. "Munir un ensemble d'une topologie" est un terme pour effrayer l'étudiant. En langage profane, cela veut simplement dire que l'on va définir *les distances*, la notion *d'être proche*. La topologie habituelle que l'on définie, et à laquelle nous sommes habitués depuis notre tendre enfance nous dit par exemple que 4.3 et plus *proche* de 4.2 que 5. Tant que nous construisons l'ensemble des nombres rationnels, nous n'avons pas besoin de ce concept, celui d'*avant* et *après* nous suffira.

22.1 Les entiers naturels $\mathbb{N}$.

Dedekind, grand mathématicien de la fin du dix-neuvième siècle disait : "Dieu inventa les nombres entiers ; tous le reste est invention humaine". La construction moderne des nombres entiers est due à Peano et ses cinq principes. De façon intuitif, on peut dire que c'est le plus simple ensemble ou chaque élément (excepté le premier qu'on appelle 0) possède *un* élément *juste avant* et *un* élément *juste après*, et cela de façon non-cyclique. Bon, bien sûr, comme nous sommes en train de faire des mathématiques, nous devons définir exactement ce que ces termes veulent dire. Voilà les axiomes de Peano.

1. $0 \in \mathbb{N}$ [1]

1. Grand débat philosophique pour savoir si il faut commencer par 0 ou par 1. Cette question n'a pas de sens tant que l'on a pas défini l'opération addition et son élément neutre. Tout ce que l'on veut ici est de définir un premier élément.

2. Chaque nombre naturel x possède un autre élément $x' = s(x) \in \mathbb{N}$ appelé son successeur (voilà pour le *juste après*).
3. 0 n'est le successeur d'aucun nombre (cela nous enlève le danger des cycles).
4. Si $s(x) = s(y)$ alors $x = y$ (cela nous enlève le problème de plusieurs nombre ayant le même successeur).
5. Axiome d'induction. Soit Q une propriété telle que

— Q est valable pour 0
— Si Q est valable pour x, alors Q est valable pour $s(x)$
— Alors Q est valable pour tous les nombres entiers (cela entre autre nous enlève le problème d'avoir plusieurs "premier élément")

Nous avons insisté sur cette construction pour souligner que tout ce dont on a besoin à cette étape est le concept d'avant et d'après.

Comme nous ne voulons pas écrire un texte très rigoureux, nous allons aller un peu plus vite à partir de là. On peut commencer par donner des noms aux divers éléments. Par exemple, le successeur de 0 sera appelé un (et noté 1), le successeur de 1 deux (2) et ainsi de suite. On peut donc noter $\mathbb{N} = \{0, 1, 2, ...\}$. Ensuite, nous allons munir notre ensemble de l'opération $+$. C'est une application qui à deux nombres entiers associe un troisième, et cela en généralisant le concept de successeur : $x + 0 = x$ et $x + s(y) = s(x + y)$. Par exemple, et $x + 1 = s(x)$ [2]. L'opération $+$ a bien sûr toutes les bonnes propriétés d'associativité, commutativité, etc. dont nous sommes habitué (exercices : les démontrer). Nous laissons au lecteur le soin d'en donner une définition rigoureuse.

Tant que nous y sommes, nous pouvons également définir l'opération $\times$ (multiplication) comme une autre application qui a chaque deux nombres entiers associe un troisième : $x \times 0 = 0$; $x \times s(y) = x \times y + x$. Bien sûr, c'est la multiplication habituelle et on aurait été plus claire si on avait noté $x \times (y + 1) = x \times y + x$. Par exemple, $x \times 1 = x$.

Enfin, ça ne mange pas de pain de définir rigoureusement les relations de comparaison $<$ et $>$, à nouveau en suivant la piste des successeurs.

Nous suggérons au lecteur de faire ces constructions en détails et de façon rigoureuse, c'est un exercice très intéressant.

Nous disposons donc maintenant d'un ensemble $\mathbb{N}$, muni des deux opérations $+$ et $\times$. En langage chique, nous dirrions que $(\mathbb{N}, +, \times)$ est un anneau commutatif.

22.2 Les ensembles $\mathbb{Z}$ et $\mathbb{Q}$.

A partir de là, nous allons commencer à élargir notre ensemble N.

