

École Normale Supérieure
Département de Physique

FIP 1^{ère} année
(2010/2011)



Mathématiques pour physiciens

UPMC
PARISUNIVERSITAS

Claude ASLANGUL
Université Pierre et Marie Curie
(Paris 6)
aslangul@lptmc.jussieu.fr

*“C’est avec la logique que nous prouvons
et avec l’intuition que nous trouvons”*

(Henri POINCARÉ)

Préambule

Ce cours ne prétend en aucune façon être un cours de mathématiques, mais vise à familiariser les étudiants en Licence de physique avec des méthodes et des techniques mathématiques sans la maîtrise desquelles la physique d’aujourd’hui tout simplement n’existerait pas. Cette ambition limitée – qui pourra sembler modeste aux yeux de certains puristes – n’en fera pas pour autant un catalogue de recettes énoncées sans démonstration : on s’efforcera de trouver l’équilibre entre la rigueur et l’intuition, n’hésitant pas à donner des versions “faibles” de vraies démonstrations de théorèmes grâce à l’utilisation d’hypothèses contraignantes, adoptées dans le seul but de maintenir la démonstration à un niveau relativement élémentaire, en tout cas largement suffisant pour les préoccupations usuelles du physicien. Il arrivera même que l’on se borne à indiquer le fil d’une démonstration sans donner le détail de celle-ci, animé par la conviction que la compréhension des idées doit prendre parfois le pas sur un exposé détaillé ralentissant le rythme sans apporter vraiment d’élément nouveau.

L’émergence et la construction des théories physiques “modernes” (par exemple la Mécanique statistique, la Relativité, la Mécanique quantique) auraient été impossibles sans le recours à des concepts ou des outils mathématiques assez élaborés et très puissants. Cette nécessité s’était déjà imposée au XIX^e siècle pour l’Électromagnétisme (et sa cohorte d’équations aux dérivées partielles) et la Mécanique statistique dont Boltzmann peut être considéré comme le père-fondateur. Au XX^e siècle, l’analyse complexe a débarqué en force : jusque là cantonnés au rôle d’outils commodes (mais à tout prendre pas absolument nécessaires), les nombres complexes sont devenus l’expression la plus naturelle de la spécificité quantique – et de ses étrangetés.

L'objectif ici poursuivi inspire en grande partie le choix des questions abordées, et ne permet pas de prétendre à l'exhaustivité pour chacune d'entre elles. Le choix résulte inévitablement d'un compromis, tout en essayant de maintenir une structure cohérente destinée à préserver l'unité de l'exposé. En ce sens, on peut dire que le cours est articulé autour de la théorie des fonctions analytiques dont les premières pages, écrites par Cauchy, sont certainement l'une des plus belles histoires mathématiques par son élégance et sa simplicité formelle. Quant à l'exhaustivité, elle ne pourrait être qu'un leurre, tant en ce qui concerne le fond que la richesse des applications. Il est souhaité que, malgré les coupes claires inévitables, un équilibre a été trouvé entre l'efficacité et le pragmatisme d'une part, la possibilité de compréhension en profondeur d'autre part.

Il est certain que Mathématiques et Physique sont deux disciplines intellectuelles très proches, comme en témoignent non seulement leur histoire et leur développement au fil des siècles, mais aussi les échanges fructueux et les enrichissements qu'elles se sont mutuellement apportés. C'est un fait que les théories physiques s'écrivent presque spontanément en termes mathématiques – au point que certains considèrent les mathématiques comme le “langage de la nature”, pour reprendre l'expression de Galilée – et que les grands principes physiques y trouvent une expression naturelle et lumineuse. Par exemple, le principe de causalité impose des propriétés analytiques remarquables aux fonctions de réponse d'un système, comme on le verra.

Toutefois, la démarche du mathématicien et celle du physicien (toutes deux fondées sur l'intuition contrairement à ce qui est parfois prétendu) sont de natures très différentes, en simple conséquence des objectifs poursuivis par l'un et l'autre. Il arrive aussi que, en dépit des apparences, la formulation d'une “même” question ne traduise pas la même interrogation, faute d'employer des concepts parfaitement définis, ou des termes dénués de toute ambiguïté. La physique est une science expérimentale qui, partant des observations, énonce des lois ayant le statut de principes à partir desquelles toute théorie physique est construite ; une théorie physique est, à un instant donné, considérée comme exacte si elle permet de rendre compte de tous les phénomènes observés et sa fécondité se situe réellement dans son aptitude à prévoir de nouveaux phénomènes non encore observés. On peut dire que la prévision théorique d'un effet inconnu jusqu'alors, la suggestion d'une expérience permettant de l'observer ... et son observation de fait constituent le menu royal du physicien.

Le mathématicien ne saurait s'en remettre à l'expérience au sens où l'entend le physicien, et introduit ou manipule des concepts sans se soucier de leur mise à l'épreuve expérimentale, et pour cause puisque la plupart d'entre eux sont par nature inaccessibles à l'expérience¹. La notion d'infini (même le plus petit d'entre eux, le cardinal des entiers naturels), la notion de nombre irrationnel, ..., ne sauraient faire l'objet d'une quelconque quête de vérification expérimentale, et c'est ce qui a tant troublé les Anciens, les Grecs notamment : on aura beau remplacer la diagonale d'un carré de côté de longueur unité par une succession de petits segments formant un escalier autour de la diagonale, la

¹Par exemple, s'agissant de prouver la véracité d'une propriété fonction du cardinal $N \in \mathbb{N}$ d'un certain ensemble, on peut faire l'expérience avec un ordinateur, qui va examiner systématiquement les valeurs successives de N . Aussi puissant que soit l'ordinateur, il ne pourra jamais considérer qu'un nombre maximum N_{\max} . Ceci ne démontrera jamais que la propriété est vraie $\forall N \in \mathbb{N}$.

somme des longueurs des segments formant l'escalier sera toujours égale à 1, alors que la diagonale a pour longueur $\sqrt{2}$. En la circonstance, ce qui est impossible à réaliser expérimentalement c'est ce que le mathématicien appelle passer à la limite. On peut dire que la notion d'infini est hors de portée pratique du physicien, ce qui ne l'empêche évidemment pas faire le saut conceptuel et de la manipuler effectivement, quand il sait (pour de bonnes raisons) qu'il peut le faire.

Si la notion d'infini échappe au physicien dans ses tests expérimentaux, il en va de même de la notion de zéro² : le physicien considère comme nulle (et non avenue) toute "perturbation" dont les effets se situent en-deçà de ses capacités observationnelles, ou toute grandeur physique pour laquelle on a su trouver expérimentalement une borne supérieure très petite : si on déclare nulle la masse du photon ou la charge du neutron (par exemple), c'est d'une part parce que les théories construites avec ces hypothèses sont, jusqu'à présent, en accord avec l'expérience, d'autre part parce que l'on a pu trouver expérimentalement des bornes supérieures incroyablement petites³. De même, dans la construction d'un modèle physique⁴, on déclare (plus ou moins explicitement) que certains effets sont négligeables, ce qui revient à les annuler strictement à zéro.

L'une des démarcations les plus indiscutables entre l'univers du physicien et celui du mathématicien tient sans doute au fait que le premier a toujours à sa disposition des échelles pour les grandeurs physiques pertinentes du problème considéré : échelles de temps, de longueur, d'énergie, etc. C'est par rapport à ces échelles que se situent le zéro et l'infini du physicien : pour la physique atomique, la taille de notre galaxie est réellement "infinie", cependant que pour les phénomènes se situant dans le domaine d'énergie⁵ de l'eV, le noyau atomique peut être considéré comme ponctuel⁶ (de rayon nul). Au contraire, pour l'astrophysicien qui étudie l'univers à grande échelle, notre galaxie est un objet "microscopique", cependant que pour l'expert en gluons et quarks, un noyau est à lui seul un véritable univers. De la même façon, le temps de (quasi-)récurrence d'un système macroscopique est exponentiellement grand, mais en toute rigueur fini : c'est pourquoi le physicien énonce le Second Principe, légitimé par le fait qu'un intervalle de temps $\sim 10^{10^{23}}$ fois l'âge de l'univers est réellement pour lui inaccessible, tout comme l'infini au sens commun est et restera toujours hors de portée⁷. La notion d'échelle est reliée à celle de résolution : toute mesure est effectuée à l'aide d'un appareil de précision limitée fournissant des valeurs dont on peut seulement dire que chacune se trouve dans un certain intervalle fini autour de la valeur retenue ; l'intervalle est communément

²À au moins une exception toutefois, lorsqu'il s'agit d'une loi de conservation fondamentale (celle de la charge électrique, par exemple), et de la symétrie sous-jacente.

³Pour la masse du photon, la borne supérieure expérimentale actuellement admise est environ 10^{-52} kg. Quant au neutron, sa charge totale réputée nulle ne doit pas dissimuler le fait qu'il possède une structure de charge (l'atome aussi est de charge nulle !).

⁴Il n'est pas exagéré de dire que tout l'art du physicien est de construire les (bons) modèles pour rendre compte des observations.

⁵1 eV $\simeq 1,6 \times 10^{-19}$ J.

⁶L'ordre de grandeur du "rayon" des noyaux est le Fermi (1 F = 10^{-15} m).

⁷De l'hypothèse de Riemann ("Tous les zéros non-triviaux de la fonction de Riemann sont sur la droite..."), le mathématicien ne dira pas qu'elle est vraie puisqu'elle n'est pas encore démontrée ; en revanche, le physicien qui, dans les bonnes unités, travaille avec des nombres $\lesssim 25$ millions la considérera comme exacte. Même pour une science dite exacte, la vérité peut être toute relative.

appelé barre d'erreur. C'est par référence à la situation expérimentale que l'on pourra ou non, et éventuellement, conceptualiser une grandeur physique comme prenant des valeurs continues⁸.

Noter aussi qu'une échelle physique donnée a une signification toute relative, dépendant du problème analysé. Par exemple, la relaxation d'un système macroscopique (de volume V) traité en thermodynamique relativiste exige la considération de l'échelle de temps $\tau \stackrel{\text{def}}{=} \frac{V^{1/3}}{c}$ qui prend en compte la vitesse finie de propagation. De la même façon, la longueur d'onde Compton $\frac{h}{m_0c}$ n'est pertinente que pour des phénomènes impliquant une échelle d'énergie $h\nu \sim mc^2$.

Il convient également de rappeler que l'horizon de la Physique est borné ; en l'état actuel des choses, où gravitation et théorie quantique ne sont pas unifiées, le physicien est aveugle au-delà des échelles de Planck⁹. Pour n'importe quel problème de Physique, elles constituent des "infiniment petits/grands" incontournables bornant inférieurement/supérieurement toute autre échelle pertinente pour le problème considéré.

Ce sont ces distinctions qui permettent de légitimer une approche pragmatique des mathématiques à l'usage des physiciens, où l'introduction de certains nouveaux concepts en tant que résultats de passage à la limite au sens du mathématicien n'est pas toujours, au vu des principes physiques, indispensable, et ne s'impose que par pure commodité technique. L'exemple le plus simple venant à l'esprit est sans doute la notion de distribution, nécessaire en toute rigueur pour la formalisation de l'Électrostatique, laquelle pourtant n'a pas dû attendre (heureusement !) les années 1950 pour trouver son achèvement, et pour la Mécanique quantique mais cela n'a pas empêché Dirac de formuler l'Électrodynamique quantique dès 1928. En fait, le bien-fondé de l'approche physique des mathématiques repose sur l'hypothèse que le physicien n'a pas perdu la raison, rassuré par la certitude que s'il fait des bêtises, il va trouver des âneries ! Par exemple, la "fonction" de Dirac n'est que l'idéalisation conceptuelle d'une vraie et bonne fonction $\delta_\varepsilon(x - x_0)$ de largeur ε très étroite (à l'aune de la bonne échelle, précisément), d'intégrale unité, et qui, associée à une fonction $f(x)$ à variation lente, extrait précisément la valeur de la fonction au point de concentration :

$$\int_{-\infty}^{+\infty} f(x) \delta_\varepsilon(x - x_0) dx \simeq f(x_0) \quad (1)$$

Ceci n'est acceptable que dans la mesure où, si Δx est l'échelle de variation significative de f , l'inégalité $\varepsilon \ll \Delta x$ est vérifiée pour le problème considéré ; cela étant, le physicien n'a pas vraiment besoin de connaître la théorie des distributions pour écrire l'égalité

⁸Pour une source radioactive jeune et macroscopique, considérer le nombre de noyaux actifs comme une variable continue n'est pas déraisonnable puisque 1 est un "infiniment petit" comparé au nombre d'Avogadro. De même, l'intensité dans un circuit macroscopique est $I = \frac{dq}{dt}$, alors que l'on sait bien que la charge électrique est *quantifiée*, de même le nombre 1 est tout petit quand il est la différence de deux nombres quantiques valant chacun des milliards de milliards (d'où la limite (asymptotique) classique de la Mécanique quantique), etc.

⁹Le temps de Planck est $\tau_P = \sqrt{\frac{\hbar G}{c^5}} \sim 3 \times 10^{-43}$ s, la longueur de Planck est $l_P = \sqrt{\frac{\hbar G}{c^3}} \sim 9 \times 10^{-35}$ m ; quant à l'énergie de Planck $E_P = \sqrt{\frac{\hbar c^5}{G}}$, elle est de l'ordre de 4×10^8 J, soit environ 2×10^{18} GeV...

approchée (1), qui est en fait une simple triviale (au moins pour lui), et se borne à admettre la règle opérationnelle de la “fonction” de Dirac $\delta(x - x_0)$:

$$\int_a^b f(x) \delta(x - x_0) dx = f(x_0) \quad (a < x_0 < b) . \quad (2)$$

Dans le même ordre d'idées, à propos des rappels sur les séries de Fourier, on “démontrera” la relation :

$$\sum_{n \in \mathbb{Z}} e^{inx} = 2\pi \sum_{p \in \mathbb{Z}} \delta(x - 2p\pi) \quad (3)$$

qui joue un rôle important dans les problèmes de diffraction (condition de von Laue, équivalente à la condition de Bragg $n\lambda = 2d \sin \theta$). On donnera néanmoins un bref aperçu de la théorie des distributions, aussi appelées fonctions généralisées – on verra pourquoi.

Autre exemple, relatif à la notion de fonction périodique. Pour le mathématicien, cette notion est définie sans ambiguïté : une fonction $f(t)$ est T -périodique s'il existe un réel¹⁰ T tel que $f(t + T) = f(t)$. En soi, cette affirmation contient le fait que la fonction f est non nulle de $t = -\infty$ à $t = +\infty$. Pour le physicien, ce simple fait est une vue de l'esprit : si t est le temps, toute fonction $g(t)$ représente un phénomène de durée forcément finie ; ainsi, les fonctions manipulées par le physicien ne sont jamais stricto sensu périodiques. Par exemple, la fonction égale à $\sin(2\pi t/T)$ si $0 \leq t \leq \tau$ et nulle ailleurs sera considérée en Physique comme périodique si sa durée τ est grande par rapport à la période T (autrement dit, on a le temps de compter un grand nombre de périodes avant l'extinction). De façon plus quantitative, la transformée de Fourier sera considérée comme quasi monochromatique¹¹ si $\delta\omega \sim \tau^{-1}$ est très petit devant la résolution en pulsation disponible (expérimentalement ou dans l'intellect du physicien) : ici encore, une échelle (la résolution spectrale) permet au physicien de faire le saut conceptuel à propos d'objets qui ne sont pas strictement ceux qu'a définis le mathématicien.

Un dernier exemple. L'équation de Newton :

$$m \frac{d^2 \vec{r}}{dt^2} = \vec{F} \quad (4)$$

n'a jamais été démontrée par personne. La seule chose avérée (expérimentalement !) est que si l'on fait des mesures entre deux instants t et $t + \delta t$, la trajectoire polygonale construite par la succession des points discrets, relevés expérimentalement, s'inscrit sur la courbe continue différentiable (i.e. lisse, sans rugosité) déduite de (4). Bien sûr, la technologie aidant, on sait diminuer δt , mais il sera toujours fini, dans toute expérience, et ne sera jamais le dt du mathématicien. Si le physicien recourt à la forme limite (4), c'est juste parce qu'il sait intégrer les équations différentielles, et qu'il est beaucoup plus

¹⁰Il existe évidemment des fonctions périodiques dont la période est un nombre complexe. La restriction n'est ici que pour la pertinence physique de l'argument.

¹¹et dans les calculs on introduira justement des fonctions de Dirac $\delta(\omega \pm \omega_0)$, avec $\omega_0 T = 2\pi$, qui ne sont que l'idéalisation de fonctions très étroites $\delta_{\delta\omega}(\omega \pm \omega_0)$.

simple techniquement de procéder ainsi plutôt que de faire les calculs avec des accroissements finis¹². C'est aussi et surtout parce qu'aucune expérience n'est venue démentir (4) au sens où, à partir d'un δt trop petit, les points expérimentaux se seraient écartés de la ligne continue déduite de (4). Si un jour on découvre qu'il faut discrétiser temps et espace¹³ en-dessous de certains δr et δt , alors il faudra – au moins à ces échelles – renoncer à des écritures différentielles comme celles employées dans (4).

La notion de limite semble bien être l'un des concepts au sujet duquel les univers du mathématicien et du physicien se chevauchent sans se superposer parfaitement : si le second peut passer à la limite sur le papier pour faciliter un développement formel, empruntant ainsi un concept défini par le premier, il sait aussi qu'en cas de difficulté, un retour en arrière est toujours possible. Notamment, on sait bien que l'ordre des limites est parfois d'importance et conditionne le résultat. Lorsque tel est le cas, et sachant que la réponse physique d'un problème bien posé doit être unique, cet ordre doit être dicté par la physique du problème. Des exemples connus sont la transition ferromagnétique¹⁴, où les deux limites sont $N \rightarrow +\infty$ et $\mathcal{B} \rightarrow 0$, le coefficient de réflexion sur une marche de potentiel d'épaisseur l , pour laquelle les deux limites $\hbar \rightarrow 0$ et $l \rightarrow 0$ ne commutent pas non plus [2], ou encore la limite classique et la limite de température nulle ; dans ce dernier cas, le bon paramètre est un rapport du genre $\frac{k_B T}{\hbar \omega}$, et il est bien clair que cette quantité peut être zéro ou l'infini selon que l'on fait tendre T vers zéro avant \hbar , ou l'inverse.

S'agissant d'exposer des méthodes pour physiciens, on s'efforcera de montrer d'emblée en quoi les concepts et/ou les outils introduits trouvent leur application naturelle en Physique, laquelle a d'ailleurs souvent ouvert des voies aux mathématiciens¹⁵. Par ailleurs, lorsqu'il s'agira d'illustrer des résultats importants (ou des curiosités), on essaiera précisément de choisir un exemple physique concret.

¹²... ce que l'on fait toujours, en revanche, quand on utilise une procédure de résolution, ou de simulation, numérique.

À l'inverse, il arrive aussi que des fonctions non dérivables surviennent en Physique, par exemple dans le mouvement Brownien (les trajectoires browniennes n'ont pas de tangente !), ou dans la physique des systèmes désordonnés (escaliers du diable).

La grande place faite à l'intuition peut conduire à des situations dangereuses, voire à des résultats contre-intuitifs s'ils ne sont pas tout simplement erronés (la "vitesse" d'une particule brownienne n'existe pas). Il est justement espéré que la surprise provoque alors une réflexion en profondeur sur le point qui y a conduit.

¹³Ce que semblait penser Schrödinger, lui qui écrivit

“Nous ne devons pas admettre la possibilité d'une observation continue. Les observations doivent être considérées comme des événements discrets, disjoints les uns des autres. Entre elles il y a des lacunes que nous ne pouvons pas combler”

(*Physique quantique et représentation du monde*, Coll. Points Sciences, S78, Éd. du Seuil, Paris, 1992).

Si de tels quantums d'espace et de temps existent, ils sont en-deçà de toutes les capacités actuelles d'observation.

¹⁴Pour la transition ferromagnétique, on dispose de scénarios théoriques justifiant la bonne procédure de limite à considérer (*brisure d'ergodicité*) : en présence d'un champ magnétique *fini*, le temps de basculement d'un domaine augmente exponentiellement avec le nombre N de spins ; à la limite $N \rightarrow +\infty$, la moitié de l'espace des phases devient inaccessible [1].

¹⁵Par exemple, pour sa formulation de la Mécanique quantique (Mécanique des matrices, 1925), Heisenberg a introduit (apparemment sans se poser de question) des espaces vectoriels de dimension infinie dénombrable, dont la théorie a été faite ultérieurement, essentiellement par Hilbert.

Les notations utilisées respecteront les usages ; par exemple, x désignera le plus souvent l'argument d'une fonction et, de ce fait, est un nombre pur (i.e. un être mathématique bien défini en l'absence de toute référence à une unité). Au contraire, toutes les grandeurs physiques ont une dimension (masse, longueur, temps, etc). Il en résulte que l'argument de toute fonction (autre que la fonction monôme) intervenant dans un problème physique doit être sans dimension¹⁶ : écrire $\sin x$ quand x est une longueur n'a pas de sens (si on passe des mètres aux millimètres, la valeur de la fonction change !). En revanche $\sin(\frac{x}{x_0})$, où x_0 est une autre longueur du problème, est parfaitement sensé. De même $e^{-\gamma t}$ où t est le temps et γ l'inverse d'un temps, est parfaitement légitime alors que, dans les mêmes circonstances, e^{-t} est absurde.

Terminons ce préambule en laissant la parole à l'un des plus grands mathématiciens, Émile Borel, qui, dans son livre *Le Hasard* [3] et à propos de la notion de probabilité, s'exprime en des termes dont chaque mot peut être pesé :

“Toute probabilité concrète est en définitive une probabilité statistique définie seulement avec une certaine approximation. Bien entendu, il est loisible aux mathématiciens, pour la commodité de leurs raisonnements et de leurs calculs, d'introduire des probabilités rigoureusement égales à des nombres simples, bien définis : c'est la condition même de l'application des mathématiques à toute question concrète ; on remplace les données réelles, toujours inexactement connues, par des valeurs approchées sur lesquelles on calcule comme si elles étaient exactes : le résultat est approché, de même que les données”.

Notes

1. Les références indiquées ci-dessous reflètent le plus souvent un goût personnel, et ne constituent nullement une liste bibliographique exhaustive. Pour la plupart, elles renvoient à des ouvrages généraux aux qualités avérées et ayant fait date pour l'enseignement des mathématiques à l'usage des physiciens.

En outre, il existe de nombreux sites très intéressants par le souci pédagogique de leur(s) auteur(s), et parfaitement accessibles au niveau L3. L'un d'entre eux, particulièrement remarquable, est celui dû à Serge MEHL : <http://www.chronomath.com/>

2. Ce polycopié, très imparfait, doit être considéré comme un document de travail. C'est peu dire combien toutes les suggestions, remarques et critiques seront les bienvenues afin de pouvoir en améliorer le fond, le contenu et la forme¹⁷.

Je remercie d'avance les lecteurs attentifs de leur coopération.

¹⁶L'usage consistant à poser $\hbar = m = c = \dots = 1$ peut être une source de confusion.

La fonction monôme est l'exception qui confirme la règle : x^λ a une dimension qui n'entre en conflit avec personne. À l'inverse, le développement en série entière de $\sin x$ (par exemple) ferait apparaître une somme de termes n'ayant pas tous la même dimension.

¹⁷En raison de modifications effectuées en temps réel par la suite, la table des matières et la liste de références ci-dessous, établies à ce jour, pourront se révéler parfois approximatives.

Table des Matières

1	Éléments de Théorie des probabilités	1
1.1	Notion de variable aléatoire	2
1.2	Notion de probabilité	7
1.2.1	Définition intuitive	8
1.2.2	Difficultés de la définition intuitive	10
1.3	Axiomes. Premières conséquences	13
1.3.1	Axiomes	14
1.3.2	Premières conséquences	15
1.4	Sur l'intégrale de Lebesgue	18
1.5	Fonction de répartition	22
1.5.1	Fonction de répartition d'une v.a. discrète	23
1.5.2	Cas général	27
1.6	Espérances mathématiques (moyennes)	39
1.7	Lois de distribution courantes	45
1.7.1	Loi binomiale	45
1.7.2	Loi de Poisson	51
1.7.3	Loi uniforme	59
1.7.4	Loi de Gauss	59

1.7.5	Loi de Cauchy	61
1.7.6	Loi de Pareto	63
1.8	Fonction caractéristique	64
1.8.1	Définitions et propriétés	64
1.8.2	Somme de variables aléatoires	69
1.8.3	Stabilité d'une loi par l'addition	71
1.9	Lois-limites. Théorème limite central	72
1.9.1	Problématique	72
1.9.2	Théorème limite central	74
2	Introduction aux fonctions généralisées (distributions)	83
2.1	Présentation intuitive	83
2.2	Les distributions en tant que fonctionnelles linéaires	86
2.2.1	Définition d'une fonction généralisée (distribution)	86
2.2.2	Suites équivalentes : exemples à propos de la distribution de Dirac	90
2.2.3	Opérations sur les distributions. Dérivation	92
2.2.4	Les fonctions ordinaires comme distributions	96
2.2.5	Quelques distributions courantes	98
2.3	Transformées de Fourier des distributions	101
2.4	Illustrations et exemples	105
2.4.1	Oscillateur harmonique	105
2.4.2	Exemples de relations de fermeture	108
2.4.3	Les représentations $-q$ et $-p$ en Mécanique quantique	111
2.5	En guise de conclusion...	114

3	Fonctions d'une variable complexe	115
3.1	Rappels des opérations élémentaires sur les nombres complexes	115
3.2	Quelques éléments de topologie de \mathbb{C}	121
3.2.1	Ensembles de points dans le plan complexe	121
3.2.2	Point à l'infini	123
3.2.3	Courbes et domaines	125
3.3	Fonction d'une variable complexe	126
3.3.1	Définitions	126
3.3.2	Limite d'une fonction $f(z)$	129
3.3.3	Dérivée d'une fonction $f(z)$. Conditions de Cauchy – Riemann	131
3.4	Fonctions élémentaires	143
3.4.1	La fonction puissance entière $z \rightarrow Z = z^n$ ($n \in \mathbb{N}^*$) et sa(ses) fonction(s) inverse(s)	143
3.4.2	La fonction exponentielle $z \rightarrow Z = e^z$	145
3.4.3	La fonction logarithme $z \rightarrow Z = \ln z$ ($z \neq 0$)	148
3.4.4	La fonction puissance généralisée $z \rightarrow Z = z^\alpha$ ($\alpha \in \mathbb{C}$)	151
4	Intégration des fonctions d'une variable complexe	153
4.1	Préliminaires	153
4.2	Théorème de Cauchy (1825)	157
4.3	Généralisation au cas d'un domaine multiplement connexe	165
4.4	Formule de Cauchy	169
4.5	Dérivées d'ordre supérieur	172
4.6	Illustrations	176
4.6.1	$n = 0, 1, 2, \dots$	177
4.6.2	$n = -1, -2, \dots$	178

5	Représentation des fonctions analytiques par des séries.	
	Théorème des résidus	183
5.1	Séries de Taylor	184
5.2	Séries de Laurent	190
5.3	Classification des singularités d'une fonction	196
5.4	Théorème des résidus	205
5.4.1	Définitions et recettes de calcul des résidus	206
5.4.2	Démonstration du théorème des résidus	209
5.5	Prolongement analytique	213
5.6	Fonctions multiformes. Coupures. Notion de surface de Riemann	220
5.7	Un exemple physique de prolongement analytique : l'émission spontanée	225
6	Applications élémentaires du théorème des résidus	227
6.1	Lemmes de Jordan	227
6.2	Calcul d'intégrales définies	230
6.2.1	Fractions rationnelles	230
6.2.2	Intégrales d'une fonction rationnelle de lignes trigonométriques	237
6.2.3	Intégrales de Fourier	240
6.2.4	Exemples divers	243
6.3	Calcul d'intégrales de fonctions multiformes	252
6.4	Calcul d'intégrales impropres	260
6.5	Calcul de la somme de séries	265
6.6	Calcul de sommes finies	272
6.7	Quelques compléments sur les fonctions méromorphes	275
6.7.1	Formule intégrale de Poisson	275
6.7.2	Pôles et zéros d'une fonction méromorphe	277
6.8	Représentation des fonctions méromorphes et des fonctions entières	279

6.8.1	Décomposition des fonctions méromorphes en série d'éléments simples	279
6.8.2	Décomposition des fonctions entières en produits infinis	281
7	Quelques applications de la théorie des fonctions d'une variable complexe	283
7.1	Les fonctions d'Euler $\Gamma(z)$ et $B(p, q)$	284
7.1.1	La fonction Gamma $\Gamma(z)$	284
7.1.2	La fonction Beta $B(p, q)$	292
7.2	La fonction $\psi(z)$ d'Euler	297
7.3	La fonction $\zeta(z)$ de Riemann	300
7.4	Méthode du col	307
7.5	Développements asymptotiques	320
7.5.1	Idées générales et définitions	320
7.5.2	Un exemple classique	329
7.6	De l'importance des singularités apparemment innocentes	335
7.7	Un exercice pour en finir	336
7.8	Fonctions génératrices	342
7.9	Transformations conformes	342
7.10	Théorie du potentiel	342
8	Analyse de Fourier	343
8.1	Rappels sur les séries de Fourier	347
8.2	Transformation de Fourier	365
8.2.1	Définition et formule d'inversion	365
8.2.2	Propriétés de la transformation de Fourier	377
8.3	Propriétés asymptotiques	391
8.4	Généralisation en dimension quelconque	393
8.5	Causalité et analytité	396
8.6	Relations de Kramers - Kronig	402

9	Transformation de Laplace	405
9.1	Présentation	405
9.2	Définition et formule d'inversion	407
9.3	Propriétés de la transformation de Laplace	418
9.4	Propriétés asymptotiques	428
9.5	Quelques applications de la transformation de Laplace	432
9.5.1	Équations différentielles linéaires à coefficients constants	432
9.5.2	Équations aux différences finies	436
9.5.3	Équations aux dérivées partielles	441
9.5.4	Équations intégro-différentielles	445
9.5.5	Réduction de l'ordre d'une équation différentielle	449
9.5.6	Une application du théorème de convolution : le pendule isochrone de Huyghens	453
10	Équations différentielles. Introduction aux fonctions de Green	455
10.1	Généralités et définitions	455
10.2	Conditions initiales. Conditions aux limites	459
10.3	Équations différentielles linéaires à coefficients constants	461
10.3.1	Rappel de quelques résultats	461
10.3.2	Résolution à l'aide de transformations intégrales	463
10.4	Équations différentielles linéaires à coefficients variables	464
10.5	Équations différentielles et équations aux différences	465
10.6	Fonctions de Green	471
10.6.1	Préliminaires	472
10.6.2	Définition des fonctions de Green	474
10.6.3	Exemples	476

Bibliographie

- [1] Nigel GOLDENFELD, *Lectures on Phase Transitions and the Renormalization Group* (Addison - Wesley Publishing Co, Reading, 1992)
- [2] Walter APPEL, *Mathématiques pour la physique et les physiciens* (H & K Éditions, Paris, 2002)
- [3] Émile BOREL, *Le Hasard* (PUF, Paris, 1948)
- [4] Jean BASS, *Cours de mathématiques* (Masson, Paris, 1968)
- [5] Carl M. BENDER et Steven A. ORSZAG, *Advanced Mathematical Methods For Scientists And Engineers* (McGraw-Hill, Singapour, 1984)
- [6] Edward Charles TITCHMARSH, *Theory of functions*, (Oxford University Press, New York, 1988)
- [7] E.T. WHITTAKER et G.N. WATSON, *A Course of Modern Analysis*, (Cambridge University Press, 1988)
- [8] Hubert KRIVINE, *Exercices de mathématiques pour physiciens* (Cassini, Paris, 2003)
- [9] Mikhaïl LAVRENTIEV et Boris CHABAT, *Méthodes de la théorie des fonctions d'une variable complexe* (Éditions Mir, Moscou, 1972)
- [10] Jon MATTHEWS et R. L. WALKER, *Mathematical Methods of Physics* (Benjamin, New York, 1965)
- [11] N. PISKOUNOV, *Calcul différentiel et intégral*, tomes I et II, (Éditions Mir, Moscou, 1970)
- [12] Laurent SCHWARTZ, *Méthodes mathématiques pour les sciences physiques* (Hermann, Paris, 1968)
- [13] Jacques HARTHONG, cours d'Analyse et de Probabilités disponibles en ligne à l'adresse : <http://moire4.u-strasbg.fr/JHbooks>

- [14] Daniel DUGUÉ, “*Probabilités (Calcul des)*”, Encyclopædia Universalis, **18**, 1018 (Paris, 1992)
- [15] Daniel GLÆSER, “*Axiomatique*”, Encyclopædia Universalis, **3**, 670 (Paris, 1992)
- [16] William FELLER, *An introduction to probability theory and its applications*, Vol. I (J. Wiley, New York, 1971)
- [17] William FELLER, *An introduction to probability theory and its applications*, Vol. II (J. Wiley, New York, 1966)
- [18] Albert RENYI, *Calcul des probabilités* (Dunod, Paris, 1966)
- [19] Crispin W. GARDINER, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* (Springer, Berlin, 1990)
- [20] A. DAHAN DALMENICO, J.-L. CHABERT et K. CHEMLA, *Chaos et déterminisme* (Seuil, Coll. Point Sciences, Paris, 1992)
- [21] Jean PERRIN, *Les atomes* (Flammarion, Paris, 1991)
- [22] Matthias HOLSCHNEIDER, “*On the Wavelet Transformation of Fractal Objects*”, J. Stat. Phys., **50**, 963 (1988)
- [23] Marek FISZ, *Probability theory and mathematical statistics* (Wiley, New York, 1963)
- [24] J.F.C KINGMAN et S.J. TAYLOR, *Introduction to measure and probability* (Cambridge Univ. Pr., Londres, New York, 1966)
- [25] J. MARCINKIEWICZ, “*Sur une propriété de la loi de Gauss*”, Math. Z. **44**, 612-18 (1938)
- [26] Eugene LUKAS, “*A Survey of the Theory of Characteristic Functions*”, J. Math. Sci. **6**, 99-122 (1976)
- [27] L.Z.LIVSHITS, I.V. OSTROVSKII et G.P. CHISTYAKOV, “*Arithmetic of probability laws*”, Adv. Appl. Prob. **4**, 1-38 (1972)
- [28] Claude ASLANGUL, “*Diffusion on a Random Comb : Distribution Function of the Survival Probability*”, J. Stat. Phys., **78**, 1403, (1995)
- [29] B.V. GNEDENKO et A.N. KOLMOGOROV, *Limit distributions for sums of independent random variables* (Addison-Wesley, Cambridge, 1954)
- [30] M.J. LIGHTHILL, *Introduction to Fourier Analysis and Generalized Functions* (Cambridge University Press, Cambridge, 1978)
- [31] V.I.SMIRNOV, *A course of Higher Mathematics. I* (Pergamon Press, Oxford, 1964)
- [32] Tristan NEEDHAM, *Visual Complex Analysis* (Clarendon Press, Oxford, 1997)
- [33] <http://mathworld.wolfram.com/Aleph-1.html>

- [34] Marcus Du SAUTOY, *La symphonie des nombres premiers*, (Héloïse d'Ormesson, Paris, 2005)
- [35] Godfrey Harold HARDY, *L'apologie d'un mathématicien* (Éditions Belin, Paris, 1985)
- [36] I.S. GRADSTEYN et I.M. RYZHIK, *Tables of Integrals, Series and Products* (Academic Press, New York, 1980)
- [37] Edward Charles TITCHMARSH, *The Theory of Riemann Zeta-function*, (Clarendon Press, Oxford, 1986)
- [38] Philip PECHUKAS et John C. LIGHT, "On the Exponential Form of Time-Displacement Operators in Quantum Mechanics", *J. Chem. Phys.*, **44**, 3897 (1966)
- [39] Edmund HLAWKA, Johannes SCHOISSENGEIER et Rudolf TASCHNER, *Geometrical and Analytic Number Theory* (Springer, Berlin, 1991)
- [40] J. Brian CONREY, "The Riemann Hypothesis", *Notices of the AMS* (mars 2003)
- [41] <http://www.zetagrid.net/zeta/math/zeta.result.100billion.zeros.html>
- [42] André LECLAIR, "Interacting Bose and Fermi gases in low dimensions and the Riemann Hypothesis", [arXiv:math-ph/0611043v3](https://arxiv.org/abs/math-ph/0611043v3) (18 septembre 2007)
- [43] Harold M. Edwards, *Riemann's Zeta Function* (Academic Press, New York, 1974)
- [44] <http://en.wikipedia.org/wiki/Renormalization>
- [45] Bertrand DELAMOTTE, "A hint of renormalization", *Amer. J. Phys.*, **72**, 170 (2004) ; [arXiv:hep-th/0212049](https://arxiv.org/abs/hep-th/0212049) (5 septembre 2003)
- [46] Henri Poincaré, *Les méthodes nouvelles de la mécanique céleste*, (Gauthier-Villars, Paris, 1892)
- [47] John W. DETTMAN, *Applied Complex Variables* (MacMillan, New York, 1965)
- [48] <http://www.math.polytechnique.fr/xups/xups91.pdf>
- [49] Richard P. FEYNMAN et Alan R. HIBBS, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965)
- [50] D.J. WALLACE, "Some applications of instantons in statistical mechanics", in *Solitons and Condensed Matter Physics*, A.R. Bishop and T. Schneider éditeurs (Springer-Verlag, 1978) p. 104-15.
- [51] Lev LANDAU et Evguéni LIFSHITZ, *Mécanique Quantique* (Mir, Moscou, 1967)

Chapitre 1

Éléments de Théorie des probabilités

“It is impossible to trap modern physics into predicting anything with perfect determinism because it deals with probabilities from the outset.”

(Sir Arthur EDDINGTON, 1882–1944)

Ce chapitre est consacré à un exposé élémentaire de la Théorie des probabilités. L'approche intuitive délibérément choisie permettra néanmoins d'obtenir des résultats importants et indispensables pour le Physicien, de caractériser les lois de répartition les plus courantes et d'énoncer le Théorème limite central.

La théorie des probabilités constitue un pan entier des mathématiques et fait appel à des notions très abstraites liées à la difficulté du sujet¹. Celle-ci peut être mesurée par le fait que, si la notion de probabilité est finalement assez ancienne, il a fallu attendre les années 1930 pour qu'en soit faite une construction axiomatique² cohérente (Kolmogorov, 1933). On touchera du doigt plus loin les subtilités qui rendent les exposés élémentaires insatisfaisants quand le but est l'édification d'une théorie dénuée d'ambiguïtés et logiquement irréfutable.

¹Selon Daniel Dugué [14] :

“Le calcul des probabilités est certainement l'une des branches les plus récentes des mathématiques, bien qu'il ait en fait trois siècles et demi d'existence.”

²D'après Georges Glæser [15] :

“Ainsi, lorsque A. Kolmogorov a formalisé, en 1933, les fondements du calcul des probabilités, il a permis à cette science de se dégager des notions vagues et subjectives de hasard³ et de chance ; l'étude des probabilités, restée longtemps stagnante, a pris ainsi un nouvel essor, devenant une partie très vivante des mathématiques.”

On s'en tiendra ici à un exposé très élémentaire, essentiellement guidé par le pragmatisme et par le souci de garder le contact avec les préoccupations usuelles du physicien. Ceci n'interdira pas, de loin en loin, de mentionner des situations un peu exotiques que l'on rencontre parfois en Physique, par exemple les mesures singulières continues à propos des systèmes désordonnés. Un ouvrage de référence en la matière est l'admirable livre de Feller [16], [17], dont la lecture est vivement recommandée.

1.1 Notion de variable aléatoire

Il semble indiscutable que la notion de variable aléatoire⁴ est inévitablement liée à la notion d'*expérience*, ou de *mesure*⁵. On entend par là qu'un dispositif expérimental ayant été mis en place, le but de l'expérience est d'obtenir les valeurs d'une certaine grandeur d'intérêt, qu'il s'agisse de l'intensité dans un circuit électrique ou de la taille des individus d'une population donnée. Il est toujours possible d'établir une correspondance (bijective) entre les résultats d'un type d'expérience et des nombres ; par exemple, s'agissant de distinguer quantitativement des boules de couleur différente, des blanches, des rouges et des noires par exemple, on peut décider d'affecter la valeur -1 à la couleur blanc, la valeur 0 à la couleur rouge, la valeur 1 à la couleur noir. De même, pour la pièce de monnaie, on pourra dire que face correspond à 1 , et pile à 0 . Pour le dé cubique (six faces), on peut décider d'attribuer à chaque face le nombre correspondant à sa valeur. Cette association est purement conventionnelle et n'est définie que dans un but opératoire mais, évidemment, une fois choisie il faut s'y tenir. En tout cas, on pourra donc toujours supposer que l'expérience (la mesure) débouche sur une valeur *numérique*. Le caractère *aléatoire* est révélé par le fait que la répétition de l'expérience, censée se reproduire à l'identique d'une fois à l'autre, ne produit pas toujours le même résultat : N jets de la pièce produiront au total n fois pile et $N - n$ fois face, sans que l'on puisse prédire, lorsque la pièce est en l'air, de quel côté elle va retomber. L'expérience est ici le jet de la pièce, le résultat de chaque expérience est soit pile, soit face ; avec la correspondance définie plus haut, les valeurs numériques obtenues sont donc soit 1 soit 0 respectivement ; avec ce cadre de travail, la variable aléatoire est le résultat d'un lancer de la pièce, et peut prendre les deux⁶ valeurs 0 et 1 . Une variable aléatoire pouvant prendre un nombre fini de valeurs distinctes, ou un nombre infini dénombrable, est dite *discrète*. Par la suite, on considérera aussi des variables aléatoires *continues*, qui prennent leurs valeurs dans \mathbb{R} ou une partie de \mathbb{R} . L'abréviation consacrée pour *variable aléatoire* est *v.a.* Il convient toujours de soigneusement distinguer d'une part la variable aléatoire abstraite, d'autre part les valeurs obtenues lors d'une expérience – tout comme on doit distinguer une fonction $x \rightarrow f(x)$ des valeurs numériques $f(x_n)$. Le plus souvent, on notera X la v.a. et x_n ses valeurs numériques possibles, ou simplement x si la v.a. est continue.

⁴Le terme vient du latin *alea*, qui désigne le jeu de dés (dictionnaire Littré).

⁵Le sens de ce terme est précisé dans les lignes qui suivent. Le même terme est employé (pas par hasard !) par les mathématiciens, pour désigner une notion abstraite dûment définie qui, dans les cas élémentaires, se réduit à celle utilisée par l'arpenteur (longueur d'un chemin, surface d'une parcelle, ...)

⁶On parle souvent alors de variable *binnaire* ou *variable de Bernoulli* ; dans le contexte des phénomènes critiques, le physicien emploie le vocable *spin*, en référence à la *Drosophile* des experts en la matière, ou encore *variable d'Ising*.

Dans le cas du jet de la pièce, la non-reproductibilité du résultat de chaque mesure tient au fait que l'expérimentateur, de fait, ne contrôle pas tous les paramètres déterminant le mouvement de la pièce. En toute rigueur, si tous ces paramètres étaient contrôlés et connus, un protocole complètement défini devrait donner toujours le même résultat, étant entendu que la pièce est un objet macroscopique obéissant aux équations déterministes de la Mécanique classique⁷. En fait, au-delà des apparences sensibles, et des effets mesurables, de nombreuses petites causes influent le mouvement de la pièce en dehors de tout contrôle, et ce sont elles qui font finalement varier le résultat de chaque lancer. L'ensemble de ces causes est ce que l'on appelle communément le *hasard* et traduit, en quelque sorte, l'ignorance inévitable d'effets certes petits mais essentiels puisqu'ils produisent une dispersion des résultats. Le hasard est ainsi l'ensemble des causes non-maîtrisées qui sont responsables de la variabilité des résultats de mesures répétées identiques, au sens où le protocole expérimental est complètement défini.

Le but de toute théorie probabiliste est non pas de tenter de décrire (de façon déterministe) les causes d'une telle variabilité (par nature le hasard échappe à tout contrôle), mais d'en prendre acte et de fournir un cadre alternatif affectant à l'issue possible de chaque mesure un nombre (sa probabilité) permettant de quantifier les résultats d'un *très grand nombre* (à la limite infini) d'expériences. En quelque sorte, l'approche probabiliste part d'un constat (l'impossibilité de décrire dans le détail les raisons premières qui produisent un résultat ou un autre) et, par pragmatisme, aborde le problème posé suivant une autre méthode⁸. En définitive, confronté à la pertinence incontournable du hasard pour un phénomène donné, on recourt à une méthodologie radicalement différente, dont les ambitions quantitatives restent à un niveau d'exigence élevé et dont les vertus prédictives sont d'une extrême richesse.

Dans le cas des boules de couleurs, le hasard réside dans le fait que l'on définit un jeu, consistant à tirer *en aveugle* les boules dans un sac. En quelque sorte, le hasard tient à la règle du jeu elle-même, car il n'y a rien d'aléatoire en soi dans la couleur d'une boule donnée : dans le sac, chaque boule a sa couleur, qui ne change pas au cours du temps. Il convient de dire dès à présent que la théorie des jeux a joué un rôle moteur dans le développement historique de la Théorie des probabilités, dès le XVIII^e siècle (D. Bernoulli, puis plus tard Laplace, Poisson, Gauss, Galton...).

En Physique, tout résultat de mesure est une variable aléatoire, puisque toute

⁷Dans le domaine quantique, il en va tout autrement : on doit admettre – sous la poussée de l'expérience ! – que les objets microscopiques ont une dynamique *foncièrement* aléatoire. D'un autre côté, contrairement à une affirmation fréquente, la Mécanique quantique est elle aussi une théorie déterministe : elle prévoit avec certitude (?) ce que sont les probabilités des événements observables. Par rapport à la Mécanique classique, le déterminisme n'a pas disparu, il a seulement changé de registre.

Le cadre quantique est caractérisé par le fait que le *hasard* y est irréductible, et ne tient en aucune façon à l'imprécision d'une expérience ou à une incapacité de formalisation détaillée – 8 et 9.

⁸C'est la même forme de renonciation qui constitue les prémisses de la Mécanique Statistique : aucun physicien doué de raison n'envisage une seconde de décrire les propriétés d'un gaz classique en se fixant comme objectif de trouver les trajectoires des atomes du gaz. Cet objectif est utopique, ses résultats (inaccessibles !) seraient inexploitablement, enfin et surtout cette approche est en contradiction flagrante avec l'expérience : pour décrire (parfaitement) les propriétés d'un gaz, quelques paramètres en nombre très réduit suffisent (pression, volume, température, pour un système fermé) : rien à voir avec les $\sim 10^{23}$ degrés de liberté mécaniques du gaz !

mesure est entachée d'imprécision : quel que soit le soin pris dans la réalisation d'une expérience et sa conduite à terme, la valeur obtenue n'est pas infiniment précise ; d'ailleurs, la *barre d'erreur* est un ingrédient essentiel pour légitimer la valeur finalement retenue pour la grandeur physique d'intérêt. Ceci est vrai même si, selon la théorie en vigueur, la grandeur que l'on cherche à mesurer n'a en principe qu'une seule valeur possible⁹, par exemple l'énergie d'un système isolé (en pratique, un système n'est *jamaïs* strictement isolé !).

En ce qui concerne les grandeurs classiques (*i.e.* non quantiques), l'imprécision a essentiellement trois causes :

1. les appareils utilisés ont une résolution finie Δ , qui permet seulement de situer les résultats de mesure sur une grille de points isolés les uns des autres, séparés de Δ . Pour toute valeur mesurée x_0 , on peut au plus dire que la valeur de la grandeur est dans l'intervalle $[x_0 - \frac{\Delta}{2}, x_0 + \frac{\Delta}{2}]$.
2. L'état de l'appareil dépend de divers paramètres, qui peuvent varier d'une expérience à l'autre (ce que l'on peut désigner – assez vaguement, il faut bien le dire¹⁰ – par le vocable de *fluctuations*). Il en résulte que la valeur effectivement mesurée peut varier d'une expérience à l'autre – clairement, la grille de mesures elle-même peut fluctuer.
3. La grandeur elle-même peut être aléatoire, faute d'une maîtrise absolue des conditions expérimentales.

Cela étant, deux situations extrêmes peuvent être envisagées :

1. la résolution Δ est très bonne, c'est-à-dire que Δ est très petit devant les variations mesurées résultant des fluctuations de toute nature. Alors, on pourra tracer un histogramme¹¹ dont les bâtons ont pour largeur Δ , qui constituera l'histogramme des valeurs mesurées. Sa moyenne et sa largeur seront caractéristiques des variations statistiques des mesures ;
2. la résolution est médiocre, ou en tout cas Δ est grand vis-à-vis des fluctuations. Alors, toutes les valeurs mesurées coïncideront les unes avec les autres, l'écart entre les deux valeurs extrêmes n'excédant pas Δ .

⁹Il existe évidemment des grandeurs *intrinsèquement* aléatoires : la longueur d'une règle a des fluctuations liées aux fluctuations thermiques, toutes les grandeurs quantiques, ...

D'un autre côté, on peut, paradoxalement à première vue, affirmer que le monde quantique est *infiniment déterminé*, dans la mesure où certaines grandeurs physiques sont quantifiées, c'est-à-dire ne peuvent prendre leurs valeurs que dans un ensemble de mesure nulle au sens du mathématicien. Par exemple, les énergies possibles d'un système lié forment un ensemble fini ou isomorphe à \mathbb{N} , $\{E_n\}_n$. Le résultat d'une mesure donnant la valeur E_{n_0} est de fait *infiniment précis* dès lors que l'incertitude liée aux appareils de mesure est plus petite que l'écart entre E_{n_0} et ses deux valeurs les plus proches. On peut ainsi comprendre pourquoi des caractéristiques atomiques (la longueur d'onde d'une transition, par exemple) peuvent servir de standard métrologique.

¹⁰L'un des enjeux est précisément de préciser et de quantifier ces variations imprévisibles.

¹¹Le mot est pris ici au sens élémentaire ; un langage plus précis (et propre aux probabilistes) appelle *histogramme* la fonction de répartition définie plus loin (section 1.5).

Finalement, affirmer que le résultat de la mesure d'une grandeur physique (classique, voir la note 9) est une variable aléatoire est d'une extrême banalité. Si X note la grandeur mesurée, les valeurs obtenues seront (dans la bonne unité de X), des nombres x_n définis par la grille discrète de pas Δ . On pourra dire que le résultat de chaque mesure est x_n , si l'on a trouvé une valeur dans l'intervalle $[x_n - \frac{\Delta}{2}, x_n + \frac{\Delta}{2}]$. Dans le premier cas ci-dessus (très bonne résolution), la grandeur X ressort comme une variable aléatoire et est distribuée suivant un certain histogramme. Dans le second cas (fluctuations d'appareil très petites et/ou résolution médiocre), la variable X apparaît comme certaine, mais il serait aventureux d'affirmer, au vu de cette série d'expériences, qu'elle est intrinsèquement une variable certaine, c'est-à-dire ne pouvant prendre qu'une et une seule valeur¹².

La règle générale est donc l'aléatoire, la variable certaine étant en fait le cas-limite idéal où toutes les valeurs sauf une sont impossibles ; cette valeur unique donne à la grandeur la qualité de variable non distribuée. Pour une grandeur physique donnée, le statut aléatoire *vs* certaine n'est pas toujours propre à cette grandeur, mais peut dépendre du cadre d'étude¹³ et/ou du degré de précision des expériences envisagées. Ici encore, le zéro du physicien n'est pas celui du mathématicien : pour le physicien, une variable X prenant une valeur x_1 avec la probabilité $10^{-10^{23}}$ et une valeur différente x_2 avec la probabilité $1 - 10^{-10^{23}}$ sera considérée comme une variable non aléatoire¹⁴ (certaine)...

Pour désigner le résultat d'une expérience, le terme conventionnel est *événement*, noté ω . Plus précisément, on se doit de définir d'abord les événements *élémentaires* ("atomiques"), c'est-à-dire dresser la liste de tous les résultats possibles de l'expérience. Pour la pièce de monnaie, il y a deux événements élémentaires :

$$\omega_1 = \text{pile (P)} \quad \text{ou} \quad \omega_2 = \text{face (F)}$$

Pour les boules à trois couleurs, les événements (élémentaires) sont

$$\omega_1 = \text{blanc} \quad \text{ou} \quad \omega_2 = \text{noir} \quad \text{ou} \quad \omega_3 = \text{rouge}$$

Pour le dé, il y a 6 événements élémentaires, etc. L'ensemble des événements élémentaires sera noté Ω et est appelé *espace des épreuves* (aussi appelé parfois *univers*). Une fois introduite une correspondance biunivoque entre les résultats possibles d'une expérience

¹²Répétons qu'une distinction est ici toutefois nécessaire. Quand on mesure des grandeurs prenant des valeurs $x \in \mathbb{R}$ *continues* – c'est toujours le cas pour les grandeurs physiques classiques –, la prise en compte de l'imprécision expérimentale, toujours inévitable, permet de dire que la mesure de toute telle grandeur fournit une variable aléatoire. Ceci n'est plus vrai pour les grandeurs physiques *quantifiées* qu'introduit la Mécanique quantique, dont l'un des postulats affirme que toute mesure ne peut fournir qu'une valeur *discrète* x_n , à prendre dans un ensemble fini ou infini $\{X\}$, mais toujours *dénombrable*. En pareil cas, si Δ est plus petit que $\min_{\{X\}} |x_n - x_{n'}|$, mais toujours forcément fini, la mesure peut bel et bien être qualifiée d'*infinitement* précise, puisqu'elle fournit alors *la* seule et unique valeur disponible dans l'intervalle de résolution de l'appareil.

¹³Pour un système isolé (microcanonique), l'énergie est une grandeur certaine ; pour le même système placé en situation canonique, son énergie est une grandeur fluctuante.

¹⁴D'où l'affirmation contenue dans le Second principe de la Thermodynamique. Sur la tombe de Boltzmann dans le cimetière central de Vienne, cette inscription en guise d'épithaphe : $S = k \ln W$ (voir p. 6).



Figure 1.1: La tombe de Boltzmann à Vienne, avec l'illustissime formule. La constante k a été en fait introduite par... Planck, une fois que celui-ci eût dépassé son aversion de l'approche statistique de Boltzmann.

(e.g. pile \leftrightarrow 0, face \leftrightarrow 1), définissant les valeurs de la variable aléatoire, on peut dire qu'une variable aléatoire est une application de l'espace des épreuves dans \mathbb{R} (voire¹⁵ dans \mathbb{C}).

Les événements élémentaires sont des “briques” permettant de définir des événements composites (“molécules”) ; par exemple, on peut décider de lancer N fois la pièce : chaque expérience peut être représentée par le mot de N lettres PPFPPPPFFPF...F, ou par un nombre binaire 110101100010...0, c'est un événement (composite), fabriqué avec les briques que constituent les événements élémentaires. Les résultats de toutes ces expériences peuvent être complètement spécifiés à l'aide des deux seuls événements élémentaires. Obtenir pile *ou* face est un événement non-élémentaire, ainsi qu'obtenir pile *et* face¹⁶. Pour les boules dans le sac, obtenir deux blanches et une noire en trois essais est un exemple d'événement non-élémentaire, qui d'ailleurs n'est ni certain ni impossible. Si nécessaire, on désignera par des capitales (A, B, \dots) des événements composites, pour bien les distinguer des événements élémentaires ω_n .

Deux événements sont dits *exclusifs* si l'obtention de l'un interdit celle de l'autre (on dit aussi *incompatibles*, comme on parle des observables incompatibles en Mécanique quantique). Deux événements élémentaires sont donc forcément exclusifs : tout jet de

¹⁵Pour le dé, on peut tout autant définir la v.a. $e^{ik\frac{2\pi}{6}}$ pour la face de valeur k ($k = 1, 2, \dots, 6$). L'image de Ω n'est plus alors l'ensemble des six points sur l'axe réel d'abscisses k , mais six points équidistants sur le cercle unité de \mathbb{C} .

¹⁶c'est un événement dit *impossible*.

la pièce fournit à chaque fois une valeur, excluant l'autre *ipso facto* – on peut même caractériser les événements élémentaires en disant que pour chacun d'entre eux, le résultat de toute expérience est soit cet événement, soit son contraire.

À l'inverse, deux événements composites distincts A et B ne sont pas forcément exclusifs ; par exemple, les deux événements :

$A =$ tirer deux boules blanches et une boule noire

$B =$ tirer trois boules de couleurs différentes

ne sont pas exclusifs. Autre exemple : B est d'obtenir 8 comme somme du lancer de deux dés, et A est d'obtenir 2 pour l'un, 6 pour l'autre. Dans ces cas on voit aussi que la réalisation de A entraîne la réalisation de B : on dit que A entraîne B , ou que A est inclus dans B – ce que l'on note $A \subset B$, ou $A \implies B$.

L'événement consistant en la réalisation de A ou de B est noté $A \cup B$; celui consistant en la réalisation de A et de B est noté $A \cap B$ (si $A \subset B$, alors $A \cap B \neq \emptyset$). Cette notation est très utile, et parle à l'esprit : on verra (sans grande surprise, compte tenu des habitudes algébriques de l'esprit) que l'union (\cup) est en relation étroite avec l'*addition* des probabilités¹⁷, que l'intersection (\cap) est associée avec la *multiplication* des probabilités¹⁸, et qu'enfin l'inclusion (\subset ou \subseteq) correspond à l'inégalité $<$ ou \leq . Pour les événements élémentaires, on a évidemment :

$$\boxed{\cup_n \omega_n = \Omega, \quad \omega_n \cap \omega_{n'} = \emptyset \quad (\forall n \neq n')} \quad (1.1)$$

En revanche, pour des événements composites, $A \cap B \neq \emptyset$, sauf si A et B sont incompatibles. Dans la suite, on utilisera l'une ou l'autre des notations, ou les deux (voir par exemple (1.18)).

L'événement *certain* est un événement qui se produit à chaque expérience ; obtenir pile ou face est l'événement certain du jeu de pile et face. Pour la complétude, on définit aussi la notion d'événement impossible : c'est un événement qui ne peut pas se produire (tirer une boule de couleur indigo). Il y a autant d'événements impossibles que l'on veut !

1.2 Notion de probabilité

Il s'agit maintenant d'aborder la description quantitative formant la base d'une méthodologie de nature probabiliste, en suivant comme annoncé une démarche essentiellement pragmatique¹⁹. Dans un premier temps, on s'en tient à une approche intuitive où il sera supposé que le nombre d'événements élémentaires est soit fini, soit dénombrable ; ceci veut dire que le cardinal de Ω est un nombre entier fini ou infini : Ω pourra toujours

¹⁷pour des événements *exclusifs* – voir plus loin.

¹⁸pour des événements *indépendants* – voir plus loin.

¹⁹Cette approche est parfois qualifiée d'un horrible néologisme, étant désignée comme l'approche *fréquentiste*. On peut préférer *approche statistique*...

être mis en bijection avec $\{1, 2, \dots, N_0\}$ ou avec \mathbb{N} , et est alors qualifié de *discret*, tout comme la v.a. correspondante (elle prend des valeurs bien séparées les unes des autres). En conséquence, l'ensemble des événements composites construits avec les $\omega_n \in \Omega$ est forcément lui aussi dénombrable. Dans un deuxième temps, mais sans pouvoir rentrer dans le détail, on pointera les difficultés inhérentes à cette approche intuitive.

1.2.1 Définition intuitive

Énoncer la liste des événements élémentaires possibles, supposés formant un ensemble dénombrable, est un premier stade, de type qualitatif. Pour donner une véritable dimension quantitative à la description théorique, il convient d'attribuer un nombre mesurant que tel ou tel événement se produit plus ou moins souvent. Pour ce faire, l'approche intuitive consiste à se reposer sur une longue série d'expériences réputées identiques²⁰, et à faire le relevé soigneux du nombre de fois où un événement donné s'est produit ; clairement, il suffit d'effectuer ce travail sur les événements élémentaires. Si \mathcal{N} est le nombre total d'expériences, et si l'événement (élémentaire) ω_n est survenu \mathcal{N}_n fois, on définit alors sa *fréquence statistique* f_n par le rapport :

$$\boxed{\text{fréquence d'observation de } \omega_n \equiv f_n \stackrel{\text{déf}}{=} \frac{\mathcal{N}_n}{\mathcal{N}}} \quad (1.2)$$

L'acte de foi consiste alors à *admettre*, sous la poussée de l'expérience, que dans la limite d'un nombre infini de tirages, chaque fréquence a une limite qui est, suivant cette définition pragmatique, la *probabilité* de l'événement ω_n :

$$\boxed{\forall \omega_n \in \Omega, \exists \lim_{\mathcal{N} \rightarrow +\infty} f_n \text{ et } p_n \stackrel{\text{déf}}{=} \lim_{\mathcal{N} \rightarrow +\infty} f_n \iff \text{Probabilité de } \omega_n = p_n} \quad (1.3)$$

Évidemment, la convergence (supposée) de la suite des f_n vers les p_n est fort subtile : en réalité, chaque série d'expériences a sa propre suite de fréquences. Si on fixe n , en se concentrant sur un événement donné, on peut tracer une trajectoire donnant f_n en fonction du nombre d'expériences \mathcal{N} . Cette trajectoire a un aspect "rugueux" et passe sans cesse d'un côté à l'autre de la valeur-limite p_n , en suivant le plus souvent un chemin très surprenant : il peut rester très longtemps d'un même et unique côté, faire une brève excursion de l'autre, avant de franchir la limite et s'en écarter en donnant l'impression de n'y plus vouloir revenir. Clairement, pour un événement donné, la trajectoire varie d'une série d'expériences à l'autre. Le point essentiel est de postuler que toutes les trajectoires relatives à un événement donné ont une limite commune – dans le langage convenu, on peut parler de *bassin d'attraction*, vers lequel convergent toutes les trajectoires (*tous les chemins mènent à Rome*). Cette régularité est parfois appelée *stabilité des fréquences statistiques*.

²⁰On appelle souvent *tirage* ou *échantillonnage*, une telle longue suite d'expériences. À ce propos, rappelons que la notion de *reproductibilité* d'une expérience est l'une des pierres angulaires de la méthodologie de la Physique.

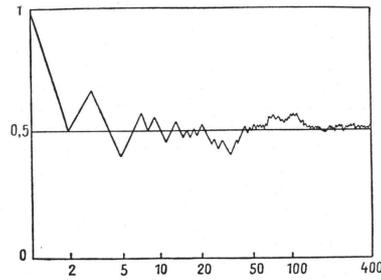


Figure 1.2: “Trajectoire” de la probabilité d’avoir *pile* lors d’un tirage de 400 coups (figure extraite de l’ouvrage de Rényi [18], p. 25).

Ici, rien n’est donc affirmé sur la loi de convergence des fréquences f_n vers leurs valeurs limites p_n : la seule hypothèse, à ce stade, est l’existence de la limite p_n pour chacune d’entre elles. La grande difficulté de l’échantillonnage statistique tient au fait qu’il faut en général un *très grand* nombre d’expériences pour que la fréquence f_n soit soit à peu près stabilisée à une valeur proche de la valeur théorique p_n .

Par leur définition (1.2), les fréquences sont des quantités positives ; leur limite p_n (la probabilité de ω_n) est donc elle aussi positive. En outre, par construction, la somme des fréquences est constante et égale à 1, quel que soit \mathcal{N} et quel que soit le tirage :

$$\sum_n f_n = \sum_n \frac{\mathcal{N}_n}{\mathcal{N}} = \frac{\sum_n \mathcal{N}_n}{\mathcal{N}} = \frac{\mathcal{N}}{\mathcal{N}} = 1 . \quad (1.4)$$

Il en va de même de la somme des p_n :

$$\boxed{\sum_n p_n = 1} \quad (1.5)$$

Par ailleurs, la fréquence d’avoir un événement ω_n ou un autre événement $\omega_{n'}$ est évidemment la *somme* de leurs fréquences respectives, somme du nombre de cas où ω_n s’est produit et du nombre de cas où $\omega_{n'}$ s’est produit :

$$\text{fréquence d’observation de } (\omega_n \text{ ou } \omega_{n'}) = f_n + f_{n'} ; \quad (1.6)$$

on a donc aussi :

$$\boxed{\text{Probabilité de } (\omega_n \text{ ou } \omega_{n'}) = p_n + p_{n'}} \quad (1.7)$$

L’événement composite : obtention de ω_n ou de $\omega_{n'}$, noté $\omega_n \cup \omega_{n'}$, conformément à ce qui a été mentionné plus haut, a donc tout naturellement une fréquence égale à la *somme* des fréquences, une propriété qui se reporte sur les probabilités correspondantes, signature du fait que deux événements élémentaires sont forcément exclusifs. D’un autre côté, la composition *et* est en un sens triviale lorsqu’il s’agit d’événements élémentaires, pour lesquels on a toujours $\omega_n \cap \omega_{n'} = \emptyset$, d’où :

$$\boxed{\text{Probabilité de } (\omega_n \text{ et } \omega_{n'}) = 0 \text{ pour tout couple d’événements élémentaires}} \quad (1.8)$$

À ce stade, on pourrait être tenté de vouloir définir la notion de probabilité sans faire référence à la notion de fréquence statistique. Par exemple, pour le dé cubique parfait, on pourrait dire d'emblée que la probabilité de chaque face est $\frac{1}{6}$, rapport entre le nombre de cas favorables et le nombre de cas possibles. En réalité, cette définition n'en est pas une, car elle suppose (implicitement) que chaque face a... la même probabilité que toutes les autres ! En effet, s'il n'en va pas ainsi (dé pipé), il est bien clair que les probabilités des différentes faces ne sont pas égales entre elles et ne valent pas toutes $\frac{1}{6}$. Il en résulte que la définition tentative ci-dessus devrait être précisée en disant que la probabilité est le rapport du nombre de cas favorables au nombre de cas possibles²¹, sachant que tous les cas possibles ont la même probabilité. Alors le caractère spécieux de cette définition saute aux yeux : pour définir la notion de probabilité, on fait appel à... la notion de probabilité. En Logique, une telle définition est souvent dite *circulaire* et, ne serait-ce que pour cette raison, cette définition ne saurait être pleinement satisfaisante et définitive. Notons toutefois que l'équiprobabilité des cas possibles est probablement (!) l'une des meilleures définitions du hasard : c'est justement parce que l'ignorance est totale que tous les cas possibles doivent être traités sur un pied d'égalité²².

1.2.2 Difficultés de la définition intuitive

L'incohérence logique qui vient d'être pointée est l'une des difficultés historiques de la théorie première des probabilités. Une autre difficulté majeure survient quand le nombre d'événements a la puissance du continu. Dans tous les exemples intuitifs présentés ci-dessus, le nombre d'événements est fini. La plupart des résultats s'étendent sans trop de difficulté au cas où Ω est équipotent à \mathbb{N} , moyennant la définition convenable de la convergence sous différentes formes ; par exemple, une distribution de probabilités p_n peut tout à fait être de la forme $p_n = \frac{C}{n^\alpha}$ ($n \in \mathbb{N}^*$) à condition que l'exposant α soit strictement supérieur à 1. En revanche, si Ω a la puissance du continu, la difficulté est profonde, et la définition intuitive extrapolée brutalement conduit aisément à des situations présentées comme paradoxales.

La notion de variable aléatoire peut se généraliser au cas où les valeurs possibles sont dans un intervalle inclus au sens large dans \mathbb{R} (ce peut être \mathbb{R} lui-même). En effet, rien dans la notion de v.a. ne s'oppose à considérer un tel cas, c'est même à la réflexion le cas le plus naturel, puisque (théoriquement) la plupart des grandeurs physiques prennent

²¹On trouve cette définition dans le livre *Théorie analytique des probabilités* de Laplace (1819) :

“La théorie des hasards consiste à réduire tous les événements du même genre à un certain nombre de cas également possibles et à déterminer le nombre de cas favorables à l'événement dont on cherche la probabilité. Le rapport de ce nombre à celui de tous les cas possibles est la mesure de cette probabilité, qui n'est ainsi qu'une fraction dont le numérateur est le nombre de cas favorables, et dont le dénominateur est le nombre de tous les cas possibles.”

²²On retrouve ici la plausibilité du postulat fondamental de la Mécanique statistique (*pour l'ensemble microcanonique, tous les états de même énergie ont la même probabilité*).

L'équiprobabilité de tous les cas possibles est intimement liée à celle d'entropie, qu'il s'agisse de l'entropie de Boltzmann, ou celle de Shannon et Wiener en Théorie de l'information (pour une introduction éclairante sur ce sujet, voir l'ouvrage de Renyi [18]).

en principe leurs valeurs dans \mathbb{R} (le cas discret est à la réflexion plutôt exceptionnel, sauf dans le domaine quantique évidemment). Le physicien peut par exemple se poser la question : quelle est la probabilité pour que la vitesse d'un atome de gaz soit inférieure à une valeur v_0 choisie à l'avance ? Ou : quelle est la probabilité pour que l'activité d'une source radioactive soit supérieure à un seuil de tolérance donné²³ ? Dans la vie de tous les jours, les variables réputées aléatoires sont souvent des variables prenant des valeurs continues : quelle est la probabilité pour que la taille d'un individu soit inférieure à une valeur donnée ?

L'un des plus célèbres paradoxes est dû à Bertrand (1888), qui a montré qu'une même (?) question – “Soit un triangle équilatéral et son cercle circonscrit. On tire une corde au hasard. Quelle est la probabilité que sa longueur soit supérieure²⁴ à celle d'un côté du triangle” ? – semble pouvoir recevoir plusieurs réponses différentes²⁵.

Une autre grave difficulté surgit à propos des événements impossibles qui, dans un espace Ω discret ont, par leur définition, une probabilité égale à zéro (puisque la somme des probabilités de tous les événements possibles vaut 1, et qu'un événement impossible est le contraire de l'événement certain), et réciproquement. Cette assimilation équivalente n'est pas acceptable dans le cas continu. Par exemple, selon la définition intuitive, la probabilité de trouver un réel compris entre 0 et 1 dans l'intervalle $[0, 1]$ est visiblement nulle. De façon plus imagée : s'agissant d'atteindre un objectif²⁶ de surface D sur une cible de rayon R , on a envie de définir intuitivement la probabilité comme le rapport $\frac{D}{\pi R^2}$; mais comme la surface d'un point est nulle, la définition produit l'absurdité : tous les événements élémentaires sont impossibles !

L'intégrale de Lebesgue permet de résoudre cette difficulté, en conduisant à une définition rigoureuse de la probabilité telle qu'un événement de probabilité nulle n'est pas forcément impossible, éclairant la nature précise de la relation logique entre impossibilité et probabilité nulle :

$$\boxed{\text{Événement impossible} \implies \text{Probabilité nulle}} \quad (1.9)$$

la réciproque étant fautive en général. L'assimilation de bon sens, pour une variable discrète, entre impossibilité et probabilité nulle ne tient plus dans le cas d'une variable continue puisqu'alors aucune valeur ne serait possible. Ce paradoxe se résout en arrivant à énoncer une définition de la probabilité telle que la probabilité qu'une v.a. continue X prenne une valeur x appartenant à un ensemble non-dénombrable (un intervalle par exemple) n'est pas l'addition des probabilités des valeurs individuelles. En effet, la mesure

²³Quoique dans ce cas, précisément, la notion de continuité au sens strict est douteuse : l'activité est le nombre de désintégrations par seconde, qui est forcément un nombre entier...

²⁴Cette façon astucieuse de poser le problème évite un piège majeur (voir ci-dessous), et préfigure la bonne formulation : on ne s'intéresse pas à la coïncidence (égalité) entre un résultat et une valeur prescrite, mais on se demande si le résultat est plus grand qu'une certaine valeur (inégalité). Cette différence fondamentale est inhérente à définition de la fonction de répartition introduite plus loin (section 1.5).

²⁵En réalité, il n'y a pas de paradoxe, car les trois raisonnements donnant trois réponses différentes correspondent en fait à trois protocoles expérimentaux différents. Pour une discussion approfondie, voir par exemple <http://www-ensps.u-strasbg.fr/enseignants/harthong/Hist/BERTRAND.HTM>

²⁶Cette formulation anodine est d'ailleurs en soi très ambiguë : s'agit-il de jouer aux fléchettes ou d'envoyer des particules α sur une (mince) feuille d'or ?

de Lebesgue μ d'un singleton²⁷ étant nulle – $\mu(\{x\}) = 0$ –, on ne peut visiblement pas écrire que la mesure d'un intervalle $A = [a, b]$ est $\lambda(A) = \sum_{x \in A} \mu(\{x\})$ – et heureusement. Au moins deux raisons à cela ; la première est que la seule valeur “raisonnable” pour une telle somme est zéro puisque chacun de ses termes est nul, ce qui donnerait une mesure nulle pour A (alors qu'elle vaut $b - a$). La seconde est que l'additivité d'une mesure est définie au plus pour les réunions *dénombrables*²⁸.

De cette définition, il résulte notamment que la probabilité qu'un nombre x tiré au hasard dans $[0, 1] \subset \mathbb{R}$ soit rationnel est nulle, et ceci bien que les rationnels inclus dans $[0, 1]$ constituent un ensemble dense dans cet intervalle. En effet, l'intégrale de Lebesgue de la fonction de Dirichlet²⁹ $\mathbf{D}(x)$ est bien définie – elle est nulle. La probabilité est nulle que le nombre trouvé soit rationnel, mais il n'est pas impossible pour un nombre d'être rationnel : il ne fait ainsi aucun doute que la relation (1.9) n'est pas à double sens. Par ailleurs, la difficulté est bien levée puisque l'ensemble des rationnels n'est pas l'espace Ω : malgré leur densité dans \mathbb{R} , les rationnels forment un ensemble de mesure nulle³⁰. En définitive, alors que pour une v.a. discrète on peut dire qu'un événement de probabilité nulle est impossible, il n'en va pas de même pour une variable continue puisqu'alors aucun événement ne serait possible.

Ces difficultés étant sommairement mentionnées, on s'en tiendra pour l'essentiel dans la suite, et s'agissant de la dimension conceptuelle, à la définition intuitive présentée ci-dessus et aux généralisations qu'elle autorise, tant pour les variables discrètes que pour les variables (réputées) continues. Au moment opportun, on reviendra brièvement sur des situations quelque peu exotiques (*escalier du diable*³¹), mentionnant d'ailleurs des situations physiques où des *monstres*³² interviennent, mais tout en gardant à l'esprit la distinction essentielle entre la notion de limite telle que le mathématicien la définit sans ambiguïté, et celle que conçoit le physicien, butant tôt ou tard sur la notion d'échelle physique (voir l'Introduction du cours). La restriction délibérée d'inspiration pragmatique est légitimée parce que, reposant fondamentalement sur la notion de fréquence statistique d'observation, elle procède d'une attitude pas vraiment incompatible avec le fait que la Physique est d'abord une science expérimentale. S'il est certain que, comme le dit Feller ([17], p. 11) “*Problems are not solved by ignoring them*”, on peut néanmoins maintenir cette attitude dans le cadre d'un enseignement de mathématiques à l'usage des physiciens.

On peut enfin ajouter que si la notion de variable continue est intrinsèque à la Physique – jusqu'à présent, on admet que les variables d'espace et de temps le sont,

²⁷L'ensemble formé par x tout seul, noté usuellement $\{x\}$, est appelé *singleton*.

²⁸Le problème discret *vs* continu n'est pas sans rapport avec les discussions autour du paradoxe de Zenon.

²⁹La fonction de Dirichlet $\mathbf{D}(x)$ est définie comme valant 1 si x est rationnel, 0 autrement.

³⁰Un tel ensemble est parfois dit *négligeable*, terminologie dénuée de toute connotation péjorative.