D'abord, nous pouvons définir l'opération soustraction $-$ comme "l'inverse" de l'addition : si $x - y = z$ alors c'est que $y + z = x$ (exercice : donner une définition rigoureuse de la soustraction). Mais cela nous pose un problème. $\mathbb{N}$ est fermé pour l'addition, c'est

2. Notez comment l'operation $+1$ devient alors le synonyme de l'opération "successeur".

à dire que l'addition de n'importe quel deux nombres est encore dans $\mathbb{N}$. Cela est loin d'être le cas pour la soustraction. Il suffit d'examiner $0-1$: si $0-1=z\in\mathbb{N}$, alors $z+1=0$, ce qui contredit violemment un des axiomes de Peano. Qu'à cela ne tienne : nous allons définir un ensemble $\mathbb{Z}$ qui contient $\mathbb{N}$ et qui est fermé pour la soustraction, $\mathbb{Z}=\{...,-2,-1,0,1,2,...\}$.

Nous pouvons également définir l'opération de division $/$comme "l'inverse" de la multiplication : si $x/y=z$ alors c'est que $x=yz$. Le même problème se pose : en général, pour un couple quelconque, x/y n'a pas de sens dans $\mathbb{N}$ ou $\mathbb{Z}$. A nouveau, on peut "agrandir" notre ensemble et définir l'ensemble des nombres rationnels $\mathbb{Q}$ qui est fermé pour l'opération de division. Notez que nous avons pas vraiment définit comment on agrandit nos ensembles, cela alourdirait trop ce texte [3].

L'ensemble $\mathbb{Q}$ est très riche. Concrètement les humains n'en sortent jamais pour faire leurs calculs. Le trait principal de cet ensemble est qu'entre n'importe quel deux nombres rationnels, on peut en trouver d'autres. Ceci dit, comme le lecteur le sait, l'ensemble $\mathbb{Q}$ reste dénombrable, et même s'il est fermé pour la division, il n'est pas *algébriquement* fermé. Par cela nous voulons dire que les racines de tous les polynômes (de coefficients entiers) ne se trouvent pas dans $\mathbb{Q}$. Par exemple, il est trivial de montrer que la racine de $x^2-2=0$ (qui représente l'hypoténuse d'un triangle rectangle de coté unité) n'est pas rationnelle.

Il suffit de suivre la même démarche et construire l'ensemble des nombre algébrique $\mathbb{A}$, la fermeture algébrique de $\mathbb{Q}$, et qui contient tous ces nombres du genre $\sqrt{2+\sqrt{3}+\sqrt[17]{253}}$. A t'on épuisé tous les nombres ou existe t'il des nombres non-algébriques qu'on appelle *transcendants* ? Est ce que par exemple, le périmètre d'un cercle de diamètre unité π, ou le nombre e sont algébriques ? La réponse à ces questions n'est venu qu'à la fin du dix-neuvième siècle.

22.3 Un peu de topologie.

Nous n'avons pas encore introduit le concept de distance entre deux nombres. La distance entre deux nombres est une application qui prend deux nombres en entrée et produit un nombre en sortie. On peut la définir sur n'importe quel corps k [4] [5] (dont par exemple le corps des rationnels). Nous demandons à cette application d'avoir un

3. Voyons rapidement la construction des rationnels. Considérons l'ensemble $A=\mathbb{N}\times\mathbb{N}$, c'est à dire l'ensemble de toutes les pairs (x,y) où x et y sont des entiers naturels. Nous définissons une relation d'équivalence $(x,y)=(x',y')$ si $xy'=x'y$. Nous définissons l'opération $+$dans A par $(a,b)+(c,d)=(ad+bc,bd)$ et l'opération $\times$ par $(a,b)\times(c,d)=(ac,bd)$. L'ensemble A partitionné par la relation d'équivalence ci-dessus et muni des deux opérations $+$ et $\times$ peut être identifié au corps des rationnels $\mathbb{Q}$. Un exercice intéressant serait de suivre les mêmes lignes pour construire les entiers relatifs à partir des entiers naturels.

4. Rappelons qu'un corps est un ensemble, muni des deux opérations $+$ et $\times$ et fermé vis à vis d'elles.

5. Bien sûr, pour définir une norme, nous n'avons pas nécessairement besoin d'un corps. Nous avons vu dès le début de ce livre comment en définir une pour l'espace vectoriel des fonctions de carré sommable.

minimum de propriétés : Pour tous $a, b, c \in k$,

1. $d(a,b) \geq 0$ et $d(a,b) = 0$ si et seulement si $a = b$.
2. $d(a,b) = d(b,a)$.
3. $d(a,b) \leq d(a,c) + d(b,c)$ (l'inégalité du triangle).