³¹En anglais : *devil's staircase*.

³²Jean Perrin emploie ce terme à propos des trajectoires browniennes dont il avait eu l'intuition géniale, suivant au microscope optique la marche erratique des grains de pollen, d'y reconnaître certains *monstres* inventés par les mathématiciens (justement, les fonctions partout continues mais nulle part dérivables. . .)

Il est bien clair que l'intuition de Perrin est un exemple typique de *passage à la limite* au sens du physicien : en tant qu'ardent défenseur de l'atomisme, il savait bien que la dissection de la trajectoire ne saurait s'approcher trop près de l'échelle d'une longueur atomique !

pour ne citer qu'un exemple –, toute vérification expérimentale d'une loi repose sur la définition d'un intervalle caractérisant la résolution de l'expérience. De ce point de vue, une grandeur continue est *de facto* assimilée à une variable discrète, et les difficultés conceptuelles évoquées ci-dessus disparaissent d'elles-mêmes ; pour en revenir à l'exemple de la cible citée ci-dessus, l'existence d'une résolution linéaire finie δ entraîne forcément que le plus petit "point" observable a en fait une surface $\propto \delta^2$, et la définition intuitive contrainte par la résolution finie inévitable fournit une réponse et une seule, parfaitement sensée. Que cette réponse dépende du protocole expérimental (la définition de la résolution δ) ne saurait ici choquer : tirer sur une plaque avec des billes ou envoyer des particules α sur une feuille d'or sont deux expériences fort différentes !

Pour en terminer avec cette introduction intuitive, il convient de rappeler le lien historique étroit entre la Théorie des probabilités et l'analyse combinatoire, elle-même liée à la Théorie des jeux. Ces deux disciplines ont joué un rôle moteur dans l'édification de la Théorie des probabilités. Fondamentalement, le lien est précisément la définition (circulaire !) énoncée plus haut, la résolution de chaque problème consistant à dénombrer les différents cas pour finalement retenir la probabilité cherchée comme le rapport du nombre de cas favorables au nombre de cas possibles, ces derniers étant réputés également probables. Dans cet exposé élémentaire, l'aspect combinatoire sera délaissé, au profit de l'introduction d'instruments d'usage courant en Physique, étant admis que les bases des techniques élémentaires de dénombrement sont connues.

Notons enfin que si la présence d'aléatoire exige le recours à un formalisme probabiliste, la réciproque est fautive : la Théorie des probabilités est un outil puissant pour l'étude de certains modèles *purement déterministes* (où, suivant la définition usuelle, tout élément aléatoire est exclu) ; c'est le cas notamment des paradigmes du chaos déterministe³³.

1.3 Axiomes. Premières conséquences

Les considérations précédentes permettent de construire un cadre théorique où, sans vouloir formaliser à l'extrême, il est bon de rassembler un nombre minimum de propositions non-contradictaires définissant un cadre cohérent (axiomatisation élémentaire), et débouchant sur des instruments de travail indispensables, dont le plus important est ce que l'on appelle la *fonction de répartition*³⁴ d'une variable aléatoire, généralement notée F , définie précisément dans la section 1.5.

³³À ce sujet, voir l'article de Sinaï dans l'ouvrage collectif [20], "*L'aléatoire du non-aléatoire*", où se révèle la distance qui sépare prédictibilité et déterminisme. Les derniers mots de Sinaï sont :

"La conclusion immédiate de tout ceci est l'absence de nécessité d'un mécanisme aléatoire extérieur pour l'applicabilité des lois de la théorie des probabilités. [...] L'instabilité d'une dynamique non-aléatoire conduit à l'apparition de l'aléatoire"

³⁴On dit aussi *loi de répartition*.

1.3.1 Axiomes

Il s'agit ici de rationaliser les constatations intuitives précédentes pour énoncer finalement les axiomes sur lesquels est construite la version élémentaire de Théorie des probabilités exposée dans ce chapitre. Les affirmations suivantes, constituant les axiomes³⁵ de Kolmogorov (1933), sont énoncées à propos de l'espace des épreuves Ω , formant l'ensemble des événements élémentaires. Elles permettent de trouver la probabilité de tout événement composite.

1. toute probabilité est un nombre compris entre 0 et 1 :

$$\boxed{0 \leq p(\omega_n) \leq 1 \quad \forall \omega_n \in \Omega} \quad (1.10)$$

2. la probabilité de l'événement certain est égale à 1, ce que l'on peut noter³⁶ :

$$\boxed{p(\Omega) = 1 \iff \sum_{n \in \mathbb{N}} p_n = 1} \quad (1.11)$$

Cette relation est le pendant de (1.5).

3. la probabilité de l'événement consistant en la réalisation de l'un ou l'autre de deux événements exclusifs est la somme de leurs probabilités :

$$\boxed{\text{Prob}[\omega_n \text{ ou } \omega_{n'}] = p_n + p_{n'}} \quad (1.12)$$

Ceci est l'affirmation de principe découlant de (1.6) et (1.7). Dans la limite où l'on considère *tous* les événements ω_n de Ω , on obtient

$$\text{Prob}[\omega_1 \text{ ou } \omega_2 \text{ ou } \dots \text{ ou } \omega_n \text{ ou } \dots] ,$$

soit la somme de toutes les probabilités p_n , qui vaut 1 d'après l'axiome 2.

◆ *Remarque*

La conjonction *ou* est parfois ambiguë : c'est pourquoi on distingue (quand c'est nécessaire) le *ou exclusif*, quand les deux possibilités s'excluent mutuellement (perdre ou gagner quand on a parié sur pile), et le *ou conjonctif* quand les deux événements ne sont pas exclusifs, et qu'il convient cependant de les distinguer (avoir sa Licence ou ne pas avoir la moyenne à chaque UE). Le *ou* entre deux événements élémentaires est forcément (et implicitement) exclusif ; le *ou disjonctif* ne peut impliquer que deux événements non-élémentaires. ◆

³⁵De façon un peu plus formelle, les trois axiomes de Kolmogorov sont cités dans l'encadré p. 15.

³⁶On suppose toujours Ω fini ou dénombrable. Toutes les sommations peuvent être notées sur \mathbb{N} : le cas où Ω possède un nombre fini N_0 événements élémentaires s'obtient en posant formellement que toute probabilité $p_{m > N_0}$ est nulle.

Les axiomes de Kolmogorov

1. Axiome 1 : Pour tout événement A , $0 \leq P(A) \leq 1$
2. Axiome 2 : $P(\Omega) = 1$
(la probabilité de l'événement certain, ou d'obtenir l'un quelconque de tous les événements possibles, est égale à 1)
3. Axiome 3 : Soit une suite dénombrable d'événements incompatibles $A_1, A_2 \dots$; on a $P(\cup_{n \in \mathbb{N}^*} A_n) = \sum_{n \in \mathbb{N}^*} P(A_n)$
Cette propriété est appelée σ -additivité.

1.3.2 Premières conséquences

Des affirmations précédentes résultent immédiatement un certain nombre de règles d'usage constant, énoncées d'abord à propos des événements élémentaires ω_n :

1. si ω_n a la probabilité p_n , le contraire de ω_n a la probabilité $1 - p_n$. En effet, l'événement contraire $\tilde{\omega}_n$ de ω_n consiste à d'obtenir n'importe lequel des autres $\omega_{n'}, n \neq n'$; d'après l'axiome 3, la probabilité d'obtenir ω_1 ou ω_2 ou ... ou ω_{n-1} ou ω_{n+1} ou ... est $\sum_{n' \neq n} p_{n'}$. D'après (1.11) :

$$p_n + \sum_{n' \neq n} p_{n'} = 1 \iff \sum_{n' \neq n} p_{n'} = 1 - p_n ; \quad (1.13)$$

d'où la règle importante :

$$\boxed{\text{Prob}[\tilde{\omega}_n \equiv \text{contraire de } \omega_n] = 1 - \text{Prob}[\omega_n]} \quad (1.14)$$

Noter que pour un événement élémentaire, toute expérience donne cet événement ou son contraire.

2. tout événement impossible a une probabilité nulle³⁷. En effet, la probabilité d'obtenir ω_n est p_n , celle d'obtenir son contraire est $1 - p_n$. Un événement impossible $A_{\text{impossible}}$ est le contraire \tilde{A} de $A = (\omega_n \text{ ou son contraire } \tilde{\omega}_n)$; ces deux derniers événements sont exclusifs ; la probabilité de A est $P(A) = p_n + (1 - p_n)$ (A est en fait certain), celle de son contraire \tilde{A} est $1 - P(A) = 0$ (la probabilité est nulle de n'obtenir aucune des valeurs possibles $\{x_n\}_n$, une quasi-évidence, ou encore, un événement impossible est le contraire de l'événement certain).

En ce qui concerne tout événement quelconque A , pas forcément élémentaire, les affirmations suivantes sont vraies :

³⁷mais un événement de probabilité nulle n'est pas forcément impossible – voir p. 11.

1. la probabilité de A , $P(A)$, est un nombre compris entre 0 et 1

$$\boxed{0 \leq P(A) \leq 1 \quad \forall A} \quad (1.15)$$

2. la probabilité d'un événement certain est égale à 1.
 3. la probabilité d'un événement impossible est égale à 0. La réalisation de A et B quand A et B sont incompatibles est un événement impossible :

$$A \cap B = \emptyset \implies P(A \cap B) = 0 . \quad (1.16)$$

4. la probabilité de \tilde{A} , contraire de A , est $P(\tilde{A}) = 1 - P(A)$
 5. si $A \subseteq B$, alors $P(A) \leq P(B)$; autrement dit :

$$\boxed{(A \implies B) \iff P(A) \leq P(B)} \quad (1.17)$$

6. la probabilité de deux événements exclusifs A et B est la somme de leurs probabilités :

$$\boxed{P(A \text{ ou } B) = P(A) + P(B) \iff P(A \cup B) = P(A) + P(B)} \quad (1.18)$$

Plus généralement, si tous les A_n sont deux à deux incompatibles :

$$P(\cup_n A_n) = \sum_n P(A_n) . \quad (1.19)$$

7. Soit A et B deux événements quelconques (pas forcément exclusifs). Pour avoir soit A , soit B , il faut avoir [A ou (disjonctif) B] ou (exclusif) [A et B], ces deux derniers événements étant clairement incompatibles ; il en résulte que, quels que soient A et B , on a $P(A \cup B) + P(A \cap B) = P(A) + P(B)$, soit :

$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)} \quad (1.20)$$

(1.18) est le cas particulier de (1.20) lorsque A et B sont incompatibles, auquel cas $P(A \cap B) = 0$.

8. Jusqu'à présent, seule l'opération d'addition des probabilités a été définie. La *multiplication* des probabilités s'introduit par l'axiome de Bayes, qui énonce un résultat majeur introduisant d'ailleurs la notion de probabilité *conditionnelle*, qui est la probabilité d'un événement sachant qu'un autre est réalisé. On note ainsi $P(A|B)$ la probabilité de A sachant que B est réalisé. L'axiome de Bayes s'énonce comme suit : la probabilité pour que A et B soient réalisés est :

$$\boxed{\text{Prob} [A \text{ et } B] \equiv P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)} \quad (1.21)$$

Il est licite de lire cette formule à l'envers :

$$\boxed{P(A|B) \stackrel{\text{déf}}{=} \frac{\text{Prob} [A \text{ et } B]}{P(B)} \equiv \frac{P(A \cap B)}{P(B)}} \quad (1.22)$$

et de la considérer comme la *définition* de la probabilité conditionnelle. Dans les dernières formules, A et B jouent évidemment des rôles strictement symétriques.

Cela étant posé, il est alors possible de définir la notion capitale d'événements *indépendants* ; A et B sont dits indépendants ssi $P(A|B)$ ne dépend pas de B ou, de façon équivalente, ssi $P(B|A)$ ne dépend pas de A :

$$\begin{array}{c} A \text{ et } B \text{ indépendants} \\ \Updownarrow \\ P(A|B) \text{ indépendant de } B \iff P(B|A) \text{ indépendant de } A \end{array} \quad (1.23)$$

Dans ces conditions, la probabilité conditionnelle $P(A|B)$ ne dépend nullement de B , et n'est autre que la probabilité attachée au seul événement A , $P(A)$, et de même pour B , les rôles étant inversés. Pour des événements indépendants, on a donc $P(A|B) = P(A)$, $P(B|A) = P(B)$ et (1.20) devient :

$$\boxed{A \text{ et } B \text{ indépendants} \iff \text{Prob}[A \text{ et } B] = P(A)P(B)} \quad (1.24)$$

Cette relation joue un rôle de tout premier plan dans les applications. On peut dire que la difficulté d'un problème augmente *infiniment* lorsque les variables aléatoires ne sont pas indépendantes³⁸, un peu comme un système de particules en interaction est en un sens infiniment plus difficile à étudier qu'un ensemble des mêmes particules sans interaction mutuelle.

9. La notion de probabilité conditionnelle joue un rôle important en Théorie quantique. Elle apparaît tout naturellement lorsque l'on compare l'évolution libre d'un système entre deux instants t_0 et t à son évolution si l'on effectue une mesure d'observable entre ces deux mêmes instants. Prenons la situation la plus simple où le système est un spin $\frac{1}{2}$, \vec{J} , dont les états propres communs à (\vec{J}^2, J_z) sont notés $|\pm 1\rangle$, et lorsque l'on fait à l'instant intermédiaire t_i , une mesure de l'observable J_x , incompatible avec J_z puisque $[J_z, J_x] = i\hbar J_y \neq 0$. Dans le premier cas, l'état à l'instant t est :

$$|\Psi(t)\rangle = U(t, t_0)|\Psi(t_0)\rangle, \quad (1.25)$$

où $U(t, t_i)$ est le propagateur ; $|\Psi(t_i)\rangle$ est l'état initial, que l'on doit connaître faute de pouvoir faire une prévision sensée physiquement, puisque l'équation de Schrödinger est du premier ordre en temps. Dans ces conditions, la probabilité de trouver la valeur ε pour J_x à l'instant t_f est $P_{z,\varepsilon}(t_f) = |\langle \varepsilon | \Psi(t_f) \rangle|^2$. En introduisant la relation de fermeture, ceci s'écrit :

$$P_{z,\varepsilon}(t_f) = \left| \sum_{\varepsilon'=\pm 1} \langle \varepsilon | U(t_f, t_i) | \varepsilon' \rangle \langle \varepsilon' | \Psi(t_i) \rangle \right|^2. \quad (1.26)$$

C'est le module au carré d'une *somme* d'amplitudes, chacune de celles-ci étant associée à un chemin (ici, il n'y en a que deux).

³⁸Tout naturellement, on dit que deux v.a. sont indépendantes si elles se réfèrent à deux événements indépendants.

Dans le second cas, il faut décomposer l'évolution en deux étapes, et considérer les deux possibilités pour la réduction du paquet d'ondes provoquée par la mesure de l'observable (incompatible) J_x . Alors, la probabilité d'observer la valeur ε pour J_z à l'instant t_f est maintenant :

$$\tilde{P}_{z,\varepsilon}(t_f) = \sum_{\varepsilon'=\pm 1} P(\varepsilon, t_f | \varepsilon', t_m) P_{x,\varepsilon'}(t_m) . \quad (1.27)$$

C'est une somme de probabilités, épuisant toutes (il n'y en a que deux) les possibilités de résultat de la mesure de J_x à l'instant intermédiaire ; chaque terme est le produit de la probabilité d'obtenir ε' pour J_x à t_m , par la probabilité conditionnelle d'obtenir ε pour J_z à t sachant que l'on a trouvé ε' pour J_x à t_m .

Fondamentalement, la différence entre les deux expressions (1.26) et (1.27) est que la première est une somme d'amplitudes dont on prend le module carré, alors que la seconde, somme de probabilités, est une somme impliquant des (produits de) modules carrés d'amplitudes. On peut dire que tout le fait quantique est dans cette distinction.

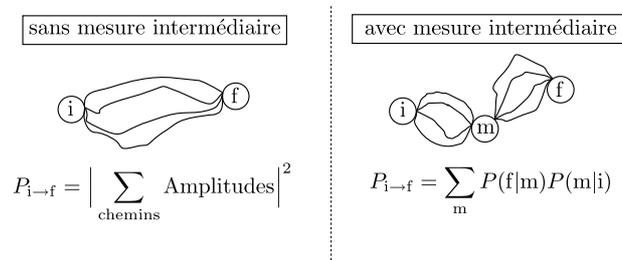


Figure 1.3: Illustration de la propagation d'un système quantique selon que l'on ne fait pas (à gauche) ou que l'on fait (à droite) une mesure intermédiaire d'observable. Toutes les probabilités s'obtiennent en effectuant la bonne somme d'amplitudes, puis en en prenant le module au carré.

1.4 Sur l'intégrale de Lebesgue

L'intégrale de Lebesgue a été brièvement mentionnée ci-dessus. Cette section a pour but d'en rappeler les idées principales en termes aussi simples que possible, étant entendu que rien ne remplace la lecture d'ouvrages consacrés d'une façon générale à la théorie de la mesure – même si leur lecture est souvent difficile pour le non-mathématicien.

À un niveau élémentaire, la théorie de l'intégration se borne à la définition de Riemann, où l'on découpe l'intervalle de définition $[a, b]$ d'une fonction $x \rightarrow f(x)$ en petits morceaux, ce qui permet de définir les sommes de Darboux $S_{D,\pm}$ et d'énoncer les conditions dans lesquelles ces deux sommes, qui encadrent un certain nombre I ,

admettent la même limite suivant un processus de limite bien défini. Lorsque c'est le cas (toujours vrai pour une fonction continue sur $[a, b]$), la limite commune s'appelle l'intégrale de f sur le segment $[a, b]$ et on note $I = \int_a^b f(x) dx$. Cette définition est susceptible de généralisations diverses (appelées *intégrales impropres*) : l'une (ou les deux) borne(s) est(sont) infinie(s), la fonction présente des discontinuités, voire diverge en certains points de l'intervalle d'intégration³⁹.

La définition de Riemann pose des difficultés théoriques de plusieurs natures. L'une d'entre elles concerne l'intégrale d'une suite de fonctions $f_n(x)$; on sait bien que si ces fonctions ont une limite $f_\infty(x)$, l'intégrale de la limite n'est pas forcément la limite de l'intégrale : pour qu'il en soit ainsi, il faut que la convergence soit *uniforme*. Par exemple⁴⁰, soit la suite $\phi_n(x) = 2^n$ si $2^{-n} \leq x \leq 2^{-(n-1)}$ et 0 ailleurs. La limite $\phi_\infty(x)$ est la fonction nulle partout, dont l'intégrale (de 0 à 1 par exemple) est nulle ; en revanche $\int_0^1 \phi_n(x) dx = 1$ quel que soit n , de sorte la limite de la suite des intégrales est égale à 1. Une autre difficulté de l'intégrale de Riemann réside dans la (possible) non-complétude d'un espace de fonctions doté d'une (semi-)norme⁴¹ définie par cette intégrale : toute suite de Cauchy n'y converge pas forcément. Par exemple, soit $x \in \mathbb{R}_+ \rightarrow f_n(x) = e^{(1-\frac{1}{n})x}$, $n \in \mathbb{N}^*$; soit par ailleurs $\|f\|_1 \stackrel{\text{déf}}{=} \int_0^{+\infty} |f(x)| dx$ la semi-norme, permettant de définir l'espace \mathcal{L}_1 des fonctions dont une telle norme est finie. La suite est bien convergente, chaque f_n a une norme finie, mais ce n'est pas le cas de f_∞ : l'espace n'est pas complet.

L'interprétation géométrique de I est évidente : c'est l'aire sous le graphe de la fonction $f(x)$; chaque somme de Darboux correspond à un encadrement du graphe par une suite de rectangles verticaux. L'idée initiale dans l'intégration de Lebesgue⁴²

³⁹Un exemple très utile en Physique est de la *partie principale de Cauchy* (qui réapparaîtra dans le chapitre 2 en tant que *distribution*, voir section 2.2.5), consistant à considérer le nombre défini par la limite suivante :

$$\lim_{\varepsilon \rightarrow 0^+} \left(\int_a^{x_0-\varepsilon} \frac{f(x)}{x} dx + \int_{x_0+\varepsilon}^b \frac{f(x)}{x} dx \right), \quad (1.28)$$

que l'on note souvent $\mathcal{P} \frac{f}{x}$, ou encore $f \frac{f(x)}{x} dx$, comme le *forte* en écriture musicale. Cette opération a une interprétation géométrique évidente : on retire *symétriquement* le point problématique, ce qui supprime deux nombres égaux en module et de signes contraires. Il n'y a aucun doute que $X - X = 0$, aussi grand que soit le nombre X . Physiquement, l'"infinitement petit" ε sera une certaine échelle physique invisible pour le problème en cours (par exemple ... le temps de Planck pour toute expérience de basse énergie !).

Cette régularisation joue un rôle important en Physique ; on la rencontre souvent en Mécanique quantique (régularisation des états propres non sommables dans \mathbb{R} – on introduit une boîte de taille L aussi grande que l'on veut et on prend au moment judicieux la limite L infinie), mais aussi en Théorie de la réponse linéaire (relations de Kramers - Kronig pour une susceptibilité, où elle exprime un principe physique immuable – celui de causalité).

En raison de la possible nécessité de devoir régulariser d'une façon ou d'une autre, il vaudrait parler *des* intégrales de Riemann, et non pas *de* l'intégrale de Riemann.

⁴⁰Dans ce cas précis, l'intégration à la Lebesgue ne fait pas mieux.

⁴¹On parle de semi-norme car deux fonctions différant en un nombre fini de points ont la même norme, alors qu'elles sont différentes, justement. La semi-norme d'une fonction non-nulle n'est pas forcément non-nulle

⁴²Lebesgue proposait l'analogie suivante : pour compter l'argent du porte-monnaie, on peut faire la somme des pièces dans l'ordre où on les prend. On peut aussi rassembler les pièces de 1 F (à l'époque, pas d'euros !), puis les pièces de 5 F, etc. La somme d'argent disponible est tout autant égale à l'addition du nombre de pièces de 1 F multiplié par 1, du nombre de pièces de 5 F multiplié par 5, etc.

consiste à découper autrement la sommation, en considérant des rectangles *horizontaux*. Pour fixer les idées, soit une fonction *positive* $f(x)$ prenant un nombre *fini* N de valeurs distinctes α_k sur des intervalles $I_k \subset \mathbb{R}$ (fonction en escalier, aussi appelée⁴³ *fonction étagée*) : $f(x) = \alpha_k$ si $x \in I_k$; supposons définie préalablement⁴⁴ la *mesure* de chaque intervalle, notée $\mu(I_k)$. L'intégration à la Lebesgue consiste à effectuer la somme :

$$\sum_{k=1}^N \alpha_k \mu(I_k) , \quad (1.29)$$

et à la définir comme l'intégrale de Lebesgue de $f(x)$, notée $\mathbf{E}(f)$, ou plus simplement $\int f$. Cela étant, la généralisation à une fonction continue (non-étagée) est possible puisque toute telle fonction peut être approchée d'autant près que l'on veut par une suite croissante de fonctions étagées.

On peut aussi dire les choses autrement. Soit $0 = \alpha_0 < \alpha_1 < \alpha_2 < \dots < \alpha_N$ une grille en ordonnées des valeurs prises par la fonction positive $f(x)$, et soit J_k l'ensemble des x tels que $f(x) < \alpha_k$. On peut alors écrire la somme dite de Lebesgue :

$$\sum_{k=0}^{N-1} (\alpha_{k+1} - \alpha_k) \mu(J_k) , \quad (1.30)$$

Si, quand $N \rightarrow +\infty$ avec $\text{Max}(\alpha_{k+1} - \alpha_k) \rightarrow 0$ la somme admet une limite finie indépendante de la grille des α_k , cette limite est l'intégrale de Lebesgue et la fonction est dite intégrable⁴⁵. Dans le cas d'une fonction de signe non fixé, elle est dite intégrable au sens de Lebesgue ssi, toujours par *définition*, $|f|$ est intégrable.

Ces définitions se transcrivent commodément en introduisant la *fonction indicatrice* d'un ensemble A , $\mathbf{1}_A(x)$, égale à 1 si $x \in A$, 0 autrement ; elle est un outil commode pour effectuer des opérations sur des ensembles, et permet même d'exporter les notations "ensemblistes" dans l'algèbre des fonctions. Par exemple, si $C = A \cap B$, on a $\mathbf{1}_C = \text{Inf}(\mathbf{1}_A, \mathbf{1}_B)$, d'où par analogie : $\mathbf{1}_C = \mathbf{1}_A \cap \mathbf{1}_B$. Pour deux fonctions $f(x)$ et $g(x)$, on peut noter $\text{Inf}(f, g) = f \cap g$; au contraire, si $C = A \cup B$, $\mathbf{1}_C = \text{Sup}(\mathbf{1}_A, \mathbf{1}_B)$ et de même $\text{Sup}(f, g) = f \cup g$. Toute fonction étagée peut alors s'écrire $f = \sum_k \alpha_k \mathbf{1}_{A_k}$, et, par définition $\mathbf{E}(f) \equiv \int f = \sum_k \alpha_k \int \mathbf{1}_{A_k}$.

Alors par définition :

$$\boxed{\mathbf{E}(\mathbf{1}_A) \equiv \int \mathbf{1}_A = \mu(A)} \quad (1.31)$$

⁴³Cette qualification est aussi employée quand l'ensemble des α_k est dénombrable.

⁴⁴D'un point de vue théorique, l'une des difficultés est précisément de définir convenablement la mesure μ , en fonction des objectifs poursuivis. C'est l'objectif atteint par la théorie de la mesure, au prix de l'introduction de concepts variés, dont la σ -additivité, la notion de tribu, etc.

La mesure de Lebesgue est une mesure parmi d'autres qui, pour deux réels $a < b$ définissant l'intervalle $I = [a, b]$, est égale à $\mu(I) = b - a$. C'est donc la longueur de l'intervalle, au sens intuitif du terme. La mesure de Lebesgue de \mathbb{Q} est nulle car les rationnels sont une réunion dénombrable de points (*singletons*), la mesure de chacun étant nulle.

⁴⁵On dit aussi *mesurable*. Revenant aux sommes de Darboux définissant l'intégrale de Riemann, il devient évident que quand les deux intégrales existent, elles sont égales.

Cela étant, on peut alors démontrer les résultats suivants :

1. si $f \geq 0$, $\mathbf{E}(f)$ est un nombre positif ou nul ou ∞ . Si f change de signe, on coupe en morceaux ; soit⁴⁶ $f_{\pm} \stackrel{\text{déf}}{=} \text{Max}(\pm f, 0)$, appelées parties positive/négative de f . Alors $\mathbf{E}(f)$ existe ssi $\mathbf{E}(f_-)$ et $\mathbf{E}(f_+)$ existent et $\mathbf{E}(f) = \mathbf{E}(f_+) - \mathbf{E}(f_-)$. Une fonction f est dite *intégrable* si $\mathbf{E}(f)$ est fini.
2. Si la suite f_n monotone croissante ($f_1 \leq f_2 \leq \dots$) a pour limite f_{∞} , alors :

$$\lim_{n \rightarrow \infty} \mathbf{E}(f_n) = \mathbf{E}(f) . \quad (1.32)$$

Autrement dit, la limite de la suite des intégrales est égale à l'intégrale de la limite de la suite des fonctions. Ce résultat s'appelle *théorème de convergence croissante de Lebesgue*.

3. En prenant la suite des sommes partielles de la série $S_N \stackrel{\text{déf}}{=} \sum_{n=1}^N f_n(x)$, ce dernier résultat s'énonce comme suit pour une série :

$$\sum_{n=1}^{\infty} \mathbf{E}(S_N) = \mathbf{E}\left(\sum_{n=1}^{\infty} S_N\right) , \quad (1.33)$$

au sens où les deux membres sont finis tous les deux, ou aucun des deux ne l'est.

4. Le théorème de convergence croissante de Lebesgue, valable pour des fonctions positives, repose sur l'hypothèse que la suite des f_n est *croissante*. On ne peut pas s'affranchir de cette hypothèse, comme le montre le contre-exemple de la suite $\phi_n(x)$ définie p. 19.

En introduisant la suite croissante $g_N = \text{Inf}_{n \geq N} f_n$, on a⁴⁷ $\lim \text{Inf} f_n = \lim g_N$, et le théorème de convergence croissante permet d'écrire $\int \lim \text{Inf} f_n = \lim \int g_N$. Quel que soit N , $\int g_N \leq \int f_n$ si $n \geq N$; il en résulte $\int g_N \leq \text{Inf}_{n \geq N} \int f_n \leq \lim \text{Inf} \int f_n$; d'où $\int \lim \text{Inf} f_n \leq \lim \int g_N \leq \lim \text{Inf} \int f_n$; c'est le lemme de Fatou, qui s'énonce : si f_n est positive et intégrable :

$$\mathbf{E}(\lim \text{Inf}(f_n)) \leq \lim \text{Inf} \mathbf{E}(f_n) , \quad (1.35)$$

En particulier, si $f_n \rightarrow f$, $\mathbf{E}(f) \leq \lim \text{Inf} \mathbf{E}(f_n)$.

5. Théorème de la convergence dominée : soit la suite f_n de fonctions intégrables convergeant *simplement* vers f_{∞} . S'il existe une fonction positive intégrable G telle que $|f_n| \leq G$ quel que soit n , alors f_{∞} est intégrable et $\mathbf{E}(f_n) \rightarrow \mathbf{E}(f_{\infty})$: on n'a plus besoin de la propriété de convergence *uniforme*.

⁴⁶Soit la fonction échelon-unité $\theta(x) = 0$ si $x > 0$, nulle si $x < 0$. Supposons que $f(x)$ est positive si $x < x_0$, négative dans le cas contraire. Alors $f_+(x) = \theta(x_0 - x)f(x)$, $f_-(x) = \theta(x - x_0)f(x)$.

⁴⁷Dans la notation introduite p. 20, on a :

$$\lim \text{Inf} f_n \equiv \bigcup_{i=1}^{+\infty} \bigcap_{j=1}^{+\infty} f_j , \quad \lim \text{Sup} f_n \equiv \bigcap_{i=1}^{+\infty} \bigcup_{j=1}^{+\infty} f_j . \quad (1.34)$$

L'hypothèse de l'existence d'une fonction G intégrable est cruciale : revenant à l'exemple de la suite ϕ_n (voir p. 19), on a $\phi_n \leq \frac{1}{x}$ sur $]0, 1]$ or justement l'ordre des limites compte : $\int \lim \neq \lim \int$; le théorème n'est tout simplement pas applicable puisque $\int_{]0, 1]} \frac{dx}{x} = +\infty$.

Pour terminer cette section au contenu forcément schématique, il est utile de comparer les deux intégrales de Riemann et de Lebesgue. Un premier résultat est que toute fonction intégrable au sens de Riemann – par exemple une fonction continue sur $[a, b]$ – l'est aussi au sens de Lebesgue (et alors les deux intégrales sont égales), mais c'est justement la réciproque qui est fautive. L'exemple classique est celui de la fonction de Dirichlet $\mathbf{D}(x)$; la mesure de \mathbb{Q} est nulle puisque l'intégrale de Lebesgue de $\mathbf{D}(x)$ est bien définie, et égale à 0. Dans le cas de Riemann, qui fait intervenir les sommes de Darboux du genre $\sum_i (x_{i+1} - x_i) f(\xi_i)$, un premier choix consiste à prendre à chaque fois $\xi_i \in \mathbb{Q}$ (possible puisque \mathbb{Q} est dense dans \mathbb{R}), auquel cas l'intégrale de Riemann sur $[a, b]$ est tout simplement égale à $b - a$, puisque tous les $f(\xi_i)$ valent 1 ; si au contraire on prend tous les ξ_i dans \mathbb{R}/\mathbb{Q} , toutes les sommes de Darboux correspondantes sont nulles. Visiblement, l'intégrale de Riemann de $\mathbf{D}(x)$ n'existe pas (les sommes n'ont pas la même limite).

Un autre exemple est instructif, celui de la fonction $x \rightarrow f(x) = \frac{\sin x}{x}$, qui illustre l'importance d'avoir défini une fonction intégrable comme une fonction dont l'intégrale de la *valeur absolue* existe. $f(x)$ n'est pas intégrable au sens de Lebesgue puisqu'elle n'est pas absolument intégrable ; en revanche, elle est intégrable sur \mathbb{R} au sens de Riemann (elle vaut π), ce qui semble d'ailleurs invalider la première affirmation du paragraphe précédent. En réalité, l'intégrale de Riemann est *impropre* et n'est définie que moyennant une prescription précise de sommation, à savoir $\lim_{X \rightarrow +\infty} \int_{-X}^{+X} \frac{\sin x}{x} dx$; il est bien clair que si l'on décidait de rassembler toutes les arches positives ensemble, le résultat serait tout autre ! Tout comme pour une série numérique non absolument convergente, toute modification de la procédure précise de sommation peut conduire à un résultat différent⁴⁸.

1.5 Fonction de répartition

La fonction de répartition (et ses produits dérivés⁴⁹) est l'instrument primordial pour la description quantitative d'une variable aléatoire. Si dans les cas "ordinaires" F n'est pas à proprement parler l'outil que l'on manipule le plus souvent (quand les densités (au sens large) existent, elles sont indiscutablement plus usitées), cette fonction demeure conceptuellement l'objet fondamental de la description exhaustive de la *dispersion* des valeurs inhérente à la notion de variable aléatoire, et à la mesure quantitative de ce que le sens commun perçoit comme leur fréquence d'apparition plus ou moins grande.

⁴⁸Un exemple classique en Physique : l'énergie de cohésion d'un réseau unidimensionnel d'ions de charges alternées.

⁴⁹sans jeu de mot !

En anglais, *fonction de répartition* se dit *distribution function*, ou, encore plus précisément, *cumulative distribution function*. En français, la dénomination *fonction de distribution* désigne le plus souvent la somme $F_{\text{at}} + F_{\text{cont}}$ dans la décomposition de Lebesgue (voir éq. (1.55)).

1.5.1 Fonction de répartition d'une v.a. discrète

Dans la suite de cette sous-section, X désigne une variable aléatoire⁵⁰ prenant une suite de valeurs discrètes x_n , chacune de ces dernières étant associée à la réalisation d'un événement particulier. Par exemple, pour la pièce de monnaie, si on associe la valeur -1 à face, la valeur $+1$ à pile, le résultat du lancer de la pièce constitue la variable aléatoire X , qui peut prendre les deux valeurs $x_1 = -1$ et $x_2 = +1$. L'ensemble des x_n est une image parfaite de l'espace des épreuves Ω ; désignant par \mathcal{V} l'application $\Omega \rightarrow \mathbb{R}$ définissant la v.a., on pourra noter au besoin :

$$\{x_1, x_2, \dots, x_m, \dots\} = \mathcal{V}(\Omega) . \quad (1.36)$$

Pour la commodité des raisonnements, on supposera que l'indexation des valeurs de l'aléatoire X a été choisie une fois pour toutes de sorte que :

$$x_1 < x_2 < \dots < x_m < \dots . \quad (1.37)$$

La fonction de répartition, notée $F(x)$, est par définition la fonction dont la valeur en x est la probabilité pour que X prenne une valeur inférieure ou égale à x . Très précisément, on définit $F(x)$ comme suit⁵¹ :

$$\boxed{F(x) \stackrel{\text{déf}}{=} \text{Prob}[X \leq x]} \quad (1.38)$$

De cette définition, résultent immédiatement les propriétés suivantes :

1. $F(x)$ est une fonction positive ou nulle (c'est une probabilité).
2. La probabilité est nulle que X prenne une valeur inférieure à $-\infty$:

$$\boxed{F(-\infty) = 0} \quad (1.39)$$

3. La probabilité que X prenne une valeur inférieure à $+\infty$ est égale à 1 :

$$\boxed{F(+\infty) = 1} \quad (1.40)$$

4. $F(x)$ est une fonction croissante au sens large⁵² puisque, si $a < b$, $X \leq a \implies X < b$; d'après (1.17), il en résulte :

$$\boxed{F(a) \leq F(b) \quad (a < b)} \quad (1.41)$$

⁵⁰ X est ici généralement une variable sans dimension dont les valeurs x_n sont des nombres purs. En pratique, X est le rapport entre une grandeur physique L et une échelle de référence l_0 faisant office d'unité.

⁵¹Attention ! Dans certains ouvrages on définit $F(x)$ comme $F(x) \stackrel{\text{déf}}{=} \text{Prob}[X < x]$.

⁵²On dit aussi *non-décroissante* [14], comme en traduction littérale de l'anglicisme *non-decreasing*.

5. La fonction $F(x)$ est partout continue à droite, c'est-à-dire que :

$$\boxed{\forall \varepsilon > 0, \lim_{\varepsilon \rightarrow 0} F(x + \varepsilon) = F(x) \iff F(x + 0) = F(x) \quad \forall x} \quad (1.42)$$

Cette propriété résulte de la définition même de la fonction $F(x)$:

$$F(x + \varepsilon) = \text{Prob}[X \leq x + \varepsilon] \implies \lim_{\varepsilon \rightarrow 0} \text{Prob}[X \leq x + \varepsilon] = \text{Prob}[X \leq x] \stackrel{\text{d\'ef}}{=} F(x) . \quad (1.43)$$

Ces propriétés sont *caractéristiques*, au sens où toute fonction $F(x)$ donnée les possédant toutes peut être considérée comme la fonction de répartition d'une variable aléatoire.

Il est maintenant possible d'établir l'important résultat suivant. Soit deux nombres a et b tels que $a < b$; la probabilité $\text{Prob}[X \leq b]$ peut se décomposer en la somme de deux événements exclusifs : (X est inférieur ou égal à a) ou (X est strictement supérieur à a et inférieur ou égal à b). En terme d'équation :

$$\text{Prob}[X \leq b] = \text{Prob}[X \leq a] + \text{Prob}[a < X \leq b] , \quad (1.44)$$

et en vertu de la définition de F , ceci se transcrit comme suit :

$$F(b) = F(a) + \text{Prob}[a < X \leq b] , \quad (1.45)$$

soit :

$$\boxed{\text{Prob}[a < X \leq b] = F(b) - F(a)} \quad (1.46)$$

La probabilité étant une quantité positive ou nulle, cette dernière équation redit bien que la fonction $F(x)$ est croissante au sens large.

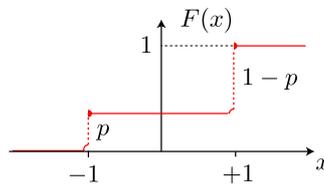


Figure 1.4: Fonction de répartition $F(x)$ pour le jeu de pile ou face, quand on est convenu des valeurs ± 1 pour les réalisations de pile et de face ; pour une pièce parfaite $p = \frac{1}{2}$. Le quart de rond à gauche de chaque saut rappelle que $F(x)$ est définie en chaque saut par sa limite à droite (continuité à droite).

Dans le cas de la pièce avec les deux valeurs conventionnelles choisies ± 1 , à quoi ressemble la fonction de répartition ? Tant que $x < -1$, plus petite valeur possible de X , F est nulle. Dès que x franchit la valeur -1 , F est égale à la probabilité associée à face, soit p (p vaut $\frac{1}{2}$ pour une pièce parfaite), puisque $\text{Prob}[-1 < X \leq a]$ avec $a < +1$ (la

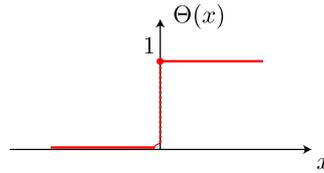


Figure 1.5: Graphe de la fonction $\Theta(x)$ définie en (1.50). Cette fonction est partout continue à droite.

valeur attachée à pile) est la probabilité d’obtenir face. Tant que x , tout en augmentant, reste inférieur à $+1$, la probabilité ne change pas, F reste donc constante et égale à p (traduisant le fait qu’il n’y a pas d’événement “intermédiaire” entre pile et face). Enfin, quand $x \geq +1$, $F(x)$ est la probabilité d’avoir ou pile ou face, et vaut donc 1.

En définitive, F est ici une fonction en *escalier* (étagée), variant par saut à chaque fois que la variable x franchit l’une des valeurs possibles pour l’aléatoire X (c’est cette fonction F que les probabilistes appellent *histogramme*, contrairement au langage du *vulgum pecus*). À chaque valeur, le saut est égal à la probabilité de l’événement correspondant (voir fig. 1.4). On doit d’ores et déjà noter que la *dérivée* de la fonction F est nulle presque partout, puisque F est une fonction constante par morceaux⁵³.

C’est le bon moment pour examiner précisément F au voisinage d’un saut, en illustration de la propriété 5 ci-dessus. Considérons le premier saut en $x_1 = -1$; par définition de F :

$$x < -1, F(x) = \text{Prob}[X \leq x < -1] = 0, \quad (1.47)$$

$$x = -1, F(x) = \text{Prob}[X \leq x = -1] = p, \quad (1.48)$$

d’où $F(-1 - 0) = 0$, $F(-1) = p$ et $F(x) = p$ si $-1 < x < +1$, d’où $F(-1 + 0) = F(-1)$, exprimant la continuité à droite. Il en va de même près de $x = +1$. Finalement, pour la pièce de monnaie et le jeu de pile ou face :

$$F(x) = \begin{cases} 0 & x < -1 \\ p & -1 \leq x < +1 \\ 1 & x \geq +1 \end{cases}. \quad (1.49)$$

Introduisant la fonction échelon $\Theta(x)$ précisément définie⁵⁴ comme :

$$\Theta(x) \stackrel{\text{déf}}{=} \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}, \quad (1.50)$$

la fonction (1.49) s’écrit :

$$\boxed{F(x) = p\Theta(x + 1) + (1 - p)\Theta(x - 1)} \quad (1.51)$$

⁵³Bien sûr, le physicien écrira $F'(x) = p\delta(x + 1) + (1 - p)\delta(x - 1)$: pour lui, la dérivée de F est ici un peigne de Dirac à deux dents ; on reviendra par la suite sur cette audace, au demeurant fort légitime.

⁵⁴ Θ diffère d’un détail de la fonction de Heaviside définie dans le contexte de la transformation de Laplace (voir chapitre 9) : Θ est définie en $x = 0$ et vaut 1 (voir fig. 1.5).

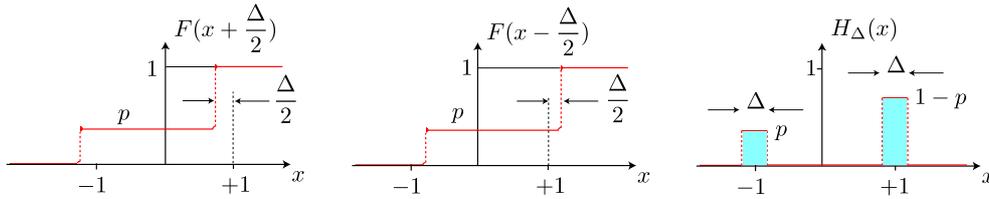


Figure 1.6: Construction (à gauche) de l’“histogramme” (à droite) donnant la probabilité de trouver une variable binaire dans l’intervalle $]x - \frac{\Delta}{2}, x + \frac{\Delta}{2}]$. Chaque bâton a une largeur Δ .

Ces éléments étant acquis, on peut par exemple trouver la probabilité pour que X soit compris dans un intervalle de mesure donnée Δ autour d’une valeur x quelconque. Plus précisément, compte tenu de (1.46) :

$$\text{Prob}[x - \frac{\Delta}{2} < X \leq x + \frac{\Delta}{2}] = F(x + \frac{\Delta}{2}) - F(x - \frac{\Delta}{2}) \equiv H_{\Delta}(x) . \quad (1.52)$$

Le graphe de $H_{\Delta}(x)$ se déduit aisément de celui de $F(x)$, en décalant chaque tracé de $\frac{\Delta}{2}$ vers la droite, qui donne $F(x - \frac{\Delta}{2})$, vers la gauche, qui donne $F(x + \frac{\Delta}{2})$ et en faisant la différence pour fabriquer $H_{\Delta}(x)$ conformément à (1.52) (voir fig. 1.6). C’est ce type de dessin que l’on appelle “histogramme” dans le langage courant, en contradiction avec la terminologie des experts en Théorie des probabilités. $H_{\Delta}(x)$ est la différence de deux fonctions continues à droite, c’est aussi une fonction continue à droite.

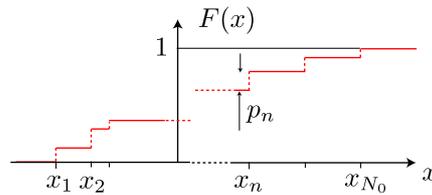


Figure 1.7: Fonction de répartition $F(x)$ pour une aléatoire discrète à N_0 valeurs (voir (1.53)).

Pour une variable aléatoire X prenant N_0 valeurs distinctes ordonnées conformément à (1.37), et telles que $\text{Prob}[X = x_n] = p_n$, la fonction de répartition $F(x)$ présente un saut à chaque fois que son argument x franchit une nouvelle valeur possible x_n . Chaque saut a une amplitude égale à la probabilité p_n de l’événement où X prend la valeur x_n ; la variation totale de F entre $\pm\infty$ est égale à la somme des sauts, soit $p_1 + p_2 + \dots + p_{N_0}$, qui est égale à 1 comme il se doit. À nouveau, la dérivée de F est nulle presque partout⁵⁵, et F est l’escalier schématisé sur la fig. 1.7. L’expression de F

⁵⁵et, *one more time*, le physicien écrira $F'(x) = \sum_n p_n \delta(x - x_n)$.

est dans ce cas :

$$F(x) = \sum_{n=1}^{N_0} p_n \Theta(x - x_n) \quad (1.53)$$

en généralisation immédiate de (1.51) (voir fig. 1.7).

1.5.2 Cas général

Jusqu'à présent, toutes les variables aléatoires considérées ont été supposées *discrètes*, au sens où le nombre d'événements élémentaires est soit fini, soit dénombrable, permettant d'imager Ω par un ensemble $\{x_1, x_2, \dots\}$ de cardinal fini ou ayant la puissance de \mathbb{N} .

Comme très sommairement expliqué dans la sous-section 1.2.2, la considération de variables aléatoires *continues* pose de grandes difficultés de fond, qu'il s'agisse de problèmes *apparemment* bien posés pouvant recevoir différentes réponses – en raison du *flou* concernant la définition ce que l'on appelle un cas favorable et/ou un cas possible (exemple, le *Paradoxe de Bertrand*, voir p. 11) –, ou de situations où la probabilité nulle est manifestement l'apanage d'événements qui ne sont manifestement pas impossibles (voir éq. (1.9)).

Quoi qu'il en soit, la fonction F reste l'outil de référence, et se prête immédiatement à la généralisation aux variables continues, permettant d'ailleurs tout naturellement la définition précise des cas exotiques où l'intuition première se trouve quelque peu malmenée (voir ci-dessous, décomposition de Lebesgue, (1.55)). En effet, rien ne s'oppose à ce que la définition (1.38) soit recopiée à l'identique pour une variable aléatoire X dont les valeurs possibles sont continues :

$$F(x) \stackrel{\text{déf}}{=} \text{Prob}[X \leq x] , \quad (1.54)$$

et ce fait même démontre la puissance de l'outil que constitue la fonction de répartition, par contraste avec les définitions plus élémentaires (où l'on ne parle d'emblée que de *densité de probabilité*). Cela étant, la définition étant strictement la même que pour une variable discrète, toutes les propriétés 1 à 5 de F énoncées p. 23 à propos d'une variable aléatoire discrète restent valides. Le point qualitatif qui de prime abord saute aux yeux est que maintenant F ne varie plus seulement par morceaux : les valeurs possibles de l'aléatoire X étant réparties de façon dense par intervalles disjoints ou non, $F(x)$ augmente graduellement. Schématiquement, F n'est plus exclusivement réduite à une fonction en escalier constante par morceaux, et peut monter continûment entre deux marches, voire ne pas présenter de marche du tout (voir figure 1.8). Cette représentation intuitive est certes plausible, mais masque une éventualité qu'il n'est pas possible de traduire par un tracé de courbe... en raison de l'épaisseur du trait.

En effet, un résultat majeur dû à Lebesgue décrit les propriétés de la fonction de répartition F pour une variable aléatoire *quelconque*, dont les valeurs possibles sont

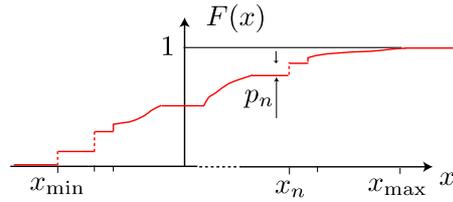


Figure 1.8: Fonction de répartition $F(x)$ pour une aléatoire prenant ses valeurs dans $[x_{\min}, x_{\max}] \subset \mathbb{R}$, dont les probabilités varient par saut pour certaines valeurs discrètes $x_{\min}, \dots, x_n, \dots$. Le graphe de F_{cont} est la juxtaposition continue des arcs disjoints ; celui de F_{at} est (ici) un escalier avec des marches en x_n . Souvent, les deux composantes F_{at} et F_{cont} sont définies sur deux intervalles disjoints.

denses par intervalles, avec en plus la possibilité de valeurs discrètes, isolées les unes des autres, noyées dans les intervalles denses ou à l'extérieur de ceux-ci. Ce résultat peut s'énoncer comme suit ; F peut toujours se décomposer au plus en la somme de *trois* composantes, notées F_{at} , F_{cont} et $F_{\text{sing cont}}$:

$$\boxed{F = F_{\text{at}} + F_{\text{cont}} + F_{\text{sing cont}}} \quad (1.55)$$

avec les propriétés suivantes :

1. chaque composante est non-décroissante (croissante au sens large).
2. F_{at} est la *fonction des sauts*, constante en dehors des points de discontinuité de F formant un ensemble *dénombrable*. Aux points de discontinuité de F , le saut de F_{at} est égal à celui de F . En d'autres termes, F_{at} représente la partie "escalier" (étagée) de F , quand elle existe. La valeur à droite de F_{at} est égale à la somme des sauts p_n . En Physique, une composante de type F_{at} est souvent appelée⁵⁶ *distribution atomique*, par référence aux états liés d'un atome.
3. La différence $F - F_{\text{at}}$ est donc une fonction continue, puisque c'est F_{at} qui contient toute la partie discontinue de F . Elle se décompose à son tour en deux morceaux :

$$F - F_{\text{at}} = F_{\text{cont}} + F_{\text{sing cont}} \quad (1.56)$$

(a) F_{cont} est une fonction *absolument continue*⁵⁷, c'est-à-dire qu'elle peut s'écrire

⁵⁶À ce propos, le mathématicien parle parfois de *fonction de masse*.

⁵⁷Soit une réunion finie d'intervalles disjoints de $[a, b]$: $[a, b] = \cup_{k=1}^N [a_k, b_k]$. Une fonction F de $[a, b]$ dans \mathbb{R} est dite *absolument continue* sur $[a, b]$ si pour toute telle partition finie ([24], § 9.2) :

$$\forall \varepsilon > 0, \exists \delta > 0, \sum_{k=1}^N (b_k - a_k) < \delta \implies \sum_{k=1}^N |F(b_k) - F(a_k)| < \varepsilon \quad (1.57)$$

En prenant $N = 1$, on retrouve la définition de la continuité usuelle : une fonction absolument continue est continue.

sous la forme d'une intégrale de la forme :

$$F_{\text{cont}}(x) = \int_{-\infty}^x \rho(x') dx' \quad (1.58)$$

en termes plus ordinaires, la dérivée de F_{cont} existe et est égale à ρ ; F_{cont} est ainsi l'intégrale de sa dérivée et relève du théorème fondamental du calcul intégral. Revenant à (1.46) et prenant $a = x$ et $b = x + \Delta x$, on en déduit dans la limite $\Delta x \rightarrow 0$ (puisque F'_{cont} existe) :

$$\exists \rho(x) \stackrel{\text{déf}}{=} \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} [F(x + \Delta x) - F(x)] = \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} [\text{Prob}[x < X \leq x + \Delta x]] ; \quad (1.59)$$

compte tenu de (1.58), on pourra ainsi écrire en pratique :

$$dF_{\text{cont}}(x) = \rho(x) dx \quad (1.60)$$

Tout naturellement, la fonction $\rho(x)$ est appelée *densité de probabilité* associée à F – noter que si X a une dimension physique, $\rho(x)$ a la dimension inverse puisque le produit $\rho(x)dx$ est un nombre pur. La densité $\rho(x)$ est positive puisque $F_{\text{cont}}(x)$ est croissante au sens large, peut avoir des sauts de première espèce (c'est pourquoi l'égalité $F'(x) = \rho(x)$ est vraie seulement presque partout) et peut même présenter des divergences ; elle doit cependant être intégrable puisque l'on doit avoir $F_{\text{cont}}(+\infty) - F_{\text{cont}}(-\infty) \leq 1$, soit $\int_{\mathbb{R}} \rho(x) dx \leq 1$. En particulier, une divergence du genre $\rho(x) \sim |x - x_0|^{-\alpha}$ est tout à fait possible⁵⁸ à condition que $\alpha < 1$; en pareil cas, $F(x) - C^{\text{ste}} \sim |x - x_0|^{1-\alpha}$.

- (b) Enfin, $F_{\text{sing cont}}$ est la différence $F - F_{\text{at}} - F_{\text{cont}}$, souvent appelée *composante singulière continue* ; c'est une fonction croissante au sens large, continue mais pas absolument continue. Elle ne varie que sur un ensemble de mesure nulle ou, si on préfère, sa dérivée est nulle presque partout. En quelque sorte, cette composante un peu exotique est ce qui reste dans F quand on lui a retranché ce à quoi on est plus ou moins habitué. Il n'y a pas de doute que si on trouve :

$$F_{\text{at}}(x > x_{\text{max}}) + F_{\text{cont}}(x > x_{\text{max}}) < 1 , \quad (1.61)$$

alors une telle composante existe.

Il n'est pas aisé de se faire une représentation mentale intuitive d'une fonction continue non dérivable ; on peut partir de l'image d'une ligne rugueuse constituée de tout petits segments de droite ; dans un deuxième temps, on décompose chaque petit segment en un grand nombre d'autre petits segments d'orientation "désordonnée" (mais de sorte que la fonction soit toujours non décroissante), et on recommence. . . Les *escaliers du diable* que l'on rencontre en Physique à propos de certains systèmes désordonnés peuvent fournir un

⁵⁸C'est aussi la raison pour laquelle la fonction d'onde de la Mécanique quantique – en dehors de toute autre considération – ne peut avoir de singularité plus violente que $|x - x_0|^{-\alpha/2}$, $\alpha < 1$.

exemple de composante singulière continue (dans la mesure où l'on accepte de diviser sans fin). L'expression tient à l'image simplifiée suivante : à l'étape zéro, on part d'un escalier avec un certain nombre de marches ; puis on effectue une transformation d'échelle qui décompose chacune des marches elle-même en un escalier, et l'on itère à l'infini, l'intervalle en abscisse entre la première et la dernière marche étant fixé une fois pour toutes.

Cette situation est typique des systèmes possédant une invariance d'échelle, traduisant le fait qu'ils ont le même aspect quand on fait un premier *zoom*, puis un second, puis un troisième, etc. : ces systèmes sont communément appelées *fractales*, et offrent à l'œil un spectacle magnifique, étrange, voire parfois inquiétant.

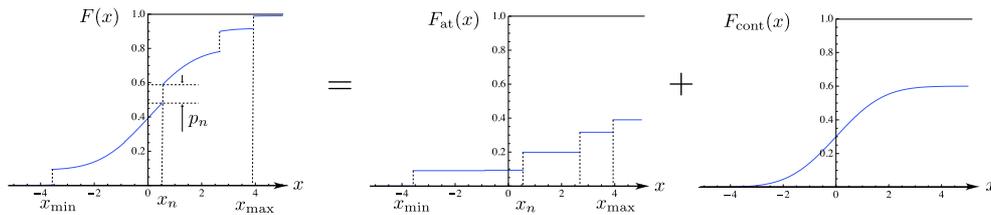


Figure 1.9: Illustration de la décomposition de Lebesgue dans le cas où il n'y a pas de composante singulière continue.

◆ *Remarques*

1. Le “théorème fondamental du calcul intégral”, énoncé pour l'intégrale de Riemann, stipule que si $f(x)$ est continue sur $[a, b]$, alors la fonction $F(x) \stackrel{\text{d'éf}}{=} \int_a^x f(x') dx'$ est dérivable et $F'(x) = f(x)$, d'où $F(x) - F(a) = \int_a^x F'(x') dx'$, autrement dit : l'intégrale de la dérivée est égale à la variation de la fonction.

En ce qui concerne l'intégrale de Lebesgue, qui d'ailleurs ne présuppose rien sur la continuité de $f(x)$, cette dernière égalité n'est pas vraie en général (même pour une fonction $F(x)$ continue et monotone). Par exemple, soit $F(x) \stackrel{\text{d'éf}}{=} x^2 \sin \frac{1}{x^2}$, dont la dérivée $F'(x) = 2x \sin \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2}$ n'est pas intégrable au sens de Lebesgue puisque $\int_0^1 |F'(x)| dx = \infty$: $F(x)$ ne saurait être la primitive d'une intégrale qui n'existe pas. Un autre exemple est la distribution de Cantor - Lebesgue $F_C(x)$ (voir p. 37), dont la dérivée $F'_C(x)$ est nulle presque partout et qui pourtant monte de 0 à 1 quand x varie de 0 à 1 : son intégrale est nulle de sorte que $1 = F_C(1) - F_C(0) \neq \int_0^1 F'_C(x) dx = 0$. La fonction $F_C(x)$ est continue mais pas absolument continue. On peut montrer qu'une fonction absolument continue est l'intégrale de Lebesgue de sa dérivée, d'où l'écriture (1.58).

2. À titre mnémotechnique, les considérations précédentes peuvent être paraphrasées en termes plus naïfs, et dans le cadre de l'intégrale de Riemann. On dit souvent que la dérivation et l'intégration sont deux opérations inverses l'une de l'autre, de

sorte que leur “produit” est l’opération identité ; sans souci de rigueur dans les écritures, on écrit ainsi, pour une fonction nulle en $x = 0$:

$$(I) : F(x) = \frac{d}{dx} \int_0^x F(x') dx' ; \quad (II) : F(x) = \int_0^x \frac{d}{dx'} F(x') dx' , \quad (1.62)$$

tendant à faire croire que la dérivée de l’intégrale (I) est toujours l’intégrale de la dérivée (II), ce qui est inexact : si (I) est vrai presque partout, (II) est sûrement faux pour une fonction dont la dérivée est presque partout nulle, comme le montre le choix $F(x) = \Theta(x - x_0)$, pour lequel on a (raisonner graphiquement) :

$$\int_0^x \Theta(x' - x_0) dx' = (x - x_0)\Theta(x - x_0) , \quad \frac{d}{dx} [(x - x_0)\Theta(x - x_0)] = \begin{cases} 0 & \text{si } x < x_0 \\ 1 & \text{si } x > x_0 \end{cases} , \quad (1.63)$$

d’où $\frac{d}{dx} \int_0^x \Theta(x' - x_0) dx' = \Theta(x - x_0)$ partout sauf en $x = x_0$. Au contraire, comme $\frac{d}{dx} \Theta(x - x_0) = 0$ partout sauf en x_0 , l’intégrale $\int_0^x \frac{d}{dx'} \Theta(x' - x_0) dx'$ est certainement nulle.

Intuitivement, la différence repose sur le fait que pour (II), on effectue la dérivation *avant* l’intégration, et on sait bien que la dérivation est une opération *dangereuse*. Au contraire, pour (I), on commence par intégrer, ce qui, le cas échéant, adoucit les anomalies (ci-dessus, l’intégration d’une fonction discontinue donne une fonction continue). Les singularités étant effacées par l’intégration, la dérivation en dernier ne saurait les faire apparaître *ex nihilo*.

3. Pour une fonction $f(x)$, il est naturel de supposer que différence $f(x + \delta x) - f(x)$ peut *a priori* être proportionnelle à δx^α , $\alpha > 0$, (à d’autres termes additifs près qui tendent vers zéro avec δx). On a ainsi, par exemple :

$$f(x + \delta x) - f(x) = (\delta x)^\alpha f^{(\alpha)}(x) + \dots ; \quad (1.64)$$

le cas usuel, mais à la réflexion assez exceptionnel, est celui où $\alpha = 1$, traduisant une propriété remarquable de la fonction : sa dérivabilité. Jean Perrin, commentant ses observations au sujet du mouvement brownien, écrit ([21], p. 166) à propos de la trajectoire d’un grain de pollen : “... , c’est un cas où il est vraiment naturel de penser à ces fonctions sans dérivée que les mathématiciens ont imaginées, et que l’on regarderait à tort comme de simple curiosités mathématiques, puisque la nature les suggère aussi bien que les fonctions à dérivée.”

Très longtemps avant B. Mandelbrot, J. Perrin citait déjà l’exemple de la représentation de la côte de Bretagne. S’agissant de définir une fonction invariante d’échelle, plusieurs choix sont *a priori* possibles (pour une autre définition, tout aussi légitime, voir (1.163)). Une définition parmi d’autres est la suivante [22]. Soit l’incrément $F_\Delta(x) \stackrel{\text{déf}}{=} f(x + \Delta) - f(x)$; la propriété d’invariante d’échelle (*self-similarity*) pour $f(x)$ se traduit par la relation :

$$F_{\lambda\Delta}(x) \rightarrow \lambda^\alpha F_\Delta(x) \text{ si } \lambda \rightarrow 0 \iff f(x + \lambda\Delta) - f(x) \rightarrow \lambda^\alpha [f(x + \Delta) - f(x)] , \quad (1.65)$$

où α est un certain exposant appelé *local scaling exponent*. Le cas $\alpha = 1$ correspond à une fonction f dérivable en x_0 . \blacklozenge

Bien évidemment, une ou deux composantes de F dans (1.55) peut(vent) être absente(s), c'est-à-dire être identiquement nulle. Dans toute la suite, mis à part l'exemple classique de Cantor (voir p. 37), on ne considèrera plus la possibilité d'une composante du type $F_{\text{sing cont}}$; pour simplifier encore les choses, on se placera le plus souvent dans deux cas :

- F n'a qu'une composante de type F_{at} : c'est en fait le cas d'une variable aléatoire discrète, qui a servi de point de départ pour la définition de la fonction de répartition (section 1.5). On parle alors de v.a. *discrète*.
- F n'a qu'une composante de type F_{cont} , c'est-à-dire que d'une part la variable aléatoire est continue, que d'autre part il existe une densité de probabilité. Compte tenu de (1.46) et de (1.58), on peut aussi écrire :

$$\text{Prob}[x < X \leq x + dx] = F(x + dx) - F(x) = F'(x)dx = \rho(x)dx . \quad (1.66)$$

Une telle v.a. est dite *absolument continue*.

En fait, le premier cas peut être considéré comme un cas particulier du second, si l'on accepte de manipuler la fonction de Dirac, en la prenant sans état d'âme⁵⁹ comme la *dérivée* de la fonction $\Theta(x)$ définie en (1.50). Alors, dans les deux situations précédentes, F admet une dérivée et tous les cas considérés dans la suite peuvent être traités en bloc, en jouant avec la densité de probabilité associée à F . Cela étant admis, on pourra dès lors toujours supposer que toute fonction de répartition sans composante singulière continue est donnée par :

$$F(x) = \sum_{n \in \mathbb{N}} p_n \Theta(x - x_n) + \int_{-\infty}^x \rho(x') dx' \equiv F_{\text{at}} + F_{\text{cont}} \quad (1.67)$$

et admet la "dérivée" $p(x) \stackrel{\text{déf}}{=} F'(x)$ donnée par⁶⁰ :

$$p(x) = \sum_{n \in \mathbb{N}} p_n \delta(x - x_n) + \rho(x) \quad (1.68)$$

En toute rigueur, avec la définition précise (1.38) d'où découle (1.60), c'est $\delta^{(+)}$ qui figure dans la composante atomique. Répétons que l'une ou l'autre de ces deux composantes peut parfaitement être identiquement nulle : si $\rho(x) = 0$, la variable X est purement discrète ; par contraste, elle est absolument continue si tous les poids p_n sont nuls. On constatera au fur et à mesure que cette méthode opératoire fonctionne bien dans tous les cas rencontrés en pratique. En particulier, en intégrant membre à membre (1.68) entre $\pm\infty$, on trouve :

$$\int_{-\infty}^{+\infty} p(x) dx = \sum_{n \in \mathbb{N}} p_n + \int_{-\infty}^{+\infty} \rho(x) dx = 1 , \quad (1.69)$$

⁵⁹Ces manipulations seront justifiées dans le chapitre 2 – voir notamment (2.55) avec $f = \delta$.

⁶⁰La sommation $n \in \mathbb{N}$ ne signifie pas qu'il y a une infinité dénombrable d'événements discrets : cette notation contient le cas où seuls des p_n en nombre fini sont non nuls.

après avoir remarqué que le premier membre est $F(+\infty) - F(-\infty) = 1$, par définition de la fonction de répartition. Il en résulte que $p(x)$ est une fonction intégrable, au sens où son intégrale entre $\pm\infty$ est nécessairement finie (et vaut 1). Pour faire court, on peut aussi écrire (1.69) sous la forme $\int_{-\infty}^{+\infty} p(x) dx = 1$. À ce propos, répétons que trouver $\int_{-\infty}^{+\infty} p(x) dx < 1$ est le symptôme de l'existence d'une composante atomique⁶¹ et/ou d'une composante singulière continue. Dans la littérature francophone, $p(x)$ s'appelle *loi de distribution*.

On rencontre souvent en Physique des v.a. "hybrides" admettant une fonction de répartition du type écrit en (1.67). C'est par exemple le cas pour l'énergie E de l'atome d'hydrogène dans un état Ψ donné qui, aux dires de la Mécanique quantique, est une variable aléatoire au sens où des mesures répétées de l'énergie de l'atome fournissent une succession de valeurs appartenant toutes à un certain *catalogue*, avec des probabilités calculables quand on connaît Ψ . Sa fonction de répartition a l'allure suivante :

$$F(E) = \underbrace{\sum_{n \in \mathbb{N}} p_n \Theta \left(E + \frac{E_1}{n^2} \right)}_{F_{\text{at}}} + \underbrace{\int_0^E \rho(E') dE'}_{F_{\text{cont}}} \quad (1.70)$$

où Θ est la fonction définie en (1.50). Les p_n et la fonction $\rho(E)$ dépendent de l'état Ψ considéré et satisfont la relation (dite de *normalisation*) $\sum_{n \in \mathbb{N}} p_n + \int_0^{+\infty} \rho(E) dE = 1$; E_1 est l'échelle typique d'énergie de la Physique atomique et vaut à peu près 13.6 eV. La partie discrète (énergies négatives, *spectre discret*) correspond aux états liés : ce sont eux qui représentent l'atome constitué (d'où, sans doute, la terminologie F_{at} pour une fonction de répartition en escalier) ; la partie continue (*spectre continu*) représente les états non-liés, décrivant quantiquement ce qui se passe quand on envoie un électron (projectile) en direction d'un proton fixe (cible). Dans le cas considéré, F contient effectivement les deux composantes, le sens physique de chacune étant parfaitement clair.

Dans toute la suite, on appellera conventionnellement variable aléatoire *ordinaire* une v.a. dont la fonction de répartition est dépourvue de composante singulière continue ; l'immense majorité des v.a. rencontrées en Physique est de cette nature. Ceci justifie de raisonner presque toujours au plus avec $p(x)$, retrouvant de ce fait le langage usuel en Physique où, il faut bien le dire, l'apparition de composante singulière est très peu fréquente.

■ *Un exemple d'escalier du diable pas vraiment diabolique*

Imaginons que la fonction de répartition $F(x)$ d'une aléatoire X , dont on sait par ailleurs qu'elle prend ses valeurs sur l'intervalle $[0, 1] \subset \mathbb{R}$, satisfait l'équation suivante⁶² :

$$F(x) = pF(\alpha x) \quad (x < 1) , \quad (1.72)$$

⁶¹Transposé en Mécanique quantique, une telle inégalité traduit la non-complétude de l'ensemble des vecteurs propres considérés. Par exemple, dans le cas du potentiel coulombien (atome d'hydrogène), il existe à la fois des états liés et des états non-liés. La somme sur les états liés ou non-liés seuls produit une inégalité de ce type.

⁶²C'est par exemple la limite – supposée exister – d'une récurrence discrète pour une v.a. dont on sait par ailleurs que sa fonction de répartition, à l'étape n , satisfait $F_{n+1}(x) = pF_n(\alpha x)$.

où p et α sont réels, avec $0 < p < 1$, et $\alpha > 1$. Avec un peu d'ingéniosité, on trouve que la solution de (1.72) satisfaisant $F(x \leq 0) = 0$ et $F(x \geq 1) = 1$ est la fonction :

$$F(x) = \sum_{k=0}^{+\infty} (1-p)p^k \Theta(x - \alpha^{-k}) \quad (1.73)$$

montrant que la variable aléatoire considérée X prend les valeurs α^{-k} , $k \in \mathbb{N}$ et que :

$$\text{Prob}[X = \alpha^{-k}] = (1-p)p^k . \quad (1.74)$$

On vérifie en outre que la somme des sauts, $(1-p)(1+p+p^2+\dots)$, est bien égale à 1.

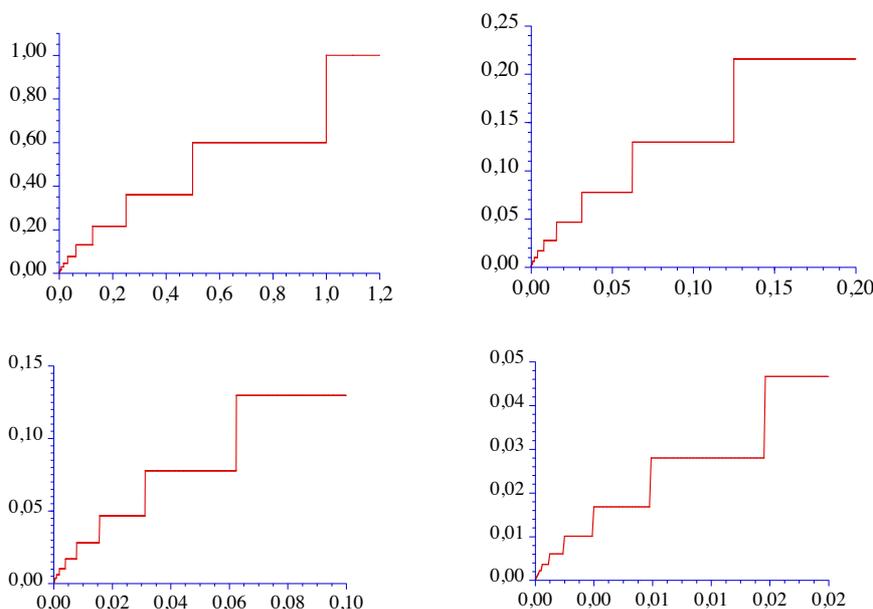


Figure 1.10: Zooms successifs sur le graphe de la fonction (1.73) (avec $q = 0.6$, $\alpha = 2$), révélant son caractère quelque peu diabolique...

Le tracé du graphe de la fonction (1.73) (voir figure 1.10) révèle toutes les curiosités de cette dernière (on ne peut pas éviter la largeur du trait de figure lié à la dimension de la pointe du crayon !) : quand on fait un *zoom*, on retrouve une figure ayant

Des situations nettement plus complexes se rencontrent dans certains systèmes désordonnés, où une variable d'Ising (de Bernoulli) est telle que l'on a :

$$F_{n+1}(x) = pF_n(\phi(x)) + (1-p)F_n(\tilde{\phi}(x)) , \quad (1.71)$$

ϕ et $\tilde{\phi}$ étant deux fonctions données. Même avec des fonctions ϕ et $\tilde{\phi}$ très simples (des applications linéaires suffisent !), la loi-limite $F(x)$ se révèle très bizarre [28] ...

le même aspect. Toutefois, le *zoom* sur une marche donnée révèle qu'elle ne se décompose pas elle-même en un escalier : l'invariance d'échelle apparente a ici ses limites. En tout cas, cet escalier n'est pas assez diabolique pour être une composante singulière continue (il est discontinu), la fonction de répartition (1.73) étant en fait du genre composante atomique, dont la densité généralisée s'écrit :

$$p(x) = (1-p) \sum_{k=0}^{+\infty} p^k \delta(x - \alpha^{-k}) . \quad (1.75)$$

La densification des points de discontinuité quand on se dirige vers $x = 0$ ne change rien au fait que la dérivée de $F(x)$ est presque partout nulle, l'ensemble des points de discontinuité restant dénombrable. $F(x)$ est un exemple de fonction croissante au sens large dont la dérivée est presque partout nulle, mais comme elle est *discontinue*, il n'y a rien de paradoxal dans la concomitance d'une fonction variable avec une dérivée nulle, et c'est finalement pourquoi il n'y a, ici, rien de vraiment diabolique.

◆ *Note*

Pour les esprits curieux, quelques indications sur les méthodes permettant de trouver la solution (1.73). Comme on sait ce que vaut F en dehors de $[0, 1]$ (0 à gauche, 1 à droite), il suffit de considérer des valeurs $x \in [0, 1]$.

Avec de la réflexion, la solution se trouve à vue, sans calculs. En effet, si x est choisi de sorte que $\alpha x > 1$, alors l'équation (1.72) s'écrit explicitement $F(x) = p \times 1$, puisque $F(\alpha x > 1) = 1$. On en déduit $F(x) = p \forall x \in [\frac{1}{\alpha}, 1]$.

Examinons maintenant $x < \frac{1}{\alpha}$; si on se restreint en fait à l'intervalle $[\frac{1}{\alpha^2}, \frac{1}{\alpha}]$, l'équation donne $F(x) = p \times p = p^2$, puisque αx étant entre $\frac{1}{\alpha}$ et 1, on sait de ce qui précède que $F(\alpha x)$ vaut alors p . Et ainsi de suite : F est la fonction en escalier montant de p^{k+1} à p^k au point $x = \alpha^{-k}$ et s'écrit donc :

$$F(x) = \sum_{k=0}^{+\infty} (p^k - p^{k+1}) \Theta(x - \alpha^{-k}) \quad (1.76)$$

C'est bien la fonction trouvée en (1.73).

Donnons d'autres façons de s'y prendre – au prix d'une anticipation –, constituant d'ailleurs un petit exercice à propos des transformées de Fourier et Laplace (voir chapitres 8 et 9).

Sans pousser le raisonnement jusqu'au bout comme ci-dessus, on remarque que, puisque $F(x \geq 1) = 1$, l'équation permet d'emblée de trouver les valeurs de $F(x)$ aux points de la forme α^{-k} , en profitant de la continuité à droite de $F(x)$. En effet, l'équation dit que :

$$F\left(\frac{1}{\alpha}\right) = p \times F(x = 1) = p , \quad F\left(\frac{1}{\alpha^2}\right) = pF\left(\frac{1}{\alpha}\right) = p^2 , \quad (1.77)$$

et ainsi de suite, suggérant que la solution doit pouvoir s'écrire $F(x) = x^\lambda$. L'idée étant admise, toute la question est de trouver l'exposant λ . Le report

de cette forme dans (1.72), on obtient :

$$1 = p\alpha^\lambda \equiv e^{\lambda \ln \alpha} , \quad (1.78)$$

dont on tire naïvement $\lambda = -\frac{\ln p}{\ln \alpha}$; ceci donne bien une sorte d'enveloppe pour $F(x)$, avec les valeurs p^k aux points $x_k = \alpha^{-k}$. En réalité, l'équation (1.78) a une *infinité* de solutions puisque tous les exposants λ_n de la forme :

$$\lambda_n = \frac{2in\pi - \ln p}{\ln \alpha} \quad (1.79)$$

conviennent tout autant. L'équation (1.72) étant linéaire, toute combinaison linéaire des solutions correspondantes est solution ; à ce stade, on peut affirmer que $F(x)$ est une série du genre :

$$F(x) = \sum_{n \in \mathbb{Z}} F_n x^{\frac{2in\pi - \ln p}{\ln \alpha}} = x^{-\frac{\ln p}{\ln \alpha}} \sum_{n \in \mathbb{Z}} F_n e^{2in\pi \frac{\ln x}{\ln \alpha}} . \quad (1.80)$$

On note que la dernière somme à droite est une série de Fourier par rapport à la variable $y = \frac{\ln x}{\ln \alpha}$, représentant une fonction 1-périodique en y , et en se souvenant que toutes les écritures doivent être comprises dans l'intervalle $[0, 1]$ pour x .