Ce n'est pas beaucoup demander, mais à partir du moment où nous disposons d'une métrique, nous pouvons faire une quantité phénoménale de choses. Essentiellement, c'est l'étape où l'on passe de l'algèbre à l'analyse, où on peut commencer à définir le concept de *continu*, de la convergence des suites, ...

Un concept étroitement lié à la notion de distance est celle de *la valeur absolue*. Supposons que nous disposons d'une valeur absolue sur un corps k avec les propriétés suivantes :

p1 : $|a| = 0$ ssi $a = 0$.

p2 : $|ab| = |a||b|$

p3 : $|a + b| \leq |a| + |b|$

alors nous pouvons facilement définir la distance entre deux éléments par $d(a,b) = |a - b|$. Nous laissons au lecteur le soin de démontrer cela. L'exemple usuel de la valeur absolue sur $\mathbb{Q}$ est $|x| = x$ si $x \geq 0$ et $-x$ sinon. Bien sûr, ce n'est pas la seule valeur absolue possible, nous en verrons des exemples plus bas.

Comme nous l'avons dit, dès que nous disposons d'une distance, nous pouvons définir la convergence des suites. Nous disons que la suite $a_n$converge vers la limite a si tous les éléments de la suite, à partir d'un certain N sont aussi *proche* de la limite que nous le souhaitons. Dans l'ensemble $\mathbb{Q}$, nous écrirons par exemple que a est la limite de a_n si pour tout $\epsilon \in \mathbb{Q}$, nous pouvons trouver N tel que si $n > N$ alors $d(a, a_n) < \epsilon$.

Un des problèmes de cette définition de la convergence est que pour savoir si une suite converge, nous devons connaître à l'avance sa limite ! Le grand Cauchy a trouvé comment y remédier : une suite converge si la distance entre deux éléments quelconques converge vers zéro au delà d'un certain N : si pour tout $\epsilon \in \mathbb{Q}$, nous pouvons trouver N tel que si $n, m > N$ alors $d(a_m, a_n) < \epsilon$ alors la suite est convergente.

Cela nous pose un nouveau problème : la limite d'une suite dans un corps k n'a aucune raison d'appartenir au même corps. Mais nous pouvons continuer notre procédure d'enrichissement et considérer un ensemble qui contient *et* le corps k *et* toutes les limites de toutes les suites convergentes. Nous verrons ci-dessous deux exemples de fermeture topologique de l'ensemble $\mathbb{Q}$: l'ensemble des nombres réels et l'ensemble des nombres $p-$adiques.

22.4 L'ensemble des nombres réels.

Munissons nous de la valeur absolue usuelle, et la distance (la métrique) qui en découle. Et considérons les suites convergentes dans $\mathbb{Q}$. Il est évident que beaucoup (vraiment beaucoup) de ces suites n'ont pas leurs limites dans $\mathbb{Q}$.

Exemple 22.1 Le nombre $1/e$, défini comme la limite de $\sum_{n=0}(-1)^n/n!$ n'est pas un nombre rationnel. Pour voir cela supposons qu'il l'est et ecrivons le comme p/q. Nous décomposons la série en une somme jusqu'au terme q et le reste :

$$\frac{p}{q} = \sum_{n=0}^{q}(-1)^n/n! + R_q$$

Comme nous avons affaire à une série alternative convergente, le reste est plus petit que le dernier terme : $|R_q| < 1/q!$. Multiplions maintenant les deux cotés par $q!$. Nous avons à gauche un entier, et à droite un entier *plus* un terme plus petit que l'unité. Le coté droit n'est donc pas un entier naturel. Notre hypothèse de rationnalité de $1/e$ est donc fausse.

Nous définissons l'ensemble des nombres réel $\mathbb{R}$ comme un ensemble qui contient l'ensemble $\mathbb{Q}$ et les limites de toutes les suites convergentes dans $\mathbb{Q}$ au sens de Cauchy. Les opérations $+$ et $\times$ se généralisent aisement par continuité. Par exemple,pour $a, b \in \mathbb{R}$, (mais pas nécessairement rationnel) $a+b = \lim(a_n+b_n)$ où a_n et b_n sont des suites dans $\mathbb{Q}$ convergeant vers a et b.