Plutôt que de chercher à obtenir directement les coefficients de Fourier F_n , montrons que la fonction obtenue d'emblée en (1.75) peut effectivement se mettre sous la forme d'un tel développement.

L'expression (1.73) peut s'écrire :

$$F(x) = p^{1+E(-\frac{\ln x}{\ln \alpha})} \quad 0 < x < 1 , \quad (1.81)$$

où $E(x)$ est la fonction partie entière⁶³ ($E(x) = n$ si $n \leq x < n+1$, $n \in \mathbb{Z}$). Maintenant, la fonction $E(y) - y$ est une fonction 1-périodique, donc toute quantité de la forme $r^{E(y)-y}$ ($0 < r < 1$) est une fonction de même période, dont la série de Fourier se trouve facilement (utiliser (8.29)) :

$$r^{E(y)-y} = -\frac{1-r}{r} \sum_{k \in \mathbb{Z}} \frac{e^{2ik\pi y}}{\ln r + 2ik\pi} . \quad (1.82)$$

Ce dernier résultat peut d'ailleurs s'obtenir directement en calculant d'abord la transformée de Laplace de $r^{E(y)}$:

$$\mathcal{L} \left[r^{E(y)} \right] (z) = \sum_{k=0}^{+\infty} \int_k^{k+1} r^k e^{-zy} dy = \frac{1}{z} \frac{1 - e^{-z}}{1 - re^{-z}} \quad (1.83)$$

puis en effectuant la transformation de Laplace inverse ; par résidus⁶⁴ on

⁶³On note parfois $[x] \equiv E(x)$.

⁶⁴ z_k est l'un des zéros de $1 - re^{-z}$.

obtient :

$$r^{E(y)} = \sum_{k=-\infty}^{+\infty} \frac{1 - e^{-z_k}}{z_k \times 1} e^{z_k y} = \sum_{k=-\infty}^{+\infty} \frac{1 - \frac{1}{r}}{\ln r + 2ik\pi} e^{(\ln r + 2ik\pi)y} = \frac{r-1}{r} r^y \sum_{k=-\infty}^{+\infty} \frac{e^{2ik\pi y}}{\ln r + 2ik\pi} \quad (1.84)$$

d'où :

$$r^{E(y)-y} = \frac{r-1}{r} \sum_{k=-\infty}^{+\infty} \frac{e^{2ik\pi y}}{\ln r + 2ik\pi} . \quad (1.85)$$

En particulier ($r \rightarrow p$, $y \rightarrow -\frac{\ln x}{\ln \alpha}$) :

$$F(x) = p^{1+E[-\frac{\ln x}{\ln \alpha}]} = p p^{-\frac{\ln x}{\ln \alpha}} \frac{p-1}{p} \sum_{k=-\infty}^{+\infty} \frac{e^{2ik\pi - \frac{\ln x}{\ln \alpha}}}{\ln p + 2ik\pi} = (p-1) p^{-\frac{\ln x}{\ln \alpha}} \sum_{k=-\infty}^{+\infty} \frac{e^{-2ik\pi \frac{\ln x}{\ln \alpha}}}{\ln p + 2ik\pi} . \quad (1.86)$$

Comme $p^{-\frac{\ln x}{\ln \alpha}} = e^{-\frac{\ln x}{\ln \alpha} \ln p} = e^{-\frac{\ln p}{\ln \alpha} \ln x} = x^{-\frac{\ln p}{\ln \alpha}}$, (1.86) s'écrit aussi :

$$F(x) = x^{-\frac{\ln p}{\ln \alpha}} \sum_{k=-\infty}^{+\infty} \frac{(p-1)}{\ln p + 2ik\pi} e^{-2ik\pi \frac{\ln x}{\ln \alpha}} = x^{-\frac{\ln p}{\ln \alpha}} \sum_{n \in \mathbb{Z}} \frac{(p-1)}{\ln p - 2in\pi} x^{\frac{2in\pi}{\ln \alpha}} . \quad (1.87)$$

Cette expression est bien de la forme (1.80), et on lit les coefficients F_n par identification. ◆ ■

■ *Un exemple d'escalier du diable vraiment diabolique : la fonction de Cantor - Lebesgue*

L'ensemble de Cantor (1883) – aussi appelé *poussière de Cantor* – est un ensemble de points dans⁶⁵ \mathbb{R} possédant des propriétés très étranges. On peut le définir de façon itérative en partant de l'intervalle $[0, 1]$; dans la première étape, on supprime l'ouvert $]\frac{1}{3}, \frac{2}{3}[$, obtenant l'ensemble $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. À la deuxième étape, on supprime le tiers médian ouvert des deux intervalles ainsi conservés, ce qui donne l'ensemble formé des quatre intervalles :

$$[0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1] ;$$

le processus est répété à l'infini et l'ensemble de Cantor \mathcal{C} est par définition l'ensemble des points conservés à toute étape de l'itération⁶⁶. \mathcal{C} ne contient aucun intervalle

⁶⁵Par un procédé analogue de suppression itérative, on peut définir des ensembles bizarres dans \mathbb{R}^2 (triangle de Cantor, carré de Cantor, *tipi* de Cantor, rideaux de Cantor, ...), dans \mathbb{R}^3 (cube de Cantor), etc.

⁶⁶La construction de \mathcal{C} peut aussi se décrire en considérant l'écriture d'un nombre en base 3.

de mesure non-nulle, et on peut même après tout se demander s'il reste seulement quelque chose au bout du compte. En fait, il faut bien réaliser que l'on supprime des intervalles ouverts, par exemple retirer $]\frac{1}{3}, \frac{2}{3}[$ laisse derrière les points $\frac{1}{3}$ et $\frac{2}{3}$. Les étapes suivantes ne suppriment pas ces points et d'autres puisque les intervalles supprimés sont toujours ouverts et à l'intérieur de ceux présents à l'étape précédente. \mathcal{C} n'est donc pas vide, et contient un ensemble infini de points.

La mesure de \mathcal{C} est nulle : à chaque étape, la mesure de l'ensemble conservé est égal aux deux tiers de celle de l'ensemble précédent ; la longueur restante est donc $\frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} \dots \infty = 0$. Une autre façon de le voir est de calculer la longueur totale retirée, qui vaut :

$$\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \dots \equiv \sum_{n \in \mathbb{N}} \frac{2^n}{3^{n+1}} = \frac{1}{3} \frac{1}{1 - \frac{2}{3}} = 1 . \quad (1.88)$$

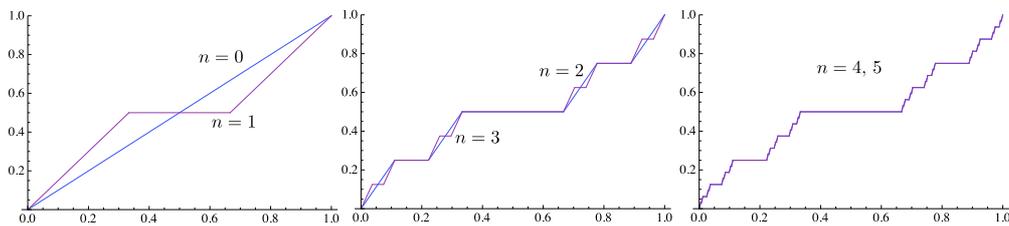


Figure 1.11: Graphes de la suite de fonctions $f_n(x)$ définie en (1.89) ; noter que dès $n \geq 4$, l'épaisseur du trait rend les courbes indiscernables les unes des autres.

La *fonction de Cantor - Lebesgue* $F_{\mathcal{C}}(x)$ peut-être définie comme la limite (le point fixe) $f_{\infty}(x)$ de la suite de fonctions f_n définie comme suit :

$$f_{n+1}(x) = \begin{cases} \frac{1}{2}f_n(3x) & \text{si } 0 \leq x \leq \frac{1}{3} \\ \frac{1}{2} & \text{si } \frac{1}{3} \leq x \leq \frac{2}{3} \\ \frac{1}{2} + \frac{1}{2}f_n(3(x - \frac{2}{3})) & \text{si } \frac{2}{3} \leq x \leq 1 \end{cases} \quad (1.89)$$

avec $f_0(x) = x$ comme "condition initiale". La fonction $F_{\mathcal{C}}(x) \stackrel{\text{def}}{=} f_{\infty}(x)$ est vraiment très bizarre, et défie l'intuition (y compris la plus chevronnée ?) : elle est partout continue, sa dérivée est nulle presque partout, et pourtant elle monte de 0 à 1 en prenant toutes les valeurs quand x parcourt $[0, 1]$. En la prolongeant à 0 pour $x < 0$ et à 1 au-delà de $x = 1$, la fonction de Cantor peut être considérée comme une fonction de répartition : c'est celle d'une v.a. uniformément distribuée sur \mathcal{C} . Elle n'a aucune composante atomique (elle est continue), et n'a pas de composante continue puisque sa dérivée est presque partout nulle ; ce dernier point est aussi évident (!?) par le fait que si une densité finie existait, l'intégration sur tout intervalle de mesure non-nulle devrait donner un nombre strictement positif, alors que $f_{\infty}(x)$ attribue à cet intervalle une probabilité nulle. Un autre exemple de tel *monstre* mathématique est la fonction *Point d'interrogation de Minkowski* (*Minkowski's question mark function*), notée souvent $?(x)$.

Il est évidemment possible de tracer les graphes des dérivées $f'_n(x)$; la figure 1.12 en montre quelques uns. On observe qu'il s'agit de pics rectangulaires, d'autant plus nombreux, fins et hauts que n est grand ; ayant tous la même hauteur, ils anticipent le fait que $F_{\mathcal{C}}(x)$ est bien la fonction de répartition d'une v.a. uniforme sur \mathcal{C} . ■

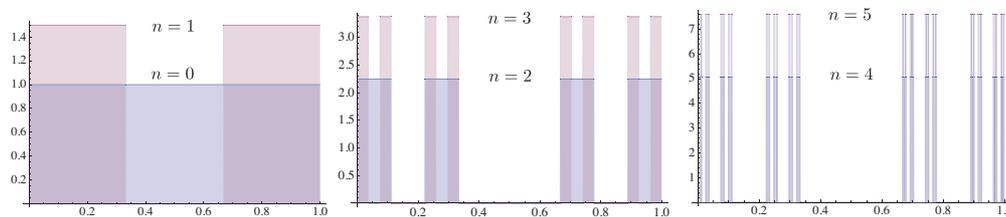


Figure 1.12: Graphes des dérivées $f'_n(x)$ de la suite définie en (1.89) ; noter l'augmentation des pics de densité avec n .

1.6 Espérances mathématiques (moyennes)

Dès qu'une description probabiliste est nécessaire (parce que les variables d'intérêt sont aléatoires), la caractérisation maximale possible est la connaissance d'une loi de répartition F ou, avec les hypothèses désormais admises, de sa dérivée $F'(x) = p(x)$ au sens large précisé ci-dessus. L'une ou l'autre de ces fonctions étant connues, il est possible de calculer *toutes* les valeurs moyennes relatives à la variable aléatoire X .

La quantité la plus naturelle pour quantifier une variable aléatoire, et aussi la plus simple, est la valeur moyenne de l'aléatoire X elle-même. Il est facile de la construire intuitivement en se disant qu'une valeur possible x_n est d'autant plus importante que sa probabilité p_n est élevée. La valeur moyenne de X , notée $\langle X \rangle$ ou $E(X)$, émerge donc naturellement, pour une variable aléatoire discrète, comme la somme pondérée :

$$E(X) \equiv \langle X \rangle = \sum_n p_n x_n \quad (1.90)$$

qui se lit aussi :

$$\langle X \rangle = \sum_n x_n \times \text{Prob}[x = x_n] . \quad (1.91)$$

Finalement, se référant à un grand nombre \mathcal{N} d'expériences et aux fréquences statistiques f_n , on voit que la moyenne n'est rien d'autre que la moyenne arithmétique des valeurs trouvées ; si on trouve \mathcal{N}_n fois la même valeur x_n , elle apparaît \mathcal{N}_n fois quand on effectue la moyenne arithmétique ; notant $x_{\text{exp}\lambda}$ la valeur obtenue lors de l'expérience λ (elle

coïncide avec l'une des valeurs possibles x_n , $x_{\text{exp}\lambda} \in \{x_n\}_n$, la moyenne arithmétique des valeurs trouvées est :

$$\frac{1}{\mathcal{N}} \sum_{\lambda=1}^{\mathcal{N}} x_{\text{exp}\lambda} = \frac{1}{\mathcal{N}} \sum_n \mathcal{N}_n x_n = \sum_n \frac{\mathcal{N}_n}{\mathcal{N}} x_n \equiv \sum_n f_n x_n . \quad (1.92)$$

La première somme court sur les résultats des \mathcal{N} expériences, les deux suivantes s'effectuent en sommant sur les valeurs différentes de la v.a. Dans la limite $\mathcal{N} \rightarrow \infty$, les f_n tendent vers les probabilités et on retrouve (1.90).

Pour une aléatoire absolument continue de densité $\rho(x)$, la moyenne se construit suivant le même argument ; on écrit :

$$\langle X \rangle = \int_{-\infty}^{+\infty} x \times \text{Prob}[x < X \leq x + dx] , \quad (1.93)$$

soit, d'après (1.66) :

$$\boxed{\langle X \rangle = \int_{-\infty}^{+\infty} x \rho(x) dx} \quad (1.94)$$

C'est à nouveau une somme pondérée : on affecte à chaque valeur x un poids égal à la probabilité élémentaire de trouver X entre x et $x + dx$, puis on fait la somme sur toutes les valeurs possibles.

À la réflexion, on voit que la moyenne dans le cas discret, (1.90), peut aussi s'écrire sous forme intégrale ; en effet, en raison de la règle opérationnelle de la fonction de Dirac :

$$\langle X \rangle = \int_{-\infty}^{+\infty} x \sum_{n \in \mathbb{N}} p_n \delta(x - x_n) dx = \sum_{n \in \mathbb{N}} p_n x_n . \quad (1.95)$$

En définitive, pour une aléatoire ordinaire décrite par une densité du type (1.68), la valeur moyenne est donnée par :

$$\langle X \rangle = \int_{-\infty}^{+\infty} x p(x) dx . \quad (1.96)$$

La moyenne $\langle X \rangle$ s'appelle aussi *espérance mathématique* ; bien évidemment, X et $\langle X \rangle$ ont la même dimension physique.

◆ *Remarque*

La moyenne $\langle X \rangle$, supposée exister, peut aussi se calculer avec la fonction de répartition F . À partir de (1.96), utilisant le fait que $p(x) = F'(x)$ et intégrant par parties, il vient :

$$\langle X \rangle = \lim_{A, B \rightarrow +\infty} \left[[xF(x)]_{-B}^A - \int_{-B}^A F(x) dx \right] . \quad (1.97)$$

Les deux contributions divergent puisque $F(x) \rightarrow 1$ si $x \rightarrow +\infty$. Dans l'hypothèse où la moyenne $\langle X \rangle$ est finie, les deux divergences doivent se compenser. Par exemple, si $F(x)$ se rapproche de sa valeur limite 0 quand $x \rightarrow -\infty$ plus vite que $\frac{1}{x}$, la limite $B \rightarrow +\infty$ est finie puisque le terme tout intégré est nul et que l'intégrale est bornée par une intégrale convergente. Du côté $x \rightarrow +\infty$, supposons pour fixer les idées que $F(x) \simeq 1 - \frac{C}{x^\alpha}$. Comme la moyenne est supposée exister, et avec $x F' \sim \frac{\alpha C}{x^\alpha}$, il faut $\alpha > 1$; cela étant, l'intégrale $\int_{-\infty}^A F(x) dx$ se comporte comme $A + \mathcal{O}(A^{-(\alpha-1)})$ quand $A \rightarrow +\infty$: la divergence $\propto A$ se compense bien avec la contribution de la borne supérieure du terme tout intégré. Quoi qu'il en soit, on devine que, sauf cas exceptionnel, la formule (1.96) est plus efficace que (1.97). ♦

Ayant identifié le mode opératoire pour calculer la valeur moyenne de X , on peut maintenant trouver la valeur moyenne de n'importe quelle fonction de X , $f(X)$, définie comme la fonction prenant les valeurs $f(x)$; pour une v.a. discrète :

$$\langle f(X) \rangle = \sum_{n \in \mathbb{N}} p_n f(x_n) \quad (1.98)$$

Un exemple très important est celui où $x \rightarrow f(x) = x^2$; la moyenne du carré de X dans le cas discret est :

$$\langle X^2 \rangle = \sum_{n \in \mathbb{N}} p_n x_n^2 . \quad (1.99)$$

À nouveau, dans le cas général d'une densité du type (1.68), la moyenne s'obtient comme :

$$\langle X^2 \rangle = \int_{-\infty}^{+\infty} x^2 p(x) dx . \quad (1.100)$$

Il faut noter dès maintenant que $\langle X^2 \rangle \geq \langle X \rangle^2$. En effet, calculons la valeur moyenne de la combinaison⁶⁷ $(X - \langle X \rangle)^2$. On a :

$$\langle (X - \langle X \rangle)^2 \rangle = \langle X^2 \rangle - 2\langle X \rangle \langle X \rangle + \langle X \rangle^2 = \langle X^2 \rangle - \langle X \rangle^2 . \quad (1.101)$$

Le premier membre est une quantité positive ou nulle : c'est l'intégrale de la fonction positive $(x - \langle X \rangle)^2 p(x)$. D'où :

$$\langle X^2 \rangle - \langle X \rangle^2 \geq 0 \iff \langle X^2 \rangle \geq \langle X \rangle^2 \quad (1.102)$$

En fait, l'égalité n'est obtenue que si la fonction $(x - \langle X \rangle)^2 p(x)$ est identiquement nulle ; ceci ne survient que dans un cas, celui où X ne prend qu'une seule valeur (et alors c'est forcément sa valeur *moyenne* !), soit quand X est en fait une variable certaine.

⁶⁷La variable $X - \langle X \rangle$ est appelée variable *centrée*, en raison du fait que sa valeur moyenne est visiblement nulle.

Une variable aléatoire X , par définition, présente une dispersion des valeurs qui “sortent” successivement d’une série d’expériences. Il est très important de pouvoir décrire quantitativement la dispersion de ces valeurs qui, lorsque le nombre d’expériences devient assez grand, se déplacent de façon erratique de part et d’autre de la valeur moyenne et représentent les *fluctuations* de X autour de la moyenne. D’après les observations précédentes, la mesure la plus naturelle de ces fluctuations est l’écart-type ΔX , définie comme :

$$\Delta X \stackrel{\text{déf}}{=} \sqrt{\langle X^2 \rangle - \langle X \rangle^2} \quad (1.103)$$

la quantité sous le radical est appelée *variance*, ou *écart quadratique moyen* ou plus simplement *écart quadratique* ; d’après (1.102), c’est bien une quantité positive, nulle dans le cas-limite d’une variable certaine. Des exemples simples montrent ce que l’on peut pressentir : la valeur de l’écart-type est d’autant plus grande que les valeurs de X sont très dispersées, c’est à dire que la distribution densité $p(x)$ est large, c’est-à-dire décroît lentement ; *a contrario*, l’écart-type Δx est nul dans le cas où la variable X est certaine, c’est-à-dire ne fluctue pas du tout. L’écart-type est donc une bonne mesure *globale* des fluctuations ; toutefois, en tant que grandeur intégrale calculée avec $p(x)$, il “efface” forcément certains détails de la densité de probabilité. On retiendra en tout cas l’équivalence :

$$\Delta X = 0 \iff X \text{ est une variable certaine} \quad (1.104)$$

En terme de distribution de probabilité, le cas certain correspond à $p_n = \delta_{n n_0}$ pour une densité atomique ; pour une densité continue $\rho(x)$, c’est formellement le cas-limite où $\rho = \lim_{\varepsilon \rightarrow 0} \delta_\varepsilon(x - x_0)$, $\delta_\varepsilon(x - x_0)$ étant un précurseur de la fonction de Dirac, ε étant commensurable avec l’écart-type. En toute généralité, la fonction de répartition d’une variable certaine (prenant la valeur x_0 avec probabilité 1) est $F(x) = \Theta(x - x_0)$.

D’une façon générale, on définit les *moments* M_α d’une variable aléatoire ; ce sont les valeurs moyennes des X^α (α n’est pas forcément entier, ni forcément positif) :

$$M_\alpha \stackrel{\text{déf}}{=} \langle X^\alpha \rangle = \int_{-\infty}^{+\infty} x^\alpha p(x) dx = \sum_n x_n^\alpha p_n + \int_{-\infty}^{+\infty} x^\alpha \rho(x) dx \quad (1.105)$$

On a ainsi par exemple $M_1 = \langle X \rangle$, $\Delta X^2 = M_2 - M_1^2$ et, bien sûr, $M_0 = 1$.

Il est bien clair que rien ne garantit d’avance que les différents moments existent et sont finis ; par exemple, lorsqu’il n’y a pas de composante atomique :

$$M_\alpha = \int_{-\infty}^{+\infty} x^\alpha \rho(x) dx \quad (1.106)$$

Si ρ est à support borné ne contenant pas l’origine, l’intégrale existe $\forall \alpha$; si l’origine est dans le support de ρ , il faut $\alpha > -1$ si $\rho(0)$ est fini (si ρ s’annule en $x = 0$, la condition sur α est moins sévère, elle l’est davantage si ρ diverge en $x = 0$). Quand le support s’étend jusqu’à l’infini, tout dépend de la nature de la décroissance de ρ à l’infini : pour que le moment M_α existe, il faut et il suffit que ρ décroisse plus vite que $|x|^{-(\alpha+1)}$. Notons que le moment d’ordre zéro existe toujours, et vaut 1 : ce n’est rien d’autre que l’intégrale

et/ou la somme de la densité $p(x)$: $\int p = 1 \forall p$. Si $\alpha < 0$, on parle souvent⁶⁸ de *moments inverses* ; pour une v.a. ne prenant que des valeurs strictement positives, $X \geq x_{\min} > 0$, tous les moments inverses existent (aucune intégrale ne diverge). Notons enfin que des inégalités variées entre les moments peuvent être trouvées en partant de :

$$\int_{-\infty}^{+\infty} (x^\alpha + \lambda x^\beta)^2 \rho(x) dx \geq 0 , \quad (1.107)$$

et en écrivant que le discriminant du trinôme est négatif :

$$M_{\alpha+\beta}^2 \leq M_{2\alpha} M_{2\beta} . \quad (1.108)$$

Il est parfois utile de s'intéresser non pas aux moments d'une v.a., mais à ses *cumulants* – voir p. 68.

La notion de variables aléatoires indépendantes a été définie dans la sous-section 1.3.2 (voir note 38 p.17). Dans le cas de variables non-indépendantes, il est utile de mesurer leur degré d'interdépendance. Un indice commun, et *a priori* très naturel, est la *covariance*. Pour deux v.a. X et Y , leur covariance C_{XY} (aussi appelée *coefficient de corrélation*) est définie comme :

$$C_{XY} \stackrel{\text{déf}}{=} \langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle \quad (1.109)$$

c'est une sorte de variance (ou d'écart quadratique) *croisé*. Il ne fait aucun doute que si les deux v.a. sont indépendantes, la moyenne du produit est égale au produit des moyennes et $C_{XY} = 0$:

$$X \text{ et } Y \text{ indépendantes} \implies C_{XY} = 0 \quad (1.110)$$

La réciproque est fautive, comme le montre l'exemple suivant. Soit un angle θ distribué uniformément entre 0 et 2π ; les deux fonctions $X \stackrel{\text{déf}}{=} \cos \theta$ et $Y \stackrel{\text{déf}}{=} \sin \theta$ sont deux variables aléatoires, dont la moyenne est d'ailleurs nulle. Ces deux variables ne sont sûrement pas indépendantes, puisque reliées l'une à l'autre par $X^2 + Y^2 = 1$; pourtant, leur covariance est égale à la valeur moyenne de $\sin \theta \cos \theta$ sur $[0, 2\pi]$ et est donc nulle. L'usage est d'appeler *variables décorrélées* des v.a. dont la covariance est nulle, mais cette terminologie doit être bien comprise : elle ne signifie pas forcément que les v.a. sont *indépendantes*. D'une façon générale, et en termes physiques, l'existence d'un lien de causalité entre deux événements entraîne inévitablement que ceux-ci ne sont pas indépendants (même si leur coefficient de corrélation est nul).

Il est bien clair que la connaissance de $p(x)$ (ou de $F(x)$) constitue l'information maximale escomptable dans un cadre probabiliste. En pratique, on est souvent bien démuni, ne connaissant d'expérience qu'un petit nombre de moments – sauf bien sûr si

⁶⁸Ces quantités apparaissent tout naturellement dans l'étude des marches au hasard en milieu aléatoire. Ainsi, si W_x est un taux de saut d'un endroit à l'autre variant localement dans l'espace, le temps de transit d'un point à l'autre est une certaine somme sur les durées des étapes intermédiaires : c'est donc une somme du genre $\sum_x \frac{1}{W_x}$, dont la valeur moyenne implique $\langle \frac{1}{W} \rangle$, premier moment inverse.

on est capable d'échafauder un modèle permettant de trouver théoriquement les fonctions $F(x)$ ou $p(x)$. Dans le premier cas, se pose alors parfois le problème difficile de reconstruire la fonction inconnue à partir des moments obtenus d'une façon ou d'une autre – visiblement d'autant plus délicat que la distribution est *large*, un caractère qui traduit la pertinence des événements rares. Il existe un théorème important, le théorème de Marcinkiewicz, que l'on énoncera plus précisément par la suite (section 1.8), et qui érige quelques garde-fous évitant les canulars (par exemple des distributions contenant des probabilités... négatives).

Il est naturel de se demander avec quelle probabilité on peut trouver une v.a. X loin de sa valeur moyenne ; l'inégalité de Tchebychev répond quantitativement à cette question. Si la moyenne est $\langle X \rangle$ et la variance ΔX^2 , cette inégalité s'énonce :

$$\varpi \stackrel{\text{déf}}{=} \text{Prob}[|X - \langle X \rangle| \geq \lambda \Delta X] \leq \frac{1}{\lambda^2} \quad (1.111)$$

où, en pratique et pour l'intérêt de la question, $\lambda \geq 1$. Pour la clarté de la démonstration, admettons que la fonction de distribution de X est une densité $p(x)$ (au sens large) définie sur \mathbb{R} . La probabilité ϖ de l'événement cherché est l'intégrale de $p(x)$ sur le complémentaire \bar{I} de l'intervalle $I \stackrel{\text{déf}}{=} [\langle X \rangle - \lambda \Delta X, \langle X \rangle + \lambda \Delta X]$; en multipliant membre à membre par $(\lambda \Delta X)^2$, on peut écrire la chaîne d'inégalités :

$$(\lambda \Delta X)^2 \varpi = \int_{\bar{I}} (\lambda \Delta X)^2 p(x) dx \leq \int_{\bar{I}} (x - \langle X \rangle)^2 p(x) dx \leq \int_{-\infty}^{+\infty} (x - \langle X \rangle)^2 p(x) dx \quad (1.112)$$

L'intégrale de droite n'est autre que ΔX^2 , qui est positif, d'où $\varpi \leq \lambda^{-2}$. Comme la variance de la variable $\frac{X - \langle X \rangle}{\Delta X}$ vaut 1, l'inégalité de Tchebychev s'écrit aussi :

$$\text{Prob}[|X - \langle X \rangle| \geq \lambda] \leq \frac{\Delta X^2}{\lambda^2} \quad (1.113)$$

On définit parfois la *médiane* comme le nombre réel m tel que :

$$\text{Prob}[X \leq m] \leq m \leq \text{Prob}[X \geq m] \quad (1.114)$$

Pour une variable discrète, cette définition a un intérêt très relatif : en raison des plateaux de $F(x)$, m est indéterminé. En revanche, pour une aléatoire continue dont la fonction de répartition est strictement croissante, $F(m) = \frac{1}{2}$ définit un réel m et un seul.

◆ *Remarque*

L'écart-type est indéniablement une bonne mesure de l'incertitude liée au fait que la variable est aléatoire. D'autres mesures complémentaires sont données par les moments d'ordre supérieur. On peut toutefois souhaiter disposer d'un indice qualifiant la distribution des probabilités elle-même, donnant un nombre à la hauteur

de l'étalement des probabilités. Cette grandeur existe, et est appelée entropie de Shannon et Wiener. La notant S_{SW} , elle s'écrit comme suit⁶⁹ pour une distribution discrète de probabilités p_n , $1 \leq n \leq N$:

$$S_{\text{SW}} \stackrel{\text{déf}}{=} - \sum_{n=1}^N p_n \ln p_n \quad (1.115)$$

C'est un nombre compris entre 0 et $\ln N$: $S_{\text{SW}} = 0$ ssi toutes les probabilités sont nulles sauf une qui vaut 1 (pas d'aléatoire en fait), $S_{\text{SW}} = \ln N$ si la distribution est plate (uniforme), soit $p_n = \frac{1}{N} \forall n$ (situation la plus incertaine ("désordonnée"), tous les événements sont équiprobables). L'expression (1.115) se retrouve en Mécanique statistique : $S = -k_{\text{B}} \sum_n p_n \ln p_n$ est de fait l'entropie (au sens physique du terme) d'un système pouvant se trouver dans différents états avec les probabilités p_n . Le postulat fondamental de la Mécanique statistique affirme que pour un système isolé, tous les états de même énergie ont la même probabilité : c'est bien ce qui assure à l'entropie S sa plus grande valeur possible, $S_{\text{max}} = k_{\text{B}} \ln N$. \blacklozenge

1.7 Lois de distribution courantes

Il s'agit ici de présenter quelques lois survenant fréquemment en Physique, en commençant par les lois discrètes.

1.7.1 Loi binomiale

La loi binomiale est une loi discrète qui survient quand on considère une population dont les individus, *agissant indépendamment les uns des autres*, peuvent être chacun dans deux états possibles, avec les probabilités p et $q = 1 - p$. Par exemple, il peut s'agir d'une source de N_0 noyaux instables au départ qui, à un certain instant t (fixé pour le moment), peuvent être soit excités ("vivants", "non désintégrés", probabilité p), soit à l'état fondamental stable ("morts", "désintégrés", probabilité $q = 1 - p$). La variable aléatoire est ici N , nombre de noyaux survivants à l'instant considéré, chaque noyau vivant sa vie indépendamment de tous les autres. Tout ceci se transpose à une vapeur atomique⁷⁰ *diluée*, constituée d'atomes à deux niveaux $|e\rangle$ (excité) et $|g\rangle$ (fondamental). Autres exemples : les boules blanches et noires que l'on tire successivement dans un sac, les différentes suites possibles de pile et face quand on jette N_0 fois la pièce.

⁶⁹Quand on substitue \log_2 au logarithme népérien, on obtient la fonction dite *Information manquante* de la Théorie de l'information (édifiée précisément par Shannon et Wiener, pour une introduction, voir l'ouvrage de Renyi [18]).

⁷⁰L'hypothèse d'une vapeur diluée assure que les atomes ne se voient pas les uns les autres ; dans le cas contraire, lorsque la distance typique entre deux atomes est de l'ordre de grandeur de la longueur d'onde émise, le phénomène de *superradiance* est possible.

En définitive, la loi binomiale est la distribution des probabilités pour N_0 expériences indépendantes (successives ou simultanées) tirant au hasard une variable de Bernoulli (variable aléatoire à deux valeurs). Plus précisément, si on affecte la valeur 1 à l'état vivant, et zéro à l'état mort, N est égal à la somme de ces N_0 variables de Bernoulli : la loi binomiale est de fait la loi de probabilité d'une *somme* de variables de Bernoulli (indépendantes) – c'est aussi la distribution du gain d'un joueur au jeu de pile et face. Ce fait sera retrouvé par la suite, notamment à propos de la fonction caractéristique d'une somme de v.a. indépendantes (voir (1.194) et section 1.8.2).

Formulée à propos de la source radioactive, la question centrale est ici de trouver la probabilité p_n :

$$\text{Probabilité d'avoir } n \text{ survivants} \equiv \text{Prob}[N = n] \stackrel{\text{d\'ef}}{=} p_n . \quad (1.116)$$

Le nombre de façons distinctes d'avoir n individus vivants est $C_{N_0}^n$. En effet, numérotons les individus de 1 à N_0 et disposons-les sur une ligne en mettant les n vivants à gauche, les $N_0 - n$ morts à droite. Il est clair que l'on peut effectuer n'importe quelle permutation parmi les vivants (il y a $n!$ permutations), ou parmi les morts ($(N_0 - n)!$ permutations), sans modifier la nature de l'événement " n vivants et $N_0 - n$ morts" ; par ailleurs, la numérotation des individus est arbitraire et il y a $N_0!$ numérotations distinctes. Au total, le nombre de configurations contenant n vivants et $N_0 - n$ morts est égal à $\frac{N_0!}{n!(N_0-n)!} = C_{N_0}^n$.

Chaque répartition exclut toute autre ; une répartition donnée a la probabilité $p^n(1-p)^{N_0-n}$. La probabilité p_n d'avoir n noyaux vivants est donc la somme des $C_{N_0}^n$ probabilités toutes égales à $p^n(1-p)^{N_0-n}$, soit :

$$p_n = C_{N_0}^n p^n (1-p)^{N_0-n} \quad (0 \leq n \leq N_0) \quad (1.117)$$

On vérifie immédiatement que la somme des p_n vaut 1 puisque $[p + (1-p)]^{N_0} = 1$. La distribution n'est pas symétrique dans la transformation $n \rightarrow N_0 - n$, sauf si $p = \frac{1}{2}$, évidemment.

À p donné, la valeur moyenne (espérance mathématique) du nombre de survivants est :

$$\langle N \rangle = \sum_{n=0}^{N_0} n p_n = \sum_{n=0}^{N_0} n p^n (1-p)^{N_0-n} C_{N_0}^n . \quad (1.118)$$

Ce type de somme se calcule aisément en introduisant la fonction génératrice des probabilités⁷¹ des probabilités $f(s)$:

$$f(s) \stackrel{\text{d\'ef}}{=} \sum_{n=0}^{N_0} s^n p_n \equiv \sum_{n=0}^{N_0} s^n p^n (1-p)^{N_0-n} C_{N_0}^n = [ps + (1-p)]^{N_0} , \quad (1.119)$$

⁷¹La terminologie vient du fait que, au facteur $n!$ près, les dérivées successives de f en $s=0$ sont précisément les probabilités : $p_n = \frac{1}{n!} \left[\frac{\partial^n}{\partial s^n} f(s; t) \right]_{s=0}$.

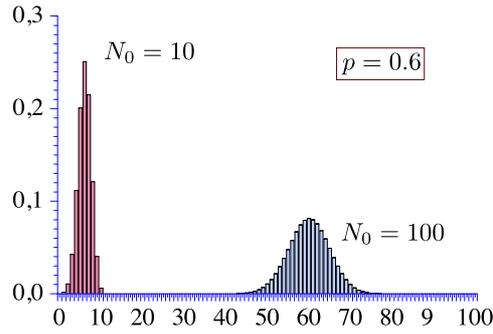


Figure 1.13: Distribution des probabilités binomiales (1.117) p_n en fonction de n , pour $p = 0.6$ et $N_0 = 10, 100$. Les probabilités sont strictement nulles si n est en dehors de $[0, N_0]$.

et en calculant ses dérivées successives en $s = 1$. Par exemple :

$$\langle N \rangle = \left[\frac{d}{ds} f(s) \right]_{s=1} = \left[\frac{d}{ds} [ps + (1-p)]^{N_0} \right]_{s=1}, \quad (1.120)$$

d'où :

$$\boxed{\langle N \rangle = N_0 p} \quad (1.121)$$

Par définition, la moyenne de $\langle N^2 \rangle$ est :

$$\langle N^2 \rangle \stackrel{\text{déf}}{=} \sum_{n=0}^{N_0} n^2 p_n = \sum_{n=0}^{N_0} n^2 p^n (1-p)^n C_{N_0}^n \quad (1.122)$$

et se calcule en dérivant $f(s)$ une nouvelle fois⁷² ; au total, on trouve que l'écart quadratique est :

$$\boxed{\Delta N^2 \stackrel{\text{déf}}{=} \langle N^2 \rangle - \langle N \rangle^2 = N_0 p (1-p)} \quad (1.123)$$

Ceci permet d'écrire l'expression des fluctuations relatives :

$$\frac{\Delta N}{\langle N \rangle} = \frac{1}{\sqrt{N_0}} \sqrt{\frac{1}{p} - 1}. \quad (1.124)$$

Il faut noter l'occurrence du facteur $N_0^{-1/2}$, qui reviendra souvent par la suite dans la considération des fluctuations.

S'agissant de la source radioactive, la probabilité pour un noyau d'être encore vivant à l'instant t est $p(t) = e^{-\lambda t}$, où λ^{-1} est la durée de vie moyenne d'un noyau à l'état actif. En effet, la probabilité d'être actif à l'instant $t + \delta t$ est le produit de la probabilité

⁷²Remarquer que $n^2 = n(n-1) + n$.

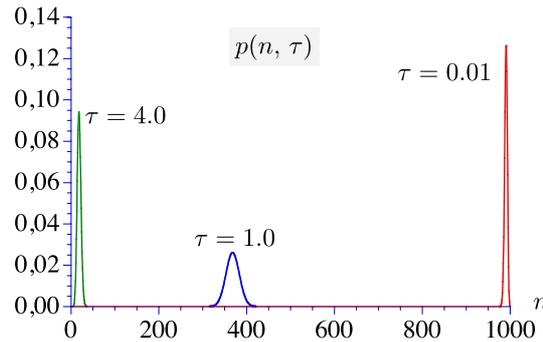


Figure 1.14: Distribution des noyaux survivants à trois âges d'une source contenant initialement $N_0 = 1000$ noyaux actifs. Chaque distribution est repérée par la valeur du paramètre $\tau = \lambda t$ (à droite : source jeune ; à gauche : source vieille).

d'être actif à t , soit $p(t)$, multipliée par la probabilité de n'avoir pas basculé entre t et $t + \delta t$, événement contraire de l'événement ayant pour probabilité $\lambda \delta t$: $p(t + \delta t) = (1 - \lambda \delta t)p(t)$. Le passage à la limite $\delta t \rightarrow 0$ donne $\dot{p}(t) = -\lambda p(t)$.

La probabilité $p(t)$ décroît toujours (plus on attend, plus on a de chances de trouver qu'un noyau donné est mort), et tend évidemment vers zéro. La dépendance en t de p se reporte sur toutes les probabilités, $p_n(t)$, donc sur toutes les espérances mathématiques. Ainsi, la moyenne de N devient une fonction du temps et (1.121) donne $\langle N \rangle(t) = N_0 e^{-\lambda t}$, expression usuelle de la loi de décroissance radioactive (qui porte en fait sur une *moyenne*, comme on omet trop souvent de le mentionner⁷³). Aux premiers instants ($t \ll \frac{1}{\lambda}$, $p \sim 1$ et $q \equiv 1 - p \sim 0$), la distribution est ramassée vers N_0 (d'ailleurs, au départ ($t = 0$) N est une variable certaine : on se donne comme information que la source contient N_0 noyaux actifs). À l'instant⁷⁴ T où $p = q$, soit $e^{-\lambda T} = \frac{1}{2}$, ou encore $T = \frac{\ln 2}{\lambda} \simeq \frac{0.693}{\lambda}$, la distribution est symétrique autour de $\frac{N_0}{2}$, il y a – en termes de probabilités – autant de morts que de vivants. Enfin, pour $t \gg \lambda^{-1}$, $p \simeq 0$ et $q \simeq 1$, toutes les probabilités sont exponentiellement petites, sauf $p_0 \lesssim 1$. Ceci ne veut pas dire qu'une source radioactive devenue vieille est inoffensive : selon (1.124), on a alors $\frac{\Delta N(t)}{\langle N \rangle(t)} \simeq \frac{1}{\sqrt{N_0}} e^{+\frac{\lambda t}{2}}$, ce qui signifie que les fluctuations relatives $\frac{\Delta N}{\langle N \rangle} \dots$ divergent exponentiellement quand $t \rightarrow +\infty$! Ainsi, même une vieille source – dont l'activité *moyenne* est devenue faible en raison de son grand âge – est capable d'émettre des *bouffées* dangereuses ; cette observation élémentaire est à la base de l'éternel débat : peut-on tolérer des faibles doses (moyennes) ou au contraire la seule dose acceptable est-elle la dose *zéro* ? – qui n'existe nulle part d'ailleurs.

⁷³Ce fait est en pratique sans importance : pour une source jeune contenant un grand nombre de noyaux, et en comptant les désintégrations pendant des intervalles de temps grands par rapport à $\frac{1}{N\lambda}$, les fluctuations sont très faibles (le *bruit* est très petit). Il n'empêche que, conceptuellement, la différence est de taille. Pouvoir confondre (au moins dans un premier temps !) valeur moyenne et variable aléatoire résulte ici, comme pour tout système macroscopique, de l'énormité du nombre d'Avogadro.

⁷⁴ T est appelé (improprement) *période* ; en ordre de grandeur, $T \sim \lambda^{-1}$.

▼ Dynamique des probabilités pour le déclin radioactif

La présentation précédente peut être reprise d'un autre point de vue, permettant au passage d'introduire la notion de *processus stochastique* (un *must* en la matière : l'ouvrage de Gardiner [19]). Sans vouloir en donner une définition générale, on peut dire qu'un processus stochastique est la succession des valeurs d'une variable aléatoire lorsqu'un certain paramètre varie, le plus souvent il s'agit du temps ou d'une coordonnée spatiale, mais ce peut être aussi l'énergie, quand un système se promène entre des états caractérisés par leur énergie. Ici, la v.a. est N , nombre de survivants à l'instant t – qui prend les valeurs entières $0 \leq n \leq N_0$ –, dont les probabilités sont les p_n données en (1.117), et qui deviennent des fonctions du temps par l'injection de la loi exponentielle pour p : $p = e^{-\lambda t}$.

On peut obtenir directement les $p_n(t)$ sans effectuer le petit dénombrement qui a conduit à (1.117), en écrivant, puis en résolvant, l'équation d'évolution des $p_n(t)$. Celle-ci s'obtient par un raisonnement⁷⁵ de *bilan*, en dressant l'inventaire des événements possibles entre deux instants t et $t + \delta t$, compte tenu du seul ingrédient physique : chaque noyau a une probabilité $\lambda \delta t$ de se désexciter entre t et $t + \delta t$, sans se soucier de ce que font les autres.

Soit donc la probabilité $p_n(t)$ de trouver n noyaux actifs à l'instant t , sachant qu'il y en avait N_0 à l'instant initial ($t = 0$). En appliquant les règles usuelles :

- événements indépendants \leftrightarrow multiplication des probabilités soit :

$$\text{Prob}[A \text{ et } B] = P(A) P(B)$$

- événements exclusifs \leftrightarrow addition des probabilités soit :

$$\text{Prob}[A \text{ ou } B] = P(A) + P(B) ,$$

et en faisant l'inventaire des processus élémentaires pouvant survenir entre t et $t + \delta t$ ⁷⁶, on obtient le système d'équations de bilan :

$$p_n(t + \delta t) = (1 - \lambda \delta t)^n p_n(t) + (n + 1) \lambda \delta t (1 - \lambda \delta t)^n p_{n+1}(t) + \mathcal{O}((\lambda \delta t)^k \geq 2) , \quad (1.125)$$

avec $n = 0, 1, \dots, N_0$ (et $p_{N_0+k} \equiv 0$ si $k \geq 1$). Le premier terme provient de la possibilité : aucun des n noyaux présents à l'instant t n'a changé d'état ; le second

⁷⁵Ce raisonnement par *bilan* a une portée universelle en Physique. C'est par ce type d'argument qu'Einstein fut conduit à introduire (1917) les coefficients A_ν et B_ν (dits *coefficients d'Einstein*) ; le second, B_ν , est relativement classique (absorption et émission stimulées). Au contraire, A_ν est de nature spécifiquement quantique et décrit l'émission *spontanée*, qui se produit même en l'absence de champ (en moyenne) et résulte des *fluctuations* (quantiques) du vide de rayonnement (champ électromagnétique dans son état fondamental – pas de photon présent).

Cet argument est aussi le fondement de l'équation-maîtresse rencontrée lors de l'étude des processus stochastiques (Mécanique statistique hors d'équilibre, Théorème- H de Boltzmann), et de l'équation de Fokker - Planck. Dans tous les cas, sous-jacente se trouve une équation de conservation "locale" de la probabilité, du genre $\partial_t \rho + \vec{\nabla} \cdot \vec{J} = 0$, exactement comme en Électricité en conséquence de la conservation de la charge.

⁷⁶Par exemple, la probabilité qu'aucun des n noyaux vivants à t ne bascule entre t et $t + \delta t$ est $(1 - \lambda \delta t)^n$.

terme signifie qu'il y avait $n + 1$ noyaux à l'instant t et que l'un de ces $n + 1$ noyaux a basculé, les n autres n'ayant pas changé d'état. Le symbole noté \mathcal{O} rassemble des processus multiples, qui sont d'ordre supérieur en δt (par exemple : plusieurs noyaux se sont désexcités durant δt). Au même ordre en δt , le système (1.125) est :

$$p_n(t + \delta t) = (1 - n\lambda\delta t)p_n(t) + (n + 1)\lambda\delta t p_{n+1}(t) + \mathcal{O}((\lambda\delta t)^k \geq 2) . \quad (1.126)$$

Après division par δt et passage à la limite $\delta t \rightarrow 0$, on obtient le système⁷⁷ différentiel suivant, de dimension $N_0 + 1$:

$$\boxed{\frac{dp_n}{dt} = (n + 1)\lambda p_{n+1}(t) - n\lambda p_n(t) \quad (n = 0, 1, 2, \dots, N_0)} \quad (1.127)$$

avec la condition aux limites $p_{N_0+1}(t) \equiv 0 \forall t$. L'intégration de ce système avec la condition initiale $p_n(t = 0) = \delta_{n, N_0}$ montre que la solution est de fait la distribution donnée en (1.117) avec $p = e^{-\lambda t}$. Noter que si $p = \frac{1}{2}$ – soit à l'instant $t = \frac{\ln 2}{\lambda} = T -$, la distribution est symétrique : $p_n = p_{N_0-n}$.

★ **Appendice** Une façon commode de résoudre le système (1.127) est la suivante. On considère à nouveau la fonction génératrice⁷⁸ $f(s; t)$:

$$f(s; t) = \sum_{n=0}^{N_0} s^n p_n(t) , \quad (1.128)$$

qui est maintenant une fonction des deux variables s et t . À partir de (1.127), on voit de suite que f satisfait l'équation aux dérivées partielles :

$$\frac{\partial f}{\partial t} = \lambda(1 - s) \frac{\partial f}{\partial s} , \quad (1.129)$$

avec la condition aux limites $f(t = 0; s) = s^{N_0}$ puisque $p_n(t = 0) = \delta_{n, N_0}$. L'inspection attentive⁷⁹ de cette équation montre que sa solution est de la forme $F[(1 - s)a(t)]$, où F est une fonction pour l'instant quelconque, et à condition que $a(t)$ satisfasse l'équation $\dot{a} = -\lambda a \iff a(t) = e^{-\lambda t}$, d'où résulte :

$$f(s; t) = F[(1 - s)e^{-\lambda t}] . \quad (1.130)$$

La condition aux limites impose $F(1 - s) = s^{N_0} \iff F(S) = (1 - S)^{N_0}$ et finalement :

$$\boxed{f(s; t) = [1 - (1 - s)e^{-\lambda t}]^{N_0} = [(1 - e^{-\lambda t}) + se^{-\lambda t}]^{N_0}} \quad (1.131)$$

⁷⁷Il est facile de vérifier que la somme des premiers membres est bien égale à zéro.

⁷⁸Tout comme en (1.119), le nombre de probabilités p_n est fini (il est égal à $N_0 + 1$), de sorte que f est un polynôme.

En posant $s = e^{it}$, on voit que $f(s = e^{it}) \stackrel{\text{déf}}{=} \phi(t)$ n'est autre que la fonction caractéristique introduite plus loin (voir section 1.8) ; c'est la transformée de Fourier discrète des p_n . $\phi(t)$ est visiblement une fonction de même nature que la fonction $f(s)$ introduite en (1.119).

⁷⁹Une procédure moins intuitive et plus systématique consiste à effectuer une transformation de Laplace de l'équation (1.129) ; en posant $F(z; s) = \int_0^{+\infty} e^{-zt} f(s; t) dt$, on trouve que $F(z; s)$ satisfait une équation différentielle ordinaire. L'intégration de celle-ci, suivie de la transformation de Laplace inverse, redonne l'expression (1.131) (faire cet exercice !).

Le développement de cette expression en puissances de s (suivant la formule du binôme), et l'identification avec le développement (1.128), reproduit l'expression (1.117) des probabilités, comme il se doit. ★

Conformément à (1.53), la fonction de répartition binomiale est :

$$F(x) = \sum_{n=0}^{N_0} C_{N_0}^n p^n (1-p)^{N_0-n} \Theta(x-n) ; \quad (1.132)$$

c'est un escalier à $N_0 + 1$ marches. ▲

1.7.2 Loi de Poisson

La loi de Poisson est relative à une variable aléatoire N prenant ses valeurs dans \mathbb{N} et s'écrit par définition avec $n \in \mathbb{N}$, $a \in \mathbb{R}_+$; on vérifie immédiatement que $\sum_{n \in \mathbb{N}} p_n = 1$. :

$$\boxed{\text{Prob}[N = n] \stackrel{\text{déf}}{=} p_n = e^{-a} \frac{a^n}{n!}} \quad (1.133)$$

avec $n \in \mathbb{N}$, $a \in \mathbb{R}_+$; on vérifie immédiatement avec $n \in \mathbb{N}$, $a \in \mathbb{R}_+$. Comme on va le voir, cette loi, qui ne dépend que du seul paramètre a , décrit notamment, dans une certaine limite qui va être précisée, ce que l'on appelle les *phénomènes d'attente*. Afin de préciser ceci, on se place dans la situation concrète suivante.

Soit un grand nombre d'individus (clients), \mathcal{N} , souhaitant se rendre en un lieu donné (un guichet) en un temps total donné, T . Chaque individu agit pour son propre compte sans se soucier des autres. Considérant la v.a. N du nombre de clients arrivant pendant le même intervalle de temps fixe et donné une fois pour toutes, noté Δt , la question est de trouver $\text{Prob}[N = n] \equiv p_n$.

Une donnée fondamentale est la probabilité p pour qu'un client arrive pendant Δt . L'hypothèse la plus naturelle (en tout cas la plus simple) consiste à poser :

$$\boxed{p = \frac{\Delta t}{T}} \quad (1.134)$$

C'est une façon de traduire le *hasard* complet, au sens où aucun instant n'est privilégié entre 0 et T , que tous les temps se valent. Sans qu'il soit réellement possible de le préciser en détail, on sent bien qu'adopter cette définition pour p contient une hypothèse de *stationnarité*.

Pendant l'intervalle de temps Δt , un individu arrive au guichet avec une certaine probabilité p , ou n'y arrive pas, avec la probabilité $1 - p$. Pour une population de \mathcal{N} clients potentiels agissant indépendamment les uns des autres, la probabilité qu'il arrive n personnes est donc la même probabilité que celle d'avoir n fois pile (par exemple) avec

\mathcal{N} lancers de la pièce : c'est donc une probabilité binomiale, donnée par (1.117) et qui ici s'écrit ($N_0 \rightarrow \mathcal{N}$) :

$$p_n = C_{\mathcal{N}}^n p^n (1-p)^{\mathcal{N}-n} . \quad (1.135)$$

Ceci fournit la réponse à la question posée, mais il est intéressant de l'analyser dans une certaine limite, où à la fois \mathcal{N} est très grand et où l'intervalle de temps de comptage Δt est très petit devant la durée totale d'observation T , ce qui constitue typiquement une analyse à grande échelle du problème, dans un modèle de quasi-*continuum*.

Il est clair que les deux limites $\mathcal{N} \gg 1$ et $\Delta t \ll T$ ne doivent pas être prises indépendamment si l'on souhaite une réponse non triviale : par exemple, si on fait tendre formellement T vers $+\infty$ en gardant \mathcal{N} fixe, la réponse est immédiate : il ne se passe rien, $p = 0$ par (1.134), et p_n est nulle, sauf si $n = 0$; c'est d'ailleurs bien ce que donne l'expression (1.135) si on y fait brutalement $p = 0$. Inversement, si on garde T fixe et que l'on fait tendre N vers l'infini, on obtient $p_n = 0$ sauf si $n = \mathcal{N}$: tout le monde arrive en même temps. Dans un cas comme dans l'autre, la réponse est triviale et sans intérêt.

La façon la plus simple de lier les deux limites est de dire que le rapport $\frac{\mathcal{N}}{T}$ garde une valeur fixe⁸⁰ :

$$\boxed{\mathcal{N} \rightarrow +\infty , \quad T \rightarrow +\infty , \quad \frac{\mathcal{N}}{T} = C^{\text{ste}} \stackrel{\text{déf}}{=} \gamma} \quad (1.136)$$

En fait, c'est aussi la façon la plus naturelle de poser le problème en termes physiques ; la grandeur constante γ , inverse d'un temps, n'est pas un nouveau paramètre du problème et a un sens physique très clair : en tant que limite du rapport $\frac{\mathcal{N}}{T}$, c'est le nombre moyen d'individus arrivés par unité de temps dans la limite d'une population infinie et pour un phénomène durant une éternité. γ représente le taux d'arrivée des clients ; c'est une grandeur susceptible d'une vérification expérimentale préalable révélant, pour le phénomène considéré, si elle est bien définie⁸¹, et si oui combien elle vaut. Un tel test statistique permet de savoir dans quelle mesure un phénomène *réel* relève ou non de la loi de Poisson.

Avec cette procédure de limite, les grandeurs T et \mathcal{N} ne sont plus deux paramètres indépendants ; reportant $T = \mathcal{N}\gamma^{-1}$ dans (1.134) :

$$p = \gamma \frac{\Delta t}{\mathcal{N}} \equiv \frac{a}{\mathcal{N}} , \quad a \stackrel{\text{déf}}{=} \gamma \Delta t . \quad (1.137)$$

La prescription (1.136) s'écrit aussi :

$$\boxed{\mathcal{N} \rightarrow +\infty , \quad p \rightarrow 0 , \quad \mathcal{N}p = C^{\text{ste}} = a} \quad (1.138)$$

⁸⁰Ce type de limite est fréquent en Physique : la limite thermodynamique la plus simple pour un gaz de N particules dans un conteneur de volume V consiste à prendre la limite $N \rightarrow +\infty$, $V \rightarrow +\infty$, le rapport $\frac{N}{V}$ ayant une valeur fixée (c'est la densité ρ).

⁸¹En particulier, s'il s'agit d'un phénomène aléatoire *stationnaire*, γ est une quantité indépendante du temps.

elle définit sans ambiguïté le processus de limite, et est une nécessité formelle imposée par le désir d'avoir une réponse non triviale dans le cas $\mathcal{N} \gg 1$. On va voir que, de fait, les probabilités binomiales p_n (1.135) admettent alors une limite, et que cette limite est la loi de Poisson donnée par (1.133).

Cela étant, il reste un peu d'algèbre à effectuer pour arriver à (1.133). Écrivons l'expression explicite de p_n à partir de (1.135) :

$$p_n = C_{\mathcal{N}}^n p^n (1-p)^{\mathcal{N}-n} = \frac{\mathcal{N}(\mathcal{N}-1)\dots(\mathcal{N}-n+1)}{n!} \left(\frac{a}{\mathcal{N}}\right)^n \left(1 - \frac{a}{\mathcal{N}}\right)^{\mathcal{N}-n}, \quad (1.139)$$

qui se réécrit comme suit :

$$p_n = \frac{1}{n!} \frac{\mathcal{N}(\mathcal{N}-1)\dots(\mathcal{N}-n+1)}{\mathcal{N}^n} a^n \left(1 - \frac{a}{\mathcal{N}}\right)^{\mathcal{N}-n}. \quad (1.140)$$

À n fixé, la grande fraction a pour limite 1 quand $\mathcal{N} \rightarrow \infty$; toujours à n fixé, on a :

$$\lim_{\mathcal{N} \rightarrow +\infty} \left(1 - \frac{a}{\mathcal{N}}\right)^{\mathcal{N}-n} = \lim_{\mathcal{N} \rightarrow +\infty} e^{(\mathcal{N}-n)\ln(1-\frac{a}{\mathcal{N}})} = e^{-a}. \quad (1.141)$$

D'où finalement, dans la limite considérée :

$$\lim p_n = \frac{a^n}{n!} e^{-a}; \quad (1.142)$$

c'est bien la loi de Poisson annoncée, (1.133). En résumé, si N dénote toujours la variable aléatoire "nombre d'individus arrivés pendant l'intervalle de référence Δt ", on a :

$$\boxed{\text{Prob}[N = n] = \frac{(\gamma \Delta t)^n}{n!} e^{-\gamma \Delta t}} \quad (1.143)$$

dans la limite d'une population \mathcal{N} infinie et d'une durée T infinie, le rapport $\frac{\mathcal{N}}{T}$ étant fixe et égal à γ . La discussion détaillée ci-dessus montre que c'est le jeu $T \rightarrow +\infty$, $\mathcal{N} \rightarrow +\infty$, $\mathcal{N}/T = C^{\text{ste}}$ qui donne une limite non-triviale, bien que $p \rightarrow 0$, conformément à (1.137). Noter que, une fois cette limite prise, rien n'interdit de considérer des valeurs de n arbitrairement grandes.

La loi de Poisson se rencontre également à propos d'un gaz de particules (supposées classiques et sans interaction mutuelle) à l'équilibre, confiné dans un récipient de volume V , quand on veut calculer la probabilité qu'un petit volume δV contienne un nombre donné de particules. Pour une seule particule, son appartenance à δV est une variable binaire, 0 si la particule est en dehors de V , 1 si elle est dedans par exemple, avec les probabilités $1-p$ et p : il s'agit donc d'une variable de Bernoulli ; avec une hypothèse de type ergodique, on a $p = \frac{\delta V}{V}$. Pour des particules sans interactions, la probabilité pour une particule d'être dans δV reste égale à p en présence des autres puisqu'elles ne se voient pas les unes les autres. Les particules étant supposées indépendantes, leur présence simultanée dans δV est un événement composite formé avec des événements indépendants ; la probabilité correspondante, p_n , est donc le produit des probabilités,

étant entendu qu'il y a $C_{\mathcal{N}}^n$ façons de répartir n particules parmi \mathcal{N} dans δV : p_n est donc à nouveau $C_{\mathcal{N}}^n p^n (1-p)^{\mathcal{N}-n}$.

Le récipient étant fermé, le nombre de particules \mathcal{N} est donné une fois pour toutes. Si $\delta V \ll V$, $p \ll 1$, mais avec δV "infinitésimal", on peut aussi assurer $\mathcal{N}p \lesssim 1$ bien que $\mathcal{N} \gg 1$: on est ainsi dans le cas où la limite ci-dessus s'applique, $\mathcal{N}p$ étant un nombre d'ordre 1 ($=a$) ; on retrouve ainsi la loi de Poisson :

$$\lim p_n = \frac{\bar{\nu}^n}{n!} e^{-\bar{\nu}} , \quad (1.144)$$

où $\bar{\nu} = \mathcal{N}p$ s'interprète ici comme le nombre moyen de particules dans le volume δV .

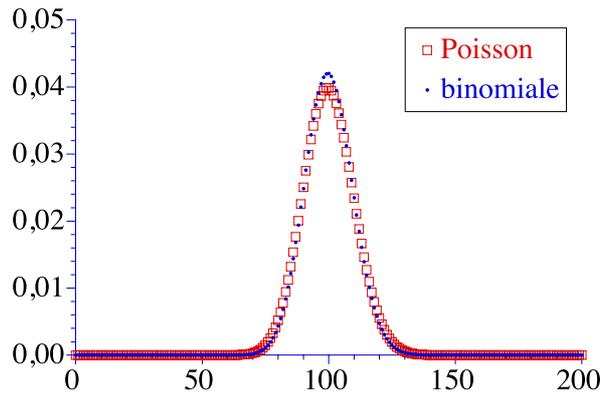


Figure 1.15: Comparaison des distributions binomiale et de Poisson. La loi binomiale est tracée avec $p = 0.1$ et $\mathcal{N} = 1000$ (noté N_0 dans (1.117)). Conformément à l'analyse conduisant à (1.145), la loi de Poisson correspondante a un paramètre a égal à $p\mathcal{N}$ soit $a = 100$.

La loi de Poisson est aussi celle qui régit des phénomènes de remplissage ; soit à remplir un segment de longueur L avec des points jetés au hasard. Il suffit, dans ce qui précède, de remplacer *individu* par *point*, T par L , le rapport noté γ est alors égal à $\frac{N}{L}$: c'est la densité de points supposée constante quand on fait tendre le nombre de points \mathcal{N} et la longueur L du segment vers l'infini.

Enfin, la loi de Poisson apparaît spontanément en Mécanique quantique, lors de l'introduction des *états cohérents* de l'oscillateur harmonique, qui servent de base pour la description d'états remarquables du champ électromagnétique.

La loi de Poisson porte aussi le nom de *Loi des petites probabilités* puisque, selon (1.137), la probabilité p tend vers zéro par le processus de limite, précisément comme $\frac{1}{\mathcal{N}}$. On retiendra :

$$\lim_{p \rightarrow 0, \mathcal{N} \rightarrow +\infty, \mathcal{N}p = C^{ste} = a} p_n^{(binomiale)}(\mathcal{N}, p) = p_n^{(Poisson)}(a = \mathcal{N}p) \quad (1.145)$$

où $p_n^{(\text{Poisson})}(a)$ désigne la loi de Poisson de paramètre a . La fig. 1.15 fournit la comparaison entre les deux lois ; si leur proximité est manifeste à l'œil, l'examen attentif révèle que, loin dans les ailes, les valeurs numériques (certes très petites) des deux lois sont en fait très différentes : ce phénomène est fréquent et se retrouve également lorsque l'on compare une loi et son approximation par sa loi attractive (voir la section 1.9). On désigne souvent par *région d'échelle* la zone centrale où la coïncidence est très satisfaisante.

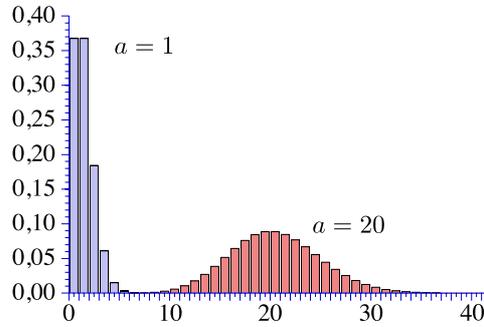


Figure 1.16: Distribution des probabilités de Poisson (1.133) pour $a = 1$ et $a = 20$. Les probabilités sont toutes non-nulles (le support est \mathbb{N}), mais deviennent très vite très petites.

La distribution de Poisson (1.133) présente des aspects très différents selon la valeur du paramètre $a = \gamma \Delta t$ (voir fig. 1.16). Lorsque $a \ll 1$, p_0 est très voisine de 1, cependant que les autres probabilités décroissent très vite avec n . C'est le cas lorsque l'intervalle de comptage des clients, Δt est très petit devant γ^{-1} , intervalle moyen séparant deux arrivées. Pour $a \lesssim 1$, les toutes premières probabilités sont voisines de 1, les autres décroissant assez vite. Pour $a = 1$, on a $p_0 = e^{-1} = p_1$, $p_n = \frac{e^{-1}}{n!}$. Quoi qu'il en soit, pour $a \sim 1$, ce sont les probabilités avec un petit n qui sont les plus importantes, le maximum de p_n survenant pour $n_{\max} \sim 1$:

$$n_{\max} \sim 1, \quad p_{n_{\max}} \lesssim 1 \quad (1.146)$$

Au contraire, si $a \gg 1$, les probabilités avec un petit n sont très petites, l'exponentielle e^{-a} l'emportant sur a^n tant que $n \lesssim \frac{a}{\ln a}$; par ailleurs, pour $n \rightarrow \infty$, $p_n \rightarrow 0 \forall a$. Il en résulte que p_n passe par un maximum pour un certain n_{\max} qui est grand⁸² devant 1. Ce maximum est petit ; en effet, en appliquant la formule de Stirling⁸³, on voit que :

$$a \gg 1 \Rightarrow n_{\max} \gg 1, \quad p_{n_{\max}} \simeq \frac{1}{\sqrt{2\pi n_{\max}}} \ll 1 \quad (1.148)$$

⁸²On va voir que $n_{\max} \simeq a$ (voir (1.149)).

⁸³La formule de Stirling est une approximation de la factorielle ; on montrera ultérieurement (ch. 7, section 7.4) que :

$$N \gg 1 : \quad N! \simeq \sqrt{2\pi N} \left(\frac{N}{e}\right)^N \quad (1.147)$$

Autrement dit, quand $a \gg 1$, la distribution des p_n est relativement “plate”, et présente un faible maximum d’ordre $(n_{\max})^{-\frac{1}{2}}$ (pour $n_{\max} = 100$, $p_{n_{\max}} \simeq 0.040$). Ceci est bien normal : supposer $a \gg 1$, c’est se placer dans la situation où l’intervalle de comptage est long devant γ^{-1} , auquel cas beaucoup d’événements (arrivée de n clients) ont une probabilité sensiblement notablement différente de zéro.

Toutes ces propriétés se confirment précisément en remarquant simplement que, selon (1.133), on a :

$$\frac{p_{n+1}}{p_n} = \frac{a}{n+1} ; \quad (1.149)$$

cette relation montre que $p_{n+1} > p_n$ tant que $n+1 < a$. Si $a \lesssim 1$, la distribution est toujours décroissante ; si $a \gg 1$, les probabilités croissent jusqu’à $n = n_{\max} \simeq a \gg 1$, puis décroissent. Cette dernière égalité approchée s’obtient aussi en dérivant formellement⁸⁴ $\ln p_n$ par rapport à n , après avoir utilisé la formule de Stirling $\ln N! \simeq N \ln N - N$ (justifié puisque $a \gg 1$) :

$$\frac{d}{dn} \ln p_n = \frac{d}{dn} (n \ln a - \ln n!) \simeq \ln a - \frac{d}{dn} (n \ln n - n) = \ln a - \ln n , \quad (1.150)$$

d’où $n_{\max} \simeq a$ quand $a \gg 1$.

Ces constatations se retrouvent également sur la moyenne de N et son écart-type.

On a :

$$\langle N \rangle = \sum_{n=0}^{+\infty} n e^{-a} \frac{a^n}{n!} = \sum_{n=1}^{+\infty} e^{-a} \frac{a^n}{(n-1)!} = a ; \quad (1.151)$$

la moyenne de N^2 est :

$$\langle N^2 \rangle = \sum_{n=0}^{+\infty} n^2 e^{-a} \frac{a^n}{n!} = \sum_{n=1}^{+\infty} [n(n-1) + n] e^{-a} \frac{a^n}{n!} = a^2 + a \equiv \langle N \rangle^2 + a , \quad (1.152)$$

d’où l’écart-type :

$$\Delta N = \sqrt{\langle N^2 \rangle - \langle N \rangle^2} = \sqrt{a} \quad (1.153)$$

et les fluctuations relatives :

$$\boxed{\frac{\Delta N}{\langle N \rangle} = \frac{1}{\sqrt{a}} \equiv \frac{1}{\sqrt{\langle N \rangle}}} \quad (1.154)$$

Si $a \lesssim 1$, la moyenne de N est près de $n = 0$, et les p_n sont très vite très petites : la distribution est ramassée près de $n \sim 1$; au contraire, si $a \gg 1$, la loi est en gros centrée en un point $n_{\max} \sim a \gg 1$, loin de l’origine, et sa largeur est égale à \sqrt{a} ; dans ce cas :

$$\frac{\Delta N}{\langle N \rangle} = \frac{1}{\sqrt{\langle N \rangle}} \ll 1 \quad (a \gg 1) \quad (1.155)$$

et on retrouve alors le comportement typique en $N^{-1/2}$ des fluctuations relatives, déjà rencontré en (1.124). On retiendra que, pour la loi de Poisson, l’écart quadratique est égal à la valeur moyenne : $\Delta N^2 = \langle N \rangle$.

⁸⁴Cette procédure peut paraître cavalière ; en fait, on peut écrire le gradient discret $f(n+1) - f(n)$, puis faire les développements limités adéquats quand $1 \ll n$. C’est d’ailleurs ainsi qu’a procédé Heisenberg lors de sa construction de la Mécanique des matrices (1925).

▼ Dynamique conduisant à une distribution de Poisson

Tout comme pour le déclin radioactif, on peut exhiber un processus stochastique produisant tout naturellement des probabilités de Poisson. Historiquement, ce problème a été traité par Schottky, dans le but d'expliquer le bruit électronique dans ces circuits, en conséquence de l'arrivée aléatoire de charges (bruit de *grenaille*, *shot noise* en anglais). L'hypothèse de base est que la probabilité d'arrivée d'une charge supplémentaire entre t et $t + \delta t$ est égale à $\lambda \delta t$, λ étant *de facto* la probabilité d'arrivée d'une charge par unité de temps⁸⁵, supposée indépendante du temps (hypothèse de *stationnarité*).

Soit $p_n(t)$ la probabilité que n charges soient arrivées entre les instants 0 et t . Le bilan des variations possibles des probabilités entre t et $t + \delta t$ s'établit comme suit :

$$p_n(t + \delta t) = (1 - \lambda \delta t) p_n(t) + \lambda \delta t p_{n-1}(t) + \mathcal{O}((\lambda \delta t)^k \geq 2) . \quad (1.156)$$

Après division par δt et passage à la limite $\delta t \rightarrow 0$, il vient :

$$\boxed{\frac{dp_n}{dt} = -\lambda p_n(t) + \lambda p_{n-1}(t) \quad (n \in \mathbb{N})} \quad (1.157)$$

Ce système d'équations se résout en introduisant à nouveau la fonction génératrice comme en (1.128) :

$$f(s; t) = \sum_{n=0}^{+\infty} s^n p_n(t) , \quad (1.158)$$

qui est ici une *série* (rien ne limite le nombre de particules arrivées), visiblement uniformément convergente $\forall |s| \leq 1$; une fois trouvée la fonction $f(s; t)$, il suffit de regarder son développement en puissances de s pour obtenir $p_n(t)$, qui est le coefficient de s^n .

À partir de (1.157), on voit vite que $f(s; t)$ satisfait l'équation aux dérivées partielles⁸⁶ :

$$\frac{\partial f}{\partial t} = \lambda(s - 1)f(s; t) \quad (1.159)$$

(comparer avec (1.129)). La solution générale est $C(s)e^{\lambda(s-1)t}$. En supposant la population nulle à $t = 0$ ($p_n(0) = \delta_{n0}$, soit $f(s, t = 0) = 1 \forall s$, on a $C(s) = 1$, d'où :

$$f(s; t) = e^{\lambda(s-1)t} ; \quad (1.160)$$

développant alors le second membre en puissances de s , on en déduit $p_n(t)$ par identification :

$$p_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} . \quad (1.161)$$

⁸⁵Le même formalisme décrit aussi une population d'individus augmentant spontanément avec l'arrivée des cigognes apportant chacune un nouveau-né, et en l'absence de décès, c'est-à-dire sur échelle de temps courte devant la durée de vie moyenne d'un individu.

⁸⁶Comme il se doit, la dérivée $\frac{\partial f}{\partial t}$ est nulle $\forall t$ si $s = 1$. Noter aussi que (1.159) est en fait une simple équation *différentielle*.

Il s'agit bien d'une loi de Poisson de paramètre $a = \lambda t$. La moyenne du nombre de charges à l'instant t est λt , et l'écart-type est $\sqrt{\lambda t}$ (voir (1.151) et (1.152)). L'allure de l'évolution temporelle de la distribution se déduit de la fig. 1.16, en remplaçant a par λt . Aux temps courts devant λ^{-1} , la distribution est ramassée près de $n = 0$. Aux grands temps, la moyenne est grande, et la fluctuation relative est $\frac{1}{\sqrt{\lambda t}}$ (voir (1.154)) : elle tend donc vers zéro, mais très lentement. ▲

D'autres lois discrètes jouent un rôle important en pratique. Par exemple, la distribution dite *hypergéométrique* apparaît dans le cas suivant. On a un ensemble de N objets et on s'intéresse à une qualité par tout ou rien de chacun d'entre eux (une pièce a des défauts, ou non, chacun a une étiquette *oui* ou une étiquette *non*) ; on sait que m d'entre eux sont "oui". La probabilité que tirant n objets, k d'entre eux soit "oui" est donnée par :

$$P(k; N, m, n) = \frac{C_m^k C_{N-m}^{n-k}}{C_N^n} \quad (1.162)$$

avec $N \in \mathbb{N}^*$, $0 \leq m \leq N$, $1 \leq n \leq N$; le support de cette loi est $\text{Max}(0, n + m - N)$; la moyenne est $\frac{nm}{N}$, la variance est $\frac{nm(N-m)(N-n)}{N^2(N-1)}$. Si $n = 1$, c'est la distribution de Bernoulli ; si N et m sont grands devant n et si $p \equiv \frac{m}{N}$ n'est pas trop près de 0 et de 1, on retombe approximativement sur la loi binômiale de paramètres n et p . Enfin, si $n \gg 1, N, m \gg n$ et p pas trop près de 0 et 1, la région centrale est à peu près gaussienne, caractérisée par une moyenne $\simeq np$ et une variance $\simeq \sqrt{np(1-p)}$.

Sur des bases purement empiriques, une autre loi, dite de Zipf, est parfois invoquée à propos de la fréquence d'apparition f_k d'un mot K dans une langue donnée ; elle stipule que les f_k décroissent comme une certaine puissance de k dans une classification où le mot le plus fréquent est en premier, le plus fréquent après celui-ci est en second, et ainsi de suite : $f_k \sim k^{-\alpha}$ (si l'on s'autorisait de considérer un nombre infini de mots (!?), il faudrait bien sûr $\alpha > 1$) ; c'est donc typiquement une loi-puissance, tout comme la loi de Pareto (sous-section 1.7.6), l'une de celles qui ont fait un débarquement en force dans de multiples domaines ces dernières décennies. Leur invariance d'échelle⁸⁷ explique leur importance pour les systèmes fractales. La distribution est :

$$P(k; N, \alpha) = \frac{k^{-\alpha}}{\sum_{p=1}^N p^{-\alpha}} \equiv \frac{k^{-\alpha}}{H_{N, \alpha}} \quad (1.164)$$

⁸⁷Une fonction $f(x)$ est dite *invariante d'échelle* si on peut trouver une fonction F telle que :

$$\frac{f(x)}{f(x')} = F\left(\frac{x}{x'}\right) \quad (1.163)$$

quels que soient x et x' ; en supposant toutes les fonctions dérivables, et en dérivant par rapport à x et par rapport à x' cette relation entraîne $x \frac{f'(x)}{f(x)} = x' \frac{f'(x')}{f(x')}$ quels que soient x et x' , ce qui veut dire que la combinaison $x \frac{f'(x)}{f(x)}$ est une certaine constante ν ; l'intégration donne immédiatement $f(x) = Kx^\nu$, montrant l'inéluctabilité de la loi-puissance dès qu'il existe une invariance d'échelle.

où $H_{N,\alpha} \stackrel{\text{déf}}{=} \sum_{p=1}^N p^{-\alpha}$ est appelé N^e nombre harmonique généralisé. Dans la limite $N \rightarrow +\infty$, ces nombres sont égaux à la fonction de Riemann $\zeta(\alpha)$, qui sera étudiée au chapitre 7.