Nous pouvons pousser un ouf de soulagement, nous sommes au bout de notre chemin (à part peut-être une extension triviale à $\mathbb{C}$). Mais est ce que c'était vraiment la peine de faire tout ce parcours ? Est ce que l'ensemble $\mathbb{R}$ est vraiment plus riche que l'ensemble des nombres algébriques ? La réponse est évidement oui, mais elle est loin d'être évidente. Jusqu'à presque la fin du dix-neuvième siècle, la réponse à cette question n'était pas connue. On a pu démontrer à cette époque avec peine que les nombre e et π ne sont pas algébriques, c'est à dire que nous ne pouvons pas trouver un polynôme de coefficients entiers dont une des racines soit un de ces nombres. Mais combien y avait il de ces nombres transcendants ? très peu, beaucoup ? La réponse, un coup de maître, est venu de Greg Cantor : les nombres algébriques forment une minorité négligeable comparée aux nombres réels. Cette démonstration a provoqué beaucoup de débats furieux à l'époque, puisque Cantor ne construisait pas *un seul* nombre transcendant. Sa démonstration se fait en deux étapes très simples : (i) les nombres algébriques sont dénombrables ; (ii) les nombres réels ne sont pas dénombrable. Voyons cela de plus près.

22.4.1 Les nombres algébriques sont dénombrables.

Comme nous l'avons dit, les nombres algébriques comprennent les racines de tous les polynômes. Les nombres rationnels sont évidement des nombres algébriques, puisque p/q est solution de l'équation $px - q = 0$.

Considérons maintenant un polynôme à coefficient entier du genre $a_0 + a_1x + ... + a_nx^n (a_n \neq 0)$. Nous appellerons *hauteur* de ce polynôme le nombre $H = n-1+|a_0|+|a_1| + ...|a_n|$. Il existe un seul polynôme de hauteur 1 : x. Pour $H = 2$, nous avons les

polynômes suivants : x^2 ; $x \pm 1$. Pour $H = 3$, x^3 ; $\pm 2x^2$; $x^2 \pm x$; $x^2 \pm 1$; $2x \pm 1$; $x \pm 2$ et ainsi de suite. Le fait intéressant est que le nombre de racines de tous les polynômes d'une hauteur H est finie (Combien y en a t il au plus ?). Nous pouvons donc ranger les nombres algébriques de façon suivante : On prend d'abord toutes les racines associées à la hauteur 1, et on les range dans l'ordre croissant, en éliminant les doublons. On prend ensuite toutes les racines associées à $H = 2$, on les ranges dans l'ordre croissant en éliminant les doublons et on continu le procédé pour $H = 3$, $H = 4$,... Cela nous donne par $\mathbb{A} = \{0\,;\, -1, 1\,;\, -2, -1/2, 1/2, 2\,;\, ...\}$ et il n'est pas difficile de voir que nous avons ainsi une procédure pour dénombrer les nombres algébriques !

22.4.2 Les nombres réels ne sont pas dénombrables.

Supposons que nous avons réussi à dénombrer tous les nombres réel entre 0 et 1. Nous les listons dans l'ordre croissant en utilisant leur représentation décimale :

$$\begin{aligned} r_0 &= 0.a_{00}a_{01}a_{02}... \\ r_1 &= 0.a_{10}a_{11}a_{12}... \\ r_2 &= 0.a_{20}a_{21}a_{22}... \\ &... \end{aligned}$$

Soit maintenant le chiffre r construit à partir des décimaux diagonaux : $r = 0.a_{00}a_{11}a_{22}...$ et construisons un nombre r' à partir de r en changeant chacun des décimaux de r d'une façon quelconque. Il est alors facile de voir que r' ne peut pas être dans la liste ci-dessus ! (Exercice : le démontrer).

22.4.3 Au delà des nombres réels : les hyper-réels.

Si on voulait donner une image de nos nombres, les rationnels seraient des points isolés dans un espace et les réels rempliraient le vide qu'il y a entre. Peut on encore inventer des nombres qui se mettraient *entre* les nombres réels ? Avant le dix-neuvième siècle, les mathématiciens avaient l'habitude de manipuler ce genre de nombres qu'ils appelaient des *infiniment petits*. Ces nombres cependant provoquaient pas mal de contradictions et ont été vite chassés du monde. Dans les années 1960, Abraham Robinson à réussi de les réintroduire de façon rigoureuse par une méthode pas trop loin de ce que nous avons vu pour la construction des nombres réels. Un infiniment petit est par exemple un nombre ϵ tel que $0 < \epsilon < 1/n$ quelque soit $n \in \mathbb{N}$. Dans l'ensemble des nombres hyper-réel, chaque réel classique est entouré d'un nuage de nombre à distance infiniment petit.