Après avoir cité quelques lois courantes pour une v.a. discrète, on considère maintenant des variables aléatoires continues dont la fonction de répartition $F(x)$ est absolument continue : $\exists F'(x)$ et $F'(x) = \rho(x)$.

1.7.3 Loi uniforme

C'est la plus simple des densités pour une variable aléatoire continue X : on suppose que la probabilité de trouver X au voisinage d'un point quelconque d'un intervalle $[a, b]$ est la même pour tous les points de cet intervalle :

$$\rho(x) = \begin{cases} \text{C}^{\text{ste}} & \text{si } a < x < b \\ 0 & \text{autrement} \end{cases} . \quad (1.165)$$

Comme l'intégrale de ρ doit être normalisée à l'unité, la C^{ste} vaut $\frac{1}{b-a}$. Par exemple, si Φ est un angle que l'on tire uniformément entre $-\pi$ et $+\pi$, $\text{C}^{\text{ste}} = \frac{1}{2\pi}$ et :

$$\langle \Phi \rangle = \int_{-\pi}^{+\pi} \frac{1}{2\pi} \phi \, d\phi = 0 , \quad \langle \Phi^2 \rangle = \int_{-\pi}^{+\pi} \frac{1}{2\pi} \phi^2 \, d\phi = \frac{\pi^2}{3} . \quad (1.166)$$

L'écart-type d'un angle tiré uniformément est donc $\frac{\pi}{\sqrt{3}} \simeq 1.81$ radians. D'une façon générale, pour une variable uniforme X , le rapport entre l'écart-type et la longueur de l'intervalle des valeurs possibles est égal à $\frac{1}{2\sqrt{3}} \simeq 29\%$.

La loi de répartition $F(x)$ associée à la densité uniforme $\rho(x)$ est nulle si $x < a$, égale à 1 si $x \geq b$ et augmente linéairement entre a et b (c'est l'intégrale d'une constante) :

$$F(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x \geq b \end{cases} ; \quad (1.167)$$

c'est bien une fonction continue.

1.7.4 Loi de Gauss

La loi de Gauss est la plus connue des lois de distribution pour une variable aléatoire continue X , sans doute parce que c'est la plus universelle, pour les raisons qui seront expliquées dans la section 1.9. La densité correspondante $\rho_{\text{Gauss}}(x)$ est :

$$\rho_{\text{Gauss}}(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(x-x_0)^2}{2\sigma^2}} \quad (1.168)$$

où x_0 et σ sont deux paramètres indépendants réels positifs. Le facteur $\frac{1}{\sqrt{2\pi}\sigma}$ assure la normalisation à 1 de l'intégrale de ρ : $\int_{-\infty}^{+\infty} \rho_{\text{Gauss}}(x) dx = 1$. On vérifie immédiatement que :

$$\langle X \rangle = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} x e^{-\frac{(x-x_0)^2}{2\sigma^2}} dx = x_0 , \quad (1.169)$$

$$\langle X^2 \rangle = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} x^2 e^{-\frac{(x-x_0)^2}{2\sigma^2}} dx = x_0^2 + \sigma^2 . \quad (1.170)$$

Ceci donne l'interprétation des paramètres apparaissant dans l'expression (1.168) de la gaussienne ; x_0 n'est autre que la valeur moyenne et σ est l'écart-type :

$$\Delta X = \sqrt{\langle X^2 \rangle - \langle X \rangle^2} = \sigma ; \quad (1.171)$$

la variance est donc égale à σ^2 . Le graphe de $\rho_{\text{Gauss}}(x)$ est une courbe en cloche, symétrique autour de x_0 , $\rho_{\text{Gauss}}(x - x_0) = \rho_{\text{Gauss}}(-x + x_0)$; le maximum est d'autant plus grand et la largeur d'autant plus fine que σ est petit⁸⁸, et inversement.

La loi de Gauss présente la particularité suivante : tous les moments M_k , $k \geq 3$, s'expriment en fonction des moments d'ordre 1 et 2, M_1 et M_2 ; ceci est une caractéristique de l'intégrale gaussienne $\int_{-\infty}^{+\infty} x^k e^{-ax^2} dx$, qui joue un rôle fondamental dans les applications⁸⁹. La loi de Gauss possède bien d'autres propriétés remarquables, qui seront données au fur et à mesure dans la suite.

La fonction de répartition gaussienne est par définition :

$$F_{\text{Gauss}}(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(x'-x_0)^2}{2\sigma^2}} dx' . \quad (1.172)$$

$F_{\text{Gauss}}(x)$ a un point d'inflexion en x_0 et s'exprime commodément à l'aide de la fonction⁹⁰ $\Phi(x)$ définie comme :

$$\Phi(x) \stackrel{\text{déf}}{=} \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy , \quad \Phi(\pm\infty) = \pm 1 ; \quad (1.173)$$

on trouve :

$$\boxed{F_{\text{Gauss}}(x) = \frac{1}{2} \left[1 + \Phi\left(\frac{x-x_0}{\sqrt{2}\sigma}\right) \right]} \quad (1.174)$$

$F_{\text{Gauss}}(x)$ se rapproche de ses valeurs limites avec un écart $\propto \frac{\sigma}{|x|} e^{-\frac{x^2}{2\sigma^2}}$. La fonction $1 - \Phi(x)$ est aussi notée $\text{erfc}(x)$.

⁸⁸La fonction de Gauss est un précurseur de la fonction de Dirac $\delta(x - x_0)$, vers laquelle elle tend dans la limite $\sigma \rightarrow 0$ – voir chapitre 2.

⁸⁹Notamment pour la description des processus stochastiques gaussiens. C'est aussi grâce à cette propriété que l'on peut "démontrer" l'équation de la diffusion et, plus généralement, établir l'équation fondamentale de Fokker - Planck.

⁹⁰souvent appelée *fonction erreur*, notée aussi $\text{Erf}(x)$, mais dont la définition précise peut varier d'un auteur à l'autre.

On verra par la suite (section 1.9) en quoi la loi de Gauss (dite *normale*) a un caractère universel et se rencontre très fréquemment dans la pratique... pas toujours d'ailleurs en raison du Théorème limite central : le profil (en fréquence) d'une raie atomique émise par une vapeur en équilibre thermodynamique est essentiellement gaussienne – trouver pourquoi –, une raison qui n'a strictement rien à voir avec le fait que la gaussienne est une loi attractive !

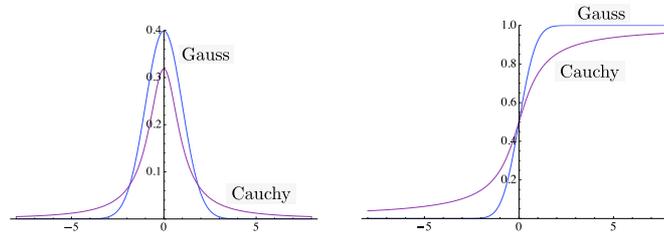


Figure 1.17: Comparaison des lois de Gauss et de Cauchy de moyennes nulles, de même largeur (prise égale à 1), cette dernière étant σ (écart-type) pour la gaussienne, a (voir (1.175)) pour la lorentzienne (dont l'écart-type est infini). À gauche : densités ; la lorentzienne est vaguement triangulaire, la gaussienne est beaucoup plus ronde et plus ramassée. À droite : fonctions de répartition ; la gaussienne est nettement plus raide que la lorentzienne et sature très vite.

1.7.5 Loi de Cauchy

Appelée *lorentzienne* par les Physiciens, cette loi a pour densité :

$$\rho_{\text{Cauchy}}(x) = \frac{1}{\pi} \frac{a}{(x - x_0)^2 + a^2} \quad (1.175)$$

avec $a \in \mathbb{R}_+$. À nouveau, $\rho_{\text{Cauchy}}(x)$ est symétrique de part et d'autre de x_0 – de toute évidence, la moyenne de X est x_0 . Cette loi a une caractéristique remarquable : la variance est *infinie* ! En effet :

$$\Delta X^2 \equiv \langle X^2 \rangle - \langle X \rangle^2 = \langle (X - \langle X \rangle)^2 \rangle = \frac{a}{\pi} \int_{-\infty}^{+\infty} \frac{(x - x_0)^2}{(x - x_0)^2 + a^2} dx = +\infty, \quad (1.176)$$

en conséquence du fait que $\rho_{\text{Cauchy}}(x)$ décroît trop lentement à l'infini⁹¹, comme x^{-2} . De surcroît, quoique l'écart-type ΔX est infini, la largeur (géométrique) est visiblement

⁹¹En fait, aucun moment n'existe à proprement parler, sauf les moments impairs de la variable centrée $X - x_0$ et à condition de les régulariser en écrivant :

$$\tilde{M}_{2k+1} \stackrel{\text{d\'ef}}{=} \langle (X - x_0)^{2k+1} \rangle = \lim_{A \rightarrow \infty} \frac{a}{\pi} \int_{-A}^{+A} \frac{y^{2k+1}}{y^2 + a^2} dy = 0; \quad (1.177)$$

c'est à ce prix qu'il est légitime de dire notamment que la moyenne de X avec la loi (1.175) est égale à x_0 ...

finie : en la définissant conventionnellement comme la largeur à mi-hauteur, elle vaut $2a$ puisque $\rho_{\text{Cauchy}}(x_0 \pm a) = \frac{1}{2}\rho_{\text{Cauchy}}(x_0)$.

La loi de Cauchy est l'exemple le plus simple de loi dite *large*, au sens où les valeurs très éloignées de la moyenne (ici égale à x_0) ont une probabilité qui n'est pas si petite que cela : d'où la *pertinence des événements rares*, qui joue un rôle important en Physique, notamment pour la relaxation de certains systèmes complexes, ou la dynamique au sein des milieux désordonnés.

La fonction de répartition de Cauchy est :

$$F_{\text{Cauchy}}(x) = \int_{-\infty}^x \frac{a}{\pi} \frac{1}{(x' - x_0)^2 + a^2} dx' = \frac{1}{2} + \frac{1}{\pi} \text{Arctan} \frac{x - x_0}{a} \quad (1.178)$$

Elle se rapproche de ses valeurs limites avec un écart variant comme $\frac{1}{|x|}$.

Exemple où l'on trouve précisément une densité de Cauchy⁹² : les valeurs de l'aléatoire $Y = \tan X$ quand X est uniformément répartie dans l'intervalle $]-\frac{\pi}{2}, +\frac{\pi}{2}[$. En effet, si F_X et F_Y sont les fonctions de répartition correspondantes, on a par définition⁹³ :

$$\text{Prob}[Y \leq y] = \text{Prob}[X \leq x] \quad \text{sachant que } y = \tan x \iff x = \text{Arctg } y, \quad (1.179)$$

soit :

$$F_Y(y) = F_X(x = \text{Arctg } y). \quad (1.180)$$

Si $F_X(x)$ est absolument continue et si la densité correspondante est $\rho_X(x)$, alors la probabilité $\text{Prob}[X \leq x]$ est l'intégrale de $\rho_X(x)$ de $-\frac{\pi}{2}$ jusqu'à y . Au total :

$$F_Y(y) = \text{Prob}[X \leq x = \text{Arctg } y] = \int_{-\frac{\pi}{2}}^{\text{Arctg } y} \rho_X(x) dx. \quad (1.181)$$

Si X est répartie *uniformément* entre $\pm\frac{\pi}{2}$, alors $\rho_X(x)$ est une constante et vaut $\frac{1}{\pi}$. En définitive :

$$F_Y(y) = \int_{-\frac{\pi}{2}}^{\text{Arctg } y} \frac{1}{\pi} dx = \frac{1}{\pi} \left(\frac{\pi}{2} + \text{Arctg } y \right). \quad (1.182)$$

Il en résulte que Y admet aussi exclusivement une fonction de répartition absolument continue, dont la densité est :

$$\rho_Y(y) = \frac{dF_Y}{dy} = \frac{1}{\pi} \frac{1}{y^2 + 1}, \quad (1.183)$$

qui est une loi de Cauchy centrée de largeur $a = 1$.

⁹²La lorentzienne est aussi le profil en fréquence d'une raie atomique quand seule la largeur naturelle est en jeu – pourquoi ?

⁹³Noter que cette relation simple résulte du fait que la fonction $x \rightarrow \tan x = y$ est *monotone non-décroissante* dans l'intervalle considéré pour x .

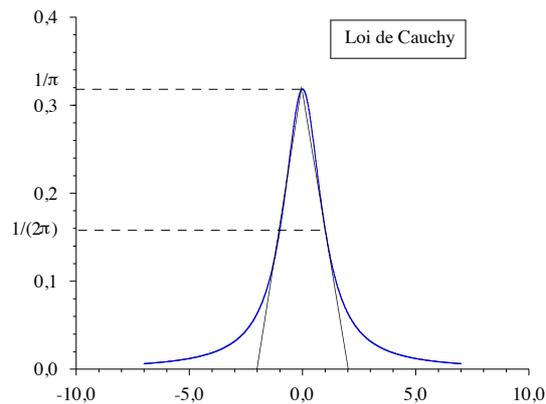


Figure 1.18: Illustration de la propriété géométrique de la densité de Cauchy ($a = 1$) : les deux tangentes à mi-hauteur (en $x = \pm a$) se coupent au sommet, et traversent l'axe des abscisses en $\pm 2a$ (le dessin est tracé avec $a = 1$).

Signalons une propriété géométrique intéressante du graphe de la loi de Cauchy, illustrée sur la fig. 1.18 : pour tracer semi-quantitativement cette courbe, on part d'un triangle isocèle de base $4a$ et de hauteur $\frac{1}{\pi a} \simeq \frac{1}{3a}$ et, partant du sommet, on fait "tangenter" la courbe aux milieux des deux autres côtés du triangle. Ceci permet de distinguer à l'œil une loi de Cauchy d'une loi de Gauss : la lorentzienne a une allure triangulaire alors que la gaussienne est plutôt ronde. Ces deux distributions se démarquent aussi fortement l'une de l'autre par l'aspect des ailes : la lorentzienne traîne en longueur alors que la gaussienne s'annule très vite.

1.7.6 Loi de Pareto

La loi de Pareto est le prototype de la loi-puissance. Elle fut proposée par Pareto pour décrire le partage de la richesse dans une population donnée ; depuis, on la retrouve dans des domaines très divers, de la "phynance" à la sociologie. L'aléatoire X étant supposée prendre des valeurs supérieures à un certain seuil x_0 , la fonction de répartition est :

$$F_{\text{Pareto}}(x) = \begin{cases} 0 & \forall x \leq x_0 \\ 1 - \left(\frac{x_0}{x}\right)^\alpha & \forall x \geq x_0 \end{cases} \quad (1.184)$$

avec $\alpha > 0$. La densité correspondante est donc :

$$\rho_{\text{Pareto}}(x) = \begin{cases} 0 & \forall x < x_0 \\ \alpha \frac{x_0^\alpha}{x^{\alpha+1}} & \forall x > x_0 \end{cases} \quad (1.185)$$

La valeur moyenne est $\frac{\alpha}{\alpha-1}x_0$, la variance est $\frac{\alpha}{(\alpha-1)^2(\alpha-2)}x_0^2$. Ces expressions attestent d'une évidence : si α est petit, la queue de la distribution fait diverger les premiers

moments (loi large). Si $\alpha \leq 2$, la variance n'existe pas ; pour $\alpha \leq 1$, il n'y a même pas de valeur moyenne.

Pour terminer cette section, mentionnons d'autres lois assez fréquentes⁹⁴ :

- loi Γ : $\rho(x) = C x^\alpha e^{-kx}$, $x \geq 0$, $\alpha > -1$;
- loi de Maxwell : $\rho(x) = C e^{-\lambda x^2}$, ainsi appelée pour des raisons évidentes ;
- loi Beta⁹⁵ : $\rho(x) = C x^\alpha (a-x)^\beta$, $0 < x < a$, $\alpha, \beta > -1$;
- loi de Laplace : $\rho(x) = C e^{-k|x|}$ (qui est essentiellement la transformée de Fourier d'une lorentzienne⁹⁶).
- loi de Gumbell : $\rho(x) = C e^{x-e^x}$, qui apparaît dans le calcul des corrections de taille finie pour l'énergie libre de certains systèmes désordonnés.

1.8 Fonction caractéristique

1.8.1 Définitions et propriétés

Par définition, et en adoptant les notations générales, la *fonction caractéristique* $\phi(t)$ d'une loi de probabilité $p(x)$ est la valeur moyenne de la quantité⁹⁷ e^{itX} :

$$\boxed{\phi(t) \stackrel{\text{d\u00e9f}}{=} E(e^{itX}) \equiv \langle e^{itX} \rangle} \quad (1.186)$$

Une première conséquence est que :

$$\boxed{\phi(0) = E(1) = 1} \quad (1.187)$$

et bien sûr $\phi(-t) = \phi^*(t)$, $\forall t \in \mathbb{R}$.

En l'absence de composante singulière continue (une hypothèse acceptable pour la plupart des besoins), et en jouant avec la règle opérationnelle de la fonction de Dirac, il est équivalent de dire que $\phi(t)$ est la transformée de Fourier de la densité généralisée $p(x)$ définie en (1.68) p. 32 :

$$\boxed{\phi(t) = \int_{-\infty}^{+\infty} p(x) e^{itx} dx \iff p \xrightarrow{\mathcal{F}} \phi = \mathcal{F}[p]} \quad (1.188)$$

⁹⁴ C désigne à chaque fois la constante de normalisation telle que l'intégrale de $\rho(x)$ soit égale à l'unité. Pour la loi Γ , on trouve $C = \frac{k^{\alpha+1}}{\Gamma(\alpha+1)}$, pour la loi de Maxwell, $C = \sqrt{\frac{\lambda}{\pi}}$, etc.

⁹⁵Pourquoi doit-on avoir $\alpha, \beta > -1$?

⁹⁶En conséquence, la fonction caractéristique (définie dans la sous-section 1.8.1) de la loi de Laplace est une lorentzienne.

⁹⁷En posant $it = -\beta H$, $\phi(t)$ prend l'allure de ce que le physicien appelle une *fonction de partition*. En cas de nécessité, la fonction $\phi(t)$ peut être prolongée analytiquement ; ici, on supposera toujours $t \in \mathbb{R}$.

Cette écriture explicite permet de voir clairement que ϕ existe $\forall t \in \mathbb{R}$ puisque $p(x)$ est une fonction positive dont l'intégrale est finie :

$$\forall t \in \mathbb{R} : |\phi(t)| \leq \int_{-\infty}^{+\infty} |e^{itx} p(x)| dx = \int_{-\infty}^{+\infty} p(x) dx = 1 ; \quad (1.189)$$

avec les conventions adoptées (voir (1.68)), l'expression explicite de $\phi(t)$ est :

$$\phi(t) = \sum_n p_n e^{itx_n} + \int_{-\infty}^{+\infty} \rho(x) e^{itx} dx ; \quad (1.190)$$

la normalisation de $p(x)$ assure la satisfaction de (1.187).

La fonction caractéristique possède des propriétés remarquables, qui en font un outil de toute première utilité (pour une revue, voir l'article de Lukas [26]). Par exemple, si $\phi(t_0) = 1$ pour un certain nombre $t_0 \in \mathbb{R}$, alors la v.a. correspondante X ne peut prendre que des valeurs multiples entiers de $\frac{2\pi}{t_0}$. En effet, si on prend la partie réelle de la relation de définition (1.186), il vient $E(1 - \cos t_0 X) = 0$; comme $1 - \cos t_0 X \geq 0$, ceci donne $\text{Prob}[1 - \cos t_0 X = 0] = 1$, soit $\text{Prob}[X \text{ est divisible par } \frac{2\pi}{t_0}] = 1$.

En outre, $\phi(t)$ est une fonction uniformément continue⁹⁸ :

$$|\phi(t_2) - \phi(t_1)| = |E(e^{it_2 X} - e^{it_1 X})| \leq E(|e^{it_2 X} - e^{it_1 X}|) = E(|e^{i(t_2-t_1)X} - 1|) . \quad (1.191)$$

Le module de la différence de deux complexes est borné par la somme des modules ; la quantité entre parenthèses est donc bornée par 2. Le théorème de la convergence dominée permet de passer à la limite sous l'intégrale, donnant quels que soient t_1 et t_2 :

$$\lim_{t_2 \rightarrow t_1} |\phi(t_2) - \phi(t_1)| = 0 . \quad (1.192)$$

La plupart des propriétés s'énoncent simplement avec l'écriture en transformée de Fourier, et s'appuient sur le fait que l'original est une fonction *positive* (et intégrable).

Il existe une bijection entre la fonction de répartition $F(x)$ et la fonction caractéristique, qui s'exprime par le *théorème d'inversion* :

$$F(x + \Delta) - F(x) = \lim_{T \rightarrow +\infty} \frac{1}{2\pi} \int_{-T}^{+T} \frac{1 - e^{-i\Delta t}}{it} e^{-itx} \phi(t) dt . \quad (1.193)$$

La démonstration peut être trouvée dans l'ouvrage de Fisz [23], p. 116 (ou dans [24], p. 324, théorème 12.7). En l'absence de composante atomique, la biunivocité est assurée par le théorème d'inversion de Fourier (voir chapitre 8).

⁹⁸Une fonction f est dite *uniformément continue* ssi, quel que soit $\varepsilon > 0$ donné, il existe $\delta > 0$ assurant que si la distance de deux points *quelconques* x_1 et x_2 est plus petite que δ , la distance des images $f(x_1)$, $f(x_2)$ est inférieure à ε ([24], § 2.3).

À titre d'exemple, les fonctions caractéristiques des lois présentées ou citées ci-dessus sont⁹⁹ :

$$\phi(t) = \begin{cases} 1 + p(e^{it} - 1) & \text{(loi de Bernoulli)} \\ [1 + p(e^{it} - 1)]^{N_0} & \text{(loi binomiale)} \\ e^{a(e^{it} - 1)} & \text{(loi de Poisson)} \\ \frac{1}{H_{N,\alpha}} \sum_{n=1}^N n^{-\alpha} e^{int} & \text{(loi de Zipf)} \\ e^{ix_0 - \frac{1}{2}t^2\sigma^2} & \text{(loi de Gauss)} \\ \frac{(C_{N-m}^n / C_N^n) {}_2F_1(-n, -m; N - m - n + 1; e^{it})}{e^{ix_0 - |t|a}} & \text{(loi hypergéométrique)} \\ \alpha(-ix_0t)^\alpha \Gamma(-\alpha, -ix_0t) & \text{(loi de Cauchy)} \\ & \text{(loi de Pareto)} \end{cases}, \quad (1.194)$$

Quant à la distribution de Cantor - Lebesgue, sa fonction caractéristique est :

$$\phi_C(t) = e^{it/2} \prod_{n \in \mathbb{N}^*} \cos \frac{t}{3^n} \quad (1.195)$$

Ce résultat peut s'obtenir en examinant le point fixe de la récurrence $\phi_n(t) \rightarrow \phi_{n+1}(t)$ déduite de (1.89).

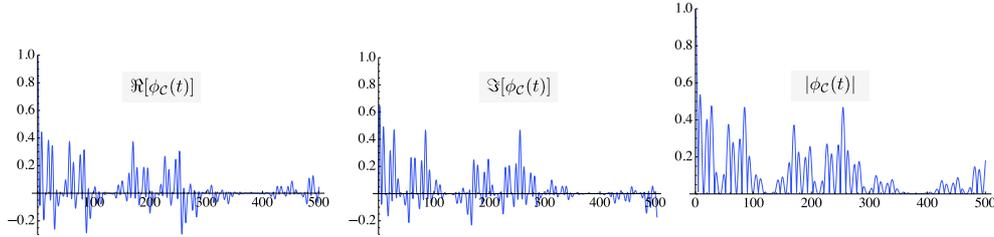


Figure 1.19: Fonction caractéristique $\phi_C(t)$ de la distribution de Cantor - Lebesgue (voir (1.195)).

Si dans la définition (1.186) on développe e^{itx} en série entière, et que l'on échange formellement sommation et moyenne, on obtient :

$$\phi(t) = \sum_{k=0}^{+\infty} \frac{(it)^k}{k!} \langle X^k \rangle ; \quad (1.196)$$

admettant provisoirement le caractère licite de ces opérations, on voit que les moyennes des puissances entières de l'aléatoire se lisent sur le développement en série entière de $\phi(t)$. Ceci justifie l'appellation de $\phi(t)$ comme *fonction génératrice des moments* (entiers), $M_k \stackrel{\text{déf}}{=} \langle X^k \rangle$, puisque quand ce développement existe, on a immédiatement :

$$M_k = i^{-k} \left[\frac{d^k}{dt^k} \phi(t) \right]_{t=0} \quad (1.197)$$

⁹⁹La notation $\Gamma(\alpha, x)$ désigne traditionnellement la fonction Gamma incomplète d'Euler, ${}_2F_1(\alpha, \beta, x)$ est une fonction hypergéométrique.

La légitimité de l'échange entre moyenne et sommation repose sur le théorème suivant. Si pour un entier n , $\langle |X|^n \rangle < \infty$, alors

$$\phi(t) = \sum_{k=0}^n \frac{(it)^k}{k!} \langle X^k \rangle + t^n \varepsilon_n(t) , \quad (1.198)$$

où $\varepsilon_n(t)$ tend vers zéro quand $t \rightarrow 0$. La démonstration peut être trouvée p. 324 de l'ouvrage [24] de Kingman et Taylor (leur théorème 12.6). Dans le cas absolument continu, et parce qu'elles forment un couple de Fourier, les fonctions $\rho(x)$ et $\phi(t)$ ont des propriétés duales l'une de l'autre. Ainsi, si $\rho(x)$ possède des moments $M_{k>0}$ de tous les ordres, alors $\phi(t)$ possède un développement en série entière centré en $t = 0$:

$$\boxed{\phi(t) = \sum_{k=0}^{+\infty} \frac{(it)^k}{k!} \int_{-\infty}^{+\infty} x^k \rho(x) dx \equiv \sum_{k=0}^{+\infty} \frac{(it)^k}{k!} M_k} \quad (1.199)$$

Il en va ainsi pour la gaussienne, pour laquelle il est maintenant facile d'établir la propriété remarquable annoncée plus haut (tous les moments s'expriment à l'aide des deux premiers). Prenons une variable gaussienne centrée (de moyenne nulle) ; le calcul direct de l'intégrale (1.188) avec $\rho(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$ donne $\phi(t) = e^{-\frac{1}{2}t^2\sigma^2}$, d'où :

$$e^{-\frac{1}{2}t^2\sigma^2} = \sum_{k=0}^{+\infty} \frac{(it)^k}{k!} M_k ; \quad (1.200)$$

en développant en série le premier membre, l'identification donne¹⁰⁰ :

$$\sum_{l=0}^{+\infty} \frac{1}{l!} \left(-\frac{1}{2}t^2\sigma^2 \right)^l = \sum_{k=0}^{+\infty} \frac{(it)^k}{k!} M_k \iff M_{2k+1} = 0 , \quad M_{2k} = \frac{(2k)!}{2^k k!} \sigma^{2k} . \quad (1.201)$$

Ainsi, pour une variable gaussienne centrée, tous les moments pairs s'expriment simplement à l'aide de $\sigma^2 \equiv M_2$. Le point important est que M_{2k} est simplement proportionnel à la k^e puissance de σ^2 : cette propriété joue un rôle majeur dans l'étude des processus stochastiques dits à *diffusion lente*¹⁰¹. Dans le cadre quantique de la Seconde quantification, cette propriété est le fondement du Théorème de Wick, qui joue un rôle de tout premier plan (Théorie quantique des champs, Problème à N -corps,...). Pour la référence ultérieure¹⁰², retenons que pour une v.a. gaussienne centrée :

$$M_2 = \sigma^2 , \quad M_4 = 3\sigma^4 , \quad M_6 = 15\sigma^6 . \quad (1.202)$$

À l'inverse, la loi de Cauchy n'a aucun moment (sauf les moments impairs, et à condition de les régulariser comme indiqué dans la note 91) ; de fait, sa fonction caractéristique

¹⁰⁰La formule de Stirling (voir (1.147)) montre que pour $k \gg 1$, $M_{2k} \simeq \sqrt{2} \left(\frac{2k}{e}\right)^k \sigma^{2k}$, qui augmente vertigineusement avec k , quel que soit σ .

¹⁰¹C'est cette propriété qui conduit à l'équation de Fokker - Planck, une des équations fondamentales permettant de décrire la dynamique hors d'équilibre d'un système macroscopique.

¹⁰²voir sous-section 1.9.2 et notamment la fig. 1.20.

$e^{-|t|^a}$ n'est pas analytique en $t = 0$: si on peut certes toujours écrire $e^{-|t|^a}$ comme la série entière $\sum_{n \in \mathbb{N}} \frac{(-a)^n}{n!} |t|^n$, il ne s'agit pas d'un développement de Taylor !

Noter cependant que si $\phi(t)$ existe toujours, ce n'est pas forcément le cas des M_k : pour la loi de Cauchy, $\phi = e^{-a|t|}$ alors qu'aucun moment n'existe. À l'inverse, s'agissant de déterminer la distribution de probabilités d'un phénomène donné, une procédure standard consiste à mesurer, ou calculer, les différents moments M_k afin de bâtir peu à peu la fonction caractéristique $\phi(t)$. Cette *reconstruction* à partir des moments est une opération délicate, surtout quand on ne dispose pas d'information *a priori* sur les propriétés analytiques de $\phi(t)$ spécifiques au problème analysé. Dans tout traitement par approximation, il convient de respecter les propriétés intrinsèques de $\phi(t)$ autant que faire se peut, afin de s'éviter des déboires : un schéma approximatif mal maîtrisé peut conduire à des ennuis, dont le plus spectaculaire est l'apparition d'une densité ... négative.

Il est aussi parfois très utile de considérer le logarithme de $\phi(t)$, $\psi(t) \stackrel{\text{déf}}{=} \ln \phi(t)$, soit¹⁰³ :

$$\boxed{e^{\psi(t)} \stackrel{\text{déf}}{=} \langle e^{itX} \rangle} \quad (1.204)$$

Quand il existe¹⁰⁴, le développement de Taylor de $\psi(t)$ fait apparaître, par définition, les *cumulants* C_k :

$$\boxed{\psi(t) = \sum_{k=1}^{+\infty} \frac{(it)^k}{k!} C_k} \quad (1.205)$$

et c'est pourquoi $\psi(t)$ est aussi appelée, dans ce cas, *fonction génératrice des cumulants*. En identifiant les développements, on voit que les C_k sont certaines fonctions des moments $M_{k'}$:

$$\boxed{C_1 = M_1, \quad C_2 = M_2 - M_1^2, \quad C_3 = M_3 - 3M_1M_2 + 2M_1^3, \dots} \quad (1.206)$$

Noter que la somme des coefficients figurant dans l'expression de $C_{k \geq 2}$ est nulle comme il se doit, puisque pour une variable certaine, tous les cumulants $C_{k \geq 2}$ sont visiblement nuls et qu'alors $M_k^i = M_1^{ki}$.

Dans certains problèmes, les cumulants sont les *bonnes* quantités décrivant globalement la dispersion des valeurs d'une variable aléatoire. Pour une distribution gaussienne, tous les cumulants sont nuls au-delà du second ($C_1 = x_0$, $C_2 = \sigma^2$, $C_{k \geq 3} = 0$) ; pour une variable certaine, tous les cumulants sont nuls (sauf le premier, égal à la valeur certaine). Si on considère le nombre de cumulants non-nuls comme une mesure de l'aléatoire perçu qualitativement, la gaussienne apparaît à cette aune comme une caractéristique de

¹⁰³De ce point de vue, *mutatis mutandis*, ψ ressemble à une énergie libre, cependant que $\langle e^{itX} \rangle$ joue le rôle d'une fonction de partition. Le logarithme complexe $\ln z$ désigne la branche qui vaut 0 si $z = 1$, et qui satisfait $\ln(z^*) = (\ln z)^*$. En effet, avec $t \in \mathbb{R}$, $\phi^*(t) = \phi(-t)$, d'où :

$$\ln \phi^*(t) = \ln \phi(-t) = \psi(-t) = \psi^*(t) \equiv (\ln \phi(t))^* . \quad (1.203)$$

¹⁰⁴Pour la loi de Cauchy (1.175), $\phi(t) = e^{ix_0 t - a|t|}$ n'a pas de développement de Taylor centré en $t = 0$.

l'aléatoire *minimum*, du plus *faible* hasard, du plus *petit* désordre : c'est la distribution de probabilité la moins "hasardeuse".

Citons enfin le théorème de Marcinkiewicz [25], qui affirme que $\psi(t)$ est soit un polynôme du second degré (pour la gaussienne), soit une *série*. Tronquer cette série dans un cas donné, en guise d'approximation dans un problème compliqué, expose à de sérieux déboires, comme par exemple trouver des probabilités négatives (par Fourier inverse de la fonction caractéristique *approchée* qui en résulte).

1.8.2 Somme de variables aléatoires

Il s'agit ici d'examiner les propriétés élémentaires de la somme de plusieurs variables aléatoires, dans l'hypothèse cruciale où celles-ci sont *indépendantes* – l'extrême importance de cette hypothèse apparaîtra très vite. Cette question permettra de mieux saisir encore la puissance de l'outil que constitue la fonction caractéristique. Soit X_1 et X_2 deux telles v.a. dont les fonctions de répartition sont F_1 et F_2 , et soit $X = X_1 + X_2$ leur somme, qui est de toute évidence une variable aléatoire. La question est de trouver la fonction de répartition de X , $F(x)$.

Afin de progresser graduellement, considérons en premier le cas de deux v.a. discrètes et, pour simplifier les écritures, supposons que chaque X_i prend N valeurs distinctes que l'on représente simplement par un entier n , $1 \leq n \leq N$. Notons $p_{i,n}$ la probabilité $\text{Prob}[X_i = n]$. La somme $X = X_1 + X_2$ prend toutes les valeurs entières entre 2 et $2N$; la probabilité pour que X soit égale à m s'obtient en faisant l'inventaire de tous les événements exclusifs donnant la valeur m pour X , et en faisant la somme de leurs probabilités. X est égal à m si X_1 prend la valeur 1 et si X_2 vaut $m - 1$, ou si X_1 prend la valeur 2 et si X_2 vaut $m - 2$, et ainsi de suite. Comme les variables X_i sont indépendantes, la probabilité de chacun des événements ($X_1 = k$, $X_2 = m - k$) est le *produit* des probabilités individuelles, soit $p_{1,k}p_{2,m-k}$. En définitive, la probabilité cherchée est :

$$\text{Prob}[X = m] = \sum_{k=1}^{m-1} p_{1,k}p_{2,m-k} \quad (2 \leq m \leq 2N) \quad (1.207)$$

Notons que ce résultat, qui a la forme d'une convolution discrète, peut s'obtenir d'une façon plus systématique en sommant sur *tous* les événements possibles de l'espace produit $\Omega^2 = \{1, 2, \dots, N\}^2$, mais en les *filtrant* avec le symbole de Kronecker afin de ne retenir que ceux donnant la valeur donnée m pour la somme $X_1 + X_2$:

$$\text{Prob}[X = m] = \sum_{n_1=1}^N \sum_{n_2=1}^N p_{1,n_1}p_{2,n_2} \delta_{n_1+n_2,m} . \quad (1.208)$$

Considérons maintenant le cas de deux variables admettant chacune une fonction de répartition absolument continue, $F_i(x)$, associée à une densité $\rho_i(x) \equiv F_i'(x)$. Suivant

une démarche analogue, la probabilité que X soit inférieure ou égale à x est égale à la somme des probabilités des événements exclusifs où à la fois $X_1 \leq x - x'$ et X_2 est compris entre x' et $x' + \delta x'$ dans la limite $\delta x' \rightarrow 0_+$:

$$\text{Prob}[X \leq x] = \text{Somme sur } x' \{ \text{Prob}[X_1 \leq x - x' \text{ et } x' < X_2 \leq x' + \delta x'] \}_{\delta x' \rightarrow 0} ; \quad (1.209)$$

cette dernière probabilité s'obtient encore par un simple produit de probabilités puisque les deux variables sont indépendantes par hypothèse ; de la sorte :

$$\text{Prob}[X \leq x] = \lim_{\delta x' \rightarrow 0_+} \sum_{x'} \{ \text{Prob}[X_1 \leq x - x'] \times \text{Prob}[x' < X_2 \leq x' + \delta x'] \} . \quad (1.210)$$

La limite $\lim_{\delta x' \rightarrow 0_+} \sum_{x'}$ n'est autre que l'intégrale de Riemann ; par ailleurs, le facteur $\text{Prob}[x' < X_2 \leq x' + \delta x'] = F_2(x' + \delta x') - F_2(x')$ donne $F_2'(x') dx'$ à la limite, d'où :

$$F(x) = \int F_1(x - x') F_2'(x') dx' \equiv (F_1 * F_2')(x) . \quad (1.211)$$

Il est clair que dans le raisonnement conduisant à (1.210), les rôles de X_1 et de X_2 peuvent être échangés : on peut tout autant dire que la probabilité pour que X soit inférieure ou égale à x est égale à la somme des probabilités des événements où à la fois $X_2 \leq x - x'$ et X_1 est compris entre x' et $x' + dx'$. Au total :

$$\boxed{F(x) = \int F_1(x - x') F_2'(x') dx' = \int F_2(x - x') F_1'(x') dx'} \quad (1.212)$$

Ainsi, F est la convolution de l'une des F_j avec la dérivée de l'autre.

Une fois trouvée la fonction de répartition $F(x)$, la question de sa densité associée se règle en considérant la limite :

$$\rho(x) = \lim_{\Delta x \rightarrow 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} ; \quad (1.213)$$

revenant à l'expression (1.211), et puisque $F_1(x)$ est continue et dérivable, la limite ci-dessus existe et n'est autre que $\int F_1'(x - x') F_2'(x') dx'$; notant $\rho(x) \equiv F'(x)$, il vient :

$$\boxed{\rho(x) = \int \rho_1(x') \rho_2(x - x') dx' \equiv (\rho_1 * \rho_2)(x)} \quad (1.214)$$

Ainsi, en raison de l'indépendance des variables, $\rho(x)$ ressort comme la convolution des densités ρ_j . Pour une somme de N variables $X = \sum_{i=1}^N X_i$, on a de même :

$$\rho(x) = \int dx_2 \int dx_3 \dots \int dx_N \rho_1(x - x_2 - x_3 - \dots - x_N) \prod_{i=2}^N \rho_i(x_i) \quad (1.215)$$

La fonction caractéristique $\phi(t)$ d'une somme de v.a. indépendantes a une forme remarquablement simple. Pour la somme $X = \sum_{i=1}^N X_i$, elle s'obtient immédiatement en partant de sa définition :

$$\phi_X(t) \stackrel{\text{déf}}{=} \langle e^{itX} \rangle = \langle e^{it(X_1 + X_2 + \dots + X_N)} \rangle = \langle e^{itX_1} e^{itX_2} \dots e^{itX_N} \rangle ; \quad (1.216)$$

par hypothèse, les variables X_j sont indépendantes, assurant que la loi conjointe à N variables est le simple produit des lois individuelles :

$$p(x_1, x_2, \dots, x_N) = p_1(x_1)p_2(x_2) \dots p_N(x_N) ; \quad (1.217)$$

il en résulte que les moyennes se factorisent dans (1.216) :

$$\phi_X(t) = \prod_{j=1}^N \langle e^{itX_j} \rangle \equiv \prod_{j=1}^N \phi_j(t) \quad (1.218)$$

En particulier, si toutes les variables indépendantes sont distribuées suivant la même loi $p_0(x)$, de fonction caractéristique $\phi_0(t)$, alors, suivant (1.218), la somme est distribuée suivant la loi ayant la fonction caractéristique $\phi_X(t)$ telle que :

$$\phi_X(t) = [\phi_0(t)]^N \quad (1.219)$$

C'est pour cette raison que, dans (1.194), la fonction caractéristique de la loi binomiale est la puissance N_0^e de la fonction caractéristique de la loi de Bernoulli : on sait qu'une variable binomiale prenant les valeurs $1, 2, \dots, N_0$ est la somme de N_0 variables de Bernoulli indépendantes prenant les valeurs 0 et 1.

Pour terminer, signalons une façon expéditive d'établir l'expression de la densité $\rho(x)$ de la somme de deux (ou plus) variables aléatoires indépendantes. Par analogie avec ce qui a conduit à (1.208), et en acceptant le jeu formel avec la fonction de Dirac jouant le rôle d'un symbole de Kronecker continu, il suffit de filtrer toutes les possibilités indépendantes de x_1 et de x_2 avec $\delta(x - x_1 - x_2)$, ne retenant ainsi que les termes correspondants à une valeur donnée x pour la somme $X = X_1 + X_2$. Appliquant la règle opérationnelle de $\delta(x)$, il vient :

$$\rho(x) = \int dx_1 \int dx_2 \rho_1(x_1)\rho_2(x_2)\delta(x - (x_1 + x_2)) = \int \rho_1(x_1)\rho_2(x - x_1) dx_1 . \quad (1.220)$$

Pour la somme de N v.a., cette façon de faire est un moyen mnémotechnique de retrouver rapidement le résultat (1.215). L'expression (1.220) est visiblement le pendant de (1.208) pour le *continuum*, la "fonction" de Dirac jouant le rôle d'un symbole de Kronecker continu.

1.8.3 Stabilité d'une loi par l'addition

Le résultat exprimé par (1.218) est d'une extrême importance (mais ne jamais oublier qu'il ne tient que pour des variables *indépendantes*). Examinons ce qu'il implique pour quelques lois présentées ci-dessus, en se référant aux expressions données en (1.194). Prenons d'abord deux variables de Poisson X_1 et X_2 , de paramètres respectifs a_1 et a_2 . La fonction caractéristique de leur somme $X = X_1 + X_2$ est :

$$\phi_X(t) = e^{a_1(e^{it}-1)} e^{a_2(e^{it}-1)} = e^{(a_1+a_2)(e^{it}-1)} . \quad (1.221)$$

C'est encore la fonction caractéristique d'une loi de Poisson. Il en résulte que la somme suit une loi de Poisson, de paramètre $a = a_1 + a_2$ – une fois établi le fait que la somme suit une loi de Poisson, on a forcément $a = a_1 + a_2$ puisque a est la valeur moyenne.

Pour deux variables gaussiennes X_1 et X_2 de moyennes respectives x_{01} et x_{02} , et d'écart-types respectifs σ_1 et σ_2 , (1.194) dit que la fonction caractéristique de la somme $X = X_1 + X_2$ est $e^{itx_{01} - \frac{1}{2}t^2\sigma_1^2} e^{itx_{02} - \frac{1}{2}t^2\sigma_2^2} = e^{it(x_{01}+x_{02}) - \frac{1}{2}t^2(\sigma_1^2+\sigma_2^2)}$: c'est encore une fonction caractéristique de gaussienne, de moyenne $x_{01}+x_{02}$ (ce qui n'est pas surprenant), et de variance égale à la somme des variances $\sigma_1^2 + \sigma_2^2$ (ce qui ne l'est pas davantage)¹⁰⁵. Enfin, pour deux variables de Cauchy, la conclusion est encore la même : leur somme obéit à une loi de Cauchy de paramètre a égal à la somme des paramètres relatifs à X_1 et X_2 . On retiendra que la convolution de deux gaussiennes est une gaussienne, la convolution de deux poissoniennes est une poissonienne, la convolution de deux lois de Cauchy est une loi de Cauchy.

Le fait remarquable est que, pour ces lois, la variable somme $X = X_1 + X_2$ obéit à une loi de *même nature* que chacune de ses "composantes" X_1 et X_2 . Cette propriété extraordinaire n'est pas vraie en général : elle ne tient visiblement pas pour deux variables de Bernoulli, ou pour deux variables binomiales. Les lois qui la possèdent sont dites *stables par l'addition* ; elles jouent un rôle de première importance en pratique, et sont par ailleurs des curiosités mathématiques justifiant l'intérêt qui leur est porté à la suite des travaux de Paul Lévy dans les années 1920. Inversement, elles suggèrent de se poser des questions intéressantes, comme par exemple : dans quelle mesure une variable aléatoire X peut-elle être décomposée en la somme de deux autres variables aléatoires, ce que l'on formule usuellement en disant : dans quelle mesure une loi est-elle *divisible* ? Le champ de recherche en la matière porte le nom d'*arithmétique des lois de probabilités*, et est encore à l'heure actuelle un domaine en pleine activité (pour une revue relativement récente, voir l'article de Livshits et al. [27]).

Il convient de savoir que la réciproque n'est vraie que pour les lois de Poisson et de Gauss : si $X_1 + X_2$ est une variable de Gauss (*resp.* de Poisson), et si X_1 et X_2 sont indépendantes, alors X_1 et X_2 sont des variables de Gauss (*resp.* de Poisson). À l'inverse, Dugué [14] signale, sans les préciser, qu'il existe des cas où $\phi_1(t)\phi_2(t) = e^{-a|t|}$ sans que X_1 et X_2 soient des variables de Cauchy.

1.9 Lois-limites. Théorème limite central

1.9.1 Problématique

Les remarques faites dans la sous-section 1.8.3 attirent l'attention sur des lois remarquables par le fait qu'elles sont *stables* par l'addition de variables aléatoires indépen-

¹⁰⁵La moyenne d'une somme de variables aléatoires est toujours la somme des valeurs moyennes, que ces variables soient indépendantes ou non. En revanche, la variance n'est égale à la somme des variances que si les variables sont indépendantes.

dantes. En termes plus précis, ceci signifie qu'étant donné des variables aléatoires indépendantes X_j obéissant à un même type de loi (de Gauss, de Poisson, etc), si on définit une suite de variables S_N (elles aussi évidemment aléatoires) :

$$S_1 = X_1, \quad S_2 = X_1 + X_2, \quad S_3 = X_1 + X_2 + X_3, \dots, S_{N+1} = S_N + X_{N+1}, \quad (1.222)$$

alors la loi de probabilité de S_N est de même nature – mais bien sûr a ses propres paramètres dépendant de N : si les X_j sont toutes gaussiennes (moyennes x_{0j} et variances σ_j^2), la gaussienne pour S_N a pour moyenne $\sum_{j=1}^N x_{0j}$ et pour variance $\sum_{j=1}^N \sigma_{X_j}^2$. En quelque sorte, ces lois remarquables sont des *points-fixes*¹⁰⁶ dans un certain espace de classes de fonctions, quand l'itération élémentaire est l'addition des v.a.

La propriété de *point fixe* est une propriété en soi ; une autre propriété pour un "point" est d'être la *limite* x_∞ d'une suite x_n de nombres proprement définie ; il peut aussi s'agir d'une suite de fonctions f_n , dont la limite est notée f_∞ . Un tel point peut être dit *attracteur* au sens où, partant d'une certaine situation, la suite des opérations conduit peu à peu inexorablement vers la limite¹⁰⁷ considérée.

Ainsi, s'agissant cette fois de la somme S_N de N variables indépendantes X_i distribuées suivant des lois $p_{X_i}(x)$ *a priori quelconques*, on peut se poser la question suivante : dans quelle mesure, et à quelle(s) condition(s), la fonction de répartition $F_N(x)$ de S_N peut-elle converger vers une certaine loi quand N tend vers l'infini ? Et si oui, quelle est cette loi ? Il est bien clair qu'une telle loi-limite, si elle existe, aura un caractère *universel* qui la fera apparaître spontanément dans une multitude de phénomènes où la somme de variables aléatoires joue un rôle de premier plan.

La réponse est oui. Une telle loi existe (dans des conditions qui seront précisées), et c'est la loi de Gauss. L'affirmation qui précise les choses porte le nom de *Théorème limite central* (TLC). Notons que s'agissant de comprendre ce qui se passe lorsque l'on considère une suite de variables aléatoires, il conviendrait au préalable de redéfinir complètement la notion même de convergence, c'est-à-dire de plonger dans la *topologie aléatoire* – un monde en soi (pour une introduction particulièrement claire, voir [14], section 1.7). On se contentera ici d'affirmations élémentaires, dont certaines dissimuleront des difficultés de fond, ou des subtilités le plus souvent inessentiels pour le physicien. Pour tout dire, des affirmations dans la suite ne seront pas seulement *élémentaires*, elles seront aussi quelque peu (ou très ?) vagues aux yeux des puristes. En tout cas, les résultats obtenus de façon parfois peu rigoureuse permettront de comprendre des faits essentiels, en particulier l'omniprésence de la loi de Gauss dans la Nature – un fait expérimental indiscutable, même si évidemment les phénomènes aléatoires n'obéissent pas tous à une telle loi, tant s'en faut.

¹⁰⁶Un point fixe est un point invariant dans une certaine loi de récursion (itération). Par exemple, soit une suite de nombres x_n définis par $x_{n+1} = f(x_n)$, x_0 donné, où f est une fonction également donnée ; un point fixe, noté x^* , est un point qui satisfait $x^* = f(x^*)$: c'est donc un point invariant par l'application f .

On connaît des itérations simplissimes aux propriétés stupéfiantes, lorsque la fonction f est *non-linéaire* (ce qui est une banalité en soi). Par exemple, l'itération dite *logistique*, définie par $f(x) = \lambda x(1-x)$, est sans doute l'exemple le plus simple de route vers le chaos quand le paramètre de contrôle λ augmente (cascade de Feigenbaum).

¹⁰⁷L'ensemble des points de départ où il en va ainsi est appelé *bassin attracteur* ou *bassin d'attraction*.

1.9.2 Théorème limite central

La question est de montrer que, dans certaines conditions qui vont apparaître d'elles-mêmes, la somme de N variables *indépendantes* :

$$S_N \stackrel{\text{déf}}{=} \sum_{j=1}^N X_j \quad (1.223)$$

tend vers une variable gaussienne lorsque N augmente indéfiniment. *A priori*, chaque variable X_j a sa propre loi de probabilité $p_{X_j}(x)$, de moyenne $\langle X_j \rangle$ et d'écart-type *finis*¹⁰⁸ σ_{X_j} ; supposant toutes les v.a. ordinaires, on suppose donc :

$$\langle X_j \rangle = \int x p_{X_j}(x) dx, \quad |\langle X_j \rangle| < +\infty, \quad \langle X_j^2 \rangle = \int x^2 p_{X_j}(x) dx < +\infty, \quad (1.224)$$

$$\sigma_{X_j}^2 = \langle X_j^2 \rangle - \langle X_j \rangle^2 < +\infty. \quad (1.225)$$

Bien sûr, l'information contenue dans la connaissance de tous les $p_{X_j}(x)$ permet de trouver p_{S_N} : il suffit de convoluer toutes les lois p_{X_j} . Formellement, la loi cherchée a pour expression :

$$p_{S_N}(x) = \int dx_1 \int dx_2 \dots \int dx_{N-1} p_{X_1}(x_1) p_{X_2}(x_2) \dots p_{X_{N-1}}(x_{N-1}) \times \\ p_{X_N}(x - x_1 - x_2 - \dots - x_{N-1}), \quad (1.226)$$

et il n'y a en principe qu'à calculer l'intégrale multiple pour obtenir p_{S_N} – dans le cas où les lois individuelles sont connues, évidemment. En général (notamment pour les petites valeurs de N), la loi p_{S_N} est ce qu'elle est, et ne présente aucun caractère d'universalité¹⁰⁹.

En réalité, le problème se présente souvent autrement : pour un enquêteur d'institut de sondage, N est de l'ordre de 1000 ; quant aux physiciens, ils sont plutôt habitués à manipuler des nombres *énormes*, dont le prototype est le nombre d'Avogadro $\mathcal{N} \sim 10^{24}$. En pareil cas, calculer l'intégrale (même avec un ordinateur puissant) serait une tâche impossible¹¹⁰, et en fait totalement inutile dans bien des cas, comme on va le voir. Par ailleurs, on ne connaît pas toujours les lois individuelles $p_{X_j}(x)$, une situation dans laquelle on ne sait strictement rien faire en l'état.

L'intrusion de *grands nombres*, voire de *très grands nombres*, est une hypothèse qualitative essentielle pour toute la suite. Elle implique que d'autres conditions doivent être réunies, notamment que *toutes* (ou presque toutes) les variables X_j doivent être réellement fluctuantes : à la limite où seul un petit nombre n d'entre elles fluctuent (les

¹⁰⁸Comme on le verra, l'existence de chaque σ_{X_j} est une hypothèse cruciale.

¹⁰⁹Encore que... Il est frappant de constater, à l'aide d'une simple calculette, que la somme d'un *petit* nombre (cinq, six) de variables indépendantes distribuées uniformément (sur $[0, 1]$ par exemple) est une variable gaussienne à une très bonne approximation (voir fig. 1.20).

¹¹⁰Si la machine met 10^{-7} seconde pour calculer une intégrale, le temps de calcul total avec \mathcal{N} est $\sim 10^{17}$ secondes, soit environ trois milliards d'années...

autres étant quasi certaines), on se retrouve de fait dans une situation avec un petit nombre de variables fluctuantes, auquel cas les affirmations énoncées ci-dessous perdent leur validité.

Dire que toutes les variables fluctuent de façon comparable, c'est affirmer que tous les écarts-types σ_{X_j} sont du même ordre de grandeur : si certains étaient gigantesques et les autres minuscules, on retrouverait le cas où, relativement, seules certaines variables fluctuent de fait. À partir du moment où tous les σ_{X_j} sont du même ordre de grandeur, on peut tous les supposer voisins de l'unité, au prix d'un recalibrage trivial commun à toutes les variables X_j . Dans toute la suite, on suppose donc :

$$\boxed{\sigma_{X_j} \sim 1 \quad \forall j} \quad (1.227)$$

Comme la variance d'une somme de variables *indépendantes* est la somme des variances, cependant que la moyenne de la somme est *toujours* la somme des moyennes¹¹¹, on peut dès maintenant affirmer que S_N est une variable aléatoire de moyenne $\sum_j \langle X_j \rangle$ et de variance égale à la somme de N nombres d'ordre 1, c'est-à-dire est d'ordre N :

$$\langle S_N \rangle = \sum_j \langle X_j \rangle, \quad \sigma_{S_N} = \sqrt{\sum_j \sigma_{X_j}^2} \sim \sqrt{N}, \quad (1.228)$$

ce qui suggère d'introduire une nouvelle variable, centrée et dont les fluctuations sont d'ordre unité ; on définit ainsi :

$$Y_N \stackrel{\text{déf}}{=} \frac{1}{\sqrt{N}} (S_N - \sum_j \langle X_j \rangle) = \frac{1}{\sqrt{N}} (S_N - \langle S_N \rangle). \quad (1.229)$$

Il s'agit maintenant de préciser la loi de distribution de Y_N , notée $p_{Y_N}(y)$ ou, de façon équivalente, de trouver une approximation de sa fonction caractéristique $\phi_{Y_N}(t)$. On a par définition :

$$\phi_{Y_N}(t) = \langle e^{itY_N} \rangle = \langle e^{it \frac{1}{\sqrt{N}} (S_N - \sum_j \langle X_j \rangle)} \rangle = e^{-i \frac{t}{\sqrt{N}} \langle S_N \rangle} \langle e^{it \frac{1}{\sqrt{N}} S_N} \rangle \equiv e^{-i \frac{t}{\sqrt{N}} \langle S_N \rangle} Z_N(t). \quad (1.230)$$

C'est la quantité $Z_N(t)$ qu'il convient d'analyser dans la limite $N \gg 1$; on a :

$$Z_N(t) = \langle e^{i \frac{t}{\sqrt{N}} (X_1 + X_2 + \dots + X_N)} \rangle = \langle \prod_j e^{i \frac{t}{\sqrt{N}} X_j} \rangle. \quad (1.231)$$

Comme les variables X_j sont indépendantes, la moyenne du produit est égal au produit des moyennes :

$$Z_N(t) = \langle \prod_j e^{i \frac{t}{\sqrt{N}} X_j} \rangle = \prod_j \langle e^{i \frac{t}{\sqrt{N}} X_j} \rangle = \prod_j \phi_j \left(\frac{t}{\sqrt{N}} \right), \quad (1.232)$$

¹¹¹Cette affirmation est un truisme lorsque le nombre N de variables additionnées est *fini*.

où l'on retrouve le produit des fonctions caractéristiques des aléatoires X_j . Compte tenu des hypothèses (existence des moyennes et des variances individuelles), on a (voir (1.198)) :

$$\phi_j(t) = 1 + it\langle X_j \rangle - \frac{t^2}{2}\langle X_j^2 \rangle + t^2\varepsilon_2^{(j)}(t) \quad \forall j, \quad (1.233)$$

où¹¹² $\lim_{t \rightarrow 0} \varepsilon_2^{(j)}(t) = 0$. Le logarithme de $Z_N(t)$ est :

$$\ln Z_N(t) = \sum_j \ln \left[1 + i\frac{t}{\sqrt{N}}\langle X_j \rangle - \frac{t^2}{2N}\langle X_j^2 \rangle + \frac{t^2}{2N}\varepsilon_2^{(j)}\left(\frac{t}{\sqrt{N}}\right) \right]. \quad (1.234)$$

On développe maintenant le logarithme suivant $\ln(1+z) = z - \frac{z^2}{2} + \mathcal{O}(z^3)$:

$$\ln Z_N(t) = \sum_j \left[i\frac{t}{\sqrt{N}}\langle X_j \rangle - \frac{t^2}{2N}\langle X_j^2 \rangle + \frac{t^2}{2N}\langle X_j \rangle^2 + \frac{t^2}{2N}\varepsilon_2^{(j)}\left(\frac{t}{\sqrt{N}}\right) \right], \quad (1.235)$$

soit, puisque $\lim_{t \rightarrow 0} \varepsilon_2^{(j)}(t) = 0$:

$$\begin{aligned} \ln Z_N(t) &= i\frac{t}{\sqrt{N}}\langle S_N \rangle - \frac{t^2}{2N} \sum_j \left[\langle X_j^2 \rangle - \langle X_j \rangle^2 - \varepsilon_2^{(j)}\left(\frac{t}{\sqrt{N}}\right) \right] \\ &\stackrel{N \gg 1}{\simeq} i\frac{t}{\sqrt{N}}\langle S_N \rangle - \frac{t^2}{2N} \sum_j \sigma_{X_j}^2. \end{aligned} \quad (1.236)$$

En définitive, dans la limite des grands N , on a¹¹³ (compte tenu de (1.228)) :

$$Z_N(t) \simeq e^{i\frac{t}{\sqrt{N}}\langle S_N \rangle - \frac{t^2}{2N}\sigma_{S_N}^2}, \quad (1.237)$$

d'où, revenant à (1.230) :

$$\phi_{Y_N}(t) \simeq e^{-i\frac{t}{\sqrt{N}}\langle S_N \rangle} e^{i\frac{t}{\sqrt{N}}\langle S_N \rangle - \frac{t^2}{2N}\sigma_{S_N}^2} = e^{-\frac{t^2}{2}\frac{\sigma_{S_N}^2}{N}}. \quad (1.238)$$

On reconnaît à nouveau la fonction caractéristique d'une gaussienne. Ainsi, la variable Y_N est asymptotiquement distribuée suivant la loi normale de moyenne nulle et de variance $\frac{1}{N}\sigma_{S_N}^2$:

$$\rho_{Y_N}(y) \simeq \frac{1}{\sqrt{2\pi}\frac{1}{\sqrt{N}}\sigma_{S_N}} e^{-\frac{y^2}{2\frac{1}{N}\sigma_{S_N}^2}} \quad (N \gg 1). \quad (1.239)$$

Entre les deux variables Y_N et S_N , prenant respectivement des valeurs notées x et y , il y a un facteur $\frac{1}{\sqrt{N}}$ (voir la définition (1.229) de Y_N), qui se retrouve au niveau des densités :

$$\rho_{Y_N}(y) dy = \rho_{S_N}(x) dx \quad \iff \quad \rho_{S_N}(x) = \rho_{Y_N}\left(\frac{x - \langle X \rangle}{\sqrt{N}}\right) \frac{dy}{dx}. \quad (1.240)$$

¹¹²À nouveau, il ne s'agit pas forcément du démarrage d'une série de Taylor. Le reste $\varepsilon_2^{(j)}(t)$ peut très bien être de la forme t^α ($\alpha \in \mathbb{R}_+$, ou $t^\alpha \ln t$, etc.).

¹¹³Il serait plus juste, à défaut d'analyse plus élaborée, de remplacer \simeq par \sim .

Comme $\frac{dy}{dx} = \frac{1}{\sqrt{N}}$, on trouve finalement :

$$\rho_{S_N}(x) \simeq \frac{1}{\sqrt{2\pi}\sigma_{S_N}} e^{-\frac{(x-\langle S_N \rangle)^2}{2\sigma_{S_N}^2}} \quad (N \gg 1) \quad (1.241)$$

où, conformément à (1.228), $\langle S_N \rangle = \sum_j \langle X_j \rangle$, $\sigma_{S_N} = \sqrt{\sum_j \sigma_j^2}$. Ainsi, la somme S_N de N variables aléatoires indépendantes X_j , de moyennes $\langle X_j \rangle$ et de variances $\sigma_{X_j}^2$, mais de lois p_{X_j} quelconques par ailleurs est, pour les très grands N , une variable à peu près gaussienne de moyenne et de variance déterminées.

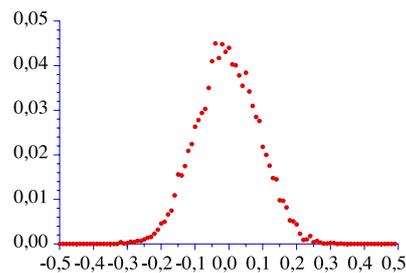


Figure 1.20: Distribution de la somme de $N = 10$ variables aléatoires indépendantes, chacun étant tirée uniformément dans $[-0.5, +0.5]$. L'aspect gaussien reconnaissable à vue se confirme en calculant les premiers moments M_k et en comparant avec les relations spécifiques des moments d'une gaussienne (voir (1.201) et (1.202)). Pour cette expérience, on a trouvé $\frac{M_4}{3M_2^2} = 0.962$, $\frac{M_6}{15M_2^3} = 0.904$, deux nombres assez voisins de 1 assurant que la somme est gaussienne à une bonne approximation. Que dire quand $N \sim 10^{23}$...

C'est ce résultat qui est appelé *Théorème limite central*, dont la qualification est largement méritée. À la réflexion, ce résultat est assez extraordinaire : il ne repose finalement que sur un petit nombre d'hypothèses, essentiellement l'indépendance des variables et l'existence d'une variance pour chacune d'entre elles – en revanche rien d'autre n'est dit sur les lois individuelles p_{X_j} des différentes variables X_j , sauf qu'elles sont toutes de variance finie. S'il s'agit de deux hypothèses très fortes, elles restent néanmoins assez peu contraignantes et, très souvent réalisées de fait en pratique, expliquent fondamentalement l'apparition si fréquente de la loi gaussienne dans la Nature.

◆ *Remarque*

Donnons une autre “démonstration”, très cavalière, de ce théorème, en supposant les variables centrées (si ce n'est pas le cas, il suffit de retrancher sa moyenne à chaque X_j). Les variables X_j étant indépendantes par hypothèse, la probabilité d'avoir une suite donnée $\{x_1, x_2, \dots, x_n\}$ est le produit des probabilités relatives à

chaque variable. La somme des $\{X_j\}$ prend la valeur X avec une densité $\rho_{S_N}(x)$ obtenue en sommant sur toutes les possibilités sous la contrainte que, précisément, la somme S_N est égale à x ; en introduisant la fonction δ de Dirac, ceci conduit à l'écriture :

$$p_{S_N}(x) = \int dx_1 \int dx_2 \dots \int dx_N p_{X_1}(x_1) p_{X_2}(x_2) \dots p_{X_N}(x_N) \delta\left(x - \sum_{n=1}^N x_n\right) . \quad (1.242)$$

Maintenant, on représente la fonction δ à l'aide de la relation connue :

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} dt , \quad (1.243)$$

ce qui transforme (1.242) en :

$$p_{S_N}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-itx} \int dx_1 \int dx_2 \dots \int dx_N p_{X_1}(x_1) p_{X_2}(x_2) \dots \times \rho_{X_N}(x_N) e^{it(x_1+x_2+\dots+x_N)} ; \quad (1.244)$$

les intégrales se factorisent et on obtient :

$$p_{S_N}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-itx} \prod_{n=1}^N \int dx_n p_{X_n}(x_n) e^{itx_n} \equiv \frac{1}{2\pi} \int_{-\infty}^{+\infty} dt e^{-itx} \prod_{n=1}^N \phi_n(t) , \quad (1.245)$$

où ϕ_n est la fonction caractéristique de p_{X_n} . Chaque ϕ_n démarre comme suit :

$$\phi_n(t) = 1 - \frac{1}{2}\sigma_n^2 t^2 + \dots ; \quad (1.246)$$

p_{S_N} apparaît alors sous la forme d'une intégrale de Fourier :

$$p_{S_N}(x) = \frac{1}{2\pi} \int dt e^{-itx} \prod_{n=1}^N \phi_n(t) = \frac{1}{2\pi} \int dt e^{-itx} \prod_{n=1}^N \left(1 - \frac{1}{2}\sigma_n^2 t^2 + \dots\right) . \quad (1.247)$$

En développant le produit et en n'écrivant que les termes en t^2 au plus, il vient¹¹⁴ :

$$p_{S_N}(x) = \frac{1}{2\pi} \int dt e^{-itx} \left[1 - \frac{t^2}{2} \sum_{n=1}^N \sigma_n^2 + \mathcal{O}(t^3)\right] . \quad (1.248)$$

Il s'agit maintenant de trouver la forme asymptotique de ρ_{S_N} quand $N \gg 1$. Comme N est très grand, et puisque l'écart-type de X est d'ordre $\sqrt{N} \gg 1$, l'intégrale est dominée¹¹⁵ par le voisinage $t \approx 0$, plus précisément par les valeurs de t telles que $|t| \lesssim \frac{1}{\sqrt{N}} \ll 1$; ceci autorise à retenir seulement les termes en t^2 . En outre, comme

¹¹⁴Si les p_n sont paires en x , le premier terme non écrit est $\mathcal{O}(t^4)$.

¹¹⁵L'argument est du type *phase stationnaire*, autre avatar de la méthode du col, voir ch. 7.

en dehors du voisinage de l'origine ainsi défini l'intégrand est quasi-nul, il n'y a aucun dommage à remplacer $1 - x^2$ par e^{-x^2} ; ainsi on trouve :

$$p_{S_N}(x) \simeq \frac{1}{2\pi} \int dt e^{-itx} e^{-\frac{t^2}{2}\sigma_{S_N}^2} . \quad (1.249)$$

Dans la limite $N \gg 1$, $\rho_{S_N}(x)$ apparaît comme la transformée de Fourier d'une gaussienne, c'est donc une gaussienne, dont la moyenne est nulle (on a centré toutes les variables X_j) et dont la variance est égale à $\sigma_{S_N}^2$. ♦

Ce théorème prend une forme encore plus simple si les variables X_j sont identiquement distribuées, suivant une seule et même loi p_X ; alors, toutes les X_j ont la même moyenne $\langle X \rangle$ et la même variance σ^2 , de sorte que $\langle S_N \rangle = N\langle X \rangle$ et $\sigma_{S_N}^2 = N\sigma^2$. Dans ces conditions, la loi gaussienne asymptotique est :

$$p_{S_N}(x) \rightarrow \rho_{S_N}(x) \simeq \frac{1}{\sqrt{2\pi} \sqrt{N} \sigma} e^{-\frac{(x - N\langle X \rangle)^2}{2N\sigma^2}} \quad (1.250)$$

Un fait doit être souligné. La fluctuation *relative* de l'aléatoire S_N est bien mesurée par le rapport entre sa fluctuation et sa valeur moyenne¹¹⁶. D'après ci-dessus, on a :

$$\frac{\sigma_{S_N}}{\langle S_N \rangle} = \frac{\sqrt{N}\sigma}{N\langle X \rangle} = \frac{1}{\sqrt{N}} \frac{\sigma}{\langle X \rangle} \propto N^{-\frac{1}{2}} \quad (1.251)$$

Le point essentiel à retenir est que la fluctuation relative décroît comme $\frac{1}{\sqrt{N}}$: elle tend certes vers zéro avec N , mais très lentement. Ce "détail" est sans importance pour le physicien qui pratique la Mécanique statistique ($N \sim 10^{24}$, ce qui donne une fluctuation relative de l'ordre de 10^{-12}), mais est sans doute un souci permanent pour les instituts de sondage ($\frac{1}{\sqrt{1000}} \simeq \frac{1}{32} \simeq 3\% \dots$).

Donnons un exemple classique d'application de ce théorème à la *marche de l'ivrogne* : une marche au hasard unidimensionnelle où, tous les δt , l'ivrogne fait un pas de longueur a soit dans un sens, soit dans l'autre, avec des probabilités respectives p (un pas vers la droite) et $1-p$ (un pas vers la gauche). On veut connaître la position typique de l'ivrogne au temps $t_N = N\delta t$.

Le problème peut se résoudre exactement par dénombrement, en supposant qu'à chaque nouveau pas l'ivrogne a complètement oublié les pas précédents (les différents mouvements élémentaires sont alors associés à des v.a. indépendantes). La probabilité $p_n(t_N)$ que l'ivrogne soit en $na + (N-n)(-a) = (2n-N)a$ après N petits pas est $C_N^n p^n (1-p)^{N-n}$, d'où l'on déduit immédiatement toutes les valeurs moyennes souhaitées.

¹¹⁶On peut dire qu'une variable certaine est le cas limite d'une variable aléatoire pour laquelle un tel rapport tend vers zéro. Par ailleurs, si la dispersion de X résulte d'une variabilité d'une série de mesures physiques, ce rapport joue le rôle d'une incertitude relative.

Si l'on s'en tient à la seule position (et à son écart-type), et en se contentant de la dynamique asymptotique (grands temps, soit $N \gg 1$), il est beaucoup plus rapide d'utiliser le Théorème limite central. La position X_N est une variable aléatoire qui est la somme des petits pas élémentaires δX_n ; chacun de ces derniers prend l'une des deux valeurs $\varepsilon_n a$, où $\varepsilon_n = +1$ avec probabilité p , -1 avec probabilité $1 - p$; si les $x_\lambda(N)$ notent les valeurs possibles de la v.a. X_N pour une trajectoire donnée ($\lambda = 2^N$), on a :

$$x_\lambda(N) = \sum_{n=1}^N \delta x_n, \quad \delta x_n = \varepsilon_n a. \quad (1.252)$$

Compte tenu de l'état de l'ivrogne (et notamment de son amnésie), les ε_n sont des v.a. indépendantes, distribuées suivant la même loi de Bernoulli ; la moyenne de ε_n est $p \times (+1) + (1 - p) \times (-1) = 2p - 1$, celle de ε_n^2 est $p \times (+1)^2 + (1 - p) \times (-1)^2 = 1$. L'écart-type σ est donc $\sqrt{4p(1 - p)}$.

La moyenne de X_N est la somme des moyennes :

$$\langle X_N \rangle = \sum_{n=1}^N \langle \varepsilon_n \rangle a = N(2p - 1)a, \quad (1.253)$$

et comme les δX_n sont des v.a. indépendantes, la variance de la somme est la somme des variances :

$$\sigma_{X_N}^2 \equiv \langle X_N^2 \rangle - \langle X_N \rangle^2 = 4p(1 - p)Na^2. \quad (1.254)$$

Ces deux résultats sont indépendants de N (petit ou grand). Maintenant, le Théorème limite central permet d'affirmer que si $N \gg 1$, alors X_N est asymptotiquement une variable gaussienne. Les valeurs de l'aléatoire X_N sont de la forme na , où n est un entier $-N \leq |n| \leq N$; pour $N \gg 1$, la loi de distribution de la position X_N est très bien approximée – au moins dans la région centrale (valeurs les plus *typiques*) – par la gaussienne :

$$P_{X_N}(n) = \frac{1}{\sqrt{2\pi}\sigma_{X_N}} e^{-\frac{(na - \langle X_N \rangle)^2}{2\sigma_{X_N}^2}} \quad (1.255)$$

Bien sûr, la vraie distribution de X_N est à support borné, compris entre $\pm Na$, alors que la gaussienne est non-nulle entre $\pm\infty$; cette distorsion dans les ailes est pratiquement sans incidence quand N est très grand, et à condition bien sûr de ne considérer que des moments d'ordre peu élevé.

Avec $t = N\delta t$, le résultat important est :

$$\sigma_{X_N} = \sqrt{4p(1 - p)Na} = \sqrt{4p(1 - p)\frac{t}{\delta t}a} \iff \sigma_{X_N} \propto \sqrt{t} \quad (1.256)$$

ce résultat – caractéristique d'un phénomène de *diffusion* – signifie que la taille typique de la région visitée par l'ivrogne croît comme la racine carrée du temps, c'est-à-dire beaucoup plus lentement que si l'ivrogne, débourré, se déplaçait toujours dans le (bon)

sens à vitesse constante (alors la distance parcourue serait $\propto t$). Il est traditionnel d'introduire la constante de diffusion, D définie (à une dimension d'espace¹¹⁷) comme :

$$\Delta X_N^2 \equiv 2Dt ; \quad (1.257)$$

on obtient ici :

$$D = 2p(1-p) \frac{a^2}{\delta t} . \quad (1.258)$$

La constante de diffusion est maximale pour $p = \frac{1}{2}$ (c'est le hasard *maximum*). À l'inverse, si p tend vers 1, la marche devient de moins en moins au hasard ; à la limite $p = 1$, la constante de diffusion s'annule, et $P_{X_N}(n)$ tend vers une fonction infiniment piquée, centrée au point d'abscisse Na : il s'agit alors d'un mouvement déterministe (abscisse = $Na = \frac{a}{\delta t}t$), l'homme à jeun se déplaçant à la vitesse constante $v = \frac{a}{\delta t}$, et explorant l'espace beaucoup plus vite que l'ivrogne qui va tantôt dans un sens, tantôt dans l'autre. Noter que la limite $p = 0$ est apparemment singulière puisque l'exposant de t vaut $\frac{1}{2}$ si $p \neq 0$ et 1 si $p = 0$. En réalité, il y a quelque part une échelle de temps dépendant de p , $\tau(p)$, au-delà de laquelle le régime en $t^{1/2}$ est effectivement réalisé (aux temps courts, le hasard se manifeste peu, et on a un régime plutôt linéaire). Que cette échelle de temps *diverge* quand $p \rightarrow 0$ n'est pas surprenant : en termes imagés, plus p est petit, plus il faut du temps pour que l'aléatoire joue pleinement son rôle.

Il convient toujours de garder à l'esprit le fait que l'approximation par une gaussienne dans la limite des grands N n'est assurément très bonne que dans la région centrale de la distribution (souvent appelée *région d'échelle*). Autrement dit, loin dans les ailes, la vraie distribution de S_N peut différer notablement de son approximation gaussienne ; les moments d'ordre élevés sont sensibles à la forme des ailes et sont parfois mal rendus lorsque l'on s'en tient à l'approximation fournie par le Théorème limite central. En particulier, si la (vraie) loi de S_N a des moments M_k infinis pour k supérieur à un certain k_0 , alors l'approximation gaussienne (qui donne des moments finis $\forall k$) peut conduire à des bêtises, même si elle reproduit bien le centre de la vraie distribution. À nouveau, les événements rares peuvent en fait jouer un rôle déterminant, alors que les événements *typiques* (les plus fréquents) semblent *a priori* plaider en faveur d'un comportement standard.

Ces remarques, nécessaires, n'altèrent la grande importance pratique du Théorème limite central que pour en délimiter l'usage raisonné, mais ne remettent pas fondamentalement en cause l'universalité qu'il sous-tend, et permet d'accepter des hypothèses plausibles comme celle consistant à affirmer que la somme d'un grand nombre de perturbations aléatoires indépendantes est une perturbation gaussienne ; en raison des propriétés remarquables de la distribution de Gauss, une telle hypothèse permet de construire des modèles traitables (*tractable models*, en anglais), sans lesquels on ne saurait pas faire grand-chose. De surcroît, comme Paul Lévy l'a montré, il est susceptible de généralisations diverses (voir par exemple l'ouvrage de Gnedenko et Kolmogorov [29]).

¹¹⁷Dans \mathbb{R}^d , on pose $\Delta \vec{r}^2 = 2dDt$, si la diffusion est isotrope. Dans (1.257), on a introduit plus naturellement $\Delta X_N \equiv \sigma_{X_N}$.

Chapitre 2

Introduction aux fonctions généralisées (distributions)

“...the source of all great mathematics is the special case, the concrete example. It is frequent in mathematics that every instance of a concept of seemingly great generality is in essence the same as a small and concrete special case.”

(Paul Richard HALMOS, 1916–2006)

Ce chapitre est une introduction élémentaire à la Théorie des distributions et vise principalement à légitimer par des arguments simples des opérations courantes en Physique dont la nature symbolique pourrait faire douter de leur validité.

2.1 Présentation intuitive

Comme on va le voir, la notion de distribution est une généralisation de celle de fonction, d'où l'appellation synonyme *fonction généralisée*. La nécessité d'une telle extension est manifeste pour la Physique, comme le montrent quelques exemples simples.

Soit une charge réputée ponctuelle¹ fixe, située au point de rayon-vecteur \vec{r}_0 . On sait que le potentiel électrostatique U créé au point \vec{R} par cette charge est donné par la

¹Bien évidemment, derrière cette qualification se dissimule un problème d'échelle physique : dans un contexte physique bien posé, on entend par *ponctuel* un objet dont la taille est minuscule devant toute autre échelle de longueur déjà présente dans le problème considéré. D'ailleurs, nul ne sait si la loi de Coulomb est vraie à toute échelle de distance ! En particulier, à grande distance, l'argument dimensionnel donnant $\frac{h}{mc}$ pour la portée d'une interaction médiée par une particule de masse m permet, connaissant la borne supérieure expérimentale de la masse du photon ($m_{\text{photon}} \lesssim 10^{-52}$ g), d'estimer à $10^9 \text{ m} = 10^{19} \text{ \AA} = 10^{24} \text{ F}$ l'échelle de distance au-delà de laquelle la loi de Coulomb non écrantée pourrait être invalide.

fonction :

$$U(\vec{R}; \vec{r}_0) = \frac{1}{4\pi\epsilon_0} \frac{q}{\|\vec{R} - \vec{r}_0\|} ; \quad (2.1)$$

c'est d'ailleurs plus précisément une fonction du seul module $\|\vec{R} - \vec{r}_0\|$.

Prenons maintenant le cas d'une répartition *continue* de charge ; on entend par là qu'il existe des charges réparties dans l'espace et que la somme des charges contenues dans un volume $V_{\vec{r}}$ autour du point \vec{r} , $Q_{V_{\vec{r}}}$, est telle que la limite du rapport $Q_{V_{\vec{r}}}/V_{\vec{r}}$ quand $V_{\vec{r}} \rightarrow 0$ existe quel que soit \vec{r} :

$$\forall \vec{r} : \exists \lim_{V_{\vec{r}} \rightarrow 0} \frac{Q_{V_{\vec{r}}}}{V_{\vec{r}}} \stackrel{\text{déf}}{=} \rho(\vec{r}) , \quad (2.2)$$

et sert de définition à la fonction² *densité de charge*, $\rho(\vec{r})$. Dans ces conditions, la charge totale de la distribution est Q :

$$Q = \int_{\mathbb{R}^3} \rho(\vec{r}) d^3r ; \quad (2.3)$$

quant au potentiel électrostatique au point \vec{R} créé par cette répartition de charge, il est donné par :

$$U(\vec{R}) = \frac{1}{4\pi\epsilon_0} \int_{\mathbb{R}^3} \frac{\rho(\vec{r})}{\|\vec{R} - \vec{r}\|} d^3r . \quad (2.4)$$

À la réflexion, on observe que le cas d'une charge *ponctuelle* q n'apparaît pas à ce stade comme un cas particulier de la distribution continue. En effet, il est impossible de définir une fonction $\rho_{\text{ponct}}(\vec{r})$ jouant le rôle d'une densité : avec la définition ci-dessus, cette "fonction" serait nulle partout sauf en un point où on ne sait pas trop quelle valeur lui donner (∞ ?). Toute intégrale impliquant un tel objet est visiblement dénuée de sens.

Cette impossibilité est ennuyeuse, mais peut être levée par un processus de limite approprié. Par exemple, on peut définir une densité de charge très "pointue", $\rho_{\delta V_{\vec{r}_0}}(\vec{r})$, définie comme une fonction prenant la valeur constante $\frac{q}{\delta V_{\vec{r}_0}}$ dans un petit volume $\delta V_{\vec{r}_0}$ centré sur \vec{r}_0 et la valeur 0 partout ailleurs ; avec cette fonction, l'intégrale $\int_{\mathbb{R}^3} \rho_{\delta V_{\vec{r}_0}}(\vec{r}) d^3r$ est parfaitement définie, vaut justement q par construction, et ce quelle que soit la valeur de $\delta V_{\vec{r}_0}$ – en particulier à la limite $\delta V_{\vec{r}_0} \rightarrow 0$.



Figure 2.1: La charge q est uniformément répartie dans la boule \mathcal{B}_n de rayon $\frac{a}{n}$, où a est une certaine longueur physique donnée.

²Par analogie avec ce qui est dit à propos de variables aléatoires prenant des valeurs continues (voir ch. 1), la répartition de charges est plus précisément *absolument* continue.

Le processus de limite peut être plus précisément défini en considérant des boules $\mathcal{B}_{\vec{r}_0, n}$ dont le rayon r_n est une fonction décroissante d'un entier, par exemple $r_n = \frac{a}{n}$, où a est une longueur donnée (fig. 2.1) ; ceci permet de définir une suite de fonctions $\rho_n(\vec{r}; \vec{r}_0)$ valant $\frac{3n^3}{4\pi a^3}q$ dans la boule, et 0 à l'extérieur. Pour chacune de ces fonctions, on a :

$$\int_{\mathbb{R}^3} \rho_n(\vec{r}; \vec{r}_0) d^3r = \int_{\mathcal{B}_{\vec{r}_0, n}} \rho_n(\vec{r}; \vec{r}_0) d^3r = q . \quad (2.5)$$

Définissant alors la fonction $\delta_n(\vec{r} - \vec{r}_0) \stackrel{\text{déf}}{=} \frac{1}{q} \rho_n(\vec{r}; \vec{r}_0)$, on a :

$$\int_{\mathbb{R}^3} \delta_n(\vec{r} - \vec{r}_0) d^3r = 1 , \quad (2.6)$$

ou encore :

$$\int_{\mathbb{R}^3} 1 \times \delta_n(\vec{r} - \vec{r}_0) d^3r = 1 . \quad (2.7)$$

La limite de la suite $\delta_n(\vec{r}; \vec{r}_0)$ n'est évidemment pas une fonction au sens ordinaire ; néanmoins, la limite de l'intégrale – qui est en fait une constante indépendante de n – est un certain nombre (égal à 1) – que l'on note ici $\langle \delta_{\vec{r}_0}, 1 \rangle$:

$$\langle \delta_{\vec{r}_0}, 1 \rangle \stackrel{\text{déf}}{=} \lim_{n \rightarrow +\infty} \int_{\mathbb{R}^3} 1 \times \delta_n(\vec{r} - \vec{r}_0) d^3r = 1 \quad (2.8)$$

le symbole δ représentant traditionnellement la *distribution de Dirac* – ici c'est plus précisément $\delta_{\vec{r}_0}$, distribution de Dirac concentrée en \vec{r}_0 . À ce stade, la notion de distribution (*fonction généralisée*) apparaît comme une recette permettant d'associer un nombre (ici 1, le second membre) à une certaine fonction ordinaire $\phi(x)$, ici la fonction valant 1 partout, présente dans l'intégrale.

Un autre exemple, plus élaboré, est fourni par la Mécanique quantique dans l'une de ses plus élémentaires applications. Soit une particule de masse m confinée dans un espace à une dimension sur un segment de longueur L mais qui, à part cette contrainte, n'est soumise à aucune force. x désignant la coordonnée, l'équation fixant les états propres $\psi(x)$ est, en représentation- q :

$$-\frac{\hbar^2}{2m} \psi''(x) = E \psi(x) , \quad 0 < x < L , \quad (2.9)$$

avec $\psi(0) = \psi(L) = 0$ et étant entendu que $\psi(x) = 0 \forall x \notin [0, L]$. Posant³ $E = \frac{\hbar^2 k^2}{2m} \in \mathbb{R}$, on obtient $\psi''(x) + k^2 \psi(x) = 0, 0 \leq x \leq L$, soit $\psi(x) = A \sin(kx + \phi)$. Les deux conditions aux limites donnent $\phi = 0$ (π), $\sin(kL + \phi) = 0$, d'où les modes propres normalisés (définis à une phase globale près) :

$$\psi_{r,L}(x) = \begin{cases} \sqrt{\frac{2}{L}} \sin \frac{r\pi x}{L} & \forall x \in [0, L] \\ 0 & \forall x \notin [0, L[\end{cases} , \quad (2.10)$$

³À ce stade, on ne dit rien sur k , qui est réel (si $E > 0$) ou imaginaire pur (si $E < 0$) ; la biunivocité de la paramétrisation exige cependant de ne retenir que l'une des deux branches de la fonction $(2mE/\hbar^2)^{1/2}$, par exemple celle qui prend des valeurs positives pour $E > 0$.

En fait, le cas $E < 0$ est exclu physiquement, la valeur propre ne pouvant être inférieure au minimum d'énergie potentielle.

avec :

$$E \in \{E_{r,L}\}_r, \quad E_{r,L} \stackrel{\text{d\u00e9f}}{=} \frac{r^2 \pi^2 \hbar^2}{2mL^2}, \quad r \in \mathbb{N}^*. \quad (2.11)$$

Les conditions aux limites donnant k de la forme $r\frac{\pi}{L}$, donc k à valeurs réelles, on trouve ainsi que les énergies sont toutes positives, ce qui est physiquement la moindre des choses, et même *strictement* positives, en conséquence des relations de Heisenberg. On vérifie sans peine que⁴ :

$$\int_{\mathbb{R}} \psi_{r,L}(x) \psi_{r',L}(x) dx = \delta_{rr'}, \quad (2.12)$$

égalité exprimant l'orthogonalité de deux fonctions propres distinctes. Par ailleurs, la densité de probabilité de trouver la particule entre deux points infiniment voisins est $P_{rL}(x) \stackrel{\text{d\u00e9f}}{=} |\psi_{r,L}(x)|^2$, de sorte que l'on a :

$$1 = \int_{\mathbb{R}} P_{rL}(x) dx. \quad (2.13)$$

Une question est maintenant la suivante : comment décrire une particule libre dans l'espace infini $x > 0$, juste soumise à une barrière impénétrable située à l'origine ? La première idée qui vient à l'esprit est d'utiliser les résultats précédents, et y prendre la limite $L \rightarrow +\infty$, mais ceci conduit visiblement à une trivialité vide de tout sens puisque toutes les fonctions propres sont identiquement nulles dans cette limite. En revanche, il est tout à fait possible de *définir* une fonction généralisée représentant la densité de probabilité de la particule confinée à \mathbb{R}_+ , P_r , par l'égalité :

$$\langle P_r, 1 \rangle \stackrel{\text{d\u00e9f}}{=} \lim_{L \rightarrow +\infty} \int_{\mathbb{R}} 1 \times P_{rL}(x) dx = 1 \quad (2.14)$$

Dans la suite, on généralise et formalise les observations précédentes, à la fois pour les préciser et pour justifier des opérations symboliques d'usage courant en Physique. Pour la simplicité, on se bornera à ne considérer que des fonctions d'une seule variable.

2.2 Les distributions en tant que fonctionnelles linéaires

2.2.1 Définition d'une fonction généralisée (distribution)

Dans le cas de fonctions d'une seule variable, x , l'analogue de (2.8) est⁵ :

$$\langle \delta_{x_0}, 1 \rangle \stackrel{\text{d\u00e9f}}{=} \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} 1 \times \delta_n(x - x_0) dx = 1. \quad (2.15)$$

⁴Dans toute la suite, la notation $\int_{\mathbb{R}}$ signifie $\int_{-\infty}^{+\infty}$.

⁵ δ_{x_0} est une notation assez courante pour désigner la distribution de Dirac concentrée en x_0 ; à défaut de cette précision, il est sous-entendu que $x_0 = 0$.

Cette écriture est parfois remplacée par une autre, dont l'aspect *symbolique* doit sauter aux yeux :

$$\boxed{\text{“ } \langle \delta_{x_0}, 1 \rangle = \int_{\mathbb{R}} 1 \times \delta(x - x_0) dx = 1 \text{ ”}} \quad (2.16)$$

où la notation traditionnelle $\delta(x - x_0)$, pas forcément très heureuse mais consacrée par l'usage, a été adoptée ; clairement, il n'existe pas de fonction au sens usuel satisfaisant cette égalité. Quoique symbolique, cette façon d'écrire fort commode est celle que l'on adopte presque toujours en Physique, les choses étant précisées. Enfin, il est bien clair que l'on peut toujours se ramener au cas $x_0 = 0$, ce glissement étant supposé fait désormais.

Jusqu'à présent, et s'agissant de la distribution de Dirac, on a associé, sous une intégrale, une certaine suite de fonctions $\delta_n(x)$ avec la fonction égale à 1 partout. Plus généralement, une distribution permet d'associer un nombre à une certaine fonction (ordinaire), appelée *fonction-test*, notée en général $T(x)$, supposée nantie des bonnes propriétés (à préciser en temps utile) et appartenant de ce fait, par définition, à un certain ensemble de fonctions.

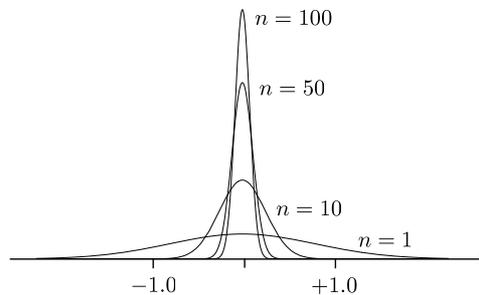


Figure 2.2: Précurseurs gaussiens $\sqrt{\frac{n}{\pi}}e^{-nx^2}$ de la distribution de Dirac.

Il est facile de préciser le mode d'action de la fonction généralisée δ sur une fonction-test $T(x)$ et en choisissant explicitement une suite de précurseurs $\delta_n(x)$, par exemple les fonctions :

$$\delta_n(x) = \begin{cases} 0 & \forall x, |x| > \frac{1}{2n} \\ n & \forall x, |x| < \frac{1}{2n} \end{cases} \quad (n \in \mathbb{N}^*) \quad (\text{créneau}) , \quad (2.17)$$

ou :

$$\delta_n(x) = \frac{1}{\pi} \frac{n}{1 + n^2 x^2} \quad (n \in \mathbb{N}^*) \quad (\text{lorentzienne}) , \quad (2.18)$$

ou encore :

$$\delta_n(x) = \sqrt{\frac{n}{\pi}} e^{-nx^2} \quad (n \in \mathbb{N}^*) \quad (\text{gaussienne}) ; \quad (2.19)$$

toutes ces fonctions sont telles que $\int_{\mathbb{R}} \delta_n(x) dx = 1$, mais leur limite n'est clairement pas une fonction. Considérons maintenant la suite d'intégrales I_n :

$$I_n \stackrel{\text{d\u00e9f}}{=} \int_{\mathbb{R}} T(x) \delta_n(x) dx , \quad (2.20)$$

dont on ne peut rien dire sans connaître les propriétés de la fonction-test $T(x)$, et dont la limite $n \rightarrow +\infty$ est, par définition, le nombre $\langle \delta, T \rangle$. Supposons que $T(x)$ admette un développement en série de Taylor dans un voisinage de l'origine de mesure supérieure à $\frac{1}{n_0}$, où n_0 est un certain entier. Avec les précurseurs en créneau, on peut écrire, $\forall n > n_0$:

$$I_n = \sum_{p=0}^{+\infty} \frac{1}{p!} T^{(p)}(0) \int_{-\frac{1}{2n}}^{+\frac{1}{2n}} nx^p dx = \sum_{q=0}^{+\infty} \frac{T^{(2q)}(0)}{(2q+1)!} 2n \frac{1}{(2n)^{2q+1}} = \sum_{q=0}^{+\infty} \frac{T^{(2q)}(0)}{(2q+1)!} \frac{1}{(2n)^{2q}} . \quad (2.21)$$

Passons à la limite $n \rightarrow \infty$; par définition, le premier membre de (2.20) tend vers $\langle \delta, T \rangle$. Par ailleurs, tous les termes de la série de (2.21) tendent vers zéro, sauf celui correspondant à $q = 0$; il vient ainsi :

$$\boxed{\langle \delta, T \rangle \stackrel{\text{déf}}{=} \lim_{n \rightarrow \infty} \int_{\mathbb{R}} \delta_n(x) T(x) dx = T(0)} \quad (2.22)$$

Visiblement, le même travail peut être effectué avec le précurseur gaussien (2.19) (mais pas avec la lorentzienne (2.18) !). À nouveau, l'usage courant fait écrire, symboliquement :

$$\boxed{\langle \delta, T \rangle \stackrel{\text{déf}}{=} \int_{\mathbb{R}} \delta(x) T(x) dx = T(0)} \quad (2.23)$$

égalité qui constitue en définitive la règle opérationnelle de la “fonction” de Dirac $\delta(x)$. Plus généralement, on a :

$$\boxed{\langle \delta(x - x_0), T \rangle \stackrel{\text{déf}}{=} \int_{\mathbb{R}} \delta(x - x_0) T(x) dx = T(x_0)} \quad (2.24)$$

où la clarté exige, dans le crochet, de préciser l'argument. Bien évidemment, avec des fonctions-tests à support borné $[a, b]$, $\beta(x)$, il convient de préciser :

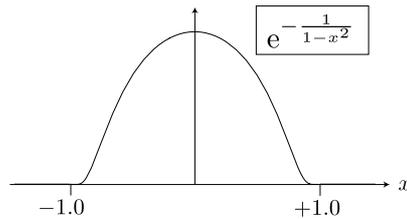
$$\langle \delta(x - x_0), \beta \rangle \stackrel{\text{déf}}{=} \int_a^b \delta(x - x_0) \beta(x) dx = \beta(x_0) , \quad a < x_0 < b \quad (2.25)$$

et $\langle \delta(x - x_0), \beta \rangle = 0$ autrement – ce qui, en toute rigueur est compris dans (2.22) si on précise que $\beta(x) \equiv 0 \forall x \notin [a, b]$.

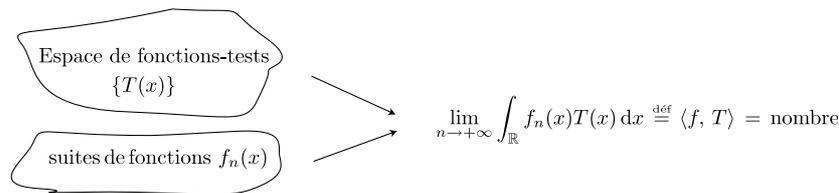
Clairement, δ apparaît bien comme une *fonctionnelle*, c'est-à-dire comme une recette permettant d'associer un *nombre* à une certaine fonction $T(x)$ nantie des bonnes propriétés, c'est-à-dire appartenant à un certain espace de fonctions dont toutes possèdent les propriétés requises. Cet espace une fois précisé, permet de définir un certain ensemble de distributions. Citons quelques espaces importants en pratique :

- l'espace \mathcal{D} des fonctions $\beta(x)$ indéfiniment dérivables à support borné (c'est-à-dire identiquement nulles au dehors d'un certain intervalle $[a, b]$). Exemple :

$$\beta(x) = \begin{cases} 0 & \forall x, |x| \geq 1 \\ e^{-\frac{1}{1-x^2}} & \forall x, |x| < 1 \end{cases} . \quad (2.26)$$

Figure 2.3: Graphe de la fonction-test $\beta(x)$ définie en (2.26).

- L'espace \mathcal{E} des fonctions indéfiniment dérivables à support quelconque.
- L'espace \mathcal{G} des “bonnes” fonctions⁶ au sens de Lighthill [30], c'est-à-dire des fonctions indéfiniment dérivables partout, et dont toutes les dérivées décroissent plus vite à l'infini que toute puissance⁷ de x . Le plus souvent dans la suite, les fonctions-tests seront prises dans \mathcal{G} et seront génériquement notées $\phi(x)$; alors, par définition : $\forall p \in \mathbb{N}, |\phi^{(p)}(x)| = \mathcal{O}(|x|^{-N}) \forall N$ (exemple : $\phi(x) = e^{-x^2}$). Les distributions définies en prenant la fonction-test dans cet espace sont appelées *distributions tempérées*.
- L'espace \mathcal{FG} des “assez bonnes” fonctions, qui sont indéfiniment dérivables et dont toutes les dérivées ne croissent pas plus vite à l'infini qu'une certaine puissance donnée⁸ de $|x|$; notant $\psi(x)$ une telle fonction, on a $|\psi^{(p)}(x)| = \mathcal{O}(|x|^N)$ (exemple : $\psi(x)$ est un polynôme de degré N).

Figure 2.4: Illustration des ingrédients nécessaires à la définition d'une fonction généralisée f ; en tant que recette permettant d'associer un nombre à des fonctions d'un certain espace, il s'agit d'une *fonctionnelle*.

En définitive, on adoptera la définition suivante : une fonction généralisée (dis-

⁶aussi parfois appelées *fonctions à décroissance rapide*. \mathcal{G} est pour *good* ; cet espace est souvent noté \mathcal{S} dans les ouvrages en français (\mathcal{S} pour Schwartz). L'espace \mathcal{S} est stable par la transformation de Fourier.

⁷La relation $f = \mathcal{O}(g)$ signifie que, dans la limite considérée, il existe une constante $C > 0$ telle que $|f| < C|g|$; ainsi, dans la limite $|x| \rightarrow \infty$, $\mathcal{O}(|x|^{-N})$ désigne un terme qui est *au plus* d'ordre $\frac{1}{|x|^N}$.

Ceci inclut la possibilité $f = \mathcal{O}(g)$, égalité signifiant que $|f| < \varepsilon|g|$ près de la limite, quel que soit $\varepsilon > 0$ choisi d'avance.

⁸Fonctions dites parfois à *croissance polynômiale*, ou à *croissance lente*. \mathcal{FG} est pour *fairly good*.

tribution) $f(x)$ est une suite de fonctions $f_n(x)$ permettant d'associer un nombre⁹ noté $\langle f, \phi \rangle$ à toute bonne fonction $\phi(x)$ suivant l'égalité :

$$\boxed{\langle f, \phi \rangle \stackrel{\text{déf}}{=} \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n(x) \phi(x) dx} \quad (2.27)$$

que, le cas échéant, l'on notera prudemment :

$$\text{“ } \langle f, \phi \rangle = \int_{\mathbb{R}} f(x) \phi(x) dx \text{ ”} \quad (2.28)$$

comme si la limite des $f_n(x)$ existait, que l'on peut commuter limite et intégration et que, enfin, il était possible de donner un sens à l'intégrale. Une autre écriture, symbolique elle aussi mais résumant bien ces définitions, est :

$$\boxed{\text{“ } f = \lim_{n \rightarrow +\infty} f_n \text{ ”}} \quad (2.29)$$

Par exemple, on s'autorisera, pour faire court et quand aucune ambiguïté n'est à craindre, à écrire :

$$\delta(x) = \lim_{n \rightarrow +\infty} \sqrt{\frac{n}{\pi}} e^{-nx^2} \quad (n \in \mathbb{N}) ; \quad (2.30)$$

l'aspect symbolique est d'autant plus manifeste que l'on peut remplacer la suite gaussienne par toute autre suite équivalente (voir section 2.2.2). Il est bien entendu que ces relations doivent être comprises comme devant être réintroduites dans des égalités précises où figurent à gauche le nombre $\langle \delta, \phi \rangle$ et à droite la limite d'une certaine intégrale.

Les définitions ci-dessus mettent clairement en évidence le caractère *linéaire* de la fonctionnelle que définit une distribution ; quelles que soient les deux constantes λ_1 et λ_2 :

$$\langle g, \lambda_1 \phi_1 + \lambda_2 \phi_2 \rangle = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} g_n(x) [\lambda_1 \phi_1(x) + \lambda_2 \phi_2(x)] dx = \lambda_1 \langle g, \phi_1 \rangle + \lambda_2 \langle g, \phi_2 \rangle \quad (2.31)$$

2.2.2 Suites équivalentes : exemples à propos de la distribution de Dirac

La distribution de Dirac a été définie comme le nombre obtenu en prenant la limite d'une suite d'intégrales impliquant d'une part une suite de fonctions $f_n(x)$ (créneau, gaussienne, etc.), d'autre part un ensemble de fonctions $T(x)$ possédant certaines propriétés. S'agissant de suites formées de bonnes fonctions, $f_n(x)$, une suite est dite *régulière* si pour toute bonne fonction $\phi(x)$ la limite suivante existe :

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} f_n(x) \phi(x) dx . \quad (2.32)$$

⁹Il s'agit bien d'une *fonctionnelle*, tout comme l'intégrale d'action $S_{\text{if}} \stackrel{\text{déf}}{=} \int_{t_i}^{t_f} L(q(t), \dot{q}(t)) dt$ est une *fonctionnelle* $S[q]$ de la fonction $q(t)$ (et de sa dérivée).

Deux suites régulières $f_n(x)$ et $g_n(x)$ sont dites *équivalentes* si les deux limites correspondantes sont égales ; ainsi les suites $f_n(x) = e^{-x^4/n^4}$ et $g_n(x) = e^{-x^2/n^2}$ sont équivalentes puisque la limite de la suite $\int_{\mathbb{R}} g_n(x)\phi(x) dx$ est égale à celle de $\int_{\mathbb{R}} f_n(x)\phi(x) dx$, toutes deux valant $\int_{\mathbb{R}} \phi(x) dx$, et définissant la fonctionnelle qui, à une fonction $\phi(x)$, lui associe son intégrale sur \mathbb{R} . L'égalité des limites de suites régulières équivalentes permet de dire qu'une fonction généralisée est la classe de toutes les suites régulières équivalentes à une suite régulière donnée.

La fonction (généralisée) de Dirac est de fait la limite de toutes les suites régulières équivalentes à la suite gaussienne $\sqrt{\frac{n}{\pi}} e^{-nx^2}$. En effet, soit $\phi(x) \in \mathcal{G}$ une bonne fonction ; comme¹⁰ $\int_{\mathbb{R}} \sqrt{\frac{n}{\pi}} e^{-nx^2} dx = 1$, on a :

$$\left| \int_{\mathbb{R}} \sqrt{\frac{n}{\pi}} e^{-nx^2} \phi(x) dx - \phi(0) \right| = \left| \int_{\mathbb{R}} \sqrt{\frac{n}{\pi}} e^{-nx^2} [\phi(x) - \phi(0)] dx \right|. \quad (2.35)$$

La différence $\phi(x) - \phi(0)$ est égale à $\int_0^x \phi'(x') dx'$, intégrale qui est majorée en module par $\max |\phi'(x)| |x|$, d'où :

$$\left| \int_{\mathbb{R}} \sqrt{\frac{n}{\pi}} e^{-nx^2} \phi(x) dx - \phi(0) \right| \leq \sqrt{\frac{n}{\pi}} \max |\phi'(x)| \left| \int_{\mathbb{R}} |x| e^{-nx^2} dx \right| ; \quad (2.36)$$

l'intégrale de droite est $2 \int_0^{+\infty} x e^{-nx^2} dx = \frac{1}{n}$, d'où :

$$\left| \int_{\mathbb{R}} \sqrt{\frac{n}{\pi}} e^{-nx^2} \phi(x) dx - \phi(0) \right| \leq \frac{1}{\sqrt{n\pi}} \max |\phi'(x)|, \quad (2.37)$$

soit finalement :

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}} \sqrt{\frac{n}{\pi}} e^{-nx^2} \phi(x) dx = \phi(0), \quad (2.38)$$

égalité qui permet d'écrire :

$$\boxed{\langle \delta, \phi \rangle = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} \sqrt{\frac{n}{\pi}} e^{-nx^2} \phi(x) dx = \phi(0)} \quad (2.39)$$

En pratique, et lorsqu'une régularisation explicite est nécessaire, la famille gaussienne est le plus souvent d'un emploi très commode. Par ailleurs, si les fonctions de la suite

¹⁰On aura à plusieurs reprises besoin du résultat :

$$\boxed{\int_{\mathbb{R}} e^{-ax^2+bx} dx = \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a}} \quad (a > 0, b \in \mathbb{C})} \quad (2.33)$$

Cette égalité se démontre directement en développant en série e^{bx} et en utilisant :

$$\int_{\mathbb{R}} x^{2n} e^{-x^2} dx = \frac{(2n-1)!!}{2^n \sqrt{\pi}} ; \quad (2.34)$$

elle sera démontrée tout autrement par la suite (ch. 6, p. 249). D'ailleurs, on montrera aussi comment, moyennant une définition préalable précise de la fonction $z^{1/2}$, la même égalité se prolonge dans tout le plan complexe pour a , à l'exception d'une certaine ligne semi-infinie partant de l'origine.

régulière sont concentrées au point x_0 , on a :

$$\langle \delta(x - x_0), \phi \rangle = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} \sqrt{\frac{n}{\pi}} e^{-n(x-x_0)^2} \phi(x) dx = \phi(x_0) \quad (2.40)$$

La considération exclusive de suites de fonctions formées avec des bonnes fonctions est assez restrictive. On admettra que la plupart des résultats énoncés ci-dessous restent vrais pour des fonctions moins “bonnes”, ce dont on peut se convaincre par des arguments intuitifs dont le manque de rigueur ne doit pas être considéré comme un défaut de fond rédhibitoire.

2.2.3 Opérations sur les distributions. Dérivation

Les définitions et résultats précédents conduisent tout naturellement à des opérations de base sur les distributions.

Opérations algébriques élémentaires

La somme de deux distributions se définit d'elle-même ; soit deux suites régulières $g_n(x)$ et $h_n(x)$ définissant les distributions g et h :

$$\langle g, \phi \rangle = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} g_n(x) \phi(x) dx, \quad \langle h, \phi \rangle = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} h_n(x) \phi(x) dx \quad (2.41)$$

La somme $g + h \stackrel{\text{déf}}{=} f$ est la distribution construite par la suite $g_n(x) + h_n(x)$:

$$f = h + g \iff \langle f, \phi \rangle = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} [g_n(x) + h_n(x)] \phi(x) dx. \quad (2.42)$$

Cela fait, on peut évidemment définir la distribution nulle comme étant la somme $f + g$, avec $g = -f$, la définition de $-f$ coulant de source. Ayant défini la distribution nulle, cela a un sens d'écrire $\Omega f = 0$ où Ω désigne une recette bien définie permettant de construire une certaine combinaison (algébrique, différentielle¹¹, ...), une fois donnée une certaine distribution. L'égalité $\Omega f = 0$ constitue ainsi une *équation* dont l'inconnue f est une distribution, et c'est pourquoi l'expression convenue est *équation au sens des distributions*.

On peut aussi définir la distribution d'une fonction linéaire, à savoir $f(ax + b)$, a et b étant des constantes réelles ; tout naturellement, on la définira comme la suite régulière des fonctions $f_n(ax + b)$, $(a, b) \in \mathbb{R}^2$. En conséquence, on a :

$$\langle f(ax + b), \phi \rangle \stackrel{\text{déf}}{=} \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n(ax + b) \phi(x) dx ; \quad (2.43)$$

¹¹On verra ci-dessous comment se définit d'elle-même la *dérivée* d'une distribution.

par un changement de variable $X = ax + b$, il vient :

$$\langle f(ax + b), \phi \rangle = \lim_{n \rightarrow +\infty} \frac{1}{|a|} \int_{\mathbb{R}} f_n(X) \phi\left(\frac{X - b}{a}\right) dX \quad ; \quad (2.44)$$

noter que c'est $|a|$ qui figure dans le facteur : si $a > 0$, les deux variables x et X varient dans le même sens ; au contraire, si $a < 0$, si x varie de $-\infty$ à $+\infty$, X varie de $+\infty$ à $-\infty$: avec la convention $\int_{\mathbb{R}} \stackrel{\text{d\'ef}}{=} \int_{-\infty}^{+\infty}$, le module de a est bien nécessaire. En particulier, avec $f = \delta$, on a l'égalité importante en pratique :

$$\langle \delta(ax + b), \phi \rangle = \lim_{n \rightarrow +\infty} \frac{1}{|a|} \int_{\mathbb{R}} \delta_n(X) \phi\left(\frac{X - b}{a}\right) dX = \frac{1}{|a|} \phi\left(-\frac{b}{a}\right) \quad (2.45)$$

que l'on peut traduire par :

$$\boxed{\delta(ax + b) = \frac{1}{|a|} \delta\left(x + \frac{b}{a}\right)} \quad (2.46)$$

Plus généralement, soit à donner un sens à $\delta(u(x))$ où $u(x)$ est une application monotone, ayant de ce fait une application inverse à une seule détermination, u^{-1} , et supposée partout dérivable. On pose naturellement :

$$\langle \delta(u(x)), \phi \rangle \stackrel{\text{d\'ef}}{=} \lim_{n \rightarrow +\infty} \int_{-\infty}^{+\infty} \delta_n(u(x)) \phi(x) dx \quad ; \quad (2.47)$$

le changement de variable $x = u^{-1}(X)$ dans l'intégrale de droite donne :

$$\langle \delta(u(x)), \phi \rangle \stackrel{\text{d\'ef}}{=} \lim_{n \rightarrow +\infty} \int_{-\infty}^{+\infty} \delta_n(X) \phi(u^{-1}(X)) |u^{-1}'(X)| dX \quad , \quad (2.48)$$

le module permettant à nouveau de maintenir l'intégrale de $-\infty$ à $+\infty$; avec $(u^{-1})' = \frac{1}{u'}$, $(u^{-1})'(X) = \frac{1}{u'(u(x))}$, d'où :

$$\boxed{\delta(u(x)) = \frac{1}{|u'(x_1)|} \delta(x - x_1) \quad , \quad u(x_1) = 0, x_1 \in \mathbb{R}} \quad (2.49)$$

Ce résultat¹² se généralise immédiatement à une application $u(x)$ monotone par morceaux en raisonnant intervalle par intervalle :

$$\boxed{\delta(u(x)) = \sum_k \frac{1}{|u'(x_k)|} \delta(x - x_k) \quad , \quad u(x_k) = 0, x_k \in \mathbb{R}} \quad (2.50)$$

permettant d'écrire par exemple :

$$\delta(\sin x) = \sum_{n \in \mathbb{Z}} |(-1)^n| \delta(x - n\pi) = \sum_{n \in \mathbb{Z}} \delta(x - n\pi) \quad , \quad (2.51)$$

¹²Bien sûr, si $u(x)$ ne s'annule pas, $\delta(u(x))$ est la distribution nulle.

$$\delta(\cos x) = \sum_{n \in \mathbb{Z}} |(-1)^n| \delta\left(x - \left(n + \frac{1}{2}\right)\pi\right) = \sum_{n \in \mathbb{Z}} \delta\left(x - \left(n + \frac{1}{2}\right)\pi\right) . \quad (2.52)$$

Noter le *truc* suivant pour mémoriser le résultat (2.50) (ou le retrouver rapidement) : dans l'intégrale (2.47), tout se passe près de $u(x) = 0$, soit près de x_1 ; dès lors, on peut remplacer la fonction $u(x)$ par sa linéarisée $u_{\text{lin}} \stackrel{\text{déf}}{=} u'(x_1)(x - x_0)$; se retrouvant avec $\delta(u'(x_1)(x - x_1))$, le résultat (2.46) donne immédiatement (2.49).

Dérivation

On définit tout naturellement la dérivée d'une distribution comme la distribution construite avec les dérivées d'une suite régulière. Partant ainsi de :

$$\langle f', \phi \rangle \stackrel{\text{déf}}{=} \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f'_n(x) \phi(x) dx , \quad (2.53)$$

et effectuant une intégration par parties, il vient :

$$\langle f', \phi \rangle = - \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n(x) \phi'(x) dx , \quad (2.54)$$

d'où :

$$\boxed{\langle f', \phi \rangle = -\langle f, \phi' \rangle} \quad (2.55)$$

et plus généralement :

$$\boxed{\langle f^{(p)}, \phi \rangle = (-1)^p \langle f, \phi^{(p)} \rangle} \quad (2.56)$$

Ainsi, la dérivée δ' est la fonction généralisée définie par les suites régulières équivalentes à $f_n(x) \equiv \delta'_n(x) = -2n\left(\frac{n}{\pi}\right)^{1/2} x e^{-nx^2}$; quelques fonctions de cette suite sont tracées sur la figure 2.5. Alors que δ est physiquement une charge, sa dérivée δ' est visiblement un *dipôle*.

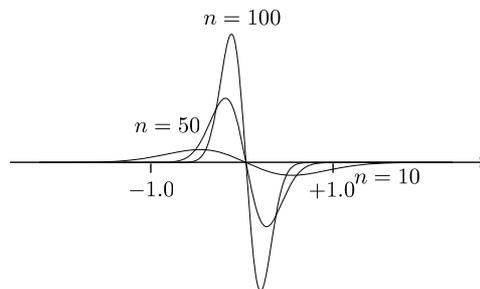


Figure 2.5: Quelques dérivées (premières) de la suite régulière gaussienne définie comme $\delta_n(x) = \sqrt{\frac{n}{\pi}} e^{-nx^2}$. Si δ évoque une charge ponctuelle, δ' est un dipôle.

Opération diverses

Clairement, les opérations *linéaires* sur les fonctions ordinaires se transcrivent à l'identique pour les distributions (exemple : $(\lambda f + \mu g)' = \lambda f' + \mu g'$). En revanche, le produit de deux distributions ne peut être défini avec les moyens ci-dessus : même si les suites $f_n(x)$ et $g_n(x)$ sont chacune régulière, leur produit $f_n(x)g_n(x)$ ne l'est pas en général¹³. Toutefois, on peut certainement définir le produit d'une distribution par une assez bonne fonction $\psi(x)$, posant pour le produit ψf :

$$\langle \psi f, \phi \rangle \stackrel{\text{déf}}{=} \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} \psi(x) f_n(x) \phi(x) dx . \quad (2.57)$$

En particulier, le produit $\psi(x)\delta(x)$ est égal à $\psi(0)\delta(x)$ puisque :

$$\langle \psi \delta, \phi \rangle \stackrel{\text{déf}}{=} \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} \psi(x) \delta_n(x) \phi(x) dx = \psi(0)\phi(0) \equiv \psi(0)\langle \delta, \phi \rangle , \quad (2.58)$$

que l'on traduit par :

$$\boxed{\psi(x)\delta = \psi(0)\delta} \quad (2.59)$$

Une conséquence immédiate, triviale mais importante, est obtenue avec $\psi(x) = x$: $x\delta = 0$, qu'il vaut la peine d'écrire plus généralement et explicitement :

$$\boxed{(x - x_0)\delta(x - x_0) = 0} \quad (2.60)$$

Cette égalité montre que l'équation $(x - x_1)f = 0$, prise au sens des distributions, admet une solution $f(x) = C_1\delta(x - x_1)$, où C_1 est une constante quelconque. Plus généralement, si $P_N(x)$ désigne un polynôme de degré N tel que $P_N(x_k) = 0$, $x_k \in \mathbb{R}$, l'équation :

$$P_N(x)f(x) = 0 \quad (2.61)$$

possède les $\nu \leq N$ solutions $\delta(x - x_k)$; comme l'équation est linéaire, toute combinaison linéaire est encore solution¹⁴ :

$$\boxed{f(x) = \sum_{k=1}^{\nu \leq N} C_k \delta(x - x_k) = 0 \implies P_N(x)f(x) = 0, P_N(x_k) = 0, x_k \in \mathbb{R}} \quad (2.62)$$

C'est ainsi que $C_+\delta(\omega - \omega_0) + C_-\delta(\omega + \omega_0)$, avec les C_{\pm} quelconques, est solution de :

$$(\omega^2 - \omega_0^2)f(\omega) = 0. \quad (2.63)$$

De telles équations, dans une vision ordinaire, seraient déclarées dépourvues de solutions, puisque les seuls objets les satisfaisant sont des "fonctions" nulles partout sauf en un

¹³Par exemple, on peut être tenté de définir le "carré" de δ en introduisant la suite $[\delta_n(x)]^2$; celle-ci est $\frac{n}{\pi}e^{-2nx^2}$, et ne fait visiblement pas ce qu'il faut...

¹⁴On n'a pas ici la prétention de démontrer que la distribution (2.62) est la solution *la plus* générale, d'où la relation logique à sens unique.

certain nombre de points où elles prennent une valeur indéterminée. C'est bien l'image la plus naïve que l'on peut se faire de la distribution δ , inventée par Dirac pour les besoins de la cause.

Le produit ψf étant maintenant introduit, on peut trouver la dérivée de ψf comme :

$$\langle (\psi f)', \phi \rangle \stackrel{\text{déf}}{=} \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} (\psi(x) f_n(x))' \phi(x) dx = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} [\psi(x) f_n'(x) + \psi'(x) f_n(x)] \phi(x) dx ; \quad (2.64)$$

la première limite définit la distribution $f' \psi$, la seconde définit $f \psi'$, d'où :

$$\boxed{(\psi f)' = \psi f' + \psi' f} \quad (2.65)$$

D'autres notions usuelles se généralisent aux distributions. Par exemple, on dira qu'une distribution f est paire (*resp.* impaire) si $\langle f, \phi \rangle = 0$ pour toute bonne fonction impaire (*resp.* paire). Selon cette définition, δ est paire, sa dérivée δ' est impaire, et ainsi de suite.

2.2.4 Les fonctions ordinaires comme distributions

Il n'est pas inutile de justifier l'appellation *fonction généralisée* en montrant comment une fonction ordinaire peut être considérée comme un cas particulier de distribution, à condition de posséder certaines propriétés qui vont être précisées. Soit $f(x)$ une fonction (ordinaire) telle qu'il existe un entier N positif assurant que $(1+x^2)^{-N} f(x)$ est absolument intégrable sur \mathbb{R} . Montrer que cette fonction ordinaire peut être considérée comme une fonction généralisée, également notée $f(x)$, c'est établir que pour toute bonne fonction $\phi(x)$, il existe une suite régulière $f_n(x)$ telle que :

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n(x) \phi(x) dx = \int_{\mathbb{R}} f(x) \phi(x) dx . \quad (2.66)$$

Le second membre existe en vertu des hypothèses sur $f(x)$: l'intégrand est le produit $(1+x^2)^{-N} f(x) \times (1+x^2)^N \phi(x) \equiv g(x) \times \psi(x)$; or $g(x)$ est absolument intégrable par hypothèse, et $\psi(x)$ est une bonne fonction. Le point est réellement d'exhiber une suite régulière $f_n(x)$ satisfaisant l'égalité (2.66), comme le fait le procédé systématique suivant.

La fonction (ordinaire) $f(x)$ étant donnée, choisissons comme suite les fonctions $f_n(x)$ définies par l'égalité :

$$f_n(x) \stackrel{\text{déf}}{=} \int_{\mathbb{R}} f(x') S(n(x' - x)) n e^{-\frac{x'^2}{n^2}} dx' , \quad (2.67)$$

où la fonction $S(X)$ est une *fonction d'arrondi*, c'est-à-dire une bonne fonction nulle pour $|X| \geq 1$, positive si $|X| < 1$ et normalisée à l'unité : $\int_{\mathbb{R}} S(X) dX = 1$. Il pourrait s'agir de la fonction (normalisée) citée à titre d'exemple de fonction à support borné, p. 88 :

$$S(X) = C e^{-1/(1-X^2)} , \quad C^{-1} = \int_{-1}^{+1} e^{-1/(1-X^2)} dX . \quad (2.68)$$

Quoi qu'il en soit, la fonction d'arrondi fait que l'intégrand dans (2.67) est non nul dans le seul intervalle $]x - \frac{1}{n}, x + \frac{1}{n}[$.

Il faut d'abord montrer que les $f_n(x)$ sont des bonnes fonctions ; la dérivée p^e de $f_n(x)$ est :

$$f_n^{(p)}(x) = \int_{\mathbb{R}} f(x') (-n)^p S^{(p)}(n(x' - x)) n e^{-\frac{x'^2}{n^2}} dx' ; \quad (2.69)$$

en raison du facteur $S^{(p)}(n(x' - x))$, l'intervalle utile pour x' est $x - \frac{1}{n} \leq x' \leq x + \frac{1}{n}$. Si $x > 0$, alors $|x - \frac{1}{n}| \leq |x'| \leq |x| + \frac{1}{n}$; si $x < 0$, x' est négatif et $|x'|$ est encadré par $| -x \pm \frac{1}{n}|$. Dans tous les cas, on a $||x| - 1| \leq |x'| \leq |x| + 1$, ce qui permet de majorer suivant :

$$|f_n^{(p)}(x)| \leq n^{p+1} [\max_{|x|<1} S^{(p)}] e^{-(|x|-1)^2/n^2} \int_{\mathbb{R}} f(x') dx' . \quad (2.70)$$

Le facteur 1 invisible dans l'intégrale est majoré par $\left(\frac{1+(|x|+1)^2}{1+x'^2}\right)^N \geq 1$; en définitive :

$$|f_n^{(p)}(x)| \leq n^{p+1} [\max_{|x|<1} S^{(p)}] e^{-(|x|-1)^2/n^2} [1 + (|x| + 1)^2]^N \int_{\mathbb{R}} \frac{|f(x')|}{(1+x'^2)^N} dx' . \quad (2.71)$$

L'intégrale existe en vertu des hypothèses ; le préfacteur est une gaussienne multipliée par un polynôme de degré $2N$; il tend donc vers zéro à l'infini plus vite que toute puissance de $|x|$, ce qui se traduit par l'existence d'un entier M tel que :

$$|f_n^{(p)}(x)| \leq \mathcal{O}(|x|^{-M}) \quad (2.72)$$

et exprime le fait que, toutes les fonctions $f_n^{(p)}(x)$ décroissant plus vite que toute puissance, $f_n(x)$ est une bonne fonction.

Cela étant établi, il faut encore montrer que la différence :

$$\int_{\mathbb{R}} f_n(x) \phi(x) dx - \int_{\mathbb{R}} f(x) \phi(x) dx \equiv I_n - I \quad (2.73)$$

tend vers zéro dans la limite $n \rightarrow +\infty$. Par définition des $f_n(x)$, après changement de variable $n(x' - x) = X$ et interversion des intégrales, la première intégrale I_n est :

$$I_n = \int_{-1}^{+1} S(X) dX \int_{\mathbb{R}} f(x') e^{-\frac{x'^2}{n^2}} \phi\left(x' - \frac{X}{n}\right) dx' . \quad (2.74)$$

La seconde intégrale I est égale à $\int_{-1}^{+1} S(X) dX \int_{\mathbb{R}} f(x) \phi(x) dx$ puisque $S(X)$ est normalisée à l'unité. La différence $I_n - I$ peut ainsi s'écrire :

$$I_n - I = \int_{-1}^{+1} S(X) dX \left[\int_{\mathbb{R}} f(x') \left[e^{-\frac{x'^2}{n^2}} \phi\left(x' - \frac{X}{n}\right) - \phi(x') \right] dx' \right] . \quad (2.75)$$

Il est commode de faire apparaître la dérivée de $\phi(x)$ en écrivant :

$$e^{-\frac{x'^2}{n^2}} \phi\left(x' - \frac{X}{n}\right) - \phi(x') = e^{-\frac{x'^2}{n^2}} \left[\phi\left(x' - \frac{X}{n}\right) - \phi(x') \right] - \phi(x') [1 - e^{-\frac{x'^2}{n^2}}] . \quad (2.76)$$

Reportant ceci, et utilisant le fait que la fonction $S(X)$ est bornée¹⁵ par 1, on a :

$$|I_n - I| \leq \max_{|X| \leq 1} \left| \int_{\mathbb{R}} f(x') e^{-\frac{x'^2}{n^2}} \left[\phi\left(x' - \frac{X}{n}\right) - \phi(x') \right] dx' - \int_{\mathbb{R}} f(x') \phi(x') (1 - e^{-\frac{x'^2}{n^2}}) dx' \right| \quad (2.77)$$

Seule la première intégrale dépend de X ; en écrivant en plus que le module de la différence est inférieur ou égal à la somme des modules, on a :

$$|I_n - I| \leq \max_{|X| \leq 1} \left| \int_{\mathbb{R}} f(x') e^{-\frac{x'^2}{n^2}} \left[\phi\left(x' - \frac{X}{n}\right) - \phi(x') \right] dx' \right| + \left| \int_{\mathbb{R}} f(x') \phi(x') (1 - e^{-\frac{x'^2}{n^2}}) dx' \right| \quad (2.78)$$

Le premier terme est borné supérieurement par $\int_{\mathbb{R}} |f(x')| \frac{1}{n} \max_{|x-x'| < 1} |\phi'(x')| dx'$; $\phi(x')$ étant une bonne fonction, il existe N tel que $\max_{|x-x'| < 1} |\phi'(x')| = \mathcal{O}(|x'|^{-2N})$, ce que l'on peut toujours écrire $\max_{|x-x'| < 1} |\phi'(x')| = A(1+x'^2)^{-N}$ où A est une constante positive. Le premier terme tend donc vers zéro $\propto n^{-1}$.

Le second terme se majore en notant que¹⁶ $\frac{1+x'^2}{n^2} > 1 - e^{-\frac{x'^2}{n^2}}$ quel que soit x' :

$$\left| \int_{\mathbb{R}} f(x') \phi(x') (1 - e^{-\frac{x'^2}{n^2}}) dx' \right| \leq \frac{1}{n^2} \int_{\mathbb{R}} |f(x') \phi(x')| (1 + x'^2) dx' . \quad (2.79)$$

Comme précédemment, on majore l'intégrale en remplaçant $|\phi(x')|$ par $\frac{A'}{(1+x'^2)^{N'}}$: le deuxième terme varie comme n^{-2} . En définitive :

$$|I_n - I| \leq \frac{1}{n} \int_{\mathbb{R}} |f(x')| \frac{A}{(1+x'^2)^N} dx' + \frac{1}{n^2} \int_{\mathbb{R}} |f(x')| \frac{A'}{(1+x'^2)^{N'}} dx' , \quad (2.80)$$

inégalité prouvant que $\lim_{n \rightarrow +\infty} I_n = I$.

2.2.5 Quelques distributions courantes

Il s'agit ici d'une part de revisiter des opérations élémentaires sur les fonctions ordinaires en les prenant comme des distributions, d'autre part de définir (ou redéfinir) les distributions les plus courantes en pratique, les reformulant comme on le fait habituellement en Physique.

Une fois que l'on a montré qu'une fonction ordinaire est un cas particulier de fonction généralisée, on peut, avec une telle fonction, appliquer les règles déjà établies pour les distributions. Par exemple, il s'avère que la fonction $\operatorname{sgn} x$, qui est bien une fonction assez ordinaire, a une dérivée non triviale au sens des distributions. On a :

$$\operatorname{sgn}(x) \stackrel{\text{déf}}{=} \begin{cases} -1 & \forall x < 0 \\ +1 & \forall x > 0 \end{cases} ; \quad (2.81)$$

¹⁵ puisqu'elle est positive et normalisée à l'unité.

¹⁶ Il suffit d'étudier la fonction $\eta(x) \stackrel{\text{déf}}{=} \frac{1+x^2}{n^2} - (1 - e^{-\frac{x^2}{n^2}})$.

revenant à (2.55), on en déduit :

$$\langle (\operatorname{sgn}(x))', \phi \rangle = -\langle \operatorname{sgn}(x), \phi' \rangle = -\int_{-\infty}^0 (-1)\phi'(x) dx - \int_0^{+\infty} (+1)\phi'(x) dx \quad (2.82)$$

le second membre étant égal à $2\phi(0)$, il vient :

$$\boxed{(\operatorname{sgn}(x))' = 2\delta(x)} \quad (2.83)$$

Exactement de la même façon, on voit que la dérivée de la fonction de Heaviside $Y(x)$:

$$Y(x) \stackrel{\text{d\u00e9f}}{=} \begin{cases} 0 & \forall x < 0 \\ +1 & \forall x > 0 \end{cases} \quad (2.84)$$

est la distribution δ :

$$\boxed{Y'(x) = \delta(x)} \quad (2.85)$$

Cette égalité se comprend bien intuitivement : la fonction $Y(x)$ est constante par morceaux ; sa dérivée est donc nulle presque partout. $Y(x)$ montant de 0 à 1 en un *point* ($x = 0$), son taux de variation y est infini. On peut aussi raisonner avec des précurseurs – quitte à sortir de l'ensemble des suites régulières – par exemple $Y_n(x) \stackrel{\text{d\u00e9f}}{=} \frac{1}{2}(1 + \tanh nx)$, dont la limite $Y_\infty(x)$ est “presque” égale à $Y(x)$, puisqu'elle est définie en $x = 0$ et y vaut $\frac{1}{2}$. Quoi qu'il en soit, $Y_n'(0) = n$ et tend bien vers l'infini avec n .

En conséquence de (2.65), on a $(Y(x)\psi(x))' = \delta(x)\psi(x) + Y(x)\psi'(x)$. Maintenant, soit la fonction (discontinue) $g(x) \stackrel{\text{d\u00e9f}}{=} \psi_1(x) + \psi_2(x)Y(x)$, où les ψ_i sont dérivables et à croissance polynômiale ; considérée comme une distribution, sa dérivée en ce sens est :

$$g'(x) = \psi_1'(x) + \psi_2'(x)Y(x) + \psi_2(x)\delta(x) = \psi_1'(x) + \psi_2'(x)Y(x) + \psi_2(0)\delta(x) ; \quad (2.86)$$

compte tenu de la définition de $g(x)$, on a $g(0_-) = \psi_1(0)$, $g(0_+) = \psi_1(0) + \psi_2(0)$, d'où $\psi_2(0) = g(0_+) - g(0_-)$. En conséquence :

$$g'(x) = \psi_1'(x) + \psi_2'(x)Y(x) + [g(0_+) - g(0_-)]\delta(x) ; \quad (2.87)$$

Les deux premiers termes coïncident avec la dérivée ordinaire de la fonction $g(x)$, obtenue en se plaçant soit à gauche, soit à droite, du point de discontinuité. Notant celle-ci $g'_{\text{reg}}(x)$, l'égalité précédente se récrit comme :

$$\boxed{g'(x) = g'_{\text{reg}}(x) + [g(0_+) - g(0_-)]\delta(x)} \quad (2.88)$$

d'où la règle : quand on dérive au sens des distributions une fonction discontinue, il apparaît, en plus des termes ordinaires, des distributions de Dirac concentrées aux points des sauts, avec un poids égal à la hauteur du saut.

En Physique (quantique, notamment), on rencontre souvent des objets comme :

$$I(\omega) \stackrel{\text{d\u00e9f}}{=} -i \int_0^{+\infty} e^{i\omega t} dt \quad (\omega \in \mathbb{R}) , \quad (2.89)$$

qui, tels quels, n'ont pas de sens, et doivent être *régularisés*. Cette régularisation doit être faite sur des bases physiques afin de la rendre d'un sens. Le plus souvent, un argument physique permet de se convaincre que l'intégrale (2.89) apparaît dans un calcul où l'inadvertance a fait délaissier un petit paramètre¹⁷ $\gamma > 0$; une fois celui-ci rétabli en amont, la bonne intégrale qui arrive typiquement est :

$$I_\gamma(\omega) \stackrel{\text{d\'ef}}{=} -i \int_0^{+\infty} e^{i(\omega+i\gamma)t} dt . \quad (2.90)$$

Pour le coup, cette intégrale est parfaitement définie et vaut :

$$I_\gamma(\omega) = -i \frac{1}{\gamma - i\omega} = \frac{1}{\omega + i\gamma} = \frac{\omega}{\omega^2 + \gamma^2} - i \frac{\gamma}{\omega^2 + \gamma^2} . \quad (2.91)$$

Supposons maintenant que $I_\gamma(\omega)$ survienne dans une intégrale sur ω en compagnie d'une bonne fonction $\phi(\omega)$, en tant qu'expression d'une certaine grandeur physique A :

$$A = \int_{\mathbb{R}} I_\gamma(\omega) \phi(\omega) d\omega = \int_{\mathbb{R}} \left(\frac{\omega}{\omega^2 + \gamma^2} - i \frac{\gamma}{\omega^2 + \gamma^2} \right) \phi(\omega) d\omega . \quad (2.92)$$

Dire que γ est un petit paramètre, c'est dire qu'il est très petit devant toute autre échelle caractérisant la variation de la fonction physique $\phi(\omega)$ qui joue ici le rôle de la bonne fonction. Autrement dit, il est licite (physiquement) à ce stade de passer à la limite $\gamma = 0_+$, et de retenir que la grandeur physique est en fait donnée par :

$$A = \lim_{\gamma \rightarrow 0_+} \int_{\mathbb{R}} \frac{\omega}{\omega^2 + \gamma^2} \phi(\omega) d\omega - i \lim_{\gamma \rightarrow 0_+} \int_{\mathbb{R}} \frac{\gamma}{\omega^2 + \gamma^2} \phi(\omega) d\omega . \quad (2.93)$$

En revenant à (2.18), on voit que la deuxième intégrale définit la distribution $\pi\delta(\omega)$. Quant à la première, l'examen du noyau $\frac{\omega}{\omega^2 + \gamma^2}$ montre qu'il a pour effet de couper symétriquement de part et d'autre de $\omega = 0$ les contributions divergentes de l'intégrand où l'on aurait fait brutalement $\gamma = 0$; par définition, c'est la régularisation en tant que *partie principale de Cauchy*, notée $\mathcal{P}\frac{1}{\omega}$:

$$\left\langle \mathcal{P}\frac{1}{\omega}, \phi \right\rangle \stackrel{\text{d\'ef}}{=} \lim_{\gamma \rightarrow 0_+} \int_{-\infty}^{+\infty} \frac{\omega}{\omega^2 + \gamma^2} \phi(\omega) d\omega \quad (2.94)$$

que l'on peut d'ailleurs tout autant définir comme :

$$\lim_{\gamma \rightarrow 0_+} \left[\int_{-\infty}^{-\gamma} \frac{1}{\omega} \phi(\omega) d\omega + \int_{+\gamma}^{+\infty} \frac{1}{\omega} \phi(\omega) d\omega \right] . \quad (2.95)$$

Au total, on obtient l'égalité importante :

$$\lim_{\gamma \rightarrow 0_+} \frac{1}{\omega + i\gamma} = \mathcal{P}\frac{1}{\omega} - i\pi\delta(\omega) \quad (2.96)$$

¹⁷C'est par exemple l'inverse d'un temps de relaxation : dans un premier temps, on aura traité le système comme un système isolé, excluant ainsi toute sorte de relaxation. Un système réel n'étant jamais isolé, il est de ce fait toujours soumis à une *dissipation*, laquelle produit une relaxation caractérisée par une échelle de temps finie, éventuellement très longue comparée à un autre temps caractéristique – ce qui aura justifié, dans un premier temps, de traiter le système comme s'il était *parfaitement* isolé, avant de tomber éventuellement sur des difficultés techniques (voir sous-section 2.4.1).

qui, comme toute égalité entre distributions, est précisément définie avec l'intervention d'une intégrale. On en déduit aussi :

$$\delta(\omega) = \frac{1}{2i\pi} \lim_{\gamma \rightarrow 0^+} \left(\frac{1}{\omega - i\gamma} - \frac{1}{\omega + i\gamma} \right) = \frac{1}{\pi} \lim_{\gamma \rightarrow 0^+} \frac{\gamma}{\omega^2 + \gamma^2} , \quad (2.97)$$

soit en revenant à l'expression intégrale (2.90) :

$$\delta(\omega) = \frac{1}{2\pi} \lim_{\gamma \rightarrow 0^+} \left(\int_{-\infty}^0 e^{i(\omega - i\gamma)t} dt + \int_0^{+\infty} e^{i(\omega + i\gamma)t} dt \right) \quad (2.98)$$

qui est une régularisation de l'écriture symbolique (2.126) ci-dessous. De la même façon, on a :

$$\mathcal{P} \frac{1}{\omega} = \frac{1}{2} \lim_{\gamma \rightarrow 0^+} \left(\frac{1}{\omega - i\gamma} + \frac{1}{\omega + i\gamma} \right) = \lim_{\gamma \rightarrow 0^+} \frac{\omega}{\omega^2 + \gamma^2} , \quad (2.99)$$

et :

$$\mathcal{P} \frac{1}{\omega} = \frac{i}{2} \lim_{\gamma \rightarrow 0^+} \left(\int_{-\infty}^0 e^{i(\omega - i\gamma)t} dt - \int_0^{+\infty} e^{i(\omega + i\gamma)t} dt \right) = \lim_{\gamma \rightarrow 0^+} \int_0^{+\infty} e^{-\gamma t} \sin \omega t dt \quad (2.100)$$

2.3 Transformées de Fourier des distributions

La transformation de Fourier sera présentée en détail dans le chapitre 8. Pour l'instant, on s'en tient aux définitions minimales en considérant exclusivement des bonnes fonctions $f(x)$. Pour une telle fonction, on peut définir¹⁸ sans difficulté sa transformée de Fourier $\mathcal{F}[f]$ par l'égalité :

$$F(k) \stackrel{\text{déf}}{=} \int_{\mathbb{R}} f(x) e^{ikx} dx , \quad k \in \mathbb{R} \quad \iff \quad f \xrightarrow{\mathcal{F}} F \quad (2.102)$$

L'existence de l'intégrale ne fait aucun doute puisque $f(x)$ est une bonne fonction. Il est facile de voir que la fonction $F(k)$ est elle aussi une bonne fonction ; en effet, en dérivant p fois, il vient :

$$F^{(p)}(k) = i^p \int_{\mathbb{R}} f(x) x^p e^{ikx} dx ; \quad (2.103)$$

une intégration par parties donne alors :

$$F^{(p)}(k) = i^p \frac{-1}{ik} \int_{\mathbb{R}} e^{ikx} \frac{d}{dx} [x^p f(x)] dx , \quad (2.104)$$

¹⁸Comme on le sait, différentes conventions existent pour définir la transformation de Fourier :

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{ikx} dx , \quad \int_{\mathbb{R}} f(x) e^{\pm 2i\pi kx} dx , \quad \text{etc.} \quad (2.101)$$

Dans ce cours, on adopte la convention usuelle des Probabilistes, donnant la fonction caractéristique $\phi = \mathcal{F}[p]$ pour une v.a. de densité $p(x)$, au sens large (voir section 1.8).

et N intégrations par parties conduisent à :

$$F^{(p)}(k) = i^p \frac{(-1)^N}{(ik)^N} \int_{\mathbb{R}} e^{ikx} \frac{d^N}{dx^N} [x^p f(x)] dx . \quad (2.105)$$

L'intégrale existe puisque $f(x)$ est une bonne fonction, de sorte que, quels que soient p et N , $|F^{(p)}(k)| = \mathcal{O}(|k|^{-N})$.

Soit maintenant la transformée de la dérivée $f'(x)$:

$$\mathcal{F}[f'] \stackrel{\text{déf}}{=} F_1(k) = \int_{\mathbb{R}} f'(x) e^{ikx} dx ; \quad (2.106)$$

une intégration par parties donne :

$$F_1(k) = -ik \int_{\mathbb{R}} f(x) e^{ikx} dx , \quad (2.107)$$

le terme tout intégré étant manifestement nul pour toute bonne fonction $f(x)$, d'où :

$$\boxed{\mathcal{F}[f'] = -ik \mathcal{F}[f]} \quad (2.108)$$

et plus généralement :

$$\boxed{\mathcal{F}[f^{(p)}] = (-ik)^p \mathcal{F}[f]} \quad (2.109)$$

Très conventionnellement, la fonction $f(x)$ est appelée *original*, cependant que $F(k)$ est appelée *image*. Cela étant posé, il est tout autant possible de définir la transformation inverse :

$$F \xrightarrow{\mathcal{F}^{-1}} f , \quad (2.110)$$

application qui s'exprime par la relation dite d'inversion :

$$\boxed{f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} F(k) e^{-ikx} dk} \quad (2.111)$$

Un moyen élémentaire pour obtenir cette formule est le suivant. Soit l'intégrale :

$$I(x) \stackrel{\text{déf}}{=} \int_{\mathbb{R}} e^{-\varepsilon k^2 + ikx} F(-k) dk = \int_{\mathbb{R}} dk \int_{\mathbb{R}} dx' e^{-\varepsilon k^2} e^{ik(x-x')} f(x') \quad (\varepsilon > 0) . \quad (2.112)$$

En permutant les deux intégrations (chaque intégrale converge uniformément), il vient :

$$I(x) = \int_{\mathbb{R}} dx' f(x') \int_{\mathbb{R}} dk e^{-\varepsilon k^2 + ik(x-x')} . \quad (2.113)$$

L'intégrale interne vaut $\sqrt{\frac{\pi}{\varepsilon}} e^{-\frac{(x-x')^2}{4\varepsilon}}$, fonction dont l'intégrale sur \mathbb{R} est $\sqrt{\frac{\pi}{\varepsilon}} \sqrt{4\pi\varepsilon} = 2\pi$; on peut ainsi écrire la différence $I(x) - 2\pi f(x)$:

$$I(x) - 2\pi f(x) = \sqrt{\frac{\pi}{\varepsilon}} \int_{\mathbb{R}} dx' e^{-\frac{(x-x')^2}{4\varepsilon}} [f(x')] - f(x)] ; \quad (2.114)$$

par le théorème des accroissements finis, $\exists \xi \in [x, x']$, $f(x') - f(x) = (x' - x)f'(\xi)$, ce qui permet la majoration :

$$|I(x) - 2\pi f(x)| \leq \sqrt{\frac{\pi}{\varepsilon}} \max |f'(x)| \int_{\mathbb{R}} e^{-\frac{(x-x')^2}{4\varepsilon}} |x - x'| dx' ; \quad (2.115)$$

l'intégrale vaut $2 \int_0^{+\infty} e^{-\frac{x^2}{4\varepsilon}} x dx = 4\varepsilon$, d'où $|I(x) - 2\pi f(x)| \leq \mathcal{O}(\sqrt{\varepsilon})$: il suffit de passer à la limite $\varepsilon \rightarrow 0$ pour obtenir (2.111).

Étant donné un couple de bonnes fonctions $((f(x), g(x)))$ et leurs transformées de Fourier $((F(k), G(k)))$, des arguments du même type permettent de démontrer l'égalité¹⁹ de Parseval - Plancherel :

$$\boxed{\int_{\mathbb{R}} F(k)G(k) dk = 2\pi \int_{\mathbb{R}} f(x)g(-x) dx} \quad (2.117)$$

Cette égalité servira ici de point de départ pour définir la transformée de Fourier $\mathcal{F}[f]$ d'une fonction généralisée f associée à la suite régulière $f_n(x)$. Soit $F_n(k)$ et $\Phi(k)$ les transformées de Fourier respectives de $f_n(x)$ et $\phi(x)$; l'égalité de Parseval - Plancherel permet d'écrire :

$$\int_{\mathbb{R}} F_n(k)\Phi(k) dk = 2\pi \int_{\mathbb{R}} f_n(x)\phi(-x) dx ; \quad (2.118)$$

prenant la limite $n \rightarrow +\infty$, on obtient au second membre la distribution $2\pi f(-x)$; le premier membre est alors, par définition, la transformée de Fourier de $f(x)$:

$$\boxed{\langle \mathcal{F}[f], \Phi \rangle \stackrel{\text{déf}}{=} \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} F_n(k)\Phi(k) dk = 2\pi \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n(x)\phi(-x) dx \equiv 2\pi \langle f(-x), \phi \rangle} \quad (2.119)$$

La première égalité sert de définition à $\mathcal{F}[f]$, la deuxième égalité montre, s'il le fallait, le lien étroit entre $\mathcal{F}[f]$ et f . En utilisant (2.109), on trouve à nouveau, mais pour des distributions :

$$\boxed{\mathcal{F}[f^{(p)}] = (-ik)^p \mathcal{F}[f]} \quad (2.120)$$

Par exemple, soit à trouver $\mathcal{F}[\delta]$; l'une des suites régulières $f_n(x)$ associée à δ est $f_n(x) \equiv \delta_n(x) = \left(\frac{n}{\pi}\right)^{1/2} e^{-nx^2}$. La transformée de Fourier de $f_n(x)$ est :

$$F_n(k) = \left(\frac{n}{\pi}\right)^{1/2} \int_{\mathbb{R}} e^{-nx^2} e^{ikx} dx = e^{-\frac{k^2}{4n}} . \quad (2.121)$$

La définition (2.119) se recopie ici comme (δ est paire) :

$$\langle \mathcal{F}[\delta], \Phi \rangle = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} e^{-\frac{k^2}{4n}} \Phi(k) dk = 2\pi \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} \left(\frac{n}{\pi}\right)^{1/2} e^{-nx^2} \phi(-x) dx \equiv 2\pi \langle \delta, \phi \rangle , \quad (2.122)$$

¹⁹On a tout autant :

$$\int_{\mathbb{R}} F(k)G(-k) dk = 2\pi \int_{\mathbb{R}} f(x)g(x) dx . \quad (2.116)$$

La suite $F_n(k)$ converge (uniformément) vers la fonction égale à 1 quelque soit $k \in \mathbb{R}$, d'où :

$$\boxed{\mathcal{F}[\delta] = 1} \quad (2.123)$$

À titre symbolique ou mnémotechnique, on peut aussi reporter $\delta(x)$ dans la relation de définition (2.102), ou encore partir d'un précurseur $\delta_n(x)$ et passer à la limite sous l'intégrale, extrayant ainsi la valeur en $x = 0$ de e^{ikx} .

Le résultat (2.123) est un exemple extrême des propriétés *duales* d'une fonction (généralisée ou non) et de sa transformée de Fourier : plus l'une est large, plus l'autre est étroite – les éléments de la suite $\delta_n(x)$ sont de plus en plus piqués quand n croît, mais leurs transformées de Fourier sont de plus en plus larges. Dit de façon plus rapide : $\delta(x)$ est une “fonction” infiniment fine et un peu exotique, sa transformée de Fourier est infiniment large et particulièrement banale. On se souvient aussi du grand écart entre la distribution de Cantor et sa fonction caractéristique (voir (1.195) et les figures ?? et 1.19).

D'une façon plus générale, on a :

$$\boxed{\mathcal{F}[\delta_{x_0}] = e^{ikx_0}} \quad (2.124)$$

De (2.123), on déduit aussi :

$$\boxed{\delta = \mathcal{F}^{-1}[1]} \quad (2.125)$$

ce qui, revenant à la formule inverse de la transformation de Fourier, fournit la représentation symbolique standard de $\delta(x)$:

$$\boxed{\delta(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ikx} dk = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ikx} dk} \quad (2.126)$$

dont une régularisation est donnée en (2.98). Enfin, en utilisant $Y'(x) = \delta(x)$, on voit que $-ik\mathcal{F}[Y] = 1$, soit :

$$\boxed{\mathcal{F}[Y](k) = -\frac{1}{ik}} \quad (2.127)$$

d'où, par la formule inverse, la représentation symbolique²⁰ de $Y(x)$:

$$\boxed{Y(x) = -\frac{1}{2i\pi} \int_{\mathbb{R}} \frac{1}{k} e^{-ikx} dk} \quad (2.129)$$

En ce qui concerne la transformée de Fourier de $\mathcal{P}\frac{1}{x}$, elle peut être obtenue de diverses façons. La première, expéditive, consiste à partir de la définition (2.95) de la

²⁰On verra plus tard comment le théorème des résidus permet de valider cette écriture en justifiant proprement l'égalité :

$$Y(x) = -\lim_{\varepsilon \rightarrow 0^+} \frac{1}{2i\pi} \int_{\mathbb{R}} \frac{1}{k + i\varepsilon} e^{-ikx} dk \quad (x \neq 0) \quad (2.128)$$

partie principale de Cauchy, en utilisant la formule de Fourier, soit :

$$\lim_{\varepsilon \rightarrow 0^+} \left(\int_{-\infty}^{-\varepsilon} \frac{1}{x} e^{ikx} dx + \int_{+\varepsilon}^{+\infty} \frac{1}{x} e^{ikx} dx \right) . \quad (2.130)$$

On sépare en parties réelle et imaginaire ; la partie réelle est nulle en tant que somme de deux intégrales visiblement opposées. La partie imaginaire est égale à $\int_{\mathbb{R}} \frac{\sin kx}{x} dx$, puisque l'intégrand est continu en $x = 0$. Cette intégrale vaut $\pi \operatorname{sgn} k$, d'où :

$$\boxed{\mathcal{F}[\mathcal{P}\frac{1}{x}](k) = i\pi \operatorname{sgn} k} \quad (2.131)$$

Une autre façon (moins désinvolte) est de repartir de la définition (2.119), en choisissant la famille de fonctions $f_{\varepsilon}(x) = \frac{1}{2} \left(\frac{1}{x+i\varepsilon} + \frac{1}{x-i\varepsilon} \right)$, ce qui oblige à trouver les transformées de Fourier de $\frac{1}{x \pm i\varepsilon}$, afin de pouvoir former $F_{\varepsilon}(k) = \frac{1}{2} [\mathcal{F}[\frac{1}{x+i\varepsilon}] + \mathcal{F}[\frac{1}{x-i\varepsilon}]]$; on verra par la suite comment les obtenir très facilement grâce au théorème des résidus, soit :

$$\mathcal{F}\left[\frac{1}{x+i\varepsilon}\right] \stackrel{\text{d\u00e9f}}{=} \int_{\mathbb{R}} \frac{1}{x+i\varepsilon} e^{ikx} dx = \begin{cases} 0 & \forall k > 0 \\ -2i\pi e^{\varepsilon k} & \forall k < 0 \end{cases} , \quad (2.132)$$

et de m\u00eame :

$$\mathcal{F}\left[\frac{1}{x-i\varepsilon}\right] \stackrel{\text{d\u00e9f}}{=} \int_{\mathbb{R}} \frac{1}{x-i\varepsilon} e^{ikx} dx = \begin{cases} +2i\pi e^{-\varepsilon k} & \forall k > 0 \\ 0 & \forall k < 0 \end{cases} . \quad (2.133)$$

La famille de fonctions $F_{\varepsilon}(k)$ est donc $i\pi[Y(k)e^{-\varepsilon k} - Y(-k)e^{\varepsilon k}]$ qui, dans la limite $\varepsilon \rightarrow 0$ donne $i\pi[Y(k) - Y(-k)]$, soit l'expression (2.131).