Concrètement, l'introduction de ces nombres n'apporte pas de nouvelles méthodes et nous ne développerons pas ce concept plus ici. Nous suggérons au lecteur intéressé de se diriger vers des livres plus spécialisés sur ces nombres et l'analyse non-standard.

22.5 Les nombres $p-$adiques.

Nous allons voir dans cette section des nombres étranges, très différents de ce que nous connaissions jusque là. La notion de *proche* et de *loin* est complètement dissociée de la notion d'avant et après, contrairement à la distance usuelle que nous avons utilisée pour construire $\mathbb{R}$ à partir de $\mathbb{Q}$. Il existe d'autres valeurs absolues, et la topologie qu'elles définissent est radicalement différente. Rappelons que la valeur absolue doit avoir les trois propriétés mentionnées à la section 22.3. Si la valeur absolue a en plus la propriété suivante :

p4 : $\qquad |x+y| \leq \max\{|x|, |y|\}$

nous l'appelons *non-archimédienne*. Notons que la propriété 4 implique la propriété 3, puisque $\max\{|x|, |y|\} \leq |x| + |y|$.

Commençons par les nombres entiers. Donnons nous un nombre premier p. N'importe quel entier n peut s'écrire de façon unique sous la forme

$$n = p^{v_p(n)} n'$$

où $p \nmid n'$ (p ne divise pas n'). Par exemple, si nous avons choisi le nombre premier 5, nous avons

$$\begin{aligned} 2 &= 5^0 2 \\ 5 &= 5^1 1 \\ 6 &= 5^0 6 \\ 150 &= 5^2 6 \end{aligned}$$

et nous avons donc $v_5(2) = v_5(6) = 0$; $v_5(5) = 1$; $v_5(150) = 2$. $v_p(n)$ est appelé la valuation $p-$adique du nombre n, et désigne la multiplicité du facteur premier p pour former le nombre n. Par convention, $v_p(0) = \infty$: on peut diviser 0 par p; le résultat étant 0, on peut encore multiplier 0 par p et cela peut continuer infiniment.

On peut étendre de façon évidente la valuation $p-$adique aux nombres rationnels : $v_p(a/b) = v_p(a) - v_p(b)$

Et ce n'est pas difficile de voir que

1. $v_p(xy) = v_p(x) + v_p(y)$
2. $v_p(x+y) \geq \min\{v_p(x), v_p(y)\}$

Si l'on compare les propriétés de v_p à une valeur absolue, nous voyons que v_p agit un peu comme un logarithme. Nous pouvons donc définir la valeur absolue $p-$adique d'un nombre x par

$$|x|_p = p^{-v_p(x)}$$

et pour revenir à l'exemple des nombres précédents, $|2|_5 = |6|_5 = 1$; $|5|_5 = |10|_5 = 1/5$; $|150|_5 = 1/25$. En utilisant notre convention, nous avons en plus $|0|_5 = 0$.

Nous devons remarquer plusieurs chose à ce niveau : (i) la valeur absolue $p-$adique d'un nombre est inférieure ou égale à 1 ; (ii) plus un nombre est divisible par p, plus sa valeur absolue est proche de 0. Nous laissons au lecteur le soin de démontrer que cette valeur absolue en est vraiment une, et qu'en plus, elle est non archimédienne. Nous pouvons en plus démontrer que si $x \neq y$, alors $|x+y|_p = \max\{|x|_p, |y_p|\}$.

Super, nous disposons d'une valeur absolue sur $\mathbb{Q}$, nous pouvons donc définir une distance : $d(x,y) = |x-y|_p$. Par exemple, pour la distance 5-adique, $d(5,6) = 1$; $d(5,10) = 1/5$; $d(5,30) = 1/125$. Notons que cette métrique a la propriété suivante : $d(x,y) \leq \max\{d(x,z), d(y,z)\} \forall x,y,z \in \mathbb{Q}$. Cette inégalité est appelé l'inégalité ultramétrique.

Notons combien cette distance est différente de la distance habituelle. Prenons par exemple trois points quelconques mais distinct x, y, z. Alors deux des distances sont égales ! Ceci découle du fait que $(x-y)+(y-z) = (x-z)$. Si $|x-y| \neq |y-z|$, alors $|x-z|$ est égale au plus grand d'entre eux.