2.4 Illustrations et exemples

2.4.1 Oscillateur harmonique

Il s'agit d'illustrer \u00e0 propos d'un syst\u00e8me tr\u00e8s \u00e9l\u00e9mentaire les affirmations \u00e9nonc\u00e9es ci-dessus \u00e0 propos des syst\u00e8mes r\u00e9els et de montrer comment l'inclusion d'une dissipation, \u00e9ventuellement tr\u00e8s faible, permet de r\u00e9gulariser des quantit\u00e9s mal d\u00e9finies pouvant survenir dans un probl\u00e8me pos\u00e9 de fa\u00e7on trop abrupte.

Soit un oscillateur harmonique : masse ponctuelle m attach\u00e9e au bout d'un ressort parfait de constante de raideur k , les oscillations se produisant le long d'un axe Ox , dont l'origine est prise \u00e0 la position d'\u00e9quilibre. Posant $k = m\omega_0^2$, l'\u00e9quation fondamentale donne pour l'\u00e9cart \u00e0 l'\u00e9quilibre $x(t)$ en l'absence d'excitation :

$$\ddot{x}(t) + \omega_0^2 x(t) = 0 . \quad (2.134)$$

Si la bille est de surcroît soumise, à partir de $t = 0$, à une force extérieure $F(t) \stackrel{\text{déf}}{=} m\phi(t)$, l'équation devient inhomogène :

$$\ddot{x}(t) + \omega_0^2 x(t) = \phi(t) \quad (t > 0) . \quad (2.135)$$

Afin que le problème soit posé de façon physique, il convient de préciser l'état de la bille au départ, c'est-à-dire sa position x_0 et sa vitesse v_0 . En outre, on peut convenir que cet instant de départ est celui où on commence à appliquer la force $m\phi(t)$; enfin rien n'interdit d'appeler $t = 0$ l'instant en question.

Soit $X(\omega)$ et $\Phi(\omega)$ les transformées de Fourier de $x(t)$ et $\phi(t)$; comme ces quantités sont réelles, on a les symétries :

$$X^*(\omega) = X(-\omega) , \quad \Phi^*(\omega) = \Phi(-\omega) . \quad (2.136)$$

Par \mathcal{F} , (2.135) devient :

$$-\omega^2 X(\omega) + \omega_0^2 X(\omega) = \Phi(\omega) . \quad (2.137)$$

En allant un peu vite en besogne, on écrit la solution :

$$X(\omega) = \frac{1}{\omega_0^2 - \omega^2} \Phi(\omega) \equiv \chi_0(\omega) \Phi(\omega) , \quad (2.138)$$

où $\chi_0(\omega)$ est par définition la susceptibilité. Cette expression est sûrement incomplète : le problème dynamique admet une solution et une seule à condition d'avoir inclus les (deux) conditions initiales, or rien dans cette expression ne les requiert.

En réalité, on peut ajouter à cette expression n'importe quelle fonction $f(\omega)$ telle que $(\omega_0^2 - \omega^2)f(\omega) = 0$, l'équation (2.137) est encore satisfaite ; on sait que la fonction f n'est autre que $\delta(\omega \pm \omega_0)$, à un facteur arbitraire près. En définitive, la solution complète est :

$$X(\omega) = C_+ \delta(\omega - \omega_0) + C_- \delta(\omega + \omega_0) + \frac{1}{\omega_0^2 - \omega^2} \Phi(\omega) \quad (2.139)$$

On peut maintenant effectuer la transformation \mathcal{F}^{-1} , pour en déduire :

$$x(t) = C'_+ e^{i\omega_0 t} + C'_- e^{-i\omega_0 t} + \mathcal{F}^{-1} \left[\frac{1}{\omega_0^2 - \omega^2} \Phi(\omega) \right] (t) \quad (C_{\pm} = 2\pi C'_{\pm}) . \quad (2.140)$$

Le terme de droite ne peut être explicité tant que l'on ne se donne pas $\phi(t)$, mais on peut affirmer qu'il est nul à $t = 0$: à cet instant précis, la bille ne sait pas encore si on va lui appliquer ou non une force, et seule peut apparaître la solution libre constituée par les deux premiers termes. Les constantes C_{\pm} s'obtiennent par calage sur les conditions initiales, lequel donne $C'_{\pm} = \frac{1}{2}(x_0 \pm i\frac{v_0}{\omega_0})$. En particulier, si la bille est initialement au repos à sa position d'équilibre, $C'_{\pm} = 0$ et la solution complète est donnée par le seul dernier terme de (2.139) ; comme c'est la source externe qui met la bille en mouvement, ce mouvement s'appelle tout naturellement *régime forcé* :

$$X_{\text{forcé, non-amorti}}(\omega) = \frac{1}{\omega_0^2 - \omega^2} \Phi(\omega) \quad (2.141)$$

La fonction $\Phi(\omega)$ est ce qu'elle est et, en particulier, n'a aucune raison d'être nulle en $\omega = \pm\omega_0$; dans ces conditions, la susceptibilité diverge pour ces valeurs de ω , et on ne voit pas bien comment revenir en arrière pour trouver $x(t)$. L'idée est maintenant de trouver la bonne régularisation ; on va voir que l'inclusion d'un frottement règle le problème, et donne un sens physique à la régularisation en tant que partie principale de Cauchy.

De fait, pour sortir de cette difficulté, il suffit de réaliser que l'équation (2.134) décrit une situation idéalisée à l'extrême, où l'énergie du système est constante. Or tout système réel n'est jamais strictement isolé et il existe toujours une cause de dissipation d'énergie, même si celle-ci peut être rendue très petite grâce au soin extrême pris pour réaliser l'expérience. Ici, toujours afin d'en rester à un niveau élémentaire, on représente la dissipation par une force de freinage, dont la version la plus élémentaire est le freinage dit *fluide*, simplement proportionnel à la vitesse. Cela étant, l'équation fondamentale nettement plus réaliste est :

$$\ddot{x}(t) + \gamma\dot{x}(t) + \omega_0^2 x(t) = \phi(t) , \quad (2.142)$$

où $\gamma > 0$ est la constante phénoménologique caractérisant le frottement. Cela étant admis, l'équation en Fourier est :

$$-\omega^2 X(\omega) - i\gamma\omega X(\omega) + \omega_0^2 X(\omega) = \Phi(\omega) . \quad (2.143)$$

qui, toujours avec la bille au repos et à l'équilibre au départ, donne la solution :

$$X_{\text{forcé, amorti}}(\omega) = \frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega} \Phi(\omega) \equiv \chi(\omega) \Phi(\omega) , \quad (2.144)$$

$\chi(\omega)$ étant maintenant la susceptibilité en présence de dissipation. Cette fois, la transformation inverse ne pose aucun problème :

$$x_{\text{forcé, amorti}}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega} \Phi(\omega) e^{-i\omega t} d\omega , \quad (2.145)$$

puisque le dénominateur ne s'annule pas pour ω réel (et en supposant en plus que $\Phi(\omega)$ est bien comme il faut et ne cause pas de souci) ; posant $\omega_{\pm} = -i\frac{\gamma}{2} \pm \sqrt{\omega_0^2 - \frac{1}{4}\gamma^2}$, l'intégrale s'écrit :

$$x_{\text{forcé, amorti}}(t) = \frac{1}{2\pi} \frac{1}{\omega_+ - \omega_-} \int_{-\infty}^{+\infty} \left(\frac{1}{\omega - \omega_-} - \frac{1}{\omega - \omega_+} \right) \Phi(\omega) e^{-i\omega t} d\omega . \quad (2.146)$$

À ce stade, on voit comment retrouver l'oscillateur non amorti en effectuant la bonne limite ; correspondant à $\gamma \rightarrow 0_+$, elle signifie que chaque intégrale est en fait très précisément :

$$\lim_{\gamma \rightarrow 0} \int_{-\infty}^{+\infty} \frac{1}{\omega + i\frac{\gamma}{2} \mp \sqrt{\omega_0^2 - \frac{1}{4}\gamma^2}} \Phi(\omega) e^{-i\omega t} d\omega . \quad (2.147)$$

Clairement, le γ^2 dans la racine carrée du dénominateur ne sert à rien ; finalement, on observe que le frottement introduit *naturellement* la régularisation en tant que partie

principale de Cauchy. En définitive, la limite de l'oscillateur non-amorti est maintenant sans ambiguïté donnée par :

$$x_{\text{forcé, amorti}}(t) = \frac{1}{4\pi\omega_0} \mathcal{P} \int_{-\infty}^{+\infty} \left(\frac{1}{\omega + \omega_0} - \frac{1}{\omega - \omega_0} \right) \Phi(\omega) e^{-i\omega t} d\omega , \quad (2.148)$$

soit :

$$x_{\text{forcé, amorti}}(t) = \frac{1}{2\pi} \mathcal{P} \int_{-\infty}^{+\infty} \frac{1}{\omega_0^2 - \omega^2} \Phi(\omega) e^{-i\omega t} d\omega \quad (2.149)$$

Ceci d'une part valide physiquement la régularisation de Cauchy, d'autre part montre que l'un des systèmes les plus élémentaires de la Physique, mais idéalisé à l'extrême, implique forcément la notion de distribution. La description limite qui en découle est la bonne représentation théorique d'un oscillateur matériel réel observé pendant une échelle de temps courte devant son temps de relaxation γ^{-1} .

2.4.2 Exemples de relations de fermeture

On rencontre souvent en Mécanique quantique des relations dites *de fermeture*, exprimant la complétude d'une base de fonctions $\psi_\alpha(x)$ pour l'espace des états d'un système. D'une façon générale, ces fonctions sont des *modes propres*, ce qui signifie qu'elles sont propres d'une observable du système qui est, très souvent, le Hamiltonien représentant l'énergie. L'indice α rappelle la valeur propre de l'observable en question et peut être discret, continu ou successivement les deux quand on balaie l'axe réel²¹.

Afin d'illustrer ceci, considérons une particule de masse m , libre de se déplacer sur un cercle de circonférence $L = 2\pi R$. Sa position est fixée par un angle θ , compté à partir d'un point conventionnellement choisi. Formellement, il s'agit d'un problème plan à symétrie cylindrique, pour lequel l'énergie cinétique est représentée par un opérateur différentiel agissant sur les coordonnées polaires (r, θ) . La distance à l'origine étant ici fixée (et égale à R), seul subsiste un opérateur agissant sur l'angle θ , et le Hamiltonien de la particule s'écrit :

$$H = -\frac{\hbar^2}{2mR^2} \frac{d^2}{d\theta^2} ; \quad (2.150)$$

l'équation pour les modes propres (ψ, E) est $H\psi = E\psi$, soit :

$$-\frac{\hbar^2}{2mR^2} \frac{d^2\psi}{d\theta^2} = E\psi , \quad (2.151)$$

la fonction $\psi(\theta)$ devant être 2π -périodique. Une solution typique est $\psi(\theta) = e^{\lambda\theta}$, avec $E = -\frac{\hbar^2}{2mR^2} \lambda^2$. La condition aux limites (ici, de périodicité) donne $\lambda \times 2\pi = i \times \text{entier } 2\pi$,

²¹On connaît quelques cas, assez singuliers, où des valeurs propres discrètes surgissent *dans* le continu.

soit $\lambda = in$, avec $n \in \mathbb{Z}$; les énergies possibles sont donc $\frac{n^2 \hbar^2}{2mR^2}$ et les fonctions propres normalisées sur le cercle suivant $\int_0^{2\pi} |\psi_n(\theta)|^2 d\theta$ sont (à une phase globale près) :

$$\boxed{\psi_n(\theta) = \frac{1}{\sqrt{2\pi}} e^{in\theta}} \quad (2.152)$$

Le produit scalaire de deux modes propres distincts est :

$$\frac{1}{2\pi} \int_0^{2\pi} (e^{in\theta})^* e^{in'\theta} d\theta = \delta_{nn'} , \quad (2.153)$$

égalité qui résulte d'une part de la normalisation de chaque fonction, d'autre part du fait que la différence $n - n'$ est un entier. Au total, (2.153) exprime l'orthonormalisation de la base de fonctions $\psi_n(\theta)$.

Considérons maintenant la somme :

$$\sum_{n=-N}^{+N} \psi_n(\theta) \psi_n^*(\theta') = \frac{1}{2\pi} \sum_{n=-N}^{+N} e^{in(\theta-\theta')} = \frac{1}{2\pi} \frac{\sin(N + \frac{1}{2})(\theta - \theta')}{\sin \frac{1}{2}(\theta - \theta')} \stackrel{\text{d'éf}}{=} \delta_N(\theta - \theta') . \quad (2.154)$$

La fonction $\delta_N(\theta)$ est 2π -périodique ; elle vaut $\frac{1}{2\pi}(2N+1)$ en $\theta = 0$ et s'annule pour la première fois en $\frac{2\pi}{2N+1}$: pour $N \gg 1$, son graphe est donc un *peigne* régulier aux dents très fines et très hautes (voir fig. 2.6). L'intégrale de $\delta_N(\theta)$ sur tout intervalle de mesure 2π vaut visiblement 1 : il suffit d'intégrer la somme finie terme à terme ; seule l'intégrale sur le terme $n = 0$ est non nulle, d'où :

$$\int_{\theta_0}^{\theta_0+2\pi} \delta_N(\theta) d\theta = 1 \quad \forall N . \quad (2.155)$$

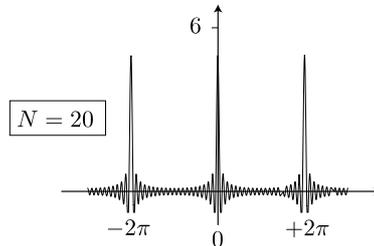


Figure 2.6: Graphe de la fonction $\delta_N(\theta)$ pour $N = 20$.

Visiblement, la suite de fonctions $\delta_N(\theta)$, si elle n'est pas régulière au sens défini ci-dessus, permet néanmoins de définir une certaine distribution, que l'on va trouver égale à δ dans la mesure où son action est strictement la même. En effet, soit la suite d'intégrales :

$$I_N \stackrel{\text{d'éf}}{=} \int_{-\pi}^{+\pi} \delta_N(\theta) \phi(\theta) d\theta , \quad (2.156)$$

où $\phi(\theta)$ est une fonction infiniment dérivable et à support borné $[-\pi, +\pi]$. Il n'est pas difficile de montrer que :

$$\lim_{N \rightarrow +\infty} \int_{-\pi}^{+\pi} \delta_N(\theta) \phi(\theta) d\theta = \phi(0) , \quad (2.157)$$

ce qui autorise à écrire l'égalité :

$$\lim_{N \rightarrow +\infty} \int_{-\pi}^{+\pi} \delta_N(\theta) \phi(\theta) d\theta = \langle \delta, \phi \rangle , \quad (2.158)$$

ou encore :

$$\boxed{\lim_{N \rightarrow +\infty} \sum_{n=-N}^{+N} \psi_n(\theta) \psi_n^*(\theta') = \delta(\theta - \theta') \quad (-\pi < \theta, \theta' < +\pi)} \quad (2.159)$$

C'est la relation de fermeture pour le problème considéré ; noter que cette forme simple est liée au fait que les ψ_n sont à la fois orthogonales entre elles, et normalisées à l'unité (voir (2.153)). On peut aussi l'écrire comme suit :

$$\lim_{N \rightarrow +\infty} \sum_{n=-N}^{+N} e^{in\theta} = 2\pi \delta(\theta) \quad (-\pi < \theta < +\pi) ; \quad (2.160)$$

cela étant, il est possible de prolonger périodiquement cette égalité sur tout l'axe réel, puisque la fonction au premier membre est 2π -périodique. On obtient ainsi :

$$\lim_{N \rightarrow +\infty} \sum_{n=-N}^{+N} e^{in\theta} = 2\pi \sum_{k \in \mathbb{Z}} \delta(\theta - k 2\pi) \quad \forall \theta \in \mathbb{R} ; \quad (2.161)$$

Utilisant $2\pi \delta(x) = \delta\left(\frac{x}{2\pi}\right)$, ceci s'écrit encore :

$$\boxed{\sum_{n \in \mathbb{Z}} e^{in\theta} = \sum_{k \in \mathbb{Z}} \delta\left(\frac{\theta}{2\pi} - k\right) \stackrel{\text{déf}}{=} \text{III}(\theta)} \quad (2.162)$$

où $\text{III}(\theta)$ est le *peigne de Dirac*, par définition, représenté par un symbole évocateur.

Donnons un autre exemple, à propos de la particule libre sur le segment de droite $[0, L]$ (puits infini de largeur L), dont les modes propres ont déjà été donnés (voir (2.10)). Comme précédemment pour le cercle, considérons maintenant la somme :

$$S_N(x, x') \stackrel{\text{déf}}{=} \sum_{n=1}^{+N} \psi_n(x) \psi_n^*(x') = \frac{2}{L} \sum_{n=1}^{+N} \sin \frac{n\pi x}{L} \sin \frac{n\pi x'}{L} , \quad (2.163)$$

Posant $k \stackrel{\text{déf}}{=} \frac{\pi}{L}$ et écrivant les sinus avec les formules d'Euler, on obtient :

$$S_N(x, x') = \frac{1}{(2i)^2} \frac{2}{L} \left[2\pi \delta_N(k(x+x')) - 1 - [2\pi \delta_N(k(x-x')) - 1] \right] , \quad (2.164)$$

où δ_N est la fonction définie en (2.154). x et x' étant strictement positifs, toutes les intégrales définissant précisément les distributions seront restreintes à l'intervalle borné $[0, L]$, et on peut donc oublier le terme $\delta_N(k(x+x'))$, qui donne à la limite la distribution $\delta(x+x')$. En définitive, on obtient :

$$S_N(x, x') = \frac{\pi}{L} \delta_N(k(x-x')) ; \quad (2.165)$$

ce qui autorise à écrire (symboliquement), utilisant $\delta(ax) = |a|^{-1} \delta(x)$ avec $a = \frac{\pi}{L}$:

$$\boxed{\lim_{N \rightarrow +\infty} \sum_{n=1}^N \psi_n(x) \psi_n^*(x') = \delta(x-x')} \quad (2.166)$$

qui est la relation de fermeture pour la particule libre confinée sur $[0, L]$. Il est possible d'établir une correspondance précise entre le cercle de longueur L et le segment de droite de même longueur dans la limite $L \rightarrow +\infty$, en définissant les conditions cycliques de Born - von Kármán, qui sont d'usage constant en Théorie des champs, en matière condensée, etc.

2.4.3 Les représentations $-q$ et $-p$ en Mécanique quantique

Il est à peine exagéré de dire que toute la Mécanique quantique peut se déduire de la relation fondatrice :

$$\boxed{qp - pq \stackrel{\text{déf}}{=} [q, p] = i\hbar \mathbf{1}} \quad (2.167)$$

où q et p sont les représentants quantiques (*observables*) de la coordonnée et du moment conjugué d'une particule se déplaçant sur \mathbb{R} , $\mathbf{1}$ désignant l'opérateur identité (qui ne fait rien). Une telle relation impliquant un produit non commutatif ne peut tenir qu'entre²² *opérateurs*. Comme l'a montré Schrödinger, une première façon de représenter cette égalité est d'admettre l'association :

$$q \longrightarrow \times q, \quad p \longrightarrow -i\hbar \frac{\partial}{\partial q} \quad (2.168)$$

puisque $[q, -i\hbar \partial_q] \psi(q) = i\hbar \psi(q)$ pour toute fonction dérivable. L'association (2.168) définit la représentation- q de la Mécanique quantique, où l'état d'un système de spin nul est entièrement décrit par une fonction d'onde $\Psi(q, t)$, dépendant de la coordonnée et du temps.

Un peu de réflexion montre que l'égalité (2.167) est tout autant satisfaite par l'association *duale* :

$$q \longrightarrow +i\hbar \frac{\partial}{\partial p}, \quad p \longrightarrow \times p, \quad (2.169)$$

puisque pour toute fonction dérivable $\phi(p)$, on a $[i\hbar \partial_p, p] \phi(p) = i\hbar \phi(p)$; dans cette façon de faire, l'état est décrit par une fonction d'onde $\Phi(p, t)$. Clairement, rien ne privilégie

²²On ignore des nombres du genre de Grassmann, qui d'ailleurs *anticommutent*.

dans l'absolu une *représentation* ou une autre, toutes deux (et n'importe quelle autre respectant (2.167)) devant avoir la même signification physique.

La question arrive alors naturellement : pour un même état physique représenté (au même instant) soit par $\Psi(q, t)$, soit par $\Phi(p, t)$ suivant l'humeur du moment (ou plutôt la commodité technique du problème), quelle est la relation précise entre ces deux fonctions ? Compte tenu du sens physique attribué à la fonction d'onde, quelle que soit la représentation adoptée, on montre (voir cours de Mécanique quantique) que les deux fonctions $\Psi(q, t)$ et $\Phi(p, t)$ sont justement reliées par la transformation de Fourier, que l'on écrit en général sous la forme symétrique :

$$\Phi(p, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{\frac{i}{\hbar}pq} \Psi(q, t) dq \iff \Psi(q, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{-\frac{i}{\hbar}pq} \Phi(p, t) dp . \quad (2.170)$$

En vertu de la linéarité de l'équation de Schrödinger $i\hbar\partial_t\bullet = H\bullet$, les fonctions $\Psi(q, t)$ et $\Phi(p, t)$ peuvent toujours être développées en combinaison linéaire (avec des coefficients dépendant du temps) de modes propres, $\psi_\alpha(q)$ et $\phi_\beta(p)$ respectivement, propres de l'équation $H\bullet = E\bullet$. Prenons le cas d'une particule libre de masse m , pour laquelle $H = \frac{p^2}{2m}$. En représentation- q , les fonctions propres satisfont :

$$\frac{1}{2m} \left(-i\hbar \frac{\partial}{\partial q} \right)^2 \psi(q) = E\psi(q) ; \quad (2.171)$$

en représentation- p , l'équation propre a la forme :

$$\frac{1}{2m} p^2 \phi(p) = E\phi(p) \iff (p^2 - 2mE) \phi(p) = 0 . \quad (2.172)$$

Cette équation est du type (2.61), et a donc pour solution :

$$\phi(p) = C_+ \delta(p - p_+) + C_- \delta(p - p_-) , \quad p_{\pm} = \sqrt{2mE} . \quad (2.173)$$

De fait, la solution la plus générale de (2.171) est une combinaison linéaire arbitraire des deux ondes planes $e^{\pm \frac{i}{\hbar} \sqrt{2mE} q}$, qui sont bien, selon (2.123), les transformées de Fourier des distributions $\delta(p - p_{\pm})$.

L'ensemble des solutions propres en représentation- q est donc constitué des ondes planes Ae^{ikq} , l'énergie E étant alors donnée par $E = \frac{\hbar^2 k^2}{2m}$ et k prenant toutes les valeurs dans \mathbb{R} . Soit maintenant la somme :

$$S_{k_{\max}}(q - q') \stackrel{\text{d\'ef}}{=} \int_{-k_{\max}}^{+k_{\max}} |A|^2 e^{ikq} e^{-ikq'} dk = 2|A|^2 \frac{\sin k_{\max}(q - q')}{q - q'} \equiv 2\pi |A|^2 \delta_{k_{\max}}(q - q') , \quad (2.174)$$

où $k_{\max} > 0$ est un vecteur d'onde de coupure, tronquant l'intégrale du côté des hautes énergies et où :

$$\delta_{k_{\max}}(q - q') \stackrel{\text{d\'ef}}{=} \frac{1}{\pi} \frac{\sin k_{\max}(q - q')}{q - q'} . \quad (2.175)$$

L'intégrale $\int_{\mathbb{R}} \frac{\sin \lambda x}{x} dx$ vaut²³ $\pi \operatorname{sgn} \lambda$ quel que soit $\lambda > 0$ (intégrale de Dirichlet) : l'intégrale de $\delta_{k_{\max}}(q)$ est donc égale à 1, et définit en fait une suite (non régulière) de la classe de δ . En effet, soit à trouver la limite :

$$\lim_{k_{\max} \rightarrow +\infty} \int_{\mathbb{R}} \delta_{k_{\max}}(q) \phi(q) dq, \quad (2.177)$$

où $\phi(q)$ est une bonne fonction ; pour cela, montrons que la différence :

$$\Delta_{k_{\max}} \stackrel{\text{déf}}{=} \int_{\mathbb{R}} \delta_{k_{\max}}(q) \phi(q) dq - \phi(0) \quad (2.178)$$

tend vers zéro. En posant $k_{\max} q = X$, elle s'écrit aussi :

$$\Delta_{k_{\max}} = \int_{\mathbb{R}} \frac{1}{\pi} \frac{\sin X}{X} \left[\phi\left(\frac{X}{k_{\max}}\right) - \phi(0) \right] dX. \quad (2.179)$$

On a $\phi\left(\frac{X}{k_{\max}}\right) - \phi(0) = \int_0^{X/k_{\max}} \phi'(q') dq'$, d'où :

$$|\Delta_{k_{\max}}| \leq \left| \int_{\mathbb{R}} \frac{1}{\pi} \frac{\sin X}{X} \max_{[0, X/k_{\max}]} \phi'(X) \frac{X}{k_{\max}} \right| = \mathcal{O}(k_{\max}^{-1}), \quad (2.180)$$

soit $\lim_{k_{\max} \rightarrow +\infty} \int_{\mathbb{R}} \delta_{k_{\max}}(q) \phi(q) dq = \phi(0)$, et finalement :

$$\lim_{k_{\max} \rightarrow +\infty} S_{k_{\max}}(q - q') = 2\pi |A|^2 \delta(q - q'). \quad (2.181)$$

Ceci montre qu'en prenant $\psi_k(q) = \frac{1}{\sqrt{2\pi}} e^{ikq}$, on obtient la relation de fermeture sous sa forme normale :

$$\lim_{k_{\max} \rightarrow +\infty} \int_{-k_{\max}}^{+k_{\max}} \psi_k(q) \psi_k^*(q') dk = \delta(q - q') \quad (2.182)$$

Quant au produit scalaire $(\psi_k, \psi_{k'})$, il ressort comme $\frac{1}{2\pi} \int_{\mathbb{R}} e^{-ikq} e^{ik'q} dq$, soit la représentation symbolique de $\delta(k - k')$. Enfin, revenant à la transformation (2.170), on obtient $\phi(p) = \sqrt{\hbar} \delta(p - \hbar k)$, d'où pour une particule libre :

$$\boxed{\text{représentation-}q : \psi_k(q) = \frac{1}{\sqrt{2\pi}} e^{ikq} \longleftrightarrow \text{représentation-}p : \phi_k(p) = \sqrt{\hbar} \delta(p - \hbar k)} \quad (2.183)$$

avec²⁴ :

$$(\psi_k, \psi_{k'}) = \delta(k - k') = (\phi_k, \phi_{k'}). \quad (2.184)$$

Ces résultats sont physiquement évidents : une particule libre d'énergie déterminée a une impulsion parfaitement déterminée (au signe près), et est donc complètement délocalisée

²³Très précisément :

$$\lim_{X \rightarrow +\infty} \int_{-X}^{+X} \frac{\sin \lambda x}{x} dx = \pi \operatorname{sgn} \lambda; \quad (2.176)$$

(faire le lien avec $\mathcal{F}[\mathcal{P}\frac{1}{x}]$).

²⁴Remarquer l'invariance des produits scalaires, exprimant que la transformation de Fourier est une opération *unitaire*.

dans tout l'espace (sa probabilité de présence est la même partout). C'est peu dire que cette affirmation, prise au sens strict, n'a pas grand sens en Physique : la bonne façon de s'exprimer est de dire que sur une échelle de longueur aussi grande que l'on veut comparée à l'échelle du phénomène considéré, la probabilité de présence est constante.

2.5 En guise de conclusion...

L'un des buts de ce chapitre était de faire le lien entre les résultats formels exposés ci-dessus et les procédés expéditifs couramment employés dans la pratique en Physique, ressemblant à première vue à des *manips*' pour le moins douteuses²⁵... On a donné ici et là des *trucs* présentés comme mnémotechniques pour retrouver certains résultats : ces manipulations doivent en fait rentrer dans les idiosyncrasies propres du Physicien.

L'une des idées-clés à retenir est que toutes les "fonctions" un peu bizarres rencontrées souvent dans la pratique doivent en fait intervenir dans des sommes, qui le plus souvent sont des sommes sur une variable continue, c'est-à-dire des intégrales ; c'est le cas notamment quand on convolue un signal d'entrée avec une fonction d'appareil pour obtenir un signal de sortie.

Intervenant dans des intégrales avec des fonctions ordinaires, ces dernières servent jouent le rôle des fonctions-tests désignées généralement par $\phi(x)$ dans les sections précédentes. Dans tous les cas pratiques, les fonctions physiquement pertinentes auront les propriétés souhaitables pour que les résultats ci-dessus, éventuellement généralisés, soient applicables. Par exemple, il pourra s'agir des états liés d'un système quantique, $\psi_n(x)$, qui, usuellement, ont un comportement typique du genre $|x|^n e^{-k|x|}$ à l'infini, et sont donc des "bonnes" fonctions. Dans d'autres cas, des considérations physiques seront toujours disponibles pour assurer une coupure permettant d'assimiler les fonctions-tests apparaissant naturellement dans le problème à des fonctions à support borné. La discussion précise pourra (*devra*), dans chaque cas, relever d'une analyse des échelles physiques intrinsèques à ce problème.

²⁵ "God does not care about our mathematical difficulties. He integrates empirically." (Albert EINSTEIN).

Chapitre 3

Fonctions d'une variable complexe

“Si la vie est complexe, c'est parce qu'elle a une partie réelle et une partie imaginaire.”

(Sophus LIE, 1842–1899)

Le but de ce chapitre est d'une part de rappeler les règles élémentaires de l'algèbre des nombres complexes, d'autre part de montrer graduellement comment des notions, familières dans le champ réel (limite d'une fonction, continuité, dérivation) peuvent se généraliser au champ complexe.

3.1 Rappels des opérations élémentaires sur les nombres complexes

Toute l'algèbre des nombres complexes (dont l'ensemble est noté \mathbb{C}) repose l'existence postulée d'un nombre fondamental noté i dont la seule et unique propriété extraordinaire est que son carré est égal à -1 :

$$\boxed{i^2 = -1} \tag{3.1}$$

Une fois admis ceci, un nombre complexe quelconque, z , est par définition un nombre construit en combinaison linéaire à coefficients réels de 1 et de i :

$$\forall z \in \mathbb{C} : z = a \times 1 + b \times i, \quad a \in \mathbb{R}, b \in \mathbb{R} \tag{3.2}$$

ce que l'on note plus simplement :

$$\boxed{z = a + ib} \tag{3.3}$$

Toutes les opérations (addition, $+$, et multiplication, \times) sont commutatives. L'addition sera précisément définie plus loin et n'a pour l'instant qu'une signification symbolique. La multiplication $a \times 1$ est l'opération usuelle des réels ; avec $b \in \mathbb{N}$, $b \times i$ est un raccourci d'écriture pour $i + i + \dots + i$ (b fois), notation généralisée immédiatement à un réel b quelconque. a est appelé partie réelle de z , b est la partie imaginaire de z , respectivement notées $\Re z \stackrel{\text{déf}}{=} a$, $\Im z \stackrel{\text{déf}}{=} b$. Deux nombres complexes sont dits égaux s'ils ont même partie réelle et même partie imaginaire :

$$\boxed{z_1 = z_2 \iff a_1 = a_2 \quad \text{et} \quad b_1 = b_2} \quad (3.4)$$

Le nombre complexe égal à zéro est celui dont les parties réelle et imaginaire sont toutes deux nulles :

$$z = 0 \iff a = 0 \quad \text{et} \quad b = 0 . \quad (3.5)$$

Un nombre complexe dont la partie réelle est nulle est dit *imaginaire pur*, un nombre complexe dont la partie imaginaire est nulle est un nombre réel "ordinaire". Pour un nombre complexe, qui est un être "bidimensionnel", la relation $z > 0$ n'a pas de sens – sauf si on sait d'avance que z est en fait un réel.

Visiblement, les nombres 1 et i jouent dans (3.2) le même rôle que deux vecteurs orthonormalisés \vec{i} et \vec{j} et l'on voit que \mathbb{C} va pouvoir être muni d'une structure d'espace vectoriel à deux dimensions sur le corps des réels : de ce point de vue, l'ensemble des nombres \mathbb{C} et l'ensemble \mathbb{R}^2 des couples de réels (x, y) sont isomorphes ; dans ce contexte on dit de \mathbb{R}^2 qu'il est le *plan complexe*. Le nombre complexe z est ainsi représenté dans le plan par un point désigné¹ par m dont les coordonnées, par rapport à un repère orthonormé, sont précisément a et b – ou plus naturellement x et y –, et on dit que z est l'*affiche* du point m . D'où la notation systématique, un repère orthonormé (Ox, Oy) étant défini :

$$z = x + iy , \quad x \in \mathbb{R}, y \in \mathbb{R} . \quad (3.6)$$

Deux nombres complexes égaux sont représentés par le même point du plan. En pratique, pour ne pas alourdir les notations et quand aucune confusion n'est possible, on assimilera le plus souvent un nombre z et son point représentatif m .

L'équation (3.3) contient l'opération $i \times b$, multiplication du nombre fondamental i et d'un réel b . Ce dernier est sur l'axe Ox à l'abscisse b alors que le nombre $i \times b$ est sur l'axe Oy à la cote b , d'où l'interprétation géométrique de la multiplication par i :

$$\boxed{\text{multiplication par } i \iff \text{rotation de } +\frac{\pi}{2} \text{ dans le plan } \mathbb{R}^2} \quad (3.7)$$

De même, le carré de i , est associé au produit de deux rotations de $\frac{\pi}{2}$, c'est donc une rotation de π – et en effet comme $i^2 = -1$, $i^2 b$ est l'affiche du point symétrique représentant le réel b , qui de l'autre côté de l'origine O . Plus généralement, $i^2 z = -z$ est l'affiche du point image de celui associé à z dans la symétrie par rapport à l'origine.

L'existence de i étant admise, et avec elle les nombres définis par (3.6), il est naturel de généraliser pour ces nombres les opérations usuelles de l'algèbre élémentaire.

¹Dans la mesure du possible, on essaiera de maintenir des notations systématiques : z est l'affiche de m , Z est celle de M , etc.

La définition de l'addition est déjà embryonnaire dans la définition (3.6), qui additionne un réel et un nombre imaginaire pur. Plus généralement, la somme (addition) de deux nombres complexes est tout naturellement ($Z = X + iY$) :

$$\boxed{Z \stackrel{\text{déf}}{=} z_1 + z_2 \iff X = x_1 + x_2 \text{ et } Y = y_1 + y_2} \quad (3.8)$$

La soustraction s'en déduit par $z_1 - z_2 \equiv z_1 + (-z_2)$. Géométriquement, si z_p est l'affixe de m_p et Z celle de M , la définition (3.8), montre que l'addition de deux complexes est en correspondance biunivoque avec la somme vectorielle :

$$\overrightarrow{OM} = \overrightarrow{Om_1} + \overrightarrow{Om_2} \quad (3.9)$$

La multiplication de deux nombres complexes s'obtient en utilisant les règles usuelles de distributivité ($z_p = x_p + iy_p$) :

$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = x_1 x_2 + i(x_1 y_2 + y_1 x_2) + i^2 y_1 y_2, \quad (3.10)$$

soit :

$$z_1 z_2 = x_1 x_2 - y_1 y_2 + i(x_1 y_2 + y_1 x_2), \quad (3.11)$$

d'où :

$$\boxed{Z \stackrel{\text{déf}}{=} z_1 z_2 \iff \Re Z = x_1 x_2 - y_1 y_2, \quad \Im Z = x_1 y_2 + y_1 x_2} \quad (3.12)$$

La multiplication étant définie, la division l'est tout autant. De toute évidence, toutes ces opérations se réduisent aux opérations élémentaires avec les nombres réels si toutes les parties imaginaires sont nulles ; en conséquence, toutes les égalités algébriques (identités remarquables, formule du binôme, etc.) se transcrivent immédiatement dans \mathbb{C} – par exemple, le carré d'un nombre complexe est $(x + iy)^2 = x^2 - y^2 + 2ixy$.

On appelle *conjugué* de z le nombre, noté² z^* , qui a même partie réelle et dont la partie imaginaire est changée de signe :

$$z = x + iy, \quad z^* = x - iy, \quad (3.13)$$

d'où résulte immédiatement $\Re z = \frac{1}{2}(z + z^*)$, $\Im z = \frac{1}{2i}(z - z^*)$. Géométriquement, z^* et z sont les affixes des points se transformant l'un dans l'autre par la symétrie-miroir définie par l'axe Ox . En posant $z_p = x_p + iy_p$ ($p = 1, 2$), on voit que :

$$(z_1 + z_2)^* = z_1^* + z_2^*, \quad (z_1 z_2)^* = z_1^* z_2^*. \quad (3.14)$$

Le produit d'un complexe et de son conjugué est remarquable :

$$z z^* = (x + iy)(x - iy) = x^2 - (iy)^2 = x^2 + y^2. \quad (3.15)$$

²Une autre notation usuelle pour le complexe conjugué z^* est \bar{z} .

C'est donc un nombre réel positif, nul seulement quand $z = 0$, *i.e.* quand $x = 0$ et $y = 0$. La quantité (positive) $\sqrt{x^2 + y^2}$ est le *module* de z , noté³ $|z|$:

$$\boxed{|z| = \sqrt{x^2 + y^2} = \sqrt{zz^*} \geq 0} \quad (3.16)$$

Le symbole $||$ généralise au champ complexe la notion familière de valeur absolue, et permet de définir une distance euclidienne dans \mathbb{C} . On vérifie sans peine que le produit $z_1 z_2$ a pour module le produit des modules : $|z_1 z_2| = |z_1| |z_2|$; en particulier $|\frac{1}{z}| = \frac{1}{|z|}$.

Géométriquement parlant, $|z|$ est la distance au sens usuel entre son point représentatif m et l'origine O du plan complexe. De la même façon, le module de la différence, $|z_1 - z_2|$, est la distance entre les deux points représentatifs m_1 et m_2 de z_1 et z_2 : c'est donc aussi la *longueur* du segment $m_1 m_2$. Par ailleurs, l'interprétation géométrique de la somme de deux complexes, (3.9), donne immédiatement l'inégalité triangulaire :

$$\boxed{|z_1 + z_2| \leq |z_1| + |z_2|} \quad (3.17)$$

qui exprime le fait que dans un triangle la longueur d'un côté est inférieure ou égale à la somme des deux autres (le module d'une somme est inférieur ou égal à la somme des modules). Plus généralement :

$$\left| \sum_{j=1}^N z_j \right| \leq \sum_{j=1}^N |z_j| \quad (3.18)$$

Une autre inégalité utile est :

$$\boxed{||z_1| - |z_2|| \leq |z_1 + z_2|} \quad (3.19)$$

En effet, par l'inégalité triangulaire, $|z_1| \equiv |z_1 + z_2 - z_2| \leq |z_1 + z_2| + |z_2|$, d'où $|z_1| - |z_2| \leq |z_1 + z_2|$, et de même $|z_2| - |z_1| \leq |z_2 + z_1|$, d'où $|z_2 + z_1| \geq \max(|z_1| - |z_2|) = ||z_1| - |z_2||$ et l'inégalité (3.19).

La décomposition $z = x + iy$ présuppose choisi un repère cartésien permettant de définir les deux coordonnées cartésiennes du point dont l'affixe est z (et on parle alors de *représentation cartésienne* de z). On peut aussi introduire les coordonnées polaires⁴ du plan, r et θ , auquel cas on obtient l'écriture polaire du nombre complexe z :

$$z = x + iy = r \cos \theta + ir \sin \theta, \quad r = |z|, \quad \theta = \arg z. \quad (3.20)$$

La combinaison $\cos \theta + i \sin \theta$ est remarquable. Toutes ses propriétés algébriques permettent de l'identifier avec une certaine fonction exponentielle, ce qui conduit à la formule d'Euler :

$$\boxed{\cos \theta + i \sin \theta = e^{i\theta}} \quad (3.21)$$

³On utilise le même symbole $|\dots|$ pour le module d'un complexe et la valeur absolue ordinaire, puisque le premier ne fait que généraliser la seconde au domaine complexe.

⁴ $r \geq 0$, $0 \leq \theta < 2\pi$ (par exemple).

En effet, considérons le premier membre de (3.21) comme une certaine fonction $E(\theta)$ et écrivons-la pour $\theta = \theta_1 + \theta_2$; par définition :

$$E(\theta_1 + \theta_2) = \cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2) . \quad (3.22)$$

Soit maintenant le produit $E(\theta_1)E(\theta_2) = \prod_{p=1}^2 (\cos \theta_p + i \sin \theta_p)$; sa partie réelle est :

$$\begin{aligned} \Re[E(\theta_1)E(\theta_2)] &= \Re \prod_{p=1}^2 (\cos \theta_p + i \sin \theta_p) = \\ &= \cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2 = \cos(\theta_1 + \theta_2) = \Re[E(\theta_1 + \theta_2)] ; \end{aligned} \quad (3.23)$$

sa partie imaginaire est :

$$\begin{aligned} \Im[E(\theta_1)E(\theta_2)] &= \Im \prod_{p=1}^2 (\cos \theta_p + i \sin \theta_p) = \\ &= \sin \theta_1 \cos \theta_2 + \sin \theta_2 \cos \theta_1 \sin(\theta_1 + \theta_2) = \Im[E(\theta_1 + \theta_2)] , \end{aligned} \quad (3.24)$$

d'où l'égalité des deux nombres complexes $E(\theta_1 + \theta_2)$ et $E(\theta_1)E(\theta_2)$:

$$\boxed{E(\theta_1 + \theta_2) = E(\theta_1)E(\theta_2)} \quad (3.25)$$

Il existe une seule fonction satisfaisant cette équation fonctionnelle : c'est la fonction exponentielle. Il en résulte que $E(\theta) = e^{a\theta}$ où a est une constante que l'on peut trouver de bien des façons, par exemple en calculant la limite $\lim_{\theta \rightarrow 0} \frac{1}{\theta}(e^{a\theta} - 1)$; par la définition de $E(\theta) \stackrel{\text{déf}}{=} \cos \theta + i \sin \theta$ et utilisant $\lim_{\theta \rightarrow 0} \frac{1}{\theta}(\cos \theta - 1) = 0$, $\lim_{\theta \rightarrow 0} \frac{\sin \theta}{\theta} = 1$, il vient $\lim_{\theta \rightarrow 0} \frac{1}{\theta}(e^{a\theta} - 1) = i$; l'identification avec le développement en série entière de la fonction exponentielle donne alors $a = i$, d'où la formule d'Euler (3.21) ; de celle-ci découlent immédiatement les deux égalités :

$$\boxed{\cos \theta = \frac{1}{2}(e^{i\theta} + e^{-i\theta}) \quad \sin \theta = \frac{1}{2i}(e^{i\theta} - e^{-i\theta})} \quad (3.26)$$

Le module du premier membre de (3.21) est $\sqrt{\cos^2 \theta + \sin^2 \theta} = 1$. On retiendra donc :

$$\boxed{|e^{i\theta}| = 1 \quad \forall \theta \in \mathbb{R}} \quad (3.27)$$

Cela étant fait, l'expression polaire d'un complexe $z = r \cos \theta + ir \sin \theta$ s'écrit aussi :

$$\boxed{z = r e^{i\theta}} \quad (3.28)$$

$r \geq 0$ est le module de z , θ est son *argument*. L'égalité de deux nombres complexes $z_p = r_p e^{i\theta_p}$ ($p = 1, 2$) s'exprime comme :

$$z_1 = z_2 \iff r_1 = r_2 \quad \text{et} \quad \theta_1 = \theta_2 \quad (2\pi) . \quad (3.29)$$

Notons que si le module du nombre complexe nul est bien défini ($r = 0$), en revanche son argument ne l'est pas, traduisant le fait géométrique que l'on peut s'approcher de l'origine suivant une direction quelconque. On définit parfois la *détermination principale* de l'argument d'un nombre complexe, notée $\text{Arg } z$, comme étant l'argument de z choisi dans l'intervalle $]-\pi, +\pi]$:

$$\boxed{-\pi < \text{Arg } z \leq +\pi} \quad (3.30)$$

La multiplication (et la division) s'expriment très simplement en représentation polaire :

$$z_1 = r_1 e^{i\theta_1}, z_2 = r_2 e^{i\theta_2} \implies z_1 z_2 = r_1 r_2 e^{i\theta_1} e^{i\theta_2} = r_1 r_2 e^{i(\theta_1 + \theta_2)} . \quad (3.31)$$

Autrement dit, $|z_1 z_2| = r_1 r_2$, $\arg(z_1 z_2) = \arg z_1 + \arg z_2 (2\pi)$. Ainsi par exemple, avec $z = r e^{i\theta}$, $\frac{1}{z} = \frac{1}{r} e^{-i\theta}$, $z^* = r e^{-i\theta}$. En représentation polaire, les puissances de $z = r e^{i\theta}$ prennent la forme :

$$z^n = (r e^{i\theta})^n = r^n e^{in\theta} = r^n (\cos n\theta + i \sin n\theta)^n , \quad (3.32)$$

où la formule d'Euler (3.21) a été utilisée pour écrire l'égalité le plus à droite. En prenant $r = 1$ et en remplaçant z au premier membre par $\cos \theta + i \sin \theta$, on obtient l'importante formule de Moivre :

$$\boxed{(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta} \quad (3.33)$$

Le développement de la puissance n^e du premier membre par la formule du binôme, et l'identification des parties réelle et imaginaire, constitue un moyen systématique pour exprimer les sin et cos d'un angle multiple. Par exemple :

$$\cos^2 \theta - \sin^2 \theta = \cos 2\theta , \quad 2 \sin \theta \cos \theta = \sin 2\theta , \quad (3.34)$$

$$\cos^3 \theta - 3 \cos \theta \sin^2 \theta = \cos 3\theta , \quad 3 \cos^2 \theta \sin \theta - \sin^3 \theta = \sin 3\theta . \quad (3.35)$$

À l'inverse, la représentation polaire permet d'exprimer simplement les n racines n^{es} d'un nombre complexe. Par exemple, soit à trouver les complexes satisfaisant :

$$z^n = \rho \quad (\rho > 0) . \quad (3.36)$$

Avec l'hypothèse $\rho > 0$, il s'agit plus précisément dans cet exemple de trouver les racines d'un réel positif noté ρ . En représentation polaire, (3.36) s'exprime comme :

$$(r e^{i\theta})^n = \rho ; \quad (3.37)$$

l'argument d'un nombre réel positif étant égal à $0 (2\pi)$, l'identification donne :

$$r^n = \rho , \quad n\theta = 0 (2\pi) \iff r = \rho^{\frac{1}{n}} , \quad \theta = k \frac{2\pi}{n} \quad (k = 1, 2, \dots, n) . \quad (3.38)$$

En particulier, les racines n^{es} de 1 sont les n nombres⁵ ω_k :

$$\boxed{\omega_k = e^{i \frac{2k\pi}{n}} \quad (k = 1, 2, \dots, n)} \quad (3.39)$$

⁵En utilisant $\sum_{k=1}^N q^k = q \frac{1-q^N}{1-q}$, $q \neq 1$, on vérifie sans peine que $\sum_{k=1}^n \omega_k = 0$ puisque $\omega_k^n = 1$.

Les points représentatifs des ω_k sont tous situés sur le cercle de rayon unité, et sont équidistants les uns des autres puisque l'on passe de l'un au suivant en tournant de $\frac{2\pi}{n}$. Les racines cubiques de l'unité jouent un rôle important en Électricité et sont souvent notées $1, j$ et j^2 :

$$n = 3 : \omega_k = e^{ik\frac{2\pi}{3}} , \quad \omega_1 = -\frac{1}{2} + i\frac{\sqrt{3}}{2} \equiv j , \quad \omega_2 = -\frac{1}{2} - i\frac{\sqrt{3}}{2} \equiv j^2 = j^* , \quad \omega_3 = 1 , \quad (3.40)$$

et satisfont :

$$1 + j + j^2 = 0 . \quad (3.41)$$

Les mêmes arguments montrent que :

$$r^n = -\rho \equiv \rho e^{i\pi} \iff r = \rho^{\frac{1}{n}} , \quad \theta = \frac{(2k+1)\pi}{n} \quad (k = 1, 2, \dots, n) . \quad (3.42)$$

Les racines de l'unité sont au cœur de sommes remarquables (les *sommes de Gauss*), que l'on retrouvera ici et là dans la suite (voir par exemple section 6.6).

3.2 Quelques éléments de topologie de \mathbb{C}

Il s'agit ici de donner quelques définitions utiles ici et là. Elles sont pour la plupart de nature topologique, et seront énoncées en termes aussi simples que possible ; la métrique du plan est euclidienne⁶, la distance entre deux complexes z_1 et z_2 étant :

$$d(z_1, z_2) \stackrel{\text{déf}}{=} |z_1 - z_2| \quad (3.43)$$

La métrique permet notamment la définition explicite de la notion de voisinage ; un nombre (ou point) z_0 étant donné, le ε -voisinage N_ε de z_0 est l'ensemble des points dont la distance à z_0 est inférieure à ε : $N_\varepsilon = \{z, |z - z_0| < \varepsilon\}$; le voisinage *pointé*, \dot{N}_ε est par définition l'ensemble $N_\varepsilon \setminus \{z_0\}$ (on ôte z_0 lui-même).

3.2.1 Ensembles de points dans le plan complexe

Toute collection de points dans \mathbb{C} est appelée ensemble de points ; usuellement, un ensemble est défini par une propriété de ses éléments. Par exemple, la partie droite du plan est l'ensemble des points dont les affixes⁷ sont à partie réelle strictement positive, ensemble que l'on pourra noter $\{z, \Re z > 0\}$; de même, $\{z, |z| < 1\}$ désigne l'intérieur du cercle unité. Cela étant posé, les notions d'appartenance, de complémentarité, de réunion, d'intersection,... introduites en théorie des ensembles sont utilisables telles quelles. Les définitions suivantes seront parfois utiles dans la suite.

⁶Une fois choisie une distance, \mathbb{C} devient un espace métrique.

⁷Pour simplifier le langage, on confond *point* et *affixe* dans la suite.

1. Un ensemble \mathcal{S} est dit *infini* s'il contient une infinité d'éléments⁸ ; il est dit *borné* s'il existe un nombre positif M tel que $|z| < M \forall z \in \mathcal{S}$.
2. Soit S un ensemble de points et S une partie non vide de \mathcal{S} . On appelle *point adhérent* à S un point $z_0 \in \mathcal{S}$ tel que l'intersection de l'un quelconque de ses voisinages avec S n'est pas vide. L'ensemble des points adhérents de S constitue l'*adhérence* de S , notée \bar{S} , qui est le plus petit fermé contenant S (toujours au sens de l'inclusion). L'adhérence est aussi appelée *fermeture*. $\bar{\mathbb{R}}$ désigne \mathbb{R} complété par le point à l'infini.
3. Soit S partie non vide de \mathcal{S} , et $z_0 \in \bar{S}$; alors :
 - ou bien il existe un voisinage N_ε de z_0 tel que $N_\varepsilon \cap S = \{z_0\}$, alors le point est *isolé*.
 - ou bien pour tout voisinage N_ε de z_0 , $(N_\varepsilon - \{z_0\}) \neq \emptyset$, alors le point z_0 est un *point d'accumulation* (ou *point limite*). Pour un tel point z_0 , tout voisinage contient au moins un point $z \neq z_0$ appartenant à S ; autrement dit, z_0 est adhérent à S sans y être isolé, ou encore z_0 est un point d'adhérence de⁹ $S - \{z_0\}$. Dans \mathbb{R} , la suite $\frac{1}{n^2}$, $n \in \mathbb{N}^*$ admet 0 pour point d'accumulation, tous ses éléments étant isolés.

En définitive, on peut aussi définir l'adhérence (fermeture) de S comme la réunion de l'ensemble des points isolés et des points d'accumulation.

Dans un espace métrique, tout voisinage d'un point limite contient une infinité de points de S ; de façon équivalente : il existe une suite de points de S convergeant vers z_0 . La précision $z \neq z_0$ est importante, faute de quoi tout point serait point d'accumulation puisqu'une suite constante est une suite convergente.

Le théorème de Bolzano - Weierstrass affirme : “*Tout ensemble infini borné de points possède au moins un point d'accumulation*”.

4. Un point est dit *intérieur* à un ensemble S s'il possède des voisinages ne contenant que des points de S . Un ensemble S est dit *ouvert* s'il ne contient que des points intérieurs (tout sous-ensemble est voisinage de chacun des points, un ouvert est égal à son intérieur) ; le disque $\{z, |z| < R\}$ est ouvert, de même que la couronne $\{z, r < |z| < R\}$. L'intérieur d'un ensemble S est le plus grand ouvert (au sens de l'inclusion) contenu dans S .
5. Un ensemble est dit *fermé* s'il contient tous ses points limites (ou s'il n'en a pas) ; le disque $\{z, |z - a| \leq R\}$ est fermé. Le complémentaire d'un ensemble ouvert est un ensemble fermé ; la couronne $\{z, r < |z - b| \leq R\}$ n'est ni ouverte, ni fermée.
6. Un point z_0 est dit *point-frontière* d'un ensemble S si tous ses voisinages contiennent des points de S et de son complémentaire. L'ensemble des points-frontières constitue la frontière de S .

⁸Puisqu'il s'agit d'énumérer les éléments, cet infini est l'infini *dénombrable*.

⁹Une autre notation courante pour $S - \{z_0\}$ est $S \setminus \{z_0\}$.

7. Un ensemble S est dit *compact* si de tout recouvrement par des ouverts on peut extraire une famille finie le recouvrant. Les compacts ont des propriétés remarquables (ils sont bornés et fermés) ; en particulier, tout sous-ensemble infini de S a au moins un point limite appartenant à S (en termes plus intuitifs : de toute suite de Cauchy, on peut extraire une sous-suite convergente). Pour \mathbb{R} , les parties compactes sont les intervalles $[a, b]$; pour \mathbb{C} , ce sont les intersections de deux disques fermés $\{z, r \leq |z - z_0| \leq R\}$.

3.2.2 Point à l'infini

Tout comme on ferme \mathbb{R} , il est utile de définir la fermeture de \mathbb{C} . Ceci peut se faire à l'aide de la transformation dite *projection stéréographique*, qui permet d'associer de façon biunivoque un point de \mathbb{C} à un point de la sphère unité de \mathbb{R}^3 . Cette transformation est définie comme suit¹⁰.

Le plan complexe est pris comme plan horizontal $\xi O \eta$ et on désigne par $O \zeta$ la troisième coordonnée, verticale. Dans \mathbb{R}^3 , la sphère de rayon $R = 1$ a pour équation cartésienne $\xi^2 + \eta^2 + \zeta^2 = 1$. Le point de coordonnées $(0, 0, 1)$ est le pôle Nord, N ; l'intersection de la sphère et du plan $\xi O \eta$ est l'équateur. Soit maintenant un point de coordonnées horizontales (x, y) , représentant le nombre complexe $z = x + iy$, affixe du point M. La transformation stéréographique associe à M le point d'intersection \mathcal{M} de la sphère avec la droite passant par M et N ; mis à part le pôle Nord, chaque point de \mathbb{C} a une image sur la sphère, et inversement à cette exception près, la correspondance est donc biunivoque. On voit sans peine que les points de \mathbb{C} situés à l'intérieur du cercle de rayon unité sont envoyés dans l'hémisphère sud, ceux situés à l'extérieur de ce cercle sont envoyés dans l'hémisphère Nord. Plus généralement, il saute aux yeux que plus un point est loin de l'origine de \mathbb{C} , plus son image est proche du pôle Nord (l'image de l'origine O est le pôle sud). En passant à la limite $|z| \rightarrow +\infty$, le pôle Nord apparaît comme l'image d'un "point" (qui n'est pas un nombre complexe : son argument est totalement indéterminé) appelé par définition *le point à l'infini*, noté $\infty_{\mathbb{C}}$ ou plus simplement ∞ quand aucune ambiguïté n'est à craindre. \mathbb{C} complété par ce point est le *plan complexe étendu*, noté $\bar{\mathbb{C}}$: $\bar{\mathbb{C}} \stackrel{\text{déf}}{=} \mathbb{C} \cup \infty_{\mathbb{C}}$. Cela étant fait, il existe une correspondance biunivoque entre $\bar{\mathbb{C}}$ et tous les points de la sphère (aussi appelée sphère de Riemann).

Il est facile d'obtenir les formules explicites de la transformation. On a :

$$\xi = \frac{2x}{r^2 + 1}, \quad \eta = \frac{2y}{r^2 + 1}, \quad \zeta = \frac{r^2 - 1}{r^2 + 1} \quad (3.44)$$

où $r = \sqrt{x^2 + y^2}$. On peut enfin définir une métrique¹¹ fondée sur la longueur de la corde reliant deux points transformés. Si (ξ_j, η_j, ζ_j) sont les coordonnées du point transformé

¹⁰D'autres définitions sont possibles. On peut par exemple introduire une sphère dont le pôle sud est tangent à l'origine du plan complexe ; c'est le choix qui est fait dans le livre de Lavrentiev et Chabat [9].

¹¹Cette appellation est légitime puisque $D(z_1, z_2)$ satisfait les conditions classiques : symétrie dans l'échange, inégalité triangulaire, positivité, nullité si les points sont confondus.

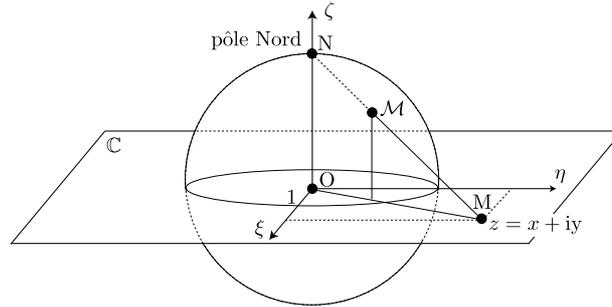


Figure 3.1: Sphère de Riemann et transformation stéréographique : le point M du plan \mathbb{C} se transforme en \mathcal{M} situé sur la sphère.

de $z_j = x_j + iy_j$ ($j = 1, 2$), la distance “cordale” de z_1 et z_2 est :

$$D(z_1, z_2) \stackrel{\text{déf}}{=} \sqrt{(\xi_1 - \xi_2)^2 + (\eta_1 - \eta_2)^2 + (\zeta_1 - \zeta_2)^2} ; \quad (3.45)$$

le calcul à partir de (3.44) donne :

$$D(z_1, z_2) = \frac{2|z_1 - z_2|}{\sqrt{1 + |z_1|^2} \sqrt{1 + |z_2|^2}} \quad (3.46)$$

Par exemple, si z_2 est le point à l’infini, $D(z_1, \infty)$ est la distance entre l’image de z_1 et le pôle Nord, soit :

$$D(z_1, \infty) = \sqrt{\xi_1^2 + \eta_1^2 + (\zeta_1 - 1)^2} = \sqrt{2(1 - \zeta_1)} = \frac{2}{\sqrt{1 + |z_1|^2}} \quad (3.47)$$

Pour un ensemble borné $\{z, |z| < M\}$, la distance ainsi définie et la distance euclidienne satisfont les inégalités :

$$D(z_1, z_2) \leq 2|z_1 - z_2| \leq (1 + M^2)D(z_1, z_2) , \quad (3.48)$$

qui découlent directement de la définition (3.45). Il en résulte que, pour les ensembles bornés, toutes les questions de convergence, de compacité,... peuvent être réglées avec l’une ou l’autre de ces distances (elles sont interchangeables). En revanche, un ensemble non borné possède le point à l’infini comme point limite vis-à-vis de D : il contient une suite infinie z_1, z_2, \dots telle que $\lim_{n \rightarrow \infty} z_n = \infty$, d’où résulte :

$$\lim_{n \rightarrow \infty} D(z_n, \infty) = \lim_{n \rightarrow \infty} \frac{2}{\sqrt{1 + |z_n|^2}} = 0 \quad (3.49)$$

Le plan complexe étendu $\bar{\mathbb{C}}$ est donc compact selon D , ce qui n’est pas surprenant puisque $\bar{\mathbb{C}}$ est borné vis-vis de D ($D(z_1, z_2) \leq 2$ quels que soient z_1 et z_2).

Il est possible de donner une interprétation du point à l’infini en introduisant la fonction $z \rightarrow Z = f(z) \stackrel{\text{déf}}{=} \frac{1}{z}$, qui est définie quel que soit $z \neq 0$; le cercle unité

est transformé en lui-même, ses points intérieurs deviennent des points extérieurs, et inversement. Les images par f des points $|z| > R$ sont dans le voisinage $N_{1/R}$ de l'origine du plan de Z . Par la transformation stéréographique, les mêmes points ont pour image les points de la sphère de Riemann tels que $\zeta > \frac{R^2-1}{R^2+1}$; ceux-ci sont donc dans un certain voisinage (selon D) du pôle Nord. On pourrait ainsi définir le point à l'infini dans le plan z comme l'antécédent de l'origine du plan de Z par f . En définitive, les deux plans étant complétés, il existe entre eux une correspondance biunivoque incluant l'origine de l'un avec le point à l'infini de l'autre et inversement.

3.2.3 Courbes et domaines

Il s'agit essentiellement de donner quelques définitions auxquelles il sera souvent fait référence dans la suite.

Courbes

On appelle *arc simple de Jordan* un ensemble de points formant une ligne (courbe) définie paramétriquement par $z(t) = x(t) + iy(t)$, où les deux fonctions continues (et à valeurs réelles) sont telles que $t_1 \neq t_2$ entraîne $z(t_1) \neq z(t_2)$, quand t varie dans un certain intervalle $[0, T]$: un tel arc ne revient donc pas deux fois au même point et n'est pas fermé sur lui-même. L'arc est dit *lisse*, si les deux fonctions $x'(t) \stackrel{\text{déf}}{=} \frac{dx}{dt}$ et $y'(t) \stackrel{\text{déf}}{=} \frac{dy}{dt}$ sont continues et si $x'^2 + y'^2 \neq 0$. Géométriquement, ceci signifie que le vecteur tangent, qui existe en tout point de la courbe en vertu de la dernière condition, tourne gentiment sans faire de sauts quand la ligne est décrite. Un arc de Jordan est *fermé* (en boucle) ssi l'égalité $z(t_1) = z(t_2)$ a lieu exclusivement dans l'un des deux cas : $t_1 = 0$ et $t_2 = T$ ou $t_1 = T$ et $t_2 = 0$.

L'exemple le plus simple d'arc fermé est le cercle de rayon unité, défini par :

$$z(t) = e^{it} = \cos t + i \sin t, \quad (3.50)$$

avec $T = 2\pi$. Il divise le plan en deux ouverts disjoints, et constitue la frontière de chacun d'entre eux ; il s'agit d'un cas particulier du théorème de Jordan affirmant que toute courbe simple fermée sépare le plan de la sorte, son intérieur étant borné, son extérieur ne l'étant pas.

Ces définitions peuvent s'étendre au cas où les conditions de continuité et de dérivabilité sont seulement satisfaites dans un nombre fini d'intervalles, aux bornes desquels toutes les limites voulues existent : on parle alors de courbes simples/lisses par morceaux. Dans toute la suite, on désignera simplement par *contour* une courbe simple lisse, éventuellement par morceaux. Le contour sera une boucle, fermé ou non, selon ce qu'est elle-même la courbe.

La longueur L de la courbe est donnée par l'intégrale de Riemann :

$$L = \int_0^T \sqrt{x'^2 + y'^2} dt \equiv \int_0^T |z'| dt , \quad (3.51)$$

ou par une somme d'intégrales sur les différents intervalles si la courbe est comme il faut seulement par morceaux.

Domaines

Un ensemble de points S est dit *connexe* si deux points quelconques de S peuvent être reliés par un arc simple de Jordan¹² tout entier situé dans S . Un *domaine* \mathcal{D} est ensemble connexe non vide : autour de tout point de \mathcal{D} on peut tracer un ε -voisinage tout entier contenu dans \mathcal{D} (tout point de \mathcal{D} a des voisinages appartenant à \mathcal{D}). Un domaine est donc *ouvert* ; la frontière sera notée $\partial\mathcal{D}$ et $\bar{\mathcal{D}} = \partial\mathcal{D} \cup \mathcal{D}$.

Un domaine \mathcal{D} est dit *simplement connexe* si l'intérieur de toute courbe de Jordan fermée de \mathcal{D} est aussi dans \mathcal{D} ; toute courbe fermée située entièrement dans ce domaine peut être contractée en un point en restant dans le domaine (on dit aussi : toute boucle est *homotope* à zéro). Le cercle est l'archétype de domaine simplement connexe. en termes plus imagés, un domaine simplement connexe "n'a pas de trous".

Soit maintenant deux domaines simplement connexes \mathcal{D}_1 et \mathcal{D}_2 avec $\mathcal{D}_1 \subset \mathcal{D}_2$. Le complémentaire de \mathcal{D}_1 dans \mathcal{D}_2 est tel que, deux points A et B y étant choisis, on peut distinguer deux sortes de chemins : ceux qui relient A à B en pénétrant dans \mathcal{D}_1 , et ceux qui, contournant \mathcal{D}_1 , ne sortent pas du complémentaire. Ce complémentaire est certes *connexe* (c'est d'ailleurs un domaine) puisqu'il existe des chemins allant d'un point à l'autre sans sortir du domaine, mais les chemins fermés ne sont pas tous homotopes à zéro : un tel domaine est dit *multiplément connexe* ("a des trous"). On verra bientôt l'importance d'une telle différenciation. Un exemple simple de domaine multiplément connexe est le domaine annulaire situé entre deux cercles de même centre (usuellement appelé *couronne*).

3.3 Fonction d'une variable complexe

3.3.1 Définitions

La notion de fonction est bien connue : c'est une correspondance qui permet, par des opérations algébriques bien définies (une *recette*), de construire un nombre (*image*) à

¹²La connexité ainsi définie est dite *connexité par arcs*. La notion de connexité peut être définie sur un plan plus formel en disant, par exemple, qu'un espace topologique est connexe s'il n'est pas la réunion de deux ouverts non vides disjoints, ou n'est pas la réunion de deux fermés non vides disjoints. La propriété d'être connexe est bien résumée par la formule : *être d'un seul tenant*.

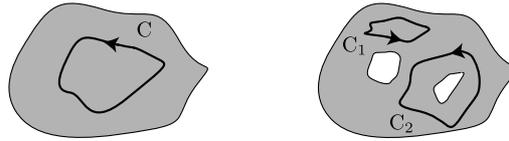


Figure 3.2: Les domaines grisés sont respectivement simplement connexe (à gauche), multiplément connexe (à droite). Les contours C et C_1 sont homotopes à zéro, C_2 ne l'est pas.

partir d'un nombre (*original*) ; ici, la seule nouveauté par rapport au champ réel est que tous les nombres impliqués sont maintenant *a priori* des nombres complexes. Cette nouveauté est moins banale qu'elle en a l'air : fondamentalement et essentiellement¹³, la différence est que les nombres original et image se déplacent dans un espace à *deux* dimensions (le plan complexe \mathbb{C} , isomorphe à \mathbb{R}^2), au lieu d'être confinés dans un domaine de l'axe réel, qui est de dimension 1. Ce degré de liberté supplémentaire est à l'origine des propriétés extraordinaires d'une classe de fonctions appelées *fonctions holomorphes*, qui sera définie en temps utile.

Ainsi, une fonction f de la variable complexe z est une opération (ou une suite d'opérations) qui, au nombre z dans un certain ensemble \mathcal{D} inclus dans le plan, associe un certain nombre complexe noté $f(z)$:

$$\forall z \in \mathcal{D} \quad \longrightarrow \quad f(z) \in \mathbb{C} . \quad (3.52)$$

\mathcal{D} est appelé l'ensemble de définition de la fonction f : c'est l'ensemble des points où l'on sait effectuer les opérations permettant de calculer $f(z)$. Le cas le plus important est celui où \mathcal{D} est un *domaine* du plan complexe. On ne considérera que des domaines dont la frontière est une suite finie d'arcs de courbes continûment différentiables par morceaux.

Quelques exemples de fonctions :

- $z \rightarrow f(z) = z^*$ est la fonction qui à chaque nombre complexe lui associe son complexe conjugué.
- $z \rightarrow f(z) = |z|$ est la fonction qui à chaque nombre complexe $z = x + iy$ lui associe son module $\sqrt{x^2 + y^2}$.
- $z \rightarrow f(z) = z^{-1}$ est la fonction qui à $z = x + iy$ associe le nombre :

$$\frac{1}{x + iy} = \frac{x - iy}{x^2 + y^2} . \quad (3.53)$$

¹³À deux dimensions, on peut *contourner* un obstacle, ce qui n'est pas possible quand on se déplace sur un fil. On se convaincra peu à peu de l'importance de la notion de *chemin* suivi par continuité (sans "lever du papier la pointe du crayon").

- $z \rightarrow f(z) = \sin z$ est la fonction qui à $z = x + iy$ associe le nombre :

$$\sin z = \frac{1}{2i} [e^{i(x+iy)} - e^{-i(x+iy)}] = \frac{1}{2i} (e^{-y}e^{ix} - e^y e^{-ix}) = \sin x \cosh y + i \sinh y \cos x ; \quad (3.54)$$

quand $z \in \mathbb{R}$, on retrouve la définition élémentaire du sinus trigonométrique "ordinaire". Par ailleurs, la dernière forme à droite montre que :

$$\sin(x + iy) = \sin x \cosh y + i \sinh y \cos x \quad (3.55)$$

qui n'est autre que la généralisation aux complexes de la formule $\sin(a + b) = \sin a \cos b + \sin b \cos a$ (en remarquant que $\sin(ia) = i \sinh a$ et $\cos(ib) = \cosh b$).

Clairement, la donnée d'une fonction $f(z)$ est équivalente à la donnée de deux fonctions à valeurs réelles $u(x, y)$ et $v(x, y)$ telles que :

$$\forall z \in \mathcal{D} , \quad f(z) = u(x, y) + iv(x, y) . \quad (3.56)$$

Par exemple pour $z \rightarrow f(z) = z^2$, on a $u(x, y) = x^2 - y^2$ et $v(x, y) = 2xy$. On ne peut évidemment pas tracer le graphe d'une fonction à valeurs complexes d'une variable complexe, comme on trace celui d'une fonction réelle d'une variable réelle : il faudrait faire des dessins dans l'espace à quatre dimensions (2 pour les deux coordonnées x et y , 2 pour les deux valeurs $u(x, y)$ et $v(x, y)$). En revanche, on peut tracer des lignes dans le plan relatives à certains attributs de $f(z)$. Par exemple, les lignes iso-module sont celles où $|f(z)|$ prend une valeur constante donnée. Les lignes iso- \Re et iso- \Im sont les courbes où les parties réelle et imaginaire sont constantes ; leurs équations cartésiennes sont respectivement :

$$u(x, y) = C^{\text{ste}} , \quad v(x, y) = C^{\text{ste}} . \quad (3.57)$$

Ainsi, pour $z \rightarrow f(z) = z^2$, la ligne où $u = 1$ est l'hyperbole d'équation $x^2 - y^2 = 1$, celle où $v = 3$ est l'hyperbole d'équation $xy = \frac{3}{2}$.

On peut se faire encore une représentation mentale de f en imaginant les surfaces obtenues en portant verticalement le long de l'axe $O\zeta$ perpendiculaire au plan xOy les valeurs (algébriques) des parties réelle et imaginaire. Les surfaces correspondantes Σ_u et Σ_v ont pour équations cartésiennes :

$$\zeta = u(x, y) , \quad \zeta = v(x, y) \quad (3.58)$$

Les lignes iso- \Re et iso- \Im sont les intersections de ces surfaces avec des plans parallèles au plan xOy .

On peut enfin tracer les images dans le plan par f de courbes remarquables décrites par la variable z , judicieusement choisies pour une raison ou une autre.

Note

Dans toute la suite, on utilisera éventuellement une notation simplifiée pour les dérivées partielles d'une fonction de plusieurs variables, $\Phi(x, y, z, \dots)$:

$$\boxed{\partial_x \Phi \stackrel{\text{d\'ef}}{=} \frac{\partial \Phi}{\partial x} \equiv \Phi'_x} \quad (3.59)$$

3.3.2 Limite d'une fonction $f(z)$

Sur l'axe réel, la notion de limite d'une fonction à valeurs réelles est bien connue. On dit qu'une fonction $f(x)$ a une limite en x_0 ssi il existe un nombre f_0 tel que¹⁴ :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall |x - x_0| < \delta \implies |f(x) - f_0| < \varepsilon \quad (3.60)$$

et on écrit :

$$\boxed{\lim_{x \rightarrow x_0} f(x) = f_0} \quad (3.61)$$

Bien noter que cette définition n'exige nullement que la fonction soit définie en x_0 .

Une fonction est dite *continue* en x_0 si¹⁵ $f_0 = f(x_0)$, et alors :

$$\boxed{\lim_{x \rightarrow x_0} f(x) = f(x_0) \iff f \text{ continue en } x_0} \quad (3.62)$$

cette écriture n'a évidemment de sens que dans la mesure où la quantité $f(x_0)$ est définie. Pour une fonction discontinue – le contraire d'une fonction continue – en un point d'abscisse x_0 , la limite n'existe pas au sens définie par (3.60) ; il arrive toutefois que la fonction $f(x)$ se rapproche d'une certaine valeur quand on vient d'un côté ou de l'autre, les deux valeurs étant distinctes. En pareil cas, il est licite de définir une limite à *gauche* et une limite à *droite*. Soit par exemple, la fonction¹⁶ $Y(x)$ définie comme : $Y(x)$ définie comme :

$$Y(x) = \begin{cases} 0 & \forall x < 0 \\ 1 & \forall x > 0 \end{cases} \quad (3.63)$$

Avec cette définition, la quantité $Y(0)$ *n'existe pas*, de sorte que $Y(x)$ n'a pas de limite en $x = 0$, mais rien n'interdit de parler de limite à gauche et de limite à droite, les deux étant distinctes (elles valent respectivement 0 et 1).

Pour une fonction d'une variable complexe, la liberté de mouvement est encore plus grande puisque, dans le plan, on peut s'approcher d'un point donné z_0 d'une *infinité* de façons. En généralisant immédiatement, on définit d'abord la limite d'une fonction $f(z)$ en un complexe z_0 comme¹⁷ :

$$\forall \varepsilon > 0, \exists \delta > 0, \forall |z - z_0| < \delta \implies |f(z) - f_0| < \varepsilon \quad (3.64)$$

et on écrit :

$$\boxed{\lim_{z \rightarrow z_0} f(z) = f_0} \quad (3.65)$$

À nouveau, cette définition ne suppose pas que f existe en z_0 ; par exemple, avec la fonction $z \rightarrow f(z) \stackrel{\text{d'éf}}{=} \frac{\sin z}{z}$, on a $\lim_{z \rightarrow 0} f(z) = 1$, mais $f(z)$ n'est pas définie en $z = 0$.

¹⁴En termes plus désinvoltes : plus x est proche de x_0 , plus f est proche de f_0 .

¹⁵Autrement dit : la recette pour calculer f à partir de x est applicable au point x_0 et donne un nombre égal à f_0 .

¹⁶aussi appelée fonction de Heaviside.

¹⁷Comme précédemment, les $||$ désignent alors précisément les *modules* des nombres complexes.

En pareil cas, on pourra tout naturellement compléter la définition de $f(z)$ en posant de plus $f(0) \stackrel{\text{déf}}{=} 1$. Par ailleurs, la même définition peut être reprise avec $z_0 = \infty$, en remplaçant la distance $|z - z_0|$ par la distance D définie dans la sous-section 3.2.2.