Comme nous disposons d'une distance, nous pouvons définir les suites et leur convergences, et fermer $\mathbb{Q}$ pour obtenir l'ensemble $\mathbb{Q}_p$. Nous pouvons développer l'analyse exactement comme nous avons fait avec les nombres réels, définir les fonctions, leurs dérivées et intégrales, ... Nous ne développons pas plus cela ici, notons simplement quelques faits inhabituels de ces ensembles :

— Pour qu'une suite a_n converge, il suffit que $|a_{n+1} - a_n| \to 0$ (c'est beaucoup plus simple que le critère de Cauchy).
— Pour que la série $\sum a_n$ converge, il suffit que $|a_n|_p \to 0$
— Si un point appartient à une boule (ouverte ou fermée), il en est le centre,
— ...

23 Bibliograhie.

Ce cours est un résumé rapide de ce que l'étudiant en physique devrait savoir. Ce cours est une introduction qui devrait permettre à l'étudiant d'attaquer les divers sujets en consultant des livres plus avancés. Ci-dessous, je liste pêle-mêle quelques livres que j'ai eu entre les mains et que j'ai trouvé particulièrement intéressant pour des étudiants de niveau L3-M2.

Jean Bass, Cours de Mathématiques. C'est un cours extrêmement complet, en plusieurs volumes, allant des mathématiques élémentaires aux sujets les plus avancées.

Nino Boccara, Analyse fonctionnelle. Le livre le plus élégant que l'auteur connaisse sur le sujet.

François Rodier, Distributions et Transformation de Fourier. L'auteur y développe les résultats essentiels de l'analyse fonctionnelle et de la théorie de la mesure avant d'exposer les distributions et les TF, avec un regard de physicien tourné vers l'expérience.

F.W. Byron & R.W. Fuller, Mathematics of classical and quantum physics. Le livre que chaque étudiant de physique devrait avoir lu.

Ondelettes et Analyse de Fourier. La première partie de ce livre expose le développement de la théorie de la mesure et des séries de Fourier tout au long du dix neuvième et du vingtième siècle.

J.G.Simmonds & J.E. Mann, A first look at Perturbation Theory. Un aperçu des divers aspects du calcul de perturbations. Le lecteur plus assoiffé se rapportera au livre de Nayfeh, "Perturbation Theory".

C. Lanczos, Linear Differential operators. Lanczos, à part avoir été un très grand scientifique, a écrit des livres d'une rare profondeur. Son livre de mécanique analytique est un pur bijou. Le livre mentionné ici traite avec une très grande rigueur et élégance des opérateurs linéaires.

H. M. Edwards, Advanced Calculus : A Differential Forms Approach. Un très beau livre sur les formes différentielles, écrit il y a une quarantaine d'année et n'ayant rien perdu de sa beauté. La plupart des livres sur les formes différentielles sont réservés aux étudiants avancés de mathématiques, où l'exposé est noyé sous des tonnes de "théorème-démonstration". Edwards fait jaillir toute la beauté de ces objets mathématiques. Pour une lecture plus avancé mais toujours aussi élégant, le lecteur pourra se rapporter au livre d'Arnold : mathématiques de la mécanique classique.

S.L. Sobolev, Partial differential equations of mathematical physics. Un grand tour des EDP de la physique, écrit par un des grands chercheurs du domaine qui sait également être très pédagogue.

David M. Bressoud : A Radical Approach to Lebesgue's Theory of Integration. Nous n'avons pas abordé l'intégrale de Lebesgue dans ce cours. Il est rare pour un physicien d'avoir besoin d'autres choses que des intégrales de Riemann (à l'exception de la théorie des équations différentielles stochastiques). Cependant, pour l'étudiant ayant des penchants pour les belles mathématiques, ce livre est un must.

Doran and Lasenby : geometric algebra for physicists. Les tenseurs (en majorité) et les formes différentielles (en minorité) sont les deux langues principales de la physique. Il existe une troisième langue, développée depuis le milieu du XIXème siècle (algèbre de Grassman) que nous appelons de nos jours l'algèbre géométrique. Les concepts de $n-$vecteurs et $n-$formes sont généralisés (on peut faire des mélanges) et la dérivation extérieure et intérieure sont unifiées. Si vous avez aimé les formes différentielles, vous allez adorer ce thème, exposé très clairement par les auteurs ci-dessus.

Autres. Beaucoup de thèmes indispensables ont été négligés dans ce manuscrit. La théorie des fonctions holomorphes et l'analyse complexe sont très bien traité chez Bass ou Byron et Fuller, nous suggérons également le livre d'Albowitz, "complex variables".

Index