De la même façon que dans le cas réel, on dit qu'une fonction $f(z)$ est *continue* si $f_0 = f(z_0)$, et alors :

$$\boxed{\lim_{z \rightarrow z_0} f(z) = f(z_0) \iff f \text{ continue en } z_0} \quad (3.66)$$

Il importe de bien saisir tout le contenu de cette définition, laquelle contient la nécessité de savoir calculer $f(z_0)$, c'est-à-dire que z_0 est un point où la recette de calcul définissant f est sans ambiguïté et conduit de ce fait à un nombre et un seul, noté $f(z_0)$ – une fois parvenu en z_0 , on ne sait plus comment on y est arrivé, ou encore : $f(z_0)$ ne porte aucune mémoire du chemin parcouru. En d'autres termes, la limite n'existe que dans la mesure où *tous* les chemins suivis pour arriver en $z_0 = x_0 + iy_0$ donnent un seul et même nombre. Clairement, cette définition est bien la généralisation de la notion de limite d'une fonction réelle continue, qui n'existe que si l'on obtient le même nombre selon que l'on arrive d'un côté ou de l'autre du point x_0 .

La condition d'existence de la limite s'exprime de façon équivalente avec les parties réelle $u(x, y)$ et imaginaire $v(x, y)$ de la fonction $f(z)$; la limite de $f(z)$ en z_0 existe si et seulement si les deux limites suivantes existent :

$$\lim_{x \rightarrow x_0, y \rightarrow y_0} u(x, y) = u_0, \quad \lim_{x \rightarrow x_0, y \rightarrow y_0} v(x, y) = v_0, \quad (3.67)$$

où les deux limites $x \rightarrow x_0, y \rightarrow y_0$ sont prises *indépendamment* l'une de l'autre. Dire que f est continue en z_0 , c'est dire que $u_0 = u(x_0, y_0)$ et $v_0 = v(x_0, y_0)$.

L'indépendance par rapport au chemin suivi du nombre obtenu quand on arrive en z_0 est d'une importance capitale. Il est facile de construire des cas où le chemin d'accès détermine la valeur finale obtenue par passage à la limite ; soit par exemple la fonction :

$$z \longrightarrow f(z) = \frac{z}{z^*} \quad (3.68)$$

et posons-nous la question de l'existence d'une limite en $z_0 = 0$. On voit géométriquement tout de suite que si l'on se dirige vers l'origine en restant sur l'axe réel, le nombre obtenu à la limite est 1, alors que si on reste confiné sur l'axe imaginaire, le nombre final est -1 . Plus généralement, avec $z = r e^{i\theta}$, on a $f(z) = e^{2i\theta}$, de sorte que si l'on arrive en O en suivant une ligne faisant l'angle θ_c avec l'axe réel, le résultat final en $z_0 = 0$ est $e^{2i\theta_c}$. Ainsi, la fonction définie en (3.68) n'a pas de limite en $z = 0$, bien que pour chaque chemin différentiable à l'origine on sache calculer par un processus de limite une valeur finale pour la fonction (qui est d'ailleurs bornée : $|f(z)| = 1 \forall z \neq 0$).

Ainsi, une fonction est dite *continue en* z_0 si elle est définie en tout voisinage de z_0 et si sa limite existe. La fonction (3.68) est bien définie dans tout voisinage de l'origine

(on sait la calculer), mais n'est pas continue en $z_0 = 0$ puisque sa limite n'existe pas (si on veut, on peut dire qu'il y a une infinité de valeurs-limites).

L'établissement de la continuité d'une fonction en un point z_0 exige, selon (3.65), de trouver un certain $\delta > 0$, une fois ε choisi. En général, δ dépend à la fois de ε et de z_0 ; c'est pourquoi on définit la notion de fonction *uniformément continue* dans un ensemble S pour qualifier une fonction dont la continuité est assurée avec δ *indépendant*¹⁸ de $z_0 \in S$. On peut démontrer les résultats suivants, qui généralisent ce que l'on sait pour les fonctions $x \rightarrow f(x)$ continues sur un intervalle fermé $[a, b]$; toute fonction $z \rightarrow f(z)$ continue dans un ensemble *borné fermé* S – par exemple un disque $|z - z_0| \leq R$:

1. est bornée, c'est-à-dire que :

$$\exists M > 0, |f(z)| < M \quad \forall z \in S, \quad (3.69)$$

2. atteint sa plus grande et sa plus petite valeur en module :

$$\exists z_{\min} \in S, \exists z_{\max} \in S, \forall z \in S, |f(z)| \geq |f(z_{\min})|, |f(z)| \leq |f(z_{\max})|, \quad (3.70)$$

3. est uniformément continue dans S .

Ces propriétés se démontrent de façon classique en raisonnant séparément avec $u(x, y)$ et $v(x, y)$.

3.3.3 Dérivée d'une fonction $f(z)$. Conditions de Cauchy – Riemann

On définit la dérivée f' d'une fonction f continue au point z_0 en recopiant la définition élémentaire dans le champ réel ; la dérivée de $z \rightarrow f(z)$ en z_0 est ainsi définie comme la limite (si elle existe) :

$$f'(z_0) \stackrel{\text{d'éf}}{=} \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0} \quad (3.71)$$

Compte tenu de la définition de la limite, le résultat est indépendant du chemin suivi pour aller de z en z_0 . Une fonction ayant cette propriété en z_0 est dite *dérivable en z_0* , ou \mathbb{C} -différentiable ; de ce fait, elle est de toute évidence continue en ce point. On notera $f'(z)$, ou encore $\frac{d}{dz}f$ la dérivée de f en z . Compte tenu de l'identité entre \mathbb{R}^2 et \mathbb{C} , on pourrait penser qu'il y a équivalence entre \mathbb{R}^2 et \mathbb{C} -différentiabilités. Il n'en est rien : la \mathbb{C} -différentiabilité est autrement plus forte.

¹⁸L'adverbe *uniformément* est utilisé au même titre que lorsque l'on définit la convergence *uniforme* d'une suite de fonctions de z_0 , pour la démarquer de la convergence *simple* (ou *en tout point*). La définition ci-dessus est équivalente à celle donnée dans la note 98 p. 65.

Une fonction dérivable en tout point d'un domaine \mathcal{D} est dite¹⁹ *holomorphe* dans ce domaine. La somme et le produit de deux fonctions holomorphes est une fonction holomorphe ; il en va de même pour le rapport $\frac{f(z)}{g(z)}$ partout où $g(z) \neq 0$. On appelle fonction *entière* toute fonction qui est holomorphe pour tout z de module fini, c'est-à-dire²⁰ $\forall z \in \mathbb{C}$. L'holomorphie entraînant la continuité, on voit qu'une fonction *entière* est bornée (voir p. 131) dans tout domaine fermé inclus dans \mathbb{C} .

Noter que souvent le domaine d'holomorphie et le domaine de définition coïncident ; c'est le cas par exemple pour $z \rightarrow \frac{1}{z}$, mais ce n'est pas le cas pour $z \rightarrow \frac{1}{z^*}$: cette fonction est définie partout sauf en $z = 0$, mais elle n'est holomorphe nulle part !

Les conditions de dérivabilité – donc notamment l'exigence d'indépendance vis-à-vis du chemin suivi pour arriver en z_0 – s'expriment par le théorème suivant, appelé conditions de Cauchy - Riemann (ou plus simplement *conditions de Cauchy*). Ces conditions seront énoncées successivement en tant que conditions nécessaires et conditions suffisantes. Le cas échéant, elles acquièrent le statut d'*équations* (aux dérivées partielles) permettant par exemple, une fonction $u(x, y)$ étant donnée, de trouver une autre fonction $v(x, y)$ telle que la combinaison $u + iv$ construite une fonction holomorphe (voir un exemple p. 141).

Conditions nécessaires de Cauchy - Riemann : $\exists f'(z_0) \implies \mathbf{C - R}$

Soit $f(z) = u(x, y) + iv(x, y)$ une fonction dérivable en $z_0 = x_0 + iy_0$. Alors, les dérivées partielles u_x, u_y, v_x et v_y existent au point (x_0, y_0) et satisfont :

$$\left(\frac{\partial u}{\partial x}\right)_{x_0, y_0} = \left(\frac{\partial v}{\partial y}\right)_{x_0, y_0} \quad \text{et} \quad \left(\frac{\partial u}{\partial y}\right)_{x_0, y_0} = -\left(\frac{\partial v}{\partial x}\right)_{x_0, y_0} . \quad (3.72)$$

Supposons que la limite $f'(z_0)$ existe et prenons d'abord le cas où z tend vers z_0 suivant un chemin parallèle à l'axe réel ; dans ces conditions, $z - z_0 = h$ où $h \in \mathbb{R}$. On peut alors écrire :

$$f'(z_0) = \lim_{h \rightarrow 0} \left[\frac{u(x_0 + h, y_0) - u(x_0, y_0)}{h} + i \frac{v(x_0 + h, y_0) - v(x_0, y_0)}{h} \right] . \quad (3.73)$$

¹⁹holomorphe signifie *même forme*. L'origine de la terminologie tient à ceci : on verra qu'une fonction holomorphe, qui est une application de \mathbb{C} dans \mathbb{C} , préserve les angles et définit de ce fait ce que l'on appelle une *transformation conforme*. Ces transformations se rencontrent souvent en Physique (effets de bord pour un condensateur plan fini, mécanique des fluides, effets de taille finie pour les systèmes critiques bidimensionnels, ...).

On signalera à plusieurs reprises que la terminologie varie dans la littérature : certains auteurs qualifient d'*analytique* une fonction ici appelée holomorphe (d'autres parlent de *fonction monogène*, etc.) ; on verra plus tard pourquoi aucune confusion n'est possible entre ces deux qualificatifs. La terminologie employée ici est également celle de Appel [2].

²⁰Une fois définies les *singularités* d'une fonction $f(z)$ (chapitre 5), on réalisera qu'une fonction entière peut avoir une singularité (même essentielle) à l'infini.

La limite au premier membre existe, il en va donc de même pour le second membre, ce qui établit d'une part l'existence des deux dérivées partielles u_x et v_x , d'autre part que $f'(z_0)$ est :

$$f'(z_0) = \left(\frac{\partial u}{\partial x}\right)_{x_0, y_0} + i \left(\frac{\partial v}{\partial x}\right)_{x_0, y_0} . \quad (3.74)$$

Prenons maintenant le cas où z tend vers z_0 suivant un chemin parallèle à l'axe imaginaire, soit $z - z_0 = ik$ avec $k \in \mathbb{R}$. Alors :

$$f'(z_0) = \lim_{k \rightarrow 0} \left[\frac{u(x_0, y_0 + k) - u(x_0, y_0)}{ik} + i \frac{v(x_0, y_0 + k) - v(x_0, y_0)}{ik} \right] , \quad (3.75)$$

d'où l'existence des deux dérivées partielles par rapport à y , et une autre écriture de la dérivée $f'(z_0)$:

$$f'(z_0) = -i \left(\frac{\partial u}{\partial y}\right)_{x_0, y_0} + \left(\frac{\partial v}{\partial y}\right)_{x_0, y_0} . \quad (3.76)$$

En rapprochant (3.74) et (3.76), on obtient la condition nécessaire :

$$\boxed{\exists f'(z_0) \implies (3.72)} \quad (3.77)$$

On voit ainsi que la *dérivabilité* est une propriété autrement plus forte que la simple continuité : non seulement les dérivées partielles du premier ordre de $u(x, y)$ et $v(x, y)$ doivent exister, mais encore faut-il qu'elles soient reliées par les conditions différentielles de Cauchy - Riemann.

Conditions suffisantes de Cauchy - Riemann : C - R + ... $\implies \exists f'(z_0)$

Dans l'autre sens, on va démontrer la proposition suivante :

Soit $f(z) = u(x, y) + iv(x, y)$ une fonction définie dans un ε -voisinage de $z_0 = x_0 + iy_0$, les dérivées partielles u_x, u_y, v_x et v_y étant continues dans ce voisinage. Si ces dérivées satisfont :

$$\left(\frac{\partial u}{\partial x}\right)_{x_0, y_0} = \left(\frac{\partial v}{\partial y}\right)_{x_0, y_0} \quad \text{et} \quad \left(\frac{\partial u}{\partial y}\right)_{x_0, y_0} = -\left(\frac{\partial v}{\partial x}\right)_{x_0, y_0} . \quad (3.78)$$

alors $f'(z_0)$ existe et :

$$f'(z_0) = u_x(x_0, y_0) + iv_x(x_0, y_0) = v_y(x_0, y_0) - iu_y(x_0, y_0) . \quad (3.79)$$

La continuité supposée des dérivées partielles des deux fonctions u et v permet d'écrire²¹ :

$$u(x_0 + h, y_0 + k) - u(x_0, y_0) = h \left(\frac{\partial u}{\partial x}\right)_{x_0, y_0} + k \left(\frac{\partial u}{\partial y}\right)_{x_0, y_0} + \alpha |\xi| , \quad (3.80)$$

²¹ Rappelons au passage que si pour une fonction $x \rightarrow f(x)$, l'existence de la dérivée en un point assure la continuité de f en ce point, cette proposition suffisante n'est pas vraie pour une fonction quelconque de plusieurs variables (quelconque signifie sans aucune référence aux conditions de Cauchy - Riemann) ;

$$v(x_0 + h, y_0 + k) - v(x_0, y_0) = h \left(\frac{\partial v}{\partial x} \right)_{x_0, y_0} + k \left(\frac{\partial v}{\partial y} \right)_{x_0, y_0} + \beta |\xi|, \quad (3.81)$$

où $z - z_0 \equiv \xi = h + ik$ et où α et β tendent vers zéro quand $|\xi| \rightarrow 0$. Il en résulte :

$$f(z_0 + \xi) - f(z_0) = \left[h \left(\frac{\partial u}{\partial x} \right)_{x_0, y_0} + k \left(\frac{\partial u}{\partial y} \right)_{x_0, y_0} \right] + i \left[h \left(\frac{\partial v}{\partial x} \right)_{x_0, y_0} + k \left(\frac{\partial v}{\partial y} \right)_{x_0, y_0} \right] + (\alpha + i\beta) |\xi|. \quad (3.82)$$

Les conditions de Cauchy étant satisfaites, on peut remplacer $\partial_y v$ par $\partial_x u$ et $\partial_y u$ par $-\partial_x v$, d'où :

$$f(z_0 + \xi) - f(z_0) = \left[\left(\frac{\partial u}{\partial x} \right)_{x_0, y_0} + i \left(\frac{\partial v}{\partial x} \right)_{x_0, y_0} \right] (h + ik) + \gamma |\xi| \equiv A(z - z_0) + \gamma |z - z_0|. \quad (3.83)$$

A est un nombre bien défini, $\frac{|z - z_0|}{z - z_0}$ est de module 1 (donc borné), cependant que $\gamma \rightarrow 0$ si $\xi \rightarrow 0$. On en conclut que, dans cette limite, le rapport $\frac{f(z_0 + \xi) - f(z_0)}{z - z_0}$ a une valeur bien définie, égale à A , d'où :

$$\boxed{(3.78) \implies \exists f'(z_0)} \quad (3.84)$$

Dans toute la suite, pour faire court, on désignera simplement par "Conditions de Cauchy - Riemann" les relations :

$$\boxed{\left(\frac{\partial u}{\partial x} \right)_{x_0, y_0} = \left(\frac{\partial v}{\partial y} \right)_{x_0, y_0} \text{ et } \left(\frac{\partial u}{\partial y} \right)_{x_0, y_0} = - \left(\frac{\partial v}{\partial x} \right)_{x_0, y_0}} \quad (3.85)$$

tout en gardant à l'esprit leur caractère nécessaire/suffisant. Par ailleurs, la terminologie est quelque peu fluctuante d'un auteur à l'autre ; notamment, les seules conditions (3.85) sont parfois prises comme *définition* du caractère holomorphe.

Enfin, il existe une autre présentation ne se référant pas à la continuité des dérivées partielles d'ordre 1, mais à la \mathbb{R}^2 -différentiabilité de $f(z)$, c'est-à-dire exigeant que dans un voisinage de z_0 on puisse écrire $f(z_0 + \xi)$ comme la somme $f(z_0) + L(\xi) + \xi \varepsilon(\xi)$, où

voici schématiquement pourquoi.

Soit une fonction $\Phi(x, y)$ ayant des dérivées (premières) dans un voisinage du point (x, y) . L'existence de la dérivée Φ'_x permet d'écrire $\Phi(x + h, y) = \Phi(x, y) + h\Phi'_x(x, y) + \varepsilon_x(h)$ où la fonction $\varepsilon_x(h)$ tend vers zéro quand $h \rightarrow 0$. De ceci on déduit $\Phi(x + h, y + k) = \Phi(x, y + k) + h\Phi'_x(x, y + k) + \varepsilon_x(h)$ (d'ailleurs, on a aussi $\Phi(x + h, y + k) = \Phi(x, y) + k\Phi'_y(x, y) + \varepsilon_y(k) + h\Phi'_x(x, y + k) + \varepsilon_x(h)$ où $\varepsilon_y(k) \rightarrow 0$ si $k \rightarrow 0$, compte tenu de l'existence de la dérivée Φ'_y). Pour que $\Phi(x, y)$ soit continue, il faut en plus que $\lim_{h \rightarrow 0, k \rightarrow 0} h\Phi'_x(x, y + k) = 0$. C'est le cas si la dérivée *seconde* $\partial_y \partial_x \Phi$ existe, assurant la continuité de $\partial_x \Phi$: ainsi, la continuité de Φ est liée à des conditions (suffisantes) sur les dérivées *secondes* de Φ , et non pas sur les seules dérivées premières.

Si en plus les dérivées Φ''_{xy} et Φ''_{yx} sont continues, alors elles sont égales (théorème de Clairaut). L'hypothèse de continuité des dérivées partielles assure que, dans les développements (3.80) et (3.81), les restes sont au moins linéaires en $|\xi|$.

Au contraire, pour une fonction holomorphe $f(z)$, dont la dérivée existe par définition, cette existence assure la continuité de $f(z)$. Comme déjà souligné p.133, la dérivabilité d'une fonction $f(z)$ est une propriété autrement plus forte que la simple continuité.

$L(\xi)$ est une forme différentielle linéaire (c'est la différentielle de la combinaison $u + iv$) et où $\varepsilon(\xi)$ tend vers zéro avec ξ . Cette condition étant réalisée, il y a alors *équivalence* entre les conditions de Cauchy - Riemann et la \mathbb{C} -différentiabilité.

Donnons pour terminer d'autres "démonstrations" plus expéditives, mais instructives. Notons d'abord que les arguments précédents se généralisent en prenant comme écart $z - z_0 = \varepsilon e^{i\theta}$; alors :

$$f'(z_0) = \lim_{\varepsilon \rightarrow 0} \frac{u(x_0 + \varepsilon \cos \theta, y_0 + \varepsilon \sin \theta) - u(x_0, y_0)}{\varepsilon e^{i\theta}} + \quad (3.86)$$

$$i \lim_{\varepsilon \rightarrow 0} \frac{v(x_0 + \varepsilon \cos \theta, y_0 + \varepsilon \sin \theta) - v(x_0, y_0)}{\varepsilon e^{i\theta}} . \quad (3.87)$$

En effectuant les développements de Taylor des numérateurs, on trouve :

$$2f'(z_0) = \left(\frac{\partial u}{\partial x}\right)_{z_0} - i \left(\frac{\partial u}{\partial y}\right)_{z_0} + i \left(\frac{\partial v}{\partial x}\right)_{z_0} + \left(\frac{\partial v}{\partial y}\right)_{z_0} + e^{-2i\theta} \left[\left(\frac{\partial u}{\partial x}\right)_{z_0} + i \left(\frac{\partial u}{\partial y}\right)_{z_0} + i \left(\frac{\partial v}{\partial x}\right)_{z_0} - \left(\frac{\partial v}{\partial y}\right)_{z_0} \right] . \quad (3.88)$$

On veut que le second membre ne dépende pas de l'angle θ ; à ce stade, une condition nécessaire et suffisante pour cela est bien la satisfaction simultanée des deux relations $\left(\frac{\partial u}{\partial x}\right)_{z_0} = \left(\frac{\partial v}{\partial y}\right)_{z_0}$ et $\left(\frac{\partial u}{\partial y}\right)_{z_0} = -\left(\frac{\partial v}{\partial x}\right)_{z_0}$.

Une dernière preuve rapide procède comme suit. Lorsque z varie de $\delta z = \delta x + i\delta y$, le taux de variation de f est $\frac{\delta f}{\delta z}$, soit :

$$\frac{\delta f}{\delta z} = \frac{\delta u + i\delta v}{\delta x + i\delta y} = \frac{\frac{\partial u}{\partial x}\delta x + \frac{\partial u}{\partial y}\delta y + i\left(\frac{\partial v}{\partial x}\delta x + \frac{\partial v}{\partial y}\delta y\right)}{\delta x + i\delta y} = \frac{\left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right)\delta x + \left(\frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}\right)\delta y}{\delta x + i\delta y} . \quad (3.89)$$

Après division par δx , dans la limite $\delta z \rightarrow 0$ il vient :

$$\frac{df}{dz} = \frac{\left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right) + \left(\frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}\right)\frac{dy}{dx}}{1 + i\frac{dy}{dx}} . \quad (3.90)$$

On veut que la fraction au second membre soit indépendante du rapport $\frac{dy}{dx}$. Pour que la fonction $\Phi(X) = \frac{a+bX}{c+dX}$ soit constante (indépendante de X), il faut et suffit que $\frac{a}{c} = \frac{b}{d}$, d'où les conditions assurant que la limite $\frac{df}{dz}$ existe :

$$\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = \frac{1}{i} \left(\frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}\right) . \quad (3.91)$$

En séparant les parties réelles et imaginaires, on retrouve bien les conditions de Cauchy.

◆ *Remarques*

1. Les conditions de Cauchy ont une interprétation géométrique simple : les lignes $\Re f(z) = C^{\text{ste}}$ et les lignes $\Im f(z) = C^{\text{ste}}$ sont orthogonales les unes aux autres.

Avant de montrer ceci, rappelons un résultat de géométrie différentielle. Soit $\phi(x, y)$ une fonction \mathbb{R}^2 -différentiable. La relation $\phi(x, y) = C^{\text{ste}}$ définit une courbe Γ dans le plan : quand x varie, y ne varie pas n'importe comment mais précisément de telle sorte que la fonction ϕ reste constante. La différentielle de ϕ , pour des variations infinitésimales *quelconques* de x et y , est :

$$d\phi = \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy . \quad (3.92)$$

Maintenant, si on veut que quand x varie de dx , la valeur de la fonction ϕ ne change pas, il faut que dy ne soit pas n'importe quoi par rapport à dx mais tel que la variation $d\phi$ soit nulle. Autrement dit, pour rester sur la courbe Γ alors que x et y ont un peu varié, il faut que les accroissements dx et dy satisfassent :

$$0 = \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy , \quad (3.93)$$

et c'est ce qui assure que l'on reste sur la courbe Γ . Cette condition fixe le rapport $\frac{dy}{dx}$, qui donne la pente de la tangente à Γ :

$$\text{pente de la tangente de } \Gamma = \frac{dy}{dx} = -\frac{\frac{\partial\phi}{\partial x}}{\frac{\partial\phi}{\partial y}} \iff \frac{dx}{\frac{\partial\phi}{\partial y}} = \frac{dy}{-\frac{\partial\phi}{\partial x}} . \quad (3.94)$$

Autrement dit, $\frac{\partial\phi}{\partial y}$ et $-\frac{\partial\phi}{\partial x}$ sont les coefficients directeurs de la tangente à la courbe Γ d'équation cartésienne $\phi(x, y) = C^{\text{ste}}$; cette tangente est donc la droite parallèle au vecteur \vec{t} de composantes :

$$\vec{t} = \left(\frac{\partial\phi}{\partial y}, -\frac{\partial\phi}{\partial x} \right) . \quad (3.95)$$

Par ailleurs, le gradient de ϕ est le vecteur du plan de composantes :

$$\vec{\nabla}\phi = \left(\frac{\partial\phi}{\partial x}, \frac{\partial\phi}{\partial y} \right) . \quad (3.96)$$

La comparaison de (3.95) et (3.96) montre que le produit scalaire des deux vecteurs \vec{t} et $\vec{\nabla}\phi$ est nul : ils sont bien orthogonaux. En d'autres termes, le vecteur $\vec{\nabla}\phi$ définit les lignes orthogonales à la courbe Γ d'équation $\phi(x, y) = C^{\text{ste}}$, ce qui est bien naturel puisque la fonction ϕ garde une valeur constante le long de Γ .

Cela étant rappelé, on voit de suite que les lignes $\Re f(z) = C^{\text{ste}}$ et $\Im f(z) = C^{\text{ste}}$ sont mutuellement orthogonales. Le vecteur tangent \vec{t}_u aux lignes $u(x, y) = C^{\text{ste}}$ a pour composantes :

$$\vec{t}_u = \left(\frac{\partial u}{\partial y}, -\frac{\partial u}{\partial x} \right) . \quad (3.97)$$

La normale aux lignes $v(x, y) = C^{\text{ste}}$ est le gradient de v , de composantes :

$$\vec{\nabla}v = \left(\frac{\partial v}{\partial x}, \frac{\partial v}{\partial y} \right) , \quad (3.98)$$

mais d'après les conditions de Cauchy, on a aussi :

$$\vec{\nabla} v = \left(-\frac{\partial u}{\partial y}, \frac{\partial u}{\partial x} \right). \quad (3.99)$$

La comparaison de (3.97) et de (3.99) montre que $\vec{t}_u = -\vec{\nabla} v$: le vecteur \vec{t}_u est parallèle au vecteur normal des lignes $\Im f(z) \equiv v(x, y) = C^{\text{ste}}$, les lignes iso- $\Re f$ sont bien orthogonales aux lignes iso- $\Im f$.

Par exemple soit $z \rightarrow f(z) = z^2$. On a $u(x, y) = x^2 - y^2$ et $v(x, y) = 2xy$. Les iso- \Re ont pour équation cartésienne $x^2 - y^2 = C$: ce sont des hyperboles équilatères, dont les asymptotes sont les deux bissectrices du repère xOy ; les iso- \Im sont définies par $2xy = C'$ et sont à nouveau de hyperboles équilatères, mais leurs asymptotes sont les deux axes de coordonnées. Ces deux familles de courbes sont bien mutuellement orthogonales.

2. Quand on utilise la représentation polaire de $z = re^{i\theta}$, $f(z)$ a pour parties réelle et imaginaire deux fonctions $U(r, \theta)$ et $V(r, \theta)$ telles que :

$$U(r, \theta) \stackrel{\text{déf}}{=} u(x, y), \quad V(r, \theta) \stackrel{\text{déf}}{=} v(x, y) \quad (3.100)$$

avec les relations de passage $r = \sqrt{x^2 + y^2}$, $\theta = \text{Arctg} \frac{y}{x} \iff x = r \cos \theta$, $y = r \sin \theta$, d'où résulte :

$$\frac{\partial r}{\partial x} = \cos \theta, \quad \frac{\partial r}{\partial y} = \sin \theta, \quad \frac{\partial \theta}{\partial x} = -\frac{\sin \theta}{r}, \quad \frac{\partial \theta}{\partial y} = \frac{\cos \theta}{r}. \quad (3.101)$$

On a alors :

$$\frac{\partial u}{\partial x} = \frac{\partial U}{\partial r} \cos \theta - \frac{\partial U}{\partial \theta} \frac{\sin \theta}{r}, \quad \frac{\partial v}{\partial y} = \frac{\partial V}{\partial r} \sin \theta + \frac{\partial V}{\partial \theta} \frac{\cos \theta}{r}, \quad (3.102)$$

$$\frac{\partial u}{\partial y} = \frac{\partial U}{\partial r} \sin \theta + \frac{\partial U}{\partial \theta} \frac{\cos \theta}{r}, \quad \frac{\partial v}{\partial x} = \frac{\partial V}{\partial r} \cos \theta - \frac{\partial V}{\partial \theta} \frac{\sin \theta}{r}. \quad (3.103)$$

En reportant dans les deux conditions de Cauchy, on obtient le système de deux équations :

$$\frac{\partial U}{\partial r} \cos \theta - \frac{\partial U}{\partial \theta} \frac{\sin \theta}{r} = \frac{\partial V}{\partial r} \sin \theta + \frac{\partial V}{\partial \theta} \frac{\cos \theta}{r}, \quad (3.104)$$

$$\frac{\partial U}{\partial r} \sin \theta + \frac{\partial U}{\partial \theta} \frac{\cos \theta}{r} = -\frac{\partial V}{\partial r} \cos \theta + \frac{\partial V}{\partial \theta} \frac{\sin \theta}{r}. \quad (3.105)$$

En formant les bonnes combinaisons linéaires, on trouve que les relations de Cauchy prennent la forme :

$$\boxed{\left(\frac{\partial U}{\partial r} \right)_{r_0, \theta_0} = \left(\frac{1}{r} \frac{\partial V}{\partial \theta} \right)_{r_0, \theta_0} \quad \left(\frac{1}{r} \frac{\partial U}{\partial \theta} \right)_{r_0, \theta_0} = -\left(\frac{\partial V}{\partial r} \right)_{r_0, \theta_0}} \quad (3.106)$$

3. Compte tenu des conditions de Cauchy (3.85), la dérivée f' en un point quelconque peut s'écrire de quatre façons différentes²² (partir par exemple de (3.74)) :

$$f'(z) = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y} = \frac{\partial u}{\partial x} - i \frac{\partial u}{\partial y} = \frac{\partial v}{\partial y} + i \frac{\partial v}{\partial x} . \quad (3.107)$$

4. À ce stade, on peut déjà soupçonner qu'une fonction holomorphe possède des dérivées de tous les ordres (ce point sera repris et démontré au chapitre 4). En effet, avec toutefois une hypothèse supplémentaire adoptée en cours de route, la dérivée f' satisfait elle-même les conditions de Cauchy. D'après (3.107), on a par exemple :

$$f'(z) = u^{(1)}(x, y) + iv^{(1)}(x, y) \text{ avec } u^{(1)}(x, y) = \frac{\partial u}{\partial x}, v^{(1)}(x, y) = \frac{\partial v}{\partial x} . \quad (3.108)$$

Si elle est satisfaite, la première condition de Cauchy pour f' s'exprime comme :

$$\frac{\partial u^{(1)}}{\partial x} \stackrel{?}{=} \frac{\partial v^{(1)}}{\partial y} \iff \frac{\partial}{\partial x} \frac{\partial u}{\partial x} \stackrel{?}{=} \frac{\partial}{\partial y} \frac{\partial v}{\partial x} . \quad (3.109)$$

À ce stade, admettons que toutes les dérivées partielles secondes existent et sont continues (fonction de classe C^2) : l'ordre des dérivations étant alors sans importance, et le modifiant au second membre ci-dessus à droite, l'égalité putative devient :

$$\frac{\partial}{\partial x} \frac{\partial u}{\partial x} \stackrel{?}{=} \frac{\partial}{\partial x} \frac{\partial v}{\partial y} \quad (3.110)$$

Or cette dernière égalité est bien vraie, puisque f est holomorphe ; on démontre de même que la deuxième condition de Cauchy est aussi satisfaite, ce qui établit que si f est holomorphe, il en va de même de sa dérivée. Le même type d'argumentation avec f' montre que f'' existe et est holomorphe, et ainsi de suite. Ce résultat assez extraordinaire – dès qu'elle est holomorphe, une fonction est *infiniment* \mathbb{C} -différentiable – sera établi au chapitre 4, section 4.5.

5. Si dans un domaine \mathcal{D} une fonction $f(z)$ a une partie réelle constante, ou une partie imaginaire constante, alors cette fonction est constante. En effet, supposons que $u(x, y)$ est une fonction constante, auquel cas $\frac{\partial u}{\partial x} = \frac{\partial u}{\partial y} = 0$, d'où par les conditions de Cauchy $\frac{\partial v}{\partial y} = \frac{\partial v}{\partial x} = 0$. Il en va de même si le module de $f(z)$ est constant et ou son argument l'est²³.

6. Définissons la fonction \tilde{f} de $\mathbb{R}^2 \rightarrow \mathbb{C}$ comme suit :

$$\tilde{f}(x, y) = f(z = x + iy) = u(x, y) + iv(x, y) . \quad (3.111)$$

Il est alors facile de montrer que les conditions de Cauchy s'écrivent comme suit :

$$\left(\frac{\partial \tilde{f}}{\partial x} \right)_{x_0, y_0} = -i \left(\frac{\partial \tilde{f}}{\partial y} \right)_{x_0, y_0} . \quad (3.112)$$

²²Voir aussi (3.88).

²³En anticipant sur la définition de la fonction logarithme (donnée ci-dessous, sous-section 3.4.3), et notant $M > 0$ la valeur constante de $|f(z)|$, on a $\ln f(z) = \ln M + i \arg f(z)$, fonction qui est constante puisque sa partie réelle l'est ; $\ln f(z)$ étant une constante, $f(z)$ l'est aussi (dans le cas où $M = 0$ dans un domaine \mathcal{D} , c'est que f est la fonction nulle dans ce domaine : elle est bien constante !)

7. Les conditions de Cauchy assurent qu'une fonction *holomorphe* $f(z)$ est bien une fonction de z seul, pas de z^* . Avec $z = x + iy$, $z^* = x - iy$, on a $x = \frac{1}{2}(z + z^*)$, $y = \frac{1}{2i}(z - z^*)$. Étant donné une fonction $f(z) = u(x, y) + iv(x, y)$, définissons les deux fonctions $\tilde{u}(z, z^*)$ et $\tilde{v}(z, z^*)$ par les relations :

$$\tilde{u}(z, z^*) = u(x, y) , \quad \tilde{v}(z, z^*) = v(x, y) , \quad (3.113)$$

de sorte que la fonction $\tilde{f} = \tilde{u} + i\tilde{v}$, qui est *a priori* une fonction de z et de z^* , coïncide partout avec la fonction $f(z)$:

$$\tilde{f}(z, z^*) = f(z) \quad \forall z \in \mathbb{C} . \quad (3.114)$$

Maintenant :

$$\frac{\partial u}{\partial x} = \frac{\partial \tilde{u}}{\partial z} \times 1 + \frac{\partial \tilde{u}}{\partial z^*} \times 1 , \quad \frac{\partial v}{\partial y} = \frac{\partial \tilde{v}}{\partial z} \times i + \frac{\partial \tilde{v}}{\partial z^*} \times (-i) . \quad (3.115)$$

La première condition de Cauchy est donc :

$$\frac{\partial \tilde{u}}{\partial z} + \frac{\partial \tilde{u}}{\partial z^*} = i \left(\frac{\partial \tilde{v}}{\partial z} - \frac{\partial \tilde{v}}{\partial z^*} \right) ; \quad (3.116)$$

un calcul analogue donne la deuxième condition de Cauchy sous la forme :

$$\frac{\partial \tilde{u}}{\partial z} - \frac{\partial \tilde{u}}{\partial z^*} = i \left(\frac{\partial \tilde{v}}{\partial z} + \frac{\partial \tilde{v}}{\partial z^*} \right) . \quad (3.117)$$

On en déduit :

$$\frac{\partial \tilde{u}}{\partial z} = i \frac{\partial \tilde{v}}{\partial z^*} , \quad \frac{\partial \tilde{u}}{\partial z^*} = -i \frac{\partial \tilde{v}}{\partial z} \iff \frac{\partial \tilde{v}}{\partial z^*} = i \frac{\partial \tilde{u}}{\partial z} , \quad (3.118)$$

d'où :

$$\frac{\partial}{\partial z^*} (\tilde{u} + i\tilde{v}) = \frac{\partial \tilde{u}}{\partial z^*} + i \cdot i \frac{\partial \tilde{u}}{\partial z^*} = 0 , \quad (3.119)$$

soit²⁴ $\frac{\partial \tilde{f}}{\partial z^*} = 0$, ce qui montre que \tilde{f} ne dépend pas de z^* ; par (3.114), il en va de même de $f(z)$.

8. Par opposition, soit une fonction $f(x, y)$ possédant une différentielle dans un certain domaine de \mathbb{R}^2 :

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy . \quad (3.121)$$

Par ailleurs $z = x + iy$ et $z^* = x - iy$ sont aussi \mathbb{R}^2 -différentiables, avec :

$$dz = dx + idy , \quad dz^* = dx - idy , \quad (3.122)$$

²⁴En sous-produit de ce calcul, on obtient aussi :

$$\frac{\partial}{\partial z} (\tilde{u} + i\tilde{v}) = \frac{\partial \tilde{u}}{\partial z} + i \cdot (-i) \frac{\partial \tilde{u}}{\partial z} = 2 \frac{\partial \tilde{u}}{\partial z} = 2i \frac{\partial \tilde{v}}{\partial z} . \quad (3.120)$$

d'où l'on déduit :

$$dx = \frac{1}{2}(dz + dz^*) , \quad dy = \frac{1}{2i}(dz - dz^*) . \quad (3.123)$$

En reportant ces expressions dans (3.121) et en factorisant selon dz et dz^* , il vient :

$$df = \frac{1}{2} \left(\frac{\partial f}{\partial x} - i \frac{\partial f}{\partial y} \right) dz + \frac{1}{2} \left(\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right) dz^* . \quad (3.124)$$

Cette expression justifie que l'on introduise deux opérateurs différentiels définis comme :

$$\partial \stackrel{\text{d'ef}}{=} \frac{\partial}{\partial z} = \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) , \quad \partial^* \stackrel{\text{d'ef}}{=} \frac{\partial}{\partial z^*} = \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \quad (3.125)$$

et alors la différentielle (3.124) prend la forme :

$$df = \frac{\partial f}{\partial z} dz + \frac{\partial f}{\partial z^*} dz^* \equiv \partial f dz + \partial^* f dz^* . \quad (3.126)$$

Sans hypothèse supplémentaire, f est une fonction de z et de z^* . Supposons maintenant que f est dérivable en un certain point z_0 ; alors $df = f'(z_0) dz$, de sorte que :

$$\left(\frac{\partial f}{\partial z} \right)_{z_0} = f'(z_0) , \quad \left(\frac{\partial f}{\partial z^*} \right)_{z_0} = 0 . \quad (3.127)$$

Il est facile de voir que la deuxième égalité dans (3.127) n'est qu'un avatar des conditions de Cauchy ; en effet, appelant u et v les parties réelle et imaginaire de f , il vient :

$$\left(\frac{\partial f}{\partial z^*} \right)_{z_0} = 0 \iff \left(\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right) (u + iv) = 0 . \quad (3.128)$$

En développant, on trouve :

$$\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right)_{z_0} + i \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right)_{z_0} = 0 \quad (3.129)$$

qui reproduit bien les conditions de Cauchy. En définitive, il y a équivalence entre les deux propriétés ci-après :

- $\forall z, z + \xi \in \mathcal{D}$, le rapport $\xi^{-1}[f(z + \xi) - f(z)]$ a une limite notée $f'(z)$ quand $\xi \rightarrow 0$ et l'application $z \rightarrow f(z)$ est dérivable (et continue)
- la fonction f satisfait $\partial^* f = 0$, l'opérateur différentiel ∂^* étant défini en (3.125).

Par exemple, la fonction $f(z) = \Re z = \frac{1}{2}(z + z^*)$ n'est pas holomorphe puisque $\partial^* f = \frac{1}{2}$. On retiendra qu'une fonction holomorphe, nécessairement, ne s'exprime qu'avec z , pas avec z^* ; c'est pourquoi les fonctions $|z| = (zz^*)^{\frac{1}{2}}$, $\Re z = \frac{1}{2}(z + z^*)$, $\Im z = \frac{1}{2i}(z - z^*)$, etc., ne sauraient être holomorphes.

Une fonction satisfaisant (3.127) est par définition une fonction holomorphe : à l'inverse, une fonction g satisfaisant :

$$\frac{\partial g}{\partial z} = 0 \quad (3.130)$$

est dite *antiholomorphe*.

Attention ! Quand une fonction satisfait (3.130), ne pas en déduire qu'elle est constante : la fonction $z \rightarrow f(z) = z^*$ est bien antiholomorphe, mais elle n'est pas constante ! La quantité $\frac{\partial f}{\partial z}$ ne représente la dérivée $f'(z)$ que si f est holomorphe (voir (3.127) à gauche).

9. La dérivée f' étant définie comme pour les fonctions de $\mathbb{R} \rightarrow \mathbb{R}$ (limite du rapport des accroissements), toutes les règles de dérivation connues restent valables ($(fg)' = f'g + fg'$, dérivation d'une fonction composée $f(g(z))$, ...)
10. Les parties réelle et imaginaire d'une fonction holomorphe vérifient l'équation de Laplace $\Delta \sharp = 0$. En effet :

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \implies \frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial x \partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x} \implies \frac{\partial^2 u}{\partial y^2} = -\frac{\partial^2 v}{\partial y \partial x}. \quad (3.131)$$

En vertu de l'égalité des dérivées croisées²⁵, il vient :

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \iff \Delta u = 0, \quad (3.132)$$

et de même $\Delta v = 0$. Une fonction f satisfaisant l'équation $\Delta f = 0$ est dite *harmonique*, une propriété possédée séparément par u et v quand $f = u + iv$ est holomorphe. On dit que v est la conjuguée harmonique de u , et réciproquement. Par exemple, si on choisit $u(x, y) = \sin x \cosh y$, on vérifie que $u_{xx} + u_{yy} = 0$, et que cette fonction est donc harmonique. De ce fait, elle peut être considérée comme la partie réelle d'une certaine fonction holomorphe $f = u + iv$. En écrivant les conditions de Cauchy et en les intégrant, on trouve que la partie imaginaire est $v(x, y) = \cos x \sinh y + C^{\text{ste}}$, de sorte que $f(z) = \sin x \cosh y + i \cos x \sinh y + C^{\text{ste}}$, soit $f(z) = \sin(x + iy) + C^{\text{ste}}$. \blacklozenge

La fonction monôme $z \rightarrow z^n$ ($n \in \mathbb{N}$) est holomorphe dans \mathbb{C} , et c'est aussi une fonction entière (il en va de même pour tout polynôme). La fonction $z \rightarrow \frac{1}{z}$ n'est holomorphe que dans $\mathbb{C}^* \equiv \mathbb{C} \setminus \{0\}$; autrement dit, $z \rightarrow \frac{1}{z}$ n'est pas définie en $z_0 = 0$ mais l'est en tout point aussi proche que l'on veut de l'origine : en pareil cas, on dit que $z_0 = 0$ est un *point singulier* ; $z \rightarrow \frac{1}{z}$ n'est pas une fonction entière.

En revanche, la fonction $z \rightarrow |z|$ n'est holomorphe nulle part car elle n'est dérivable nulle part. En effet, avec $z = re^{i\theta}$ et $\xi = \varepsilon e^{i\phi}$, on a :

$$\frac{f(z + \xi) - f(z)}{\xi} = \frac{|r + \varepsilon e^{i(\phi - \theta)}| - r}{\varepsilon e^{i\phi}} ; \quad (3.133)$$

²⁵On suppose les fonctions de classe C^2 . On verra par la suite qu'une fonction holomorphe est *infinitement* dérivable ; l'hypothèse de la continuité des dérivées secondes n'est pas en fait nécessaire.

au numérateur, le module se développe comme suit :

$$|r + \varepsilon e^{i(\phi-\theta)}| = \left([r + \varepsilon e^{i(\phi-\theta)}][r + \varepsilon e^{-i(\phi-\theta)}] \right)^{\frac{1}{2}} = (r^2 + 2\varepsilon r \cos(\phi - \theta) + \mathcal{O}(\varepsilon^2))^{\frac{1}{2}} . \quad (3.134)$$

En sortant $r > 0$ de la racine carrée et en utilisant $(1+x)^{\frac{1}{2}} = 1 + \frac{x}{2} + \mathcal{O}(x^2)$ ($x \ll 1$), il vient :

$$[r^2 + 2\varepsilon r \cos(\phi - \theta) + \mathcal{O}(\varepsilon^2)]^{\frac{1}{2}} = r + \varepsilon \cos(\theta - \phi) + \mathcal{O}(\varepsilon^2) , \quad (3.135)$$

d'où :

$$\frac{f(z + \xi) - f(z)}{\xi} = \frac{\varepsilon \cos(\theta - \phi) + \mathcal{O}(\varepsilon^2)}{\varepsilon e^{i\phi}} = e^{-i\phi} \cos(\theta - \phi) + \mathcal{O}(\varepsilon) ; \quad (3.136)$$

le nombre obtenu à la limite $\varepsilon \rightarrow 0$ dépend donc du chemin suivi pour arriver en z . Il en va de même pour les fonctions $z \rightarrow \Re z$, $z \rightarrow \Im z$.

D'une façon plus générale, une fonction $f(z)$ (non constante) dont les valeurs sont réelles ($f(z) \in \mathbb{R}$) n'est pas dérivable. En effet, ε étant un nombre réel, on a par définition de la dérivée :

$$f'(z) = \lim_{\varepsilon \rightarrow 0} \frac{f(z + \varepsilon) - f(z)}{\varepsilon} . \quad (3.137)$$

Par hypothèse, si $f'(z)$ existe, alors c'est un nombre réel (le numérateur et le dénominateur au second membre sont tous deux des quantités réelles). Mais on doit aussi avoir, en prenant par exemple un accroissement imaginaire pur :

$$f'(z) = \lim_{\varepsilon \rightarrow 0} \frac{f(z + i\varepsilon) - f(z)}{i\varepsilon} . \quad (3.138)$$

Maintenant, la fraction au second membre est imaginaire pure (le numérateur est réel puisque f est à valeurs réelles). Le rapprochement de (3.137) et de (3.138) permet de conclure que ou bien $f'(z) = 0$ (par la règle d'égalité de deux complexes, c'est le seul cas où un nombre réel est égal à un nombre imaginaire pur), ou bien $f'(z)$ n'existe pas.

On montrera ultérieurement que toute fonction holomorphe est *analytique* et réciproquement²⁶. En raison de cette équivalence entre analyticit  et holomorphie, la terminologie employ e dans la litt erature peut donner l'impression d'un certain flottement : ainsi, on pourra trouver des ouvrages o  une fonction satisfaisant les conditions de Cauchy est dite *par d efinition* analytique. En tout  tat de cause, les deux qualificatifs *analytique* et *holomorphe* doivent  tre consid er s comme synonymes²⁷, bien qu'ils se rattachent   des propri t s diff erentes : l'holomorphie fait r f rence aux conditions de Cauchy, alors que le qualificatif analytique affirme l'existence d'un d veloppement en

²⁶Une fonction $f(z)$ est dite analytique en z_0 ssi, dans le voisinage de z_0 , elle admet un d veloppement en s rie enti re de puissances positives, $f(z) = \sum_{n \in \mathbb{N}} f_n \cdot (z - z_0)^n$.

²⁷De m me, le terme *holomorphe* est parfois pris dans une acception un peu sp ciale (par exemple Lavrentiev & Chabat), d signant une fonction qui n'a pas de singularit    distance finie, ce que la plupart des auteurs qualifient de fonction *enti re*.

puissances entières positives²⁸. La synonymie est justifiée par le fait qu'il y a équivalence (au sens logique du terme) entre holomorphe et analytique.

3.4 Fonctions élémentaires

Il s'agit de montrer comment les fonctions élémentaires de l'analyse réelle, $f(x)$ avec $x \in \mathbb{R}$, se généralisent sans difficulté au cas où l'argument est un nombre complexe z , et de préciser éventuellement le sens géométrique de l'application $z \rightarrow f(z) \equiv Z$. Dans cette section, on pose le cas échéant :

$$z = x + iy = re^{i\theta} , \quad Z = X + iY = \rho e^{i\phi} . \quad (3.139)$$

De surcroît, certains exemples permettront d'introduire la notion essentielle de *fonction multiforme*, qui sera reprise et précisée par la suite.

3.4.1 La fonction puissance entière $z \rightarrow Z = z^n$ ($n \in \mathbb{N}^*$) et sa(s) fonction(s) inverse(s)

On a ici :

$$\rho = r^n , \quad \phi = n\theta . \quad (3.140)$$

La transformation géométrique associée à cette fonction est une rotation de $(n-1)\theta$ et une dilatation²⁹ par le facteur r^{n-1} . Cette fonction est visiblement holomorphe dans tout le plan puisque la limite suivante existe :

$$\lim_{\xi \rightarrow 0} \frac{(z + \xi)^n - z^n}{\xi} = \lim_{\xi \rightarrow 0} \frac{z^n + nz^{n-1}\xi + \dots - z^n}{\xi} = nz^{n-1} \quad (3.141)$$

et le résultat est visiblement indépendant du chemin suivi par ξ pour regagner l'origine. C'est un exercice facile de montrer (à l'aide du développement du binôme) que les conditions de Cauchy sont satisfaites : on écrit $(x+iy)^n = \sum_{p=0}^n C_n^p x^{n-p} (iy)^p$, d'où l'on déduit les parties réelle $u(x, y)$ et imaginaire $v(x, y)$, sur lesquelles on vérifie ces conditions.

Les images de deux complexes z_1 et z_2 de même module et d'argument différant de $\frac{2\pi}{n}$ sont confondues puisque :

$$(e^{ik\frac{2\pi}{n}})^n = e^{nik\frac{2\pi}{n}} = 1 \quad \forall k \in \mathbb{Z} . \quad (3.142)$$

Ceci montre que la fonction inverse n'est pas définie à ce stade, au sens où il existe plusieurs z_k donnant le même Z . Plus précisément :

$$z = Z^{\frac{1}{n}} \iff r = \rho^{\frac{1}{n}} , \quad \theta = \frac{\phi}{n} + k\frac{2\pi}{n} , \quad (3.143)$$

²⁸On sait d'ailleurs de l'Analyse élémentaire qu'une fonction donnée par une série *entièr*e est infiniment dérivable dans le disque de convergence : la convergence *uniforme* de la série dans ce disque assure *ipso facto* la convergence de toutes les séries obtenues par dérivation terme à terme.

²⁹C'est de fait une contraction si $r < 1$.

soit :

$$z \in \left\{ \rho^{\frac{1}{n}} e^{i\frac{\phi}{n}} e^{ik\frac{2\pi}{n}} \right\} \quad (k = 1, 2, \dots, n) \quad (3.144)$$

Ainsi, la notation $Z^{\frac{1}{n}}$ ne désigne pas un seul et unique nombre mais n nombres distincts, tous de même module mais dont les arguments sont des multiples entiers de $\frac{2\pi}{n}$: pour cette raison, la fonction $Z^{\frac{1}{n}}$ est dite *multiforme* ; les différentes solutions de la relation inverse constituent les différentes *déterminations* (ou *branches*) de cette fonction multiforme. Une fonction ne présentant pas cette particularité est parfois dite *uniforme* par opposition, mais ce qualificatif est source de confusion (il caractérise aussi une fonction à valeurs réelles dont la dérivée a un signe constant).

À titre d'exemple précis, traitons le cas $n = 2$, qui permettra d'illustrer explicitement la notion importante de fonction multiforme à propos de la racine carrée. On a ici, dans les mêmes notations :

$$\rho = r^2, \quad \phi = 2\theta. \quad (3.145)$$

L'argument du point image M (d'affixe Z) varie deux fois plus vite que celle du point original m (associé à z) : quand m ne fait qu'un demi-tour, M fait un tour complet. Considérons maintenant la fonction inverse $z = Z^{\frac{1}{2}}$ et ses deux déterminations z_1 et z_2 :

$$z = Z^{\frac{1}{2}} \iff r = \rho^{\frac{1}{2}}, \quad \theta = \frac{\phi}{2} + k\frac{2\pi}{2}, \quad (3.146)$$

soit :

$$z_1 = \rho^{\frac{1}{2}} e^{i\frac{\phi}{2}}, \quad z_2 = \rho^{\frac{1}{2}} e^{i\frac{\phi}{2} + i\pi} = -\rho^{\frac{1}{2}} e^{i\frac{\phi}{2}}. \quad (3.147)$$

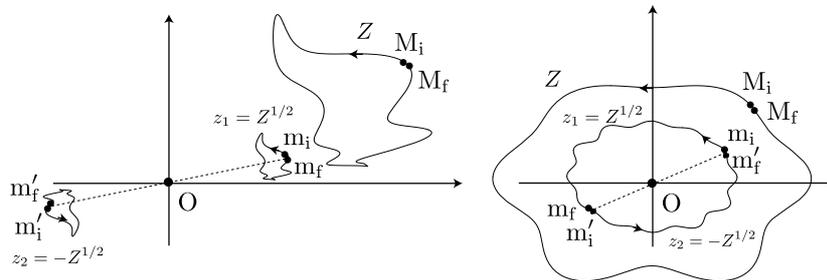


Figure 3.3: Lignes décrites par Z et par les deux déterminations $z_1 = Z^{1/2}$ et $z_2 = -Z^{1/2}$, selon que la boucle décrite par Z enserme (à droite) ou n'enserme pas (à gauche) l'origine de \mathbb{C} .

Comme l'argument du point image m varie deux fois moins vite que celle du point original M , deux cas doivent être considérés quand M décrit une courbe fermée Γ (fig. 3.3) :

1. la courbe Γ ne contient pas l'origine en son intérieur. Ceci veut dire que l'argument de Z part d'une certaine valeur ϕ et y revient ; il reprend la même valeur au point de départ M_i et au point d'arrivée M_f . Il en va de même pour l'argument de z , qui part de $\theta = \frac{1}{2}\phi$ et y revient ; alors, les deux points m_i et m_f sont confondus.
2. la courbe Γ contient l'origine en son intérieur. Alors l'argument de Z part de ϕ et arrive à $\phi + 2\pi$: ses valeurs diffèrent de 2π entre les points de départ et d'arrivée ; l'argument de z passe ainsi de $\frac{1}{2}\phi$ à $\frac{1}{2}(\phi + 2\pi)$, variant seulement de π : maintenant les deux points m_i et m_f sont distincts (leurs arguments différant de π , ces points sont symétriques l'un de l'autre par rapport à l'origine).

Ainsi, suivant que l'origine est ou non comprise dans la courbe fermée décrite par l'original Z , la courbe décrite par l'image $z \equiv Z^{\frac{1}{2}}$ est ouverte ou fermée, traduisant le fait que ses extrémités sont différentes ou confondues. L'origine apparaît comme un point remarquable pour la fonction $Z^{\frac{1}{2}}$, appelé³⁰ *point de branchement*.

Soit par exemple la détermination notée z_1 en (3.147) ; si on part d'un point situé sur le demi-axe réel négatif (l'argument de Z vaut alors π), pour lequel $z_1 = \rho^{\frac{1}{2}}e^{i\frac{\pi}{2}} = i\rho^{\frac{1}{2}}$, et que l'on tourne autour de l'origine dans le sens des aiguilles d'une montre pour revenir infiniment près du point de départ mais *en-dessous* du demi-axe réel négatif (où l'argument de Z vaut alors $-\pi$), on a en ce dernier point $z_1 = \rho^{\frac{1}{2}}e^{i\frac{-\pi}{2}} = -i\rho^{\frac{1}{2}}$. Ainsi, alors que les deux points originaux sont identiques ($M_i = M_f$, d'affixe $-\rho < 0$), les valeurs de la fonction z_1 sont distinctes ; en tout point du demi-axe réel négatif, la fonction est discontinue : ce demi-axe apparaît *dans ces conditions* comme une ligne continue de discontinuités, mais en dehors de cette ligne, la fonction z_1 est holomorphe.

Le domaine d'analyticit  de la d termination z_1 peut ainsi  tre choisi comme le plan \mathbb{R}^2 priv  (*coup *) du demi-axe r el n gatif ; cette ligne est appel e *coupure*, et doit  tre per ue comme un mur infranchissable tant que le caract re holomorphe doit  tre pr serv . Clairement, on pourrait prendre n'importe quelle autre ligne (d'ailleurs pas forc ment une demi-droite) issue de l'origine : le choix de la coupure est pure affaire de commodit , seul le³¹ point de branchement (ici l'origine) est d fini en soi, en-dehors de toute convention choisie pour le confort des calculs ou l'aisance du raisonnement g om trique.

En raisonnant trop rapidement, partant de z_1 comme ci-dessus mais en se disant que $-\pi$ ou $+\pi$ sont  gaux, on aurait trouv  $e^{i\frac{+\pi}{2}}$ pour z_1 , alors qu'il s'agit malencontreusement de la valeur de l'autre d termination not e z_2 : l'inadvertance aurait fait passer subrepticement d'une branche   l'autre. La r gle absolue pour  viter l'erreur de passer sans s'en rendre compte d'une d termination   une autre, est de suivre par continuit  des *chemins* dans le plan sans jamais "d coller" du plan.

³⁰L'origine de la terminologie appara tra clairement dans la suite.

³¹ou *les* points de branchement, car il peut en exister plusieurs. Par exemple, la fonction $(z^2 + a^2)^{\frac{1}{2}}$, $a \in \mathbb{R}_+$, a deux points de branchement en $\pm ia$.

3.4.2 La fonction exponentielle $z \rightarrow Z = e^z$

On a :

$$Z = e^{x+iy} = e^x e^{iy} = e^x (\cos y + i \sin y) . \quad (3.148)$$

Ici, $u(x, y) = e^x \cos y$ et $v(x, y) = e^x \sin y$, fonctions qui ont des dérivées continues partout dans \mathbb{C} (ouvert). On vérifie sans peine que $u_x = e^x \cos y = v_y$ et $u_y = -e^x \sin y = -v_x$. D'un autre côté, le calcul direct de la dérivée donne³² :

$$\frac{d}{dz} e^z = \lim_{\xi \rightarrow 0} \frac{e^{z+\xi} - e^z}{\xi} = \lim_{\xi \rightarrow 0} \frac{e^z e^\xi - e^z}{\xi} = e^z \lim_{\xi \rightarrow 0} \frac{e^\xi - 1}{\xi} ; \quad (3.149)$$

Sans devoir s'appuyer sur l'existence du développement limité de e^ξ en $\xi = 0$ mais en utilisant uniquement les résultats d'analyse réelle, on peut écrire :

$$e^\xi \equiv e^{h+ik} = e^h e^{ik} = e^h (\cos k + i \sin k) = [1+h+\mathcal{O}(h^2)][1+\mathcal{O}(k^2)+i[k+\mathcal{O}(k)]] , \quad (3.150)$$

d'où $e^\xi = 1+h+ik+\mathcal{O}(|\xi|^2) \equiv 1+\xi+\mathcal{O}(|\xi|^2)$, tout comme dans le champ réel. Revenant à (3.151), il vient :

$$\boxed{\frac{d}{dz} e^z = e^z} \quad (3.151)$$

La fonction exponentielle $z \rightarrow e^z \equiv f(z)$ est donc holomorphe dans \mathbb{C} et, tout comme dans le champ réel, satisfait l'équation différentielle caractéristique³³ :

$$\boxed{f' = f} . \quad (3.152)$$

Réciproquement, les seules solutions de cette équation homogène sont de la forme Ce^z où C est une constante arbitraire. Dans un problème de physique bien posé, une information supplémentaire permettra toujours de trouver C ; par exemple, on se donne (ou on connaît par ailleurs) la valeur f_0 de la fonction cherchée en un point donné z_0 . Alors la solution est unique et vaut $f_0 e^{z-z_0}$. Plus généralement, si $f(z)$ est holomorphe :

$$\boxed{\frac{d}{dz} e^{f(z)} = f'(z) e^{f(z)}} \quad (3.153)$$

et en particulier :

$$\frac{d}{dz} e^{az} = a e^{az} \quad \forall a \in \mathbb{C} \quad (3.154)$$

Pour mémoire, rappelons que la fonction exponentielle satisfait aussi l'équation fonctionnelle *caractéristique* :

$$f(z_1 + z_2) = f(z_1) f(z_2) \iff f(z) = e^{az} \quad (3.155)$$

³²Dans (3.151), la quantité $\mathcal{O}(\xi^2)$ est de la forme $\xi \varepsilon(\xi)$ où $\varepsilon(\xi) \rightarrow 0$ si $\xi \rightarrow 0$.

³³Par *caractéristique*, on entend que e^z est la seule et unique fonction satisfaisant cette équation.

où α est un nombre quelconque. L'équation n'étant pas linéaire et homogène (le premier est linéaire en f , le second est quadratique), aucune constante multiplicative arbitraire n'apparaît.

La fonction exponentielle est périodique, de (plus petite) période imaginaire pure égale à $2i\pi$:

$$e^{z+2in\pi} = e^z e^{2in\pi} = e^z \quad \forall n \in \mathbb{Z} \iff f(z+2in\pi) = f(z) . \quad (3.156)$$

La fonction e^z permet de généraliser simplement les fonctions trigonométriques circulaires et hyperboliques familières au cas où leur argument est complexe. On *définit*³⁴ ainsi :

$$\boxed{\sin z \stackrel{\text{d\u00e9f}}{=} \frac{1}{2i}(e^{iz} - e^{-iz}) \quad \cos z \stackrel{\text{d\u00e9f}}{=} \frac{1}{2}(e^{iz} + e^{-iz})} . \quad (3.157)$$

Toutes les relations trigonométriques ordinaires se généralisent alors sans peine par les règles de l'algèbre élémentaire. Par exemple :

$$\sin^2 z + \cos^2 z = \frac{1}{4i^2}(e^{2iz} - 2 + e^{-2iz}) + \frac{1}{4}(e^{2iz} + 2 + e^{-2iz}) = 1 . \quad (3.158)$$

Toutefois, des résultats bien connus pour des arguments réels, par exemple les bornes élémentaires du genre $|\sin x| < 1$, etc., ne sont plus vraies :

$$|\sin z|^2 = \sin z (\sin z)^* = \frac{1}{4}(e^{ix}e^{-y} - e^{-ix}e^y)(e^{-ix}e^{-y} - e^{ix}e^y) = \frac{1}{2}(\cosh 2y - \cos 2x) ; \quad (3.159)$$

Il est bien évident que cette dernière expression peut être aussi grande que l'on veut. De même, une équation du genre $\sin z = a$ ou $a \in \mathbb{R}$, $a > 1$, a des solutions, en fait une infinité, données par $-i \ln(a \pm \sqrt{a^2 - 1}) + (2k + \frac{1}{2})\pi$.

On vérifie immédiatement à partir de (3.157) que $\sin z$ et $\cos z$ sont des fonctions 2π -périodiques, comme dans le champ réel :

$$\sin(z + 2n\pi) = \sin z , \quad \cos(z + 2n\pi) = \cos z \quad \forall n \in \mathbb{Z} . \quad (3.160)$$

De même, on définit :

$$\tan z \stackrel{\text{d\u00e9f}}{=} \frac{\sin z}{\cos z} = -i \frac{e^{iz} - e^{-iz}}{e^{iz} + e^{-iz}} = -i \frac{1 - e^{-2iz}}{1 + e^{-2iz}} , \quad \cot z \stackrel{\text{d\u00e9f}}{=} \frac{1}{\tan z} , \quad (3.161)$$

expressions montrant que :

$$\tan(z + n\pi) = \tan z , \quad \cot(z + n\pi) = \cot z \quad \forall n \in \mathbb{Z} . \quad (3.162)$$

Quant aux dérivées, elles sont la simple généralisation au champ complexe des relations établies quand l'argument est réel :

$$\frac{d}{dz} \sin z = \frac{d}{dz} \left(\frac{1}{2i}(e^{iz} - e^{-iz}) \right) = \frac{1}{2i}(ie^{iz} + ie^{-iz}) = \cos z , \quad (3.163)$$

³⁴Il s'agit en réalité d'un prolongement analytique élémentaire et légitime de (3.26).

$$\frac{d}{dz} \cos z = -\sin z, \quad \frac{d}{dz} \tan z = \frac{1}{\cos^2 z} \quad \text{etc ;} \quad (3.164)$$

ces relations peuvent aussi se vérifier en prenant la limite $\lim_{\xi \rightarrow 0} \frac{f(z+\xi) - f(z)}{\xi}$.

On définit aussi les fonctions hyperboliques d'un argument complexe comme suit :

$$\boxed{\sinh z \stackrel{\text{def}}{=} \frac{1}{2}(e^z - e^{-z}) \quad \cosh z \stackrel{\text{def}}{=} \frac{1}{2}(e^z + e^{-z})} \quad (3.165)$$

$$\tanh z \stackrel{\text{def}}{=} \frac{\sinh z}{\cosh z} = \frac{e^z - e^{-z}}{e^z + e^{-z}} = \frac{1 - e^{-2z}}{1 + e^{-2z}}, \quad \coth z = \frac{1}{\tanh z}. \quad (3.166)$$

Ces fonctions sont aussi périodiques, mais leur période est imaginaire pure ; en effet, puisque $e^{2in\pi} = 1 \forall n \in \mathbb{Z}$:

$$\sinh(z + 2in\pi) = \sinh z, \quad \cosh(z + 2in\pi) = \cosh z. \quad (3.167)$$

De même, comme $e^{in\pi} = (-1)^n \forall n \in \mathbb{Z}$:

$$\sinh(z + in\pi) = (-1)^n \sinh z, \quad \cosh(z + in\pi) = (-1)^n \cosh z, \quad (3.168)$$

et :

$$\tanh(z + in\pi) = \tanh z, \quad \coth(z + in\pi) = \coth z \quad (3.169)$$

Noter qu'en passant de z à iz , on bascule des fonctions circulaires aux fonctions hyperboliques³⁵ :

$$\sin(iz) = i \sinh z, \quad \cos(iz) = \cosh z, \quad \tan(iz) = i \tanh z; \quad (3.170)$$

relations vraies quel que soit z , en particulier si $z = x \in \mathbb{R}$.

3.4.3 La fonction logarithme $z \rightarrow Z = \ln z$ ($z \neq 0$)

À un niveau élémentaire, il est possible de définir la fonction logarithme comme la fonction inverse de l'exponentielle :

$$\forall X \in \mathbb{R}, \quad x = e^X \in \mathbb{R}_+ \iff \forall x \in \mathbb{R}_+, \quad X = \ln x \in \mathbb{R}. \quad (3.171)$$

Cette procédure est sans ambiguïté quand $x \in \mathbb{R}$ puisque l'exponentielle est biunivoque. C'est cette définition que l'on étend au champ complexe, en écrivant (gardant la même notation, \ln , que pour le logarithme réel) :

$$\forall Z \in \mathbb{C}, \quad z = e^Z \in \mathbb{C} \iff \forall z \in \mathbb{C}, \quad Z = \ln z \in \mathbb{C}. \quad (3.172)$$

Cela étant posé, on réalise qu'il n'y a pas une fonction logarithme complexe, mais une infinité : en effet, tous les Z différant de $k \times 2i\pi$ ($k \in \mathbb{Z}$) donnent le même z . Autrement dit, $\ln z$ est défini modulo $2i\pi$, et possède une infinité de déterminations.

³⁵Pour trouver la bonne place de i , penser à $\sinh z \sim z$ si $|z| \ll 1$, etc.

Une autre façon de voir ceci consiste à écrire z sous forme polaire et à utiliser la propriété caractéristique du logarithme $\ln(zz') = \ln z + \ln z'$ (qui résulte simplement du fait que l'exponentielle de la somme est le produit des exponentielles). Alors³⁶ :

$$\ln z = \ln(re^{i\theta}) = \ln r + \ln(e^{i\theta}) = \ln r + i\theta \quad (r \neq 0) . \quad (3.173)$$

Alors, l'infinité des valeurs possibles pour le premier membre saute aux yeux puisque l'angle θ est lui-même défini à 2π près. Ceci admis, $\ln r$ au second membre de (3.173) peut toujours être considéré comme le logarithme (complexe) du nombre (réel) $re^{i\theta}|_{\theta=0}$: le $ik'2\pi$ que l'on serait tenté de rajouter peut toujours être absorbé dans le $ik'2\pi$ inévitable lié à θ . En tout cas, si on a bien $e^{i(\theta+2n\pi)} = e^{i\theta}$ quel que soit $n \in \mathbb{Z}$, en revanche, pour un angle θ donné (12345.6789 radians, par exemple), $\ln[i(\theta + 2n\pi)] \neq \ln(i\theta)$: tout comme pour la fonction puissance, il existe plusieurs formes admissibles en tant que fonction inverse, en fait autant qu'on veut³⁷, la relation $Z = e^z$ donnant une infinité de valeurs pour z . $\ln z$ est ainsi un autre exemple de fonction multiforme.

La considération de courbes fermées décrites par le point représentatif de l'original z , contenant ou non l'origine O , conduit à des conclusions analogues à celles tirées pour la racine carrée (section 3.4.1) : par exemple, si z fait un tour complet dans le sens positif autour de O , la valeur finale du logarithme est augmentée de $2i\pi$ par rapport à sa valeur de départ, quelle que soit la détermination choisie ; si la courbe laisse l'origine à l'extérieur, le logarithme prend à l'arrivée la même valeur qu'au départ³⁸. O est donc le point de branchement de la fonction logarithme, et la coupure nécessaire pour préserver le caractère holomorphe est une ligne partant l'origine et filant à l'infini dans n'importe quelle direction.

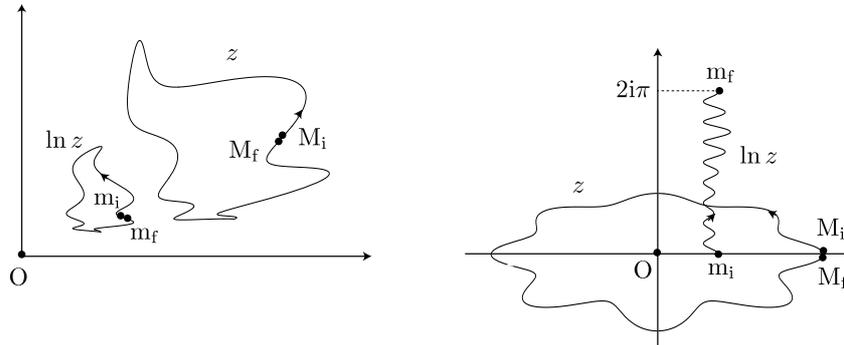


Figure 3.4: Lignes décrites par z et par la détermination $\ln z$, $0 \leq \text{Arg} z < 2\pi$, selon que la boucle décrite par z enferme (à droite) ou n'enferme pas (à gauche) l'origine de \mathbb{C} . À droite : si $x_0 \in \mathbb{R}$ est l'affixe commune des deux points M_i et M_f , l'affixe de m_i est $\ln x_0$, celle de m_f est $\ln x_0 + 2i\pi$.

³⁶De toute évidence, $\ln z$ n'est pas défini en $z = 0$, ne serait-ce que parce que $\ln r$ ne l'est pas non plus.

³⁷Pour la fonction inverse de $Z = z^n$, il n'y a que n valeurs distinctes pour z .

³⁸On note aussi que si z décrit une certaine courbe fermée ne contenant pas l'origine, toutes les déterminations, différant de entier $\times 2i\pi$ les unes des autres, décrivent des courbes translattées de 2π les unes des autres, formant un réseau à une dimension parallèle à l'axe imaginaire.

Un usage fréquent (mais pas systématique) est de noter $\text{Ln}z$ la détermination suivante :

$$z = r e^{i\theta} : \text{Ln}z = \ln r + i\theta \quad (r \neq 0, -\pi < \theta < +\pi), \quad (3.174)$$

où $\ln r$ est le logarithme élémentaire, parfois appelée³⁹ *valeur principale du logarithme* ; la coupure est alors le demi-axe réel négatif. Avec cette définition, $\text{Ln}x$ pour $x \in \mathbb{R}_+$ coïncide avec la fonction logarithme définie à un niveau élémentaire :

$$\boxed{\text{Ln}x = \ln x \quad (x > 0)}, \quad (3.175)$$

alors que pour $x < 0$:

$$\text{Ln}(x + i0) = \ln|x| + i\pi, \quad \text{Ln}(x - i0) = \ln|x| - i\pi, \quad (3.176)$$

ce que l'on peut aussi écrire ($z = r e^{i\theta}$, avec $\theta = \pm(\pi - 0)$) :

$$\boxed{\text{Ln}(z = x \pm i0) = \ln|x| \pm i\pi \quad (x < 0)} \quad (3.177)$$

Ainsi, pour deux points situés juste au-dessus ou juste au-dessous de la coupure, la fonction $\text{Ln}z$ a un saut ; si sa partie réelle est la même (pour un x négatif donné), sa partie imaginaire vaut $+i\pi$ d'un côté, $-i\pi$ de l'autre : tout comme pour la fonction $z^{\frac{1}{2}}$, la coupure est une ligne continue du plan où la fonction est discontinue. Ne pas pouvoir franchir cette ligne si l'on veut préserver le caractère holomorphe devient une évidence : la seule façon de faire est de rebrousser chemin. Pour cette détermination, on a :

$$\boxed{\text{Ln}(z^*) = (\text{Ln}z)^*} \quad (3.178)$$

qui exprime une certaine *symétrie* de la branche particulière choisie. Il arrive que le problème physique analysé impose ce genre de symétrie, auquel cas le choix de la coupure s'impose de lui-même et ne peut être fait arbitrairement.

La dérivée de $\ln z$ se trouve par le procédé habituel, en faisant le rapport des accroissements et en en prenant la limite. Par ailleurs, il est évident que *toutes* les déterminations du logarithme donnent le même résultat puisque dans toute soustraction $\ln(z + \xi) - \ln z$, le $2ik\pi$ disparaît : toutes les déterminations de $\ln z$ ont la même dérivée (une évidence d'ailleurs, puisqu'elles ne diffèrent que d'une constante additive). En raisonnant par exemple avec la détermination définie en (3.174), il vient :

$$\frac{d}{dz} \text{Ln}z = \lim_{\xi \rightarrow 0} \frac{\text{Ln}(z + \xi) - \text{Ln}z}{\xi}. \quad (3.179)$$

Le numérateur a une valeur sans ambiguïté : sa partie imaginaire est la variation de l'argument quand on passe de z au nombre $z + \xi$ infiniment voisin, qui est infiniment petite⁴⁰. On peut donc l'écrire comme le Ln du rapport $\frac{z+\xi}{z}$, d'où :

$$\frac{d}{dz} \text{Ln}z = \lim_{\xi \rightarrow 0} \frac{\text{Ln}\left(\frac{z+\xi}{z}\right)}{\xi} = \lim_{\xi \rightarrow 0} \frac{\text{Ln}\left(1 + \frac{\xi}{z}\right)}{\xi}. \quad (3.180)$$

³⁹... qui n'a rien à voir avec la partie principale de Cauchy permettant de régulariser une intégrale du type $\int_a^b \frac{f(x)}{x} dx$ où f est comme il faut en $x = 0$ et avec $a < 0 < b$.

⁴⁰Noter qu'il n'y a jamais aucune ambiguïté, quelle que soit la branche choisie, $\text{Ln}z$ ou n'importe quelle autre détermination : dans tous les cas, la variation de l'argument est la même !

En utilisant maintenant⁴¹ $\text{Ln}(1+z) = z + \mathcal{O}(z^2)$, et puisque toutes les déterminations ont la même dérivée, il vient :

$$\boxed{\frac{d}{dz} \ln z = \frac{1}{z}} \quad (3.182)$$

Une fois définie la notion de *primitive* d'une fonction $f(z)$ (voir chapitre 4), on pourra dire que, tout comme dans le champ réel, $\ln z$ est une primitive de $\frac{1}{z}$, quelle que soit la détermination choisie, en agrément avec le fait que toutes les primitives d'une même fonction ne diffèrent les unes des autres que par des constantes additives.

Si l'on veut de surcroît vérifier les conditions de Cauchy, leur forme polaire (3.106) est recommandée ; par ailleurs, puisqu'il s'agit de relations différentielles, il suffit de considérer une détermination, la détermination principale (3.174) par exemple. On a ici :

$$U(r, \theta) = \ln r \quad , \quad V(r, \theta) = \theta \quad (-\pi < \theta \leq +\pi) \quad , \quad (3.183)$$

de sorte que les conditions s'écrivent (voir (3.106)) :

$$\frac{\partial U}{\partial r} = \frac{1}{r} \quad , \quad \frac{\partial V}{\partial \theta} = 1 \quad , \quad \frac{\partial V}{\partial r} = 0 \quad , \quad \frac{\partial U}{\partial \theta} = 0 \quad . \quad (3.184)$$

Les conditions de Cauchy sont donc trivialement satisfaites $\forall \theta \in]-\pi, +\pi[$ (en revanche, $\text{Ln}z$ n'est pas une fonction *continue* dans \mathbb{C} puisque si z est sur le demi-axe réel négatif, alors $\theta = +\pi$, tandis que si z est infiniment proche de ce demi-axe, mais en-dessous, $\theta = -\pi$). La fonction n'est donc analytique que dans le domaine $\theta \in]-\pi, +\pi[$, qui est plus petit que le domaine de définition $\theta \in]-\pi, +\pi]$. Le même phénomène se produit pour toutes les autres déterminations.

On peut s'y prendre autrement ; pour la détermination principale $\text{Ln}z$ définie en (3.174), on a explicitement :

$$\text{Ln}z = \ln \sqrt{x^2 + y^2} + i\theta \quad , \quad -\pi < \theta < +\pi \quad (3.185)$$

Ici, $v(x, y) = \text{Arctan} \frac{y}{x} + C$, où C est une constante à ajuster selon que $|\theta|$ est plus petit ou plus grand que $\frac{\pi}{2}$. On en déduit $u_x = \frac{x}{x^2+y^2}$ et $v_y = \frac{x}{x^2+y^2}$, d'où $u_x = v_y$, et de même $u_y = -v_x$. En reprenant l'une des formes possibles pour la dérivée (voir (3.107)), on peut écrire $(\text{Ln}z)' = \frac{x-iy}{x^2+y^2}$, c'est-à-dire $(\text{Ln}z)' = \frac{z^*}{zz^*} = \frac{1}{z}$, comme il se doit.

⁴¹Si on ne souhaite pas utiliser la généralisation au champ complexe du développement de $\ln(1+x)$ (après tout, il n'a pas encore été introduit), on peut raisonner comme suit. Soit $\ln(1+\varepsilon e^{i\theta})$; par la définition du logarithme, on a :

$$\ln(1+\varepsilon e^{i\theta}) = \ln \sqrt{(1+\varepsilon \cos \theta)^2 + (\varepsilon \sin \theta)^2} + i \arctg \frac{\varepsilon \sin \theta}{1+\varepsilon \cos \theta} \quad (3.181)$$

Dans la limite $\varepsilon \rightarrow 0$, le log de la racine carrée se comporte comme $\frac{1}{2}(2\varepsilon \cos \theta)$, et l'arctg comme $\varepsilon \sin \theta$, d'où $\ln(1+\varepsilon e^{i\theta}) \simeq \varepsilon \cos \theta + i\varepsilon \sin \theta = \varepsilon e^{i\theta}$.

On peut encore s'y prendre autrement ; posant $\text{Ln}(1+z) = Z$, il vient $1+z = e^Z$. Si $|z| \ll 1$, $1+z \simeq 1$, donc $e^Z \simeq 1$, soit $|Z| \ll 1$. Selon (3.150), $e^Z = 1+Z+\mathcal{O}(Z^2)$, d'où $1+z = 1+Z+\mathcal{O}(Z^2)$, soit $z = Z+\mathcal{O}(Z^2)$ et finalement $\text{Ln}(1+z) = z+\mathcal{O}(z^2)$.

En raison du caractère multiforme de $\ln z$, et pour être en mesure de lever toute ambiguïté en pratique, il suffit de comprendre que la question “Combien vaut $\ln z$?” n’a pas de réponse (unique) tant que l’on n’a pas précisé la branche considérée de la fonction multiforme. En revanche, la question “Combien vaut $\ln z$ pour la branche égale à $\ln r$ en $\theta = 0$?” admet une réponse unique. L’information supplémentaire permet bien de calculer univoquement $\ln z$. En pratique, pour éviter de raisonner de travers, le plus sûr est sans doute de garder en tête la notion de continuité, en suivant mentalement un chemin du plan allant du point pour lequel l’information supplémentaire est fournie au point où l’on veut calculer la fonction.

3.4.4 La fonction puissance généralisée $z \rightarrow Z = z^\alpha$ ($\alpha \in \mathbb{C}$)

Soit maintenant à généraliser la fonction puissance au cas où l’exposant n’est plus un entier ou un rationnel, mais un nombre complexe quelconque α :

$$Z = z^\alpha \quad (\alpha \in \mathbb{C}) . \quad (3.186)$$

Prenant le logarithme des deux membres, on a $\ln Z = \alpha \ln z$; prenant maintenant l’exponentielle de cette dernière relation, il vient $Z = e^{\alpha \ln z}$, égalité qui constitue la définition naturelle de la puissance généralisée :

$$\boxed{z^\alpha \stackrel{\text{def}}{=} e^{\alpha \ln z} \quad (\alpha \in \mathbb{C})} \quad (3.187)$$

Cette fonction est visiblement multiforme en général, puisque $\ln z$ l’est et que α est quelconque⁴². Quand α est réel, les différentes branches s’expriment explicitement comme ($z = re^{i\theta}$) :

$$z^\alpha = e^{\alpha \ln r + i\alpha(\theta + 2k\pi)} = e^{\alpha \ln r + i\alpha\theta} e^{2i\alpha k\pi} = r^\alpha e^{i\alpha\theta} e^{2i\alpha k\pi} \quad (k \in \mathbb{Z}) . \quad (3.188)$$

C’est seulement dans le cas où $\alpha \in \mathbb{Z}$ que l’on retrouve une fonction à une seule détermination, puisque le produit αk est alors un entier (le dernier facteur à droite vaut alors 1 quel que soit k). Si α est un réel rationnel, $\alpha = \frac{p}{q}$, p et q premiers entre eux, il n’y a que q branches distinctes ; si α est irrationnel, z^α possède une infinité de branches distinctes. L’exemple de la fonction $(z - 1)^{1/2}$ est illustré sur la fig. 3.5 ; noter qu’il s’agit seulement de la fonction $z^{1/2}$ déjà rencontrée en 3.4.1 à propos de la fonction inverse de $z \rightarrow f(z) = z^2$, simplement décalée de +1 le long de l’axe réel.

Si $\alpha = a + ib$, alors :

$$z^{a+ib} = e^{(a+ib)[\ln r + i(\theta + 2k\pi)]} = e^{a \ln r - b(\theta + 2k\pi)} e^{i[b \ln r + a(\theta + 2k\pi)]} \quad (k \in \mathbb{Z}) . \quad (3.189)$$

Par exemple, toujours avec a et b réels :

$$\underline{z^a = e^{a \ln r} e^{ia(\theta + 2k\pi)} = r^a e^{ia(\theta + 2k\pi)}}, \quad \underline{z^{ib} = e^{-b(\theta + 2k\pi)} e^{ib \ln r}} \quad (k \in \mathbb{Z}) . \quad (3.190)$$

⁴²de sorte que les différentes branches du logarithme donnent le facteur $e^{2i\alpha k\pi}$, qui ne vaut pas 1 puisque $\alpha \neq$ entier.

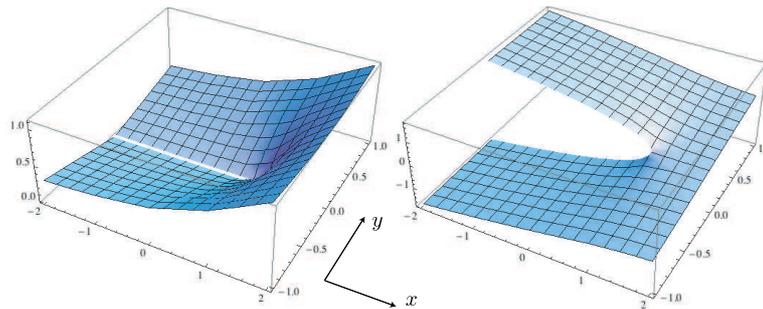


Figure 3.5: Parties réelle (à gauche) et imaginaire (à droite) de la fonction $(z - 1)^{\frac{1}{2}}$. La coupure est visible par la discontinuité de $\Im(z - 1)^{\frac{1}{2}}$ le long du demi-axe réel $z \in \mathbb{R}, z < 1$.

Par exemple :

$$\boxed{i^k = e^{-i(4k+1)\frac{\pi}{2}} \quad (k \in \mathbb{Z})} \quad (3.191)$$

Notons que pour $\alpha = 0$, (3.187) montre que $z^0 = 1$. Enfin, la même définition permet d'écrire successivement, utilisant (3.153) au passage :

$$\frac{d}{dz} z^a = \frac{d}{dz} e^{a \ln z} = \frac{a}{z} e^{a \ln z} = \frac{a}{z} z^a = a z^{a-1} ; \quad (3.192)$$

ce résultat généralise $(z^n)' = n z^{n-1}$ au champ complexe pour l'exposant.

Chapitre 4

Intégration des fonctions d'une variable complexe

“Les hommes meurent, mais leurs travaux restent.”
(dernières paroles de Augustin-Louis Cauchy, 1789–1857)

Après avoir montré comment la dérivation se généralise à \mathbb{C} , il s'agit ici de présenter quelques résultats fondamentaux sur l'intégration des fonctions d'une variable complexe.

4.1 Préliminaires

La notion d'intégrale de Riemann d'une fonction à valeurs réelles sur un domaine $I \subset \mathbb{R}$ de sa variable est bien connue. Très schématiquement, pour une fonction $f(x)$ continue¹ dans l'intervalle $[a, b] \subset \mathbb{R}$, on définit une grille de points d'abscisses x_n dans l'intervalle $[a, b]$ ($x_0 = a, x_N = b$) et on considère les sommes du genre (dites *sommes de Darboux*) :

$$S_N = \sum_{n=0}^{N-1} f(\xi_n)(x_{n+1} - x_n) , \quad (4.1)$$

où $x_n \leq \xi_n \leq x_{n+1}$. Si la limite de S_N existe quels que soient les ξ_n quand $N \rightarrow \infty$ et que $\sup |x_{n+1} - x_n|$ tend vers zéro, cette limite est appelée l'intégrale de f sur le segment $[a, b]$. Dans la notation usuelle, on écrit :

$$\lim_{N \rightarrow \infty} \lim_{\sup |x_{n+1} - x_n| \rightarrow 0} S_N = \int_a^b f(x) dx . \quad (4.2)$$

¹La généralisation à une fonction continue *par morceaux* est immédiate.

Le symbole dx désigne une variation infiniment petite de la variable x autour d'une valeur x parcourant l'intervalle $[a, b]$. Dans ce cas, et de façon un peu pédante, on peut dire que le segment de droite d'extrémités a et b est un *chemin* allant de a à b .

C'est cette définition qu'il s'agit de généraliser au cas d'une fonction continue à valeurs complexes $f(z)$, z étant lui-même un complexe. La première notion à préciser est le chemin suivi dans le plan pour aller d'une borne à l'autre de l'intégrale, c'est-à-dire d'un point A à un point B du plan, d'affixes respectives z_A et z_B . Dans toute la suite, on ne considérera que des chemins parcourus dans un certain sens (orientés) et constitués d'arcs de courbe une fois continûment différentiables (arcs lisses de Jordan), éventuellement par morceaux. Un tel chemin C , d'extrémités A et B fixées étant précisé, on le tronçonne en petits arcs délimités par des complexes z_n , $0 \leq n \leq N$ ($z_0 = z_A$, $z_N = z_B$) et on écrit la somme :

$$\mathcal{S}_N = \sum_{n=0}^{N-1} f(\xi_n)(z_{n+1} - z_n) , \quad (4.3)$$

où ξ_n est l'affixe d'un point situé sur C et appartenant au petit arc de courbe délimité par les points d'affixes z_n et z_{n+1} . On définit alors l'intégrale le long de C comme la limite :

$$\int_C f(z) dz = \lim_{N \rightarrow \infty} \lim_{\sup |z_{n+1} - z_n| \rightarrow 0} \mathcal{S}_N . \quad (4.4)$$

dz désigne une variation infinitésimale du nombre complexe z autour de la valeur z (associé à un arc de chemin infiniment petit – que l'on peut noter dC – dont les affixes sont z et $z + dz$) ; ce sont tous les petits arcs dC qui, mis bout à bout, reconstituent le chemin C . La limite (4.4) existe si f est une fonction bornée continue par morceaux et si le chemin C est une fois continûment différentiable par morceaux. En posant $f = u + iv$, $z = x + iy$ (et donc $dz = dx + idy$), on a :

$$\int_C f(z) dz = \int_C [u(x, y) + iv(x, y)](dx + idy) = \int_C (u dx - v dy) + i \int_C (u dy + v dx) . \quad (4.5)$$

Ainsi, l'intégrale d'une fonction complexe peut toujours s'exprimer à l'aide des deux intégrales réelles ci-dessus. De telles intégrales sont dites *curvilignes*, puisque les variables x et y peuvent toujours être interprétées comme les coordonnées cartésiennes d'un point M se déplaçant sur la courbe C .

Précisons la signification d'une intégrale curviligne. La courbe C a une équation cartésienne $\Lambda(x, y) = 0$, mais il est toujours possible en principe d'en trouver une représentation paramétrique $(x(t), y(t))$ où t est un nombre réel : quand t varie, le point d'affixe $z(t) = x(t) + iy(t)$ décrit la courbe C du point A (quand $t = t_A$) au point B (quand $t = t_B$).

Par exemple, la droite du plan d'équation $y = a(x - x_0) + y_0$ admet la représentation paramétrique² $x = x_0 + t$, $y = at + y_0$, les points de cette droite ayant les affixes

²C'est manifestement la représentation la plus simple, mais il en existe autant qu'on veut : on peut prendre $x = f(t) + x_0$ où f est une fonction uniforme croissante appliquant \mathbb{R} sur \mathbb{R} , et alors $y = af(t) + y_0$. La représentation paramétrique d'une courbe dépend évidemment du choix de l'origine dans le plan.

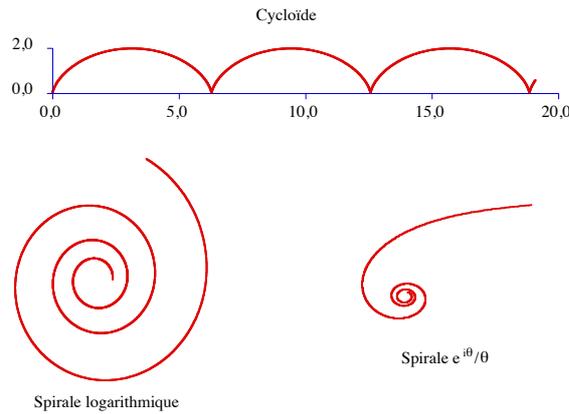


Figure 4.1: Exemples de courbes données sous forme paramétrique ou polaire. La spirale logarithmique est $x = e^{a\theta} \cos \theta$, $y = e^{a\theta} \sin \theta$: c'est le lieu des points dont l'affixe est $e^{(a+i)\theta}$; la spirale en bas à droite a pour équation polaire $r = \frac{e^{i\theta}}{\theta}$.

$z(t) = x_0 + t + i(at + y_0) = x_0 + iy_0 + (1 + ia)t$. De même, le cercle centré en $z_0 = x_0 + iy_0$ et de rayon R est l'ensemble des points de coordonnées $x = x_0 + R \cos t$, $y = y_0 + R \sin t$, et l'affixe de chacun de ces points est le complexe $z(t) = z_0 + Re^{it}$. Avec cette représentation, t a une signification très simple : c'est l'angle polaire du rayon joignant le centre du cercle au point de coordonnées $x(t)$, $y(t)$. Autre exemple : la courbe décrite par une mouche accrochée à une roue de bicyclette de rayon R qui roule sans glisser (cycloïde) a pour représentation paramétrique $x(t) = R(t - \sin t)$, $y(t) = R(1 - \cos t)$, si la mouche est "au départ" ($t = 0$) à l'origine. t désigne alors l'angle entre la verticale passant par le centre de la roue et le point où est agrippée la mouche endormie. L'affixe d'un point courant de la cycloïde est $z(t) = R(t + i) - iRe^{-it}$.

Dès que l'on connaît une représentation paramétrique de C , une intégrale curviligne s'explique aisément. Par exemple, en utilisant $dx = x'(t)dt$ et $dy = y'(t)dt$, la première intégrale de (4.5) s'écrit :

$$\int_C (u dx - v dy) = \int_{t_A}^{t_B} \{u[x(t), y(t)] x'(t) - v[x(t), y(t)] y'(t)\} dt \equiv \int_{t_A}^{t_B} \gamma(t) dt, \quad (4.6)$$

où $\gamma(t)$ est une certaine fonction de t , bien déterminée ; il en va de même pour la partie imaginaire de l'intégrale, $\int_C (u dy + v dx)$. En définitive, posant $z'(t) = x'(t) + iy'(t)$, d'où $dz = [x'(t) + iy'(t)]dt$, il vient :

$$\int_C f(z) dz = \int_{t_A}^{t_B} f[z(t)] z'(t) dt \equiv \int_{t_A}^{t_B} \Gamma(t) dt, \quad (4.7)$$

où $\Gamma(t)$ est encore une autre fonction. En résumé, l'intégrale d'une fonction complexe sur sa variable complexe peut s'écrire des différentes façons :

$$\int_C f(z) dz = \int_C [u(x, y) + iv(x, y)](dx + idy) = \int_C (u dx - v dy) + i \int_C (u dy + v dx) = \int_{t_A}^{t_B} f[z(t)] z'(t) dt \quad (4.8)$$

La courbe C étant choisie, le calcul complet est en principe possible, en tout cas l'intégrale $\int_C f(z) dz$ est définie sans ambiguïté.

Rappelons que pour un arc simple de Jordan (voir p. 125), la fonction $z(t)$ est biunivoque : $t_1 \neq t_2 \iff z(t_1) \neq z(t_2)$. La longueur dL d'un arc élémentaire est égale au module de la variation infinitésimale de z soit $dL = |dz|$; une fois choisie une représentation paramétrique, on peut ainsi écrire $dL = \sqrt{x'^2(t) + y'^2(t)} |dt| = |z'(t)| dt$. Pour un arc dont les extrémités sont définies par $t = \alpha$ et $t = \beta$ ($\alpha < \beta$), la longueur de l'arc est L :

$$L = \int_C |dz| = \int_{\alpha}^{\beta} |z'(t)| dt . \quad (4.9)$$

Pour finir, montrons que, avec les hypothèses admises, l'intégrale $\int_C f(z) dz$ existe. En effet, f étant bornée³, et comme le module d'une somme est toujours inférieur ou égal à la somme des modules⁴, on a :

$$\left| \int_C f(z) dz \right| \leq \int_C |f(z) dz| = \int_C |f(z)| |dz| \leq \int_C M |dz| = ML . \quad (4.10)$$

Les propriétés habituelles des intégrales s'étendent immédiatement aux curvilignes et restent donc également vraies pour $\int_C f(z) dz$:

1. linéarité de l'intégrand :

$$\int_C [af(z) + bg(z)] dz = a \int_C f(z) dz + b \int_C g(z) dz , \quad (4.11)$$

où a et b sont des constantes.

2. Additivité des chemins :

$$\int_{C_1 \cup C_2} f(z) dz = \int_{C_1} f(z) dz + \int_{C_2} f(z) dz . \quad (4.12)$$

3. Si $-C$ désigne le contour C parcouru en sens inverse :

$$\int_{-C} f(z) dz = - \int_C f(z) dz \quad (4.13)$$

³ M existe toujours puisque f est supposée continue (éventuellement par morceaux) et définie sur le contour. La fonction $f[z(t)]$ est donc continue sur le fermé borné $t \in [a, b]$ et y possède toujours un maximum, c'est précisément M .

⁴L'intégrale est la limite d'une somme, et c'est en vertu de ceci que $|\int_C f(z) dz| \leq \int_C |f(z) dz|$.

4.2 Théorème de Cauchy (1825)

Une fonction f étant choisie, l'intégrale $\int_C f(z) dz$ dépend en général du chemin⁵ C . La question que l'on se pose maintenant est la suivante : existe-t-il une classe de fonctions remarquables telles que, les extrémités A et B de C étant fixées, l'intégrale d'une fonction donnée prend la *même* valeur pour *tous* les chemins reliant A à B ?

Le théorème de Cauchy répond à cette question : cette classe est l'ensemble des fonctions holomorphes. D'une importance capitale, ce théorème est la clé de voûte de l'intégration dans le plan complexe. Il est dû à Cauchy et peut s'énoncer de plusieurs façons équivalentes ; l'une d'entre elles est [9] :

Si $f(z)$ est holomorphe dans un domaine simplement connexe \mathcal{D} , l'intégrale $\int_C f(z) dz$ prend la même valeur pour tous les chemins C inclus dans \mathcal{D} et ayant les mêmes extrémités.

Pour démontrer ce théorème de façon élémentaire, on ajoute l'hypothèse que $f'(z)$ est *continue* (alors que, en toute rigueur et par définition, le caractère holomorphe assure seulement l'*existence* de la dérivée – voir note 21 p. 133). Avec $f = u + iv$, l'intégrale de f le long de C est :

$$\int_C f(z) dz = \int_C (u dx - v dy) + i \int_C (u dy + v dx) . \quad (4.14)$$

D'un autre côté, soit l'intégrale curviligne :

$$\int_{\Gamma} (P dx + Q dy) , \quad (4.15)$$

où Γ est un contour dans le plan \mathbb{R}^2 reliant deux points fixes A et B, et où $P(x, y)$ et $Q(x, y)$ sont des fonctions données ; cette intégrale existe si P et Q sont continues sur Γ (éventuellement par morceaux). On sait que l'intégrale curviligne ne dépend pas du chemin suivi ssi la forme différentielle constituant l'intégrand est une différentielle totale. En d'autres termes, il existe une fonction⁶ $\Phi(x, y)$ telle que $P dx + Q dy$ est égale à la différentielle de Φ , $d\Phi$. Il en est ainsi, par définition, ssi, P (*resp.* Q) est la dérivée

⁵Tout comme le travail d'une force \vec{F} ne dérivant pas d'un potentiel, $W_C = \int_C \vec{F} \cdot d\vec{l}$, dépend du chemin suivi dans l'espace par le point d'application de la force. Pour un champ de forces à deux dimensions (dans \mathbb{R}^2), s'il existe une fonction $U(x, y)$ telle que $F_x = -\frac{\partial U}{\partial x}$, alors W_C ne dépend que des extrémités de C . Le contexte permet alors d'affirmer que F_x et $-F_y$ peuvent être considérées comme les parties réelle et imaginaire d'une fonction holomorphe $\mathcal{F}(z) = F_x(x, y) - iF_y(x, y)$; alors, le travail est la partie réelle de $\int_C \mathcal{F}(z) dz$ (tout ceci suppose que U est une fonction continue à dérivées continues).

⁶En Thermodynamique, c'est une fonction d'état, c'est-à-dire une fonction qui ne dépend que des variables thermodynamiques (d'équilibre) des états initial et final. L'affirmation de l'*existence* de l'énergie en tant que *fonction* des variables d'état constitue le Premier principe. La banalité des mots occulte souvent la profondeur d'une telle affirmation.

partielle de Φ par rapport à x (*resp.* y). En définitive, il faut d'une part que la fonction Φ existe, d'autre part que l'on ait précisément :

$$\exists \Phi(x, y) : \quad P(x, y) = \frac{\partial}{\partial x} \Phi(x, y) , \quad Q(x, y) = \frac{\partial}{\partial y} \Phi(x, y) . \quad (4.16)$$

Dans ces conditions, l'intégrale (4.14) est simplement égale à $\Phi(B) - \Phi(A)$, *différence* des valeurs de Φ aux extrémités fixées, et donc indépendante⁷ du chemin suivi pour relier l'une à l'autre.

Maintenant, un théorème classique d'Analyse (théorème de Clairaut, voir par exemple [31], p. 407) établit que si une fonction $\Phi(x, y)$ et ses dérivées Φ'_x , Φ'_y , Φ''_{xy} et Φ''_{yx} sont définies et continues en un point, alors les dérivées croisées sont égales :

$$\frac{\partial}{\partial x} \frac{\partial \Phi}{\partial y} = \frac{\partial}{\partial y} \frac{\partial \Phi}{\partial x} . \quad (4.18)$$

Autrement dit, pour une fonction continue à dérivées secondes croisées continues, peu importe l'ordre dans lequel on effectue les deux dérivations croisées. Comme P et Q sont les dérivées premières de Φ , les conditions sur Φ''_{xy} et Φ''_{yx} se reportent sur les dérivées premières de P et Q . C'est ici qu'intervient l'hypothèse de continuité sur $f'(z)$ faite au début, puisque u et v vont jouer des rôles analogues à ceux de P et Q , dérivées premières de Φ , et que la dérivée $f'(z)$ s'exprime (de quatre façons d'ailleurs) à l'aide de u'_x , v'_y , ... (voir chapitre 3). La continuité des dérivées jusqu'à l'ordre 2 inclus pour Φ se reporte sur les dérivées jusqu'à l'ordre 1 inclus de u et v : la continuité de $f'(z)$ assure bien que les quatre dérivées u'_x , v'_y , u'_y et v'_x sont continues, permettant l'application du théorème de Clairaut.

Cela étant, partant de la condition (4.16) assurant que $P dx + Q dy$ est une différentielle totale, et utilisant (4.18), on arrive à l'égalité :

$$\boxed{\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}} \quad (4.19)$$

Revenant à (4.14), on veut que l'intégrale ne dépende pas du chemin suivi, une condition qui se reporte séparément et indépendamment sur la partie réelle et sur la

⁷C'est bien ce qui caractérise une différentielle totale, combinaison linéaire d'accroissements infinitésimaux dont les coefficients sont les dérivées partielles d'une certaine fonction des variables. Quel que soit le découpage, la somme de toutes les différences $\delta\Phi = \Phi(M_{i+1}) - \Phi(M_i)$ se simplifie ($A=M_0$, $B=M_N$) :

$$\sum \delta\Phi = \sum_{i=0}^{N-1} \Phi(M_{i+1}) - \Phi(M_i) = \Phi(M_1) - \Phi(M_0) + \Phi(M_2) - \Phi(M_1) + \dots + \Phi(M_N) - \Phi(M_{N-1}) = \Phi(B) - \Phi(A) . \quad (4.17)$$

À la limite, la somme des petits accroissements est l'intégrale.

partie imaginaire de l'intégrale (4.14). Pour la partie réelle, la quantité $u dx - v dy$ doit être une différentielle totale ; (4.19) avec $P = u$ et $Q = -v$ donne :

$$-\frac{\partial v}{\partial x} = \frac{\partial u}{\partial y} . \quad (4.20)$$

De même pour la partie imaginaire $u dy + v dx$, avec $Q = u$ et $P = v$ (4.19) donne :

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} . \quad (4.21)$$

Les deux conditions (4.20) et (4.21) ne sont rien d'autre que les conditions de Cauchy ; complétées par les hypothèses de continuité des dérivées premières de u et v elles sont équivalentes à la propriété d'holomorphic. Ainsi, il y a équivalence entre l'holomorphic de $f(z)$ et la propriété pour les intégrands des parties réelle et imaginaire de (4.14) d'être des différentielles totales. Au total, l'holomorphic de $f(z)$ assure bien que l'intégrale $\int_C f(z) dz$ ne dépend pas du chemin suivi : le théorème de Cauchy est ainsi démontré.

◆ *Remarque*

La démonstration donnée ci-dessus, fort classique, est une démonstration *faible*, s'appuyant sur l'hypothèse supplémentaire de la continuité de $f'(z)$. Dès lors, il y a équivalence entre les conditions nécessaires et suffisantes de Cauchy - Riemann et l'indépendance de l'intégrale vis-à-vis du chemin suivi pour relier les extrémités fixées. Tant que l'on en reste au strict énoncé du théorème de Cauchy – qui se démontre aussi sans cette hypothèse supplémentaire –, sa réciproque reste à démontrer : c'est le théorème de Morera (voir p. 178).

◆

Il en résulte qu'une intégrale le long d'un certain contour peut aussi être notée $\int_{z_A}^{z_B} f(z) dz$, où z_A et z_B sont les affixes des extrémités A et B du chemin (seules les extrémités sont à préciser puisque l'intégrale ne dépend pas du chemin choisi pour aller de l'une à l'autre). On pourra retenir l'image suivante : finalement, pour une fonction holomorphic dans \mathcal{D} , le chemin est un élastique fixé par deux punaises dans le plan, déformable à souhait (mais en restant toujours dans \mathcal{D}) sans pour autant que les déformations de l'élastique modifient la valeur de l'intégrale.

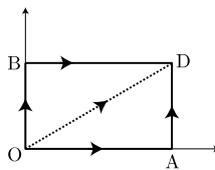


Figure 4.2: Contours utilisés pour le calcul de $\int_O^D z^2 dz$.

Traisons un exemple trivial illustrant le théorème de Cauchy. Soit la fonction $f(z) = z^2$, qui est holomorphic dans le plan entier, et calculons son intégrale le long de

plusieurs chemins reliant l'origine au point D de coordonnées (a, b) (voir fig. 4.2), sans utiliser la notion de *primitive* d'une fonction à valeurs complexes (qui sera définie plus loin⁸ – voir p. 165). Tout d'abord on a :

$$\int z^2 dz = \int (x^2 - y^2 + 2ixy)(dx + idy) = \int [(x^2 - y^2)dx - 2xydy] + i \int [(x^2 - y^2)dy + 2xydx] . \quad (4.23)$$

Prenons en premier le contour OAD. Le long de OA, on a $dy = 0$ et $y = 0$, de sorte que :

$$\int_{\text{O}}^{\text{A}} z^2 dz = \int_0^a x^2 dx = \frac{a^3}{3} . \quad (4.24)$$

Le long de AD, on a $x = a$ et $dx = 0$:

$$\int_{\text{A}}^{\text{D}} z^2 dz = \int_0^b -2aydy + i \int_0^b (a^2 - y^2)dy = -ab^2 + i(a^2b - \frac{b^3}{3}) . \quad (4.25)$$

Au total :

$$\int_{\text{OAD}} z^2 dz = \frac{a^3}{3} - ab^2 + i(a^2b - \frac{b^3}{3}) . \quad (4.26)$$

De la même façon, on trouve facilement :

$$\int_{\text{O}}^{\text{B}} z^2 dz = i \int_0^b -y^2 dy = -i\frac{b^3}{3} , \quad (4.27)$$

$$\int_{\text{B}}^{\text{D}} z^2 dz = \int_0^a (x^2 - b^2)dx + i \int_0^a 2bxdx = \frac{a^3}{3} - ab^2 + iba^2 . \quad (4.28)$$

Au total, la somme des seconds membres de (4.27) et (4.28) redonne bien (4.26). Maintenant, le long de la diagonale OD du rectangle, on a $y = \frac{b}{a}x$ et $dy = \frac{b}{a}dx$, de sorte que :

$$\int_{\text{diagonale}} z^2 dz = \int_0^a [(x^2 - \frac{b^2}{a^2}x^2)dx - 2\frac{b^2}{a^2}x^2dx] + i \int_0^a [(x^2 - \frac{b^2}{a^2}x^2)\frac{b}{a}dx + 2\frac{b}{a}x^2dx] . \quad (4.29)$$

Tous calculs faits, on retrouve encore (4.26).

Une première conséquence immédiate du théorème de Cauchy est la suivante⁹ :

⁸Une fois cette notion définie, on saura écrire :

$$\int_{\text{O}}^{\text{D}} z^2 dz = \left| \frac{z^3}{3} \right|_{\text{O}}^{\text{D}} = \frac{1}{3}(a + ib)^3 - 0^3 = \frac{1}{3}(a^3 + 3ia^2b - 3ab^2 - ib^3) , \quad (4.22)$$

qui est bien le résultat (4.26).

⁹Cette affirmation est aussi communément appelée *théorème de Cauchy*.

Si une fonction $f(z)$ est holomorphe dans un domaine simplement connexe \mathcal{D} , son intégrale $\int_C f(z) dz$ pour tout cycle C situé dans \mathcal{D} est nulle :

$$\boxed{\int_C f(z) dz = 0} \quad (4.30)$$

En effet, soit deux contours C_1 et C_2 distincts mais ayant les mêmes extrémités A et B. D'après le théorème de Cauchy, on a :

$$\int_{C_1} f(z) dz = \int_{C_2} f(z) dz . \quad (4.31)$$

À l'aide de ces deux contours, on peut faire un cycle C formé par la réunion de C_1 parcouru de A à B, et de C_2 parcouru de B à A, ce que l'on a déjà noté $-C_2$. L'intégrale le long du cycle C est la somme :

$$\int_{\text{cycle}} f(z) dz = \int_{C_1} f(z) dz + \int_{-C_2} f(z) dz , \quad (4.32)$$

mais $\int_{-C_2} f(z) dz = -\int_{C_2} f(z) dz$, de sorte que :

$$\int_{\text{cycle}} f(z) dz = \int_{C_1} f(z) dz - \int_{C_2} f(z) dz . \quad (4.33)$$

D'après (4.31), les deux intégrales au second membre sont égales, d'où le résultat exprimé par (4.30). Cette égalité peut d'ailleurs être presque considérée comme une formulation alternative du théorème de Cauchy, tant elle en est une immédiate conséquence.

Une autre démonstration donne d'ailleurs l'occasion de rappeler le théorème de Green, qui s'énonce :

$$\int_C \phi(x, y) dx + \psi(x, y) dy = \int \int_{\mathcal{D}} \left(\frac{\partial \psi}{\partial x} - \frac{\partial \phi}{\partial y} \right) dx dy , \quad (4.34)$$

où \mathcal{D} est le domaine du plan délimité par la courbe fermée simple¹⁰ C définissant le cycle, les deux fonctions ψ et ϕ étant continues ainsi que leurs dérivées premières dans \mathcal{D} et sur sa frontière C (essentiellement, cette formule s'énonce en disant que la circulation du gradient est égale au flux du rotationnel). Ceci permet de récrire (4.14) comme suit :

$$\int_{\text{cycle}} f(z) dz = \int \int_{\mathcal{D}} \left(-\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) dx dy + i \int \int_{\mathcal{D}} \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) dx dy . \quad (4.35)$$

En vertu des conditions de Cauchy, chaque intégrand est nul, d'où (4.30).

¹⁰Toute droite parallèle aux axes de coordonnées coupe C en au plus deux points.

◆ *Remarque*

Il est possible de renverser l'ordre des propositions, en commençant par démontrer le théorème de Cauchy sous la forme exprimée en (4.30) ; ceci peut se faire en quadrillant l'intérieur du contour C (voir [7], § 5.2). Cela étant fait, on en déduit immédiatement l'invariance de l'intégrale pour un arc ouvert en coupant la boucle en deux. ◆

L'articulation des idées essentielles est la suivante. Pour un domaine simplement connexe, *tout* chemin fermé peut être réduit à un point par déformation continue (toute boucle est homotope à zéro) ; quand $f(z)$ est holomorphe, la valeur de l'intégrale est indépendante du cycle, or l'intégrale sur un cycle réduit à un point est visiblement nulle (la fonction est bornée et le chemin est de longueur nulle !). Au total, l'intégrale sur tout cycle est nulle.

Par ailleurs, quand la fonction $f(z)$ est continue jusque sur la frontière de son domaine d'analyticit , le r sultat (4.30) reste vrai quand le cycle C est pr cis ment cette fronti re : gr ce   la continuit , l'int grale le long de la fronti re est  gale   l'int grale le long de tout cycle int rieur infiniment proche – pour lequel (4.30) est vrai.

En guise d'illustration, et pour montrer un premier exemple de la puissance de ce th or me, consid rons l'int grale I_n de la fonction (holomorphe) z^n ($n \in \mathbb{N}$) le long d'un cercle centr    l'origine et de rayon r . Celui-ci peut ainsi  tre param tr  suivant $z = re^{i\theta}$, r constant, $0 \leq \theta \leq 2\pi$, de sorte que l'int grale cherch e est :

$$I_n \stackrel{\text{d f}}{=} \int_{\text{Cercle de centre O et de rayon } r} z^n dz = \int (re^{i\theta})^n d(re^{i\theta}) = \int_0^{2\pi} (r^n e^{in\theta}) (ir e^{i\theta}) d\theta . \quad (4.36)$$

Au total, l'int grale vaut $ir^{n+1} \int_0^{2\pi} e^{i(n+1)\theta} d\theta = ir^{n+1} \int_0^{2\pi} [\cos(n+1)\theta + i \sin(n+1)\theta] d\theta$; cette derni re int grale est nulle $\forall n \in \mathbb{Z} \setminus \{-1\}$, puisqu'elle int gre sur un nombre entier de p riodes des fonctions p riodiques de moyenne nulle¹¹. En d finitive, l'int grale I_n est nulle $\forall n \in \mathbb{N}$, en accord avec le th or me de Cauchy. On verra dans la suite comment comprendre ce r sultat,  galement vrai pour $n < 0$ mais diff rent de -1 (voir sous-section 4.6.2).

Mais ce th or me dit infiniment plus : l'int grale est toujours nulle quand, au lieu de prendre le cercle ci-dessus, on choisit *n'importe quel contour ferm * du plan ($z^{n \in \mathbb{N}}$ est holomorphe dans \mathbb{C}). On peut bien d placer le cercle o  on veut et le d former comme on veut, le r sultat ne change pas¹². On mesurera mieux par la suite l'extraordinaire puissance du th or me de Cauchy et ses innombrables applications.

Citons une autre cons quence du th or me de Cauchy, qui permet de d finir pr cis ment la primitive d'une fonction holomorphe :

¹¹Une fois d finie la notion de primitive, on pourra aussi consid rer la primitive $\frac{1}{i(n+1)} e^{i(n+1)\theta}$, qui reprend les m mes valeurs en 0 et en 2π puisque n est entier.

¹²C'est ici vrai car $z^{n \in \mathbb{N}}$ est holomorphe dans *tout* le plan \mathbb{C} . On verra par la suite que les singularit s  ventuelles de la fonction   int grer limitent les d formations possibles du contour.

Si une fonction $f(z)$ est holomorphe dans un domaine simplement connexe \mathcal{D} , alors la fonction $F(z)$ définie par l'intégrale :

$$F(z) \stackrel{\text{déf}}{=} \int_{z_0}^z f(z') dz' \quad (4.37)$$

considérée comme fonction de sa borne supérieure est une fonction holomorphe dans \mathcal{D} , et de plus :

$$\boxed{F'(z) \stackrel{\text{déf}}{=} \frac{d}{dz} \int_{z_0}^z f(z') dz' = f(z)} \quad (4.38)$$

En effet, par la définition de la dérivée, on a :

$$F'(z) = \lim_{\xi \rightarrow 0} \frac{1}{\xi} [F(z + \xi) - F(z)] = \lim_{\xi \rightarrow 0} \frac{1}{\xi} \left[\int_{z_0}^{z+\xi} f(z') dz' - \int_{z_0}^z f(z') dz' \right] = \lim_{\xi \rightarrow 0} \frac{1}{\xi} \int_z^{z+\xi} f(z') dz' . \quad (4.39)$$

En posant $z' = z + \xi'$, la dernière intégrale est $\int_0^\xi f(z + \xi') d\xi'$, d'où :

$$F'(z) = \lim_{\xi \rightarrow 0} \frac{1}{\xi} \int_0^\xi f(z + \xi') d\xi' . \quad (4.40)$$

f étant holomorphe, c'est une fonction continue¹³ et sa dérivée existe ; si on introduit la différence $\eta(z, \xi) \stackrel{\text{déf}}{=} f(z + \xi) - f(z)$, soit $f(z + \xi) = f(z) + \eta(z, \xi)$, la fonction $\eta(z, \xi)$ est bornée car continue¹⁴ sur $[0, \xi]$ et tend vers zéro avec ξ :

$$f(z + \xi') = f(z) + \eta(z, \xi') , \quad \lim_{\xi' \rightarrow 0} \eta(z, \xi') = 0 ; \quad (4.41)$$

d'où :

$$F'(z) = \lim_{\xi \rightarrow 0} \frac{1}{\xi} \int_0^\xi [f(z) + \eta(z, \xi')] d\xi' . \quad (4.42)$$

La première intégrale est simplement $f(z) \int_0^\xi d\xi' = f(z)\xi$, d'où :

$$F'(z) = f(z) + \lim_{\xi \rightarrow 0} \frac{1}{\xi} \int_0^\xi \eta(z, \xi') d\xi' ; \quad (4.43)$$

$\eta(z, \xi')$ étant continue, soit $\max_{[0, \xi]} |\eta(z, \xi')|$ le maximum de son module ; l'intégrale est bornée en module par $\max_{[0, \xi]} |\eta(z, \xi')| \int_0^\xi d\xi' = |\xi \max_{[0, \xi]} \eta(z, \xi')|$, d'où :

$$\lim_{\xi \rightarrow 0} \left| \frac{1}{\xi} \int_0^\xi \eta(z, \xi') d\xi' \right| \leq \lim_{\xi \rightarrow 0} \left| \frac{1}{\xi} \right| |\xi \max_{[0, \xi]} \eta(z, \xi')| = \lim_{\xi \rightarrow 0} \max_{[0, \xi]} |\eta(z, \xi')| = 0 . \quad (4.44)$$

¹³Rappelons que l'existence de la dérivée $f'(z)$ en un point assure la continuité en z de $f(z)$: ce qui n'est pas (forcément) vrai pour une fonction de deux variables, l'est toujours pour une fonction holomorphe.

¹⁴ $\eta(z, \xi)$ est continue puisque $f(z)$ l'est. De plus, comme $f'(z)$ existe, on peut même affirmer que $\eta(z, \xi)$ est de la forme $\xi\varepsilon(z, \xi)$ où $\varepsilon(z, \xi)$ tend aussi vers zéro.

Cela étant, (4.43) dit alors que la limite $F'(z)$ existe et vaut $f(z)$: $F(z)$ est donc holomorphe, et sa dérivée n'est autre que $f(z)$. Comme en Analyse réelle, une fonction telle que F est appelée *primitive* de f . Noter que si l'on a utilisé en apparence la seule continuité de la fonction $f(z)$, le fait qu'elle est holomorphe est essentiel (sinon, l'écriture $\int_{z_0}^z f(z') dz'$ est ambiguë puisque le chemin n'est pas précisé).

La notion de primitive étant généralisée, il est possible d'énoncer les importants résultats suivants :

1. deux primitives F_1 et F_2 d'une même fonction f diffèrent d'une constante :

$$F_2(z) - F_1(z) = \text{constante} . \quad (4.45)$$

En effet, soit la différence $\Phi(z) = F_2(z) - F_1(z) \equiv U(x, y) + iV(x, y)$. D'après (4.38) $F_1'(z) = f(z) = F_2'(z)$, d'où :

$$\Phi'(z) = 0 . \quad (4.46)$$

Φ est une fonction holomorphe¹⁵ (c'est la différence de deux fonctions holomorphes) ; comme sa dérivée $\Phi'(z)$ est nulle, cette fonction est constante. En effet, comme on l'a vu au chapitre précédent, il existe quatre façons d'écrire la dérivée d'une fonction holomorphe, et notamment :

$$\Phi'(z) = \frac{\partial U}{\partial x} + i \frac{\partial V}{\partial x} = \frac{\partial V}{\partial y} - i \frac{\partial U}{\partial y} . \quad (4.47)$$

Compte tenu de (4.46), il vient :

$$\frac{\partial U}{\partial x} + i \frac{\partial V}{\partial x} = 0 \quad \text{et} \quad \frac{\partial V}{\partial y} - i \frac{\partial U}{\partial y} = 0 , \quad (4.48)$$

soit :

$$\frac{\partial U}{\partial x} = 0 , \quad \frac{\partial U}{\partial y} = 0 , \quad \frac{\partial V}{\partial x} = 0 , \quad \frac{\partial V}{\partial y} = 0 , \quad (4.49)$$

d'où l'on conclut que les fonctions U et V sont constantes.

2. Si $F(z)$ est une primitive quelconque d'une fonction holomorphe $f(z)$, alors l'intégrale entre deux points n'est autre que la variation de cette primitive :

$$\boxed{\int_{z_0}^z f(z') dz' = F(z) - F(z_0)} \quad (4.50)$$

tout comme l'intégrale curviligne de la différentielle d'une fonction est égale à la variation de cette fonction entre les deux extrémités. En effet, soit $F_{z_0}(z)$ la primitive nulle quand $z = z_0$:

$$F_{z_0}(z) = \int_{z_0}^z f(z') dz' , \quad F_{z_0}(z_0) = 0 . \quad (4.51)$$

¹⁵En particulier, Φ ne dépend que de z , pas de z^* (voir la remarque 8 p.139 dans le chapitre 3 au sujet d'une fonction *antiholomorphe*.)

Soit $F(z)$ une primitive *quelconque* de f . D'après (4.45), $F(z)$ et $F_{z_0}(z)$ diffèrent d'une constante, K :

$$F(z) = F_{z_0}(z) + K \equiv \int_{z_0}^z f(z') dz' + K ; \quad (4.52)$$

l'intégrale est nulle si $z = z_0$, d'où $F(z_0) = K$ soit $F(z) = \int_{z_0}^z f(z') dz' + F(z_0)$ et la relation (4.50). Comme toujours, la constante additive arbitraire incluse dans la définition de toute primitive $F(z)$ disparaît dans la différence au second membre de (4.50).

Un autre théorème utile est le suivant : soit $f(z, \xi)$ une fonction holomorphe dans \mathcal{D} et continue sur sa frontière C , ainsi que sa dérivée partielle $\partial_z f(z, \xi)$. Alors $F(z) \stackrel{\text{def}}{=} \int_C f(z, \xi) d\xi$ est holomorphe dans \mathcal{D} et sa dérivée est :

$$F'(z) = \int_C \frac{\partial f}{\partial z} d\xi \quad (4.53)$$

La démonstration procède très classiquement en écrivant les parties réelles U et V de F , et en les exprimant formellement sous la forme d'intégrales curvilignes sur le chemin convenablement paramétré. Le théorème vaut évidemment si f est holomorphe.

4.3 Généralisation au cas d'un domaine multiplement connexe

Quand la fonction n'est holomorphe que dans un domaine *multiplement* connexe, le théorème de Cauchy n'est pas applicable. De fait, en pareil cas, il existe des chemins ne donnant pas toujours la même valeur pour l'intégrale¹⁶.

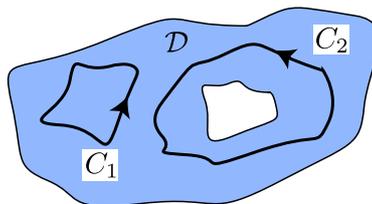


Figure 4.3: Le contour C_1 est homotope à zéro, C_2 ne l'est pas. Le théorème de Cauchy s'applique à tous les contours du type C_1 , mais pas à ceux de la classe C_2 .

¹⁶Le théorème reste évidemment vrai pour tous les contours homotopes à zéro situés dans une composante simplement connexe d'un domaine multiplement connexe (contours du type C_1 sur la fig. 4.3).

Le plus simple est de s'en convaincre sur un exemple¹⁷. Soit la fonction $z \rightarrow f(z) = \frac{1}{z}$; elle est holomorphe $\forall z \neq 0$ puisque la limite :

$$\lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{1}{z+h} - \frac{1}{z} \right) \quad (4.54)$$

existe¹⁸ $\forall z \neq 0$ (et vaut $-\frac{1}{z^2}$). Holomorphe dans le plan privé de l'origine $\mathbb{C}^* \stackrel{\text{déf}}{=} \mathbb{C} \setminus \{0\}$, elle est holomorphe dans un domaine délimité par deux contours fermés autour de l'origine ne se recoupant pas, que l'on peut toujours prendre comme deux cercles centrés en O et de rayons r et R ($r < R$) – délimitant une *couronne* circulaire, l'exemple même de domaine multiplement connexe.

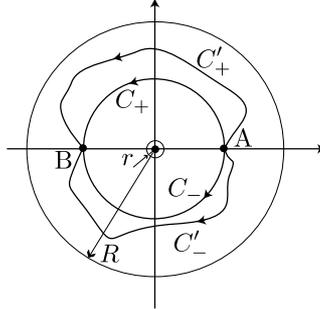


Figure 4.4: Les deux cercles concentriques de rayons respectifs r et R ($r < R$) définissent une couronne circulaire. Pour la fonction $\frac{1}{z}$, les intégrales le long des contours C_+ (resp. C_-) et C'_+ (resp. C'_-) sont égales, celles le long de C_+ et C_- sont différentes.

Comme premier contour, soit le demi-cercle supérieur C_+ de rayon ρ ($r < \rho < R$) parcouru dans le sens positif de A vers B, qui donne la même valeur à l'intégrale que tout contour C'_+ de mêmes extrémités et lui aussi situé dans la demi-couronne supérieure (voir fig. 4.4), puisqu'ils peuvent être superposés par une déformation continue les laissant tout entiers dans le domaine où $\frac{1}{z}$ est holomorphe. Le cercle de rayon ρ est décrit par le point dont l'affixe est :

$$z = \rho e^{i\theta} \quad , \quad 0 \leq \theta \leq 2\pi \quad . \quad (4.55)$$

Pour le demi-cercle supérieur, θ varie de 0 à π ; l'intégrale le long de C'_+ , égale à celle le long de C_+ , s'écrit :

$$\int_{C'_+} \frac{1}{z} dz = \int_{C_+} \frac{1}{z} dz = \int_0^\pi \frac{1}{\rho e^{i\theta}} d(\rho e^{i\theta}) = \int_0^\pi \frac{1}{\rho e^{i\theta}} i \rho e^{i\theta} d\theta = i \int_0^\pi d\theta = i\pi \quad . \quad (4.56)$$

Prenons maintenant des contours dans la demi-couronne inférieure, C'_- et le demi-cercle inférieur C_- , ayant toujours les mêmes extrémités mais forcément décrits dans le

¹⁷Un seul contre-exemple suffit.

¹⁸ $z = 0$ est donc un point remarquable pour $\frac{1}{z}$ (on s'en sera déjà douté !). C'est un point *singulier* et constitue une *singularité* de cette fonction. Les singularités d'une fonction $f(z)$ seront précisément définies dans le chapitre 5, et y feront l'objet d'une classification.

sens *néгатif*, afin que A reste le point de départ et B celui d'arrivée ; maintenant, θ varie de 0 à $-\pi$:

$$\int_{C'_-} \frac{1}{z} dz = \int_{C_-} \frac{1}{z} dz = \int_0^{-\pi} \frac{1}{\rho e^{i\theta}} d(\rho e^{i\theta}) = \int_0^{-\pi} \frac{1}{\rho e^{i\theta}} i\rho e^{i\theta} d\theta = i \int_0^{-\pi} d\theta = -i\pi . \quad (4.57)$$

Ainsi, pour deux contours ayant pourtant les mêmes extrémités A et B, l'intégrale prend deux valeurs différentes. En particulier, l'intégrale sur le cycle $C_+ \cup (-C_-)$ n'est pas nulle et vaut $i\pi - (-i\pi) = 2i\pi$; en désignant par C le cercle de rayon ρ (quelconque) centré à l'origine, on a donc le résultat très important :

$$\int_C \frac{1}{z} dz = 2i\pi . \quad (4.58)$$

Il est bien clair que le contour ne peut être ici réduit à un point tout en restant dans le domaine d'holomorphie de la fonction $\frac{1}{z} \equiv z^{-1}$.

D'après le théorème de Cauchy, les intégrales ne changent pas si on remplace C_+ par C'_+ et C_- par C'_- ; en raboutant C'_+ et $-C'_-$, on construit un contour fermé (cycle) de forme arbitraire (mais tout entier dans la couronne et ceinturant l'origine). D'où le résultat *fondamental*¹⁹ :

$$\boxed{\int_{\Gamma} \frac{1}{z} dz = 2i\pi} \quad (4.60)$$

où Γ est *n'importe quel contour fermé entourant une fois et une seule*²⁰ l'origine et décrit dans le sens inverse des aiguilles d'une montre. Ce résultat est évidemment à rapprocher du fait qu'une primitive de $\frac{1}{z}$ est $\ln z$ et que la variation du logarithme sur une boucle dépend de la position de l'origine par rapport à cette boucle : nulle si l'origine est à l'extérieur, égale $2i\pi$ si elle est à l'intérieur.

L'idée essentielle à retenir est donc celle-ci : le domaine d'analyticit  \mathcal{D} étant précisé et les extrémités du chemin étant fixées (et situées dans \mathcal{D} !), ce dernier peut être déformé continûment *ad libitum* à condition de rester tout entier dans \mathcal{D} . Il est manifeste que dans l'exemple ci-dessus, le demi-cercle inférieur ne peut pas être superposé au demi-cercle supérieur par une déformation continue qui le laisse dans la couronne comprise entre les deux cercles de rayon r et R . Le point capital est la notion de déformation *continue* du contour dans le plan en le laissant tout entier dans une partie connexe, une opération dont le sens intuitif doit être clair. Autrement dit, toutes les déformations de l'élastique le laissant en contact avec le plan (les frontières des parties simplement connexes étant des murs infranchissables), laissent invariante la valeur de l'intégrale : dans l'exemple ci-dessus, il ne faut pas faire "décoller" le chemin pour lui permettre

¹⁹Clairément, ceci se généralise immédiatement en :

$$\int_{\Gamma_{z_0}} \frac{1}{z - z_0} dz = 2i\pi , \quad (4.59)$$

où Γ_{z_0} est une boucle autour de z_0 .

²⁰D'après le calcul précédent, il est bien évident que si on tourne deux fois, on trouve $4i\pi$, etc.

de sauter par dessus le cercle intérieur (de rayon aussi petit que l'on veut d'ailleurs) contenant l'origine. On se convaincra peu à peu que le calcul intégral avec des fonctions holomorphes est essentiellement un *jeu de l'élastique* d'un genre particulier.

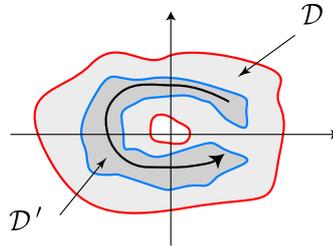


Figure 4.5: \mathcal{D} n'est pas simplement connexe, mais \mathcal{D}' l'est. $z^{-|n|}$ est holomorphe dans \mathcal{D}' .

Par référence à la fonction monôme $z \rightarrow z^n$ où n est cette fois un entier quelconque (par forcément positif), il apparaît que le cas $n = -1$ est très particulier. Reprenons le même calcul que précédemment avec n entier négatif mais différent de -1 . La fonction est holomorphe dans la couronne $r < |z| < R$; l'intégrale le long de C_+ est :

$$\int_{C'_+} \frac{1}{z^{|n|}} dz = \int_{C_+} \frac{1}{z^{|n|}} dz = \int_0^\pi \frac{1}{(\rho e^{i\theta})^{|n|}} d(\rho e^{i\theta}) = \frac{\rho^{n+1}}{n+1} [e^{i(n+1)\pi} - 1], \quad (4.61)$$

celle le long de C'_- vaut $\frac{\rho^{n+1}}{n+1} [e^{-i(n+1)\pi} - 1]$. Au total, l'intégrale sur le cycle $C'_+ \cup (-C'_-)$ vaut zéro, quoique z^n n'est pas holomorphe dans le domaine simplement connexe \mathcal{D} .

En réalité, la couronne $r < |z| < R$ étant donnée, on peut lui trouver des composantes simplement connexes où la fonction est holomorphe (c'est le cas de \mathcal{D}' , voir fig. 4.5), et toute la question est de savoir si la primitive de z^n est multiforme ou non. Elle ne l'est pas si $n = -1$, elle l'est $\forall n \in \mathbb{Z} \setminus \{-1\}$. Les mêmes arguments permettent d'étendre la notion de primitive au cas d'une fonction holomorphe dans un domaine non simplement connexe, à condition de considérer une composante simplement connexe de ce dernier.

4.4 Formule de Cauchy

La formule de Cauchy est, avec le théorème de Cauchy, l'un des résultats majeurs de ce chapitre. Elle s'énonce comme suit :

Soit une fonction $f(z)$ holomorphe dans un domaine simplement connexe \mathcal{D} . Alors $\forall z \in \mathcal{D}$, on a :

$$f(z) = \frac{1}{2i\pi} \int_C \frac{f(\xi)}{\xi - z} d\xi \quad (4.62)$$

où C est un chemin contenu dans D et tournant une fois autour de z dans le sens positif²¹.

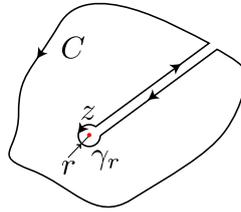


Figure 4.6: Contour C utilisé pour établir la formule de Cauchy (4.62).

Pour démontrer cette formule, considérons l'intégrale de $f(z)$ sur le contour C représenté sur la figure 4.6 : deux segments identiques parcourus en sens inverse, une grande boucle C et un petit cercle γ_r de rayon r autour du point d'affixe z ; la fonction $\frac{f(\xi)}{\xi-z}$ étant holomorphe dans le domaine privé du point z , l'intégrale est nulle d'après le théorème de Cauchy. La décomposant en ses différentes contributions et omettant d'emblée les deux segments qui se compensent mutuellement, on a :

$$\int_C \frac{f(\xi)}{\xi-z} d\xi + \int_{-\gamma_r} \frac{f(\xi)}{\xi-z} d\xi = 0 \iff \int_C \frac{f(\xi)}{\xi-z} d\xi = \int_{\gamma_r} \frac{f(\xi)}{\xi-z} d\xi . \quad (4.63)$$

Posant $\xi = z + re^{i\theta}$ dans le membre de droite, il s'écrit $\int_0^{2\pi} \frac{f(z+re^{i\theta})}{re^{i\theta}} d(re^{i\theta})$, d'où l'égalité :

$$\int_C \frac{f(\xi)}{\xi-z} d\xi = i \int_0^{2\pi} f(z + re^{i\theta}) d\theta . \quad (4.64)$$

La fonction $f(\xi)$ étant holomorphe, elle est continue : on peut donc écrire :

$$f(z + re^{i\theta}) = f(z) + \eta(r, z, \theta) , \quad \lim_{r \rightarrow 0} \eta(r, z, \theta) = 0 . \quad (4.65)$$

Le report dans l'intégrale au second membre de (4.64) transforme cette égalité en :

$$\int_C \frac{f(\xi)}{\xi-z} d\xi = 2i\pi f(z) + \int_0^{2\pi} \eta(r, z, \theta) d\theta . \quad (4.66)$$

Le premier membre est visiblement indépendant de la valeur r du petit cercle autour de z : il vaut ce qu'il vaut, que l'on complète ou non la grande boucle pour former le contour C ; de même, le premier terme au second membre ne dépend aucunement de r . Il en résulte que l'intégrale sur $\eta(r, z, \theta)$ est également indépendante de r et est, en particulier, égale à sa valeur dans la limite $r \rightarrow 0$; celle-ci est nulle puisque $\eta(r, z, \theta)$ tend vers zéro. La limite du second membre est donc $2i\pi f(z)$, ce qui établit²² la formule de Cauchy (4.62) – formule assez extraordinaire : elle montre que les valeurs de la fonction

²¹En particulier, si f est continue sur la frontière ∂D de D , on peut prendre $C = \partial D$.

²²Pour une autre démonstration, très instructive, voir [7], § 5.21.

$f(z)$ dans son domaine d'analyticité ne dépendent que de ses valeurs sur la grande boucle C , qui peut d'ailleurs être la frontière $\partial\mathcal{D}$ de \mathcal{D} (si f y est continue).

On verra que cette formule a de nombreuses conséquences très importantes. Une toute première est ce que l'on appelle *la formule de la moyenne*, qui s'exprime comme suit :

$$f(z) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} f(z + re^{i\theta}) d\theta \quad (4.67)$$

Pour établir cette formule, on peut repartir de la formule de Cauchy (4.62), prenant pour C le cercle γ_r de centre z et de rayon r , où $\xi = z + re^{i\theta}$; il vient :

$$f(z) = \frac{1}{2i\pi} \int_{\gamma_r} \frac{f(z + re^{i\theta})}{re^{i\theta}} ire^{i\theta} d\theta = \frac{1}{2\pi} \int_{-\pi}^{+\pi} f(z + re^{i\theta}) d\theta, \quad (4.68)$$

d'où la formule de la moyenne (4.67). Comme $e^{i\theta}$ est 2π -périodique, ce résultat reste vrai pour n'importe quel intervalle d'amplitude 2π ; θ_0 désignant un angle quelconque, on a donc tout autant :

$$f(z) = \frac{1}{2\pi} \int_{\theta_0}^{\theta_0+2\pi} f(z + re^{i\theta}) d\theta. \quad (4.69)$$

Le second membre de (4.67) peut bien être considéré comme une *moyenne* : c'est la moyenne des valeurs de la fonction f sur le cercle de centre z , l'angle θ étant tiré au hasard uniformément²³ sur l'intervalle $[-\pi, +\pi[$. On obtient ainsi un résultat assez extraordinaire : la valeur de f au centre du cercle est la moyenne des valeurs sur le cercle, et ce quel qu'en soit le rayon²⁴ !

En écrivant les choses comme en théorie des probabilités, et désignant par $P(\theta)$ la densité de probabilité (ici uniforme, c'est-à-dire "plate"), on a :

$$\text{moyenne de } f(z + re^{i\theta}) = \int_{-\pi}^{+\pi} P(\theta) f(z + re^{i\theta}) d\theta = \int_{-\pi}^{+\pi} \frac{1}{2\pi} f(z + re^{i\theta}) d\theta, \quad (4.70)$$

d'où une autre écriture de ce résultat :

$$f(z) = \langle f(z + re^{i\theta}) \rangle \quad (4.71)$$

Comme toujours, une telle formule remarquable permet d'établir des résultats non triviaux. Par exemple, prenons $f(z) = e^z$; alors (4.68) donne :

$$e^z = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^z e^{re^{i\theta}} d\theta \iff \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{re^{i\theta}} d\theta = 1 ; \quad (4.72)$$

²³*uniformément réparti* signifie que la densité de probabilité $P(\theta)$ est *constante* sur cet intervalle ; la normalisation ("somme des probabilités = 1") s'écrit $\int_{-\pi}^{+\pi} P(\theta) d\theta = 1$ et donne $P(\theta) = \frac{1}{2\pi}$.

²⁴Ce résultat deviendra *a posteriori* évident par mise en relation avec le développement de Taylor (infini) de $f(z)$ laquelle, en tant que fonction holomorphe, est infiniment dérivable, comme on le verra ci-dessous. En admettant la convergence uniforme de ce développement de Taylor (on démontrera au chapitre 5 qu'il en est bien ainsi), on obtient une série de terme général $\int_0^{2\pi} (re^{i\theta})^n \frac{1}{n!} f^{(n)}(0) d\theta$, $n \in \mathbb{N}$; toutes les intégrations sur θ donnent zéro, sauf pour le terme $n = 0$.

cette dernière égalité n'a rien d'évident à première vue²⁵, pas plus d'ailleurs que les deux égalités qui en découlent immédiatement en séparant les parties réelle et imaginaire.

Il y a bien plus ; la formule de la moyenne lève le voile sur un résultat autrement plus profond, appelé *Principe du maximum* : le module d'une fonction holomorphe dans \mathcal{D} et continue sur la frontière prend sa valeur maximum quelque part sur la frontière de \mathcal{D} . Avant de donner le fil de l'argument permettant de comprendre intuitivement pourquoi il en est ainsi (pour plusieurs démonstrations, voir [6], § 5.1), rappelons un lemme utile.

Soit $\phi(x)$, $x \in \mathbb{R}$, une fonction continue et telle que $\phi(x) \leq A$; si $\frac{1}{b-a} \int_a^b \phi(x) dx \geq A$, alors la fonction $\phi(x)$ est constante et égale à A . En effet, supposons que pour une valeur x quelconque $\phi(x) < A$; alors il existe un $\varepsilon > 0$ et un intervalle $[x - \delta, x + \delta]$ où $\phi(x) \leq A - \varepsilon$. Dans ces conditions, séparant l'intégrale en deux morceaux (sur $[x - \delta, x + \delta]$ et son complémentaire) :

$$\int_a^b \phi(x) dx \leq 2\delta(A - \varepsilon) + (b - a - 2\delta)A = (b - a)A - 2\delta\varepsilon, \quad (4.73)$$

d'où $\frac{1}{b-a} \int_a^b \phi(x) dx \leq A - \frac{2\delta\varepsilon}{b-a}$, en contradiction avec l'hypothèse. Il ne peut donc exister aucun $x \in [a, b]$ pour lequel $\phi(x) < A$, permettant d'affirmer que $\phi(x) \geq A$ quel que soit x . Comme on a aussi $\phi(x) \leq A$ pour tout x , la conclusion est que $\phi(x)$ est constante et égale à A .

Soit maintenant une fonction $f(z)$ holomorphe dans \mathcal{D} simplement connexe, et C un contour fermé situé dans \mathcal{D} . $|f(z)|$ possédant un maximum M quelque part sur C , $|f(z)| \leq M$, on va voir que $|f(z)| < M$ quel que soit z à l'intérieur de C . Pour cela, supposons au contraire qu'il existe z_0 dans C tel que $|f(z_0)|$ est supérieur ou égal à toute autre valeur de $|f(z)|$; C_0 étant un cercle de centre z_0 contenu dans \mathcal{D} , la formule de Cauchy est $f(z_0) = \frac{1}{2i\pi} \int_{C_0} \frac{f(z)}{z - z_0} dz$. Posant $f(z) = \rho e^{i\phi} f(z_0)$, elle se récrit comme :

$$1 = \frac{1}{2i\pi} \int_{C_0} \frac{\rho e^{i\phi}}{z - z_0} dz = \frac{1}{2\pi} \int_0^{2\pi} \rho e^{i\phi} d\theta, \quad (4.74)$$

ρ et ϕ étant des fonctions de θ . Prenant les modules, $1 \leq \frac{1}{2\pi} \int_0^{2\pi} \rho d\theta$. Par hypothèse, $\rho \stackrel{\text{d'ef}}{=} \frac{|f(z)|}{|f(z_0)|} \leq 1$; en vertu du lemme ci-dessus, on en déduit $\rho = 1$. De surcroît, prenant la valeur réelle de (4.74), il vient $1 = \frac{1}{2\pi} \int_0^{2\pi} \cos \phi d\theta$; comme $\cos \phi \leq 1$, le même lemme permet de conclure que $\cos \phi = 1$, d'où au total $f(z) = f(z_0)$, quel que soit z . En définitive, supposer qu'il existe un point quelconque z_0 où le module de $f(z)$ prend une valeur supérieure ou égale à toute autre de ses valeurs conduit à la conclusion que $|f(z)|$ est la fonction constante. Si l'on exclut ce cas sans intérêt, en tout point quelconque z_0 à l'intérieur de C , $|f(z_0)|$ ne peut être supérieur ou égal à toutes ses autres valeurs, et en particulier à M , d'où $|f(z_0)| < M$, maximum du module sur le cercle.

²⁵Sauf si on développe l'intégrand en série, on invoque la convergence uniforme et on utilise le fait que $\int_0^{2\pi} e^{in\theta} d\theta = 2\pi\delta_{n,0}$.

Récrivons maintenant la formule de la moyenne (4.67) en notant z_0 le centre du cercle C_0 d'équation $z_0 + r_0 e^{i\theta}$:

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + r_0 e^{i\theta}) d\theta ; \quad (4.75)$$

on a $|f(z_0)| = \frac{1}{2\pi} \left| \int_0^{2\pi} f(z_0 + r_0 e^{i\theta}) d\theta \right| \leq \frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + r_0 e^{i\theta})| d\theta$. En appelant M_1 le maximum de $|f(z)|$ sur le cercle C_0 , on a donc :

$$|f(z_0)| \leq \frac{1}{2\pi} \int_0^{2\pi} M_1 d\theta = \frac{1}{2\pi} M_1 \times 2\pi = M_1 ; \quad (4.76)$$

ainsi, le module de $f(z_0)$ est plus petit (ou égal) à sa plus grande valeur sur le cercle. Supposons que le maximum M_1 soit atteint au point z_1 ; on peut alors introduire le cercle C_1 de centre z_1 et de rayon r_1 et, utilisant la formule de la moyenne, affirmer que le module de f en z_1 est plus petit que la valeur maximum M_2 atteinte quelque part (en z_2) sur le cercle C_1 , sur sa portion de circonférence *extérieure* à C_0 , puisqu'en tout point intérieur à C_0 , $|f(z)| < M_1$. Répéter ce processus met en évidence une famille de cercles et une suite monotone croissante de bornes supérieures M_n pour $|f(z)|$. Jusqu'où peut-on aller ? Visiblement jusqu'à la frontière du domaine où f est holomorphe, et en supposant $f(z)$ continue sur cette frontière, d'où le Principe du maximum.

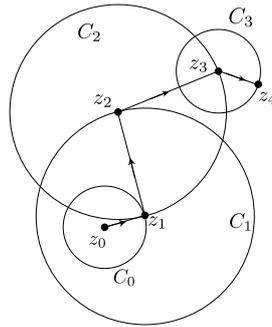


Figure 4.7: Construction de la famille de cercles C_n de centres z_n ; le maximum relatif M_n est sur le cercle C_{n-1} (voir le texte).

Cela étant compris, on peut de plus affirmer que si $f(z)$ ne s'annule pas dans \mathcal{D} , alors le *minimum* de $|f(z)|$ est également sur la frontière de \mathcal{D} . En effet, si $f(z) \neq 0 \forall z \in \mathcal{D}$, la fonction $\frac{1}{f(z)}$ est holomorphe dans \mathcal{D} ; le maximum de son module est donc sur la frontière, et c'est aussi le minimum de $|f(z)|$.

Les mêmes propriétés valent aussi pour les parties réelle et imaginaire de f : il suffit de raisonner avec les fonctions holomorphes $F(z) \stackrel{\text{d'éf}}{=} e^{f(z)}$ et $G(z) \stackrel{\text{d'éf}}{=} e^{if(z)}$. L'inexistence d'extremum pour $u(x, y)$ et $v(x, y)$ a déjà été relevé en tant que conséquence des conditions de Cauchy - Riemann, qui entraînent $\Delta u = \Delta v = 0$.

4.5 Dérivées d'ordre supérieur

La propriété d'analyticité d'une fonction $f(z)$ fait de celle-ci un objet très robuste, et a des conséquences innombrables, dont les plus importantes viennent d'être données. Une autre propriété remarquable peut être énoncée comme suit²⁶ :

Si une fonction $f(z)$ est holomorphe dans un domaine simplement connexe \mathcal{D} et continue sur la frontière²⁷ $\partial\mathcal{D}$, elle possède en chaque point de \mathcal{D} des dérivées de tous les ordres ; la dérivée d'ordre n est donnée par la formule :

$$f^{(n)}(z) = \frac{n!}{2i\pi} \int_{\partial\mathcal{D}} \frac{f(\xi)}{(\xi - z)^{n+1}} d\xi \quad (4.77)$$

La propriété pour une fonction holomorphe d'être *infinitement* dérivable a été suspectée dès le chapitre 3, une fois établies les conditions de Cauchy - Riemann. L'existence de $f^{(n)}$ va être prouvée en démontrant la formule (4.77), en déroulant un raisonnement par récurrence.

Prenons d'abord $n = 1$; par la définition de la dérivée :

$$f'(z) = \lim_{h \rightarrow 0} \frac{f(z+h) - f(z)}{h} ; \quad (4.78)$$

à l'aide de la formule de Cauchy, le second membre s'écrit successivement :

$$\lim_{h \rightarrow 0} \frac{1}{h} \frac{1}{2i\pi} \int_{\partial\mathcal{D}} \left[\frac{f(\xi)}{\xi - (z+h)} - \frac{f(\xi)}{\xi - z} \right] d\xi = \frac{1}{2i\pi} \lim_{h \rightarrow 0} \int_{\partial\mathcal{D}} \frac{f(\xi)}{\xi - z} \frac{1}{\xi - z - h} d\xi . \quad (4.79)$$

Le terme $\frac{1}{\xi - z - h}$ tend uniformément vers $\frac{1}{\xi - z}$ quand $h \rightarrow 0$. En effet, z et $|h|$ étant fixés, $z + h$ décrit le cercle de centre z et de rayon $|h|$. Soit ξ_D l'affixe du point de la frontière le plus proche de ce cercle. Alors, pour tout point ξ de la frontière, on a $|\xi - z| \geq |\xi_D - z| \equiv AD$ et $|\xi - (z+h)| \geq |\xi_D - z| - |h| \equiv BD$. Il en résulte que :

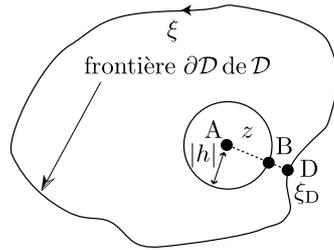
$$\left| \frac{1}{\xi - (z+h)} - \frac{1}{\xi - z} \right| = \left| \frac{h}{[\xi - (z+h)](\xi - z)} \right| \leq \frac{h}{(|\xi_D - z| - |h|)|\xi_D - z|} . \quad (4.80)$$

Cette dernière quantité est une borne indépendante de ξ , et tend vers zéro avec h , ce qui établit la convergence uniforme en ξ de la quantité au premier membre. On peut donc

²⁶Pour se souvenir des détails de (4.77), penser à l'homogénéité, en imaginant que z n'est pas un nombre pur.

²⁷ $\bar{\mathcal{D}}$ désigne le domaine fermé formé par \mathcal{D} et sa frontière $\partial\mathcal{D}$.

²⁷Référence est faite à la frontière pour obtenir le résultat dans le domaine le plus vaste possible. Le même résultat vaut pour tout contour fermé C entourant le point d'affixe z , l'essentiel étant que C contienne z en son intérieur.



$$|\xi - (z + h)| \geq BD = |\xi_D - z| - |h|$$

Figure 4.8: ξ_D est l'affixe de D, point de la frontière $\partial\mathcal{D}$ de \mathcal{D} le plus proche du cercle de centre z et de rayon $|h|$.

intervertir limite et intégration (la limite de l'intégrale est égale à l'intégrale de la limite, $[\lim, \int] = 0$) ; prenant dans (4.79) la limite sous l'intégrale :

$$f'(z) = \frac{1}{2i\pi} \int_{\partial\mathcal{D}} \frac{f(\xi)}{(\xi - z)^2} d\xi, \quad (4.81)$$

établissant la formule annoncée pour $n = 1$, ce qui constitue la “condition initiale” du raisonnement par récurrence.

Supposons maintenant que $f^{(n)}$ est donnée par l'expression (4.77) ; alors, par définition de la dérivée :

$$f^{(n+1)}(z) = \lim_{h \rightarrow 0} \frac{f^{(n)}(z + h) - f^{(n)}(z)}{h}, \quad (4.82)$$

soit, par hypothèse :

$$f^{(n+1)}(z) = \lim_{h \rightarrow 0} \frac{1}{h} \frac{n!}{2i\pi} \int_{\partial\mathcal{D}} f(\xi) \left[\frac{1}{[\xi - (z + h)]^{n+1}} - \frac{1}{(\xi - z)^{n+1}} \right] d\xi. \quad (4.83)$$

En réduisant au même dénominateur le grand crochet, et en utilisant la formule du binôme pour arranger le numérateur qui en résulte, il vient :

$$f^{(n+1)}(z) = \frac{n!}{2i\pi} \lim_{h \rightarrow 0} \frac{1}{h} \int_{\partial\mathcal{D}} f(\xi) \frac{(n+1)h(\xi - z)^n + \mathcal{O}(h^2)}{(\xi - z - h)^{n+1}(\xi - z)^{n+1}} d\xi = \frac{(n+1)!}{2i\pi} \int_{\partial\mathcal{D}} \frac{f(\xi)}{(\xi - z)^{n+2}} d\xi, \quad (4.84)$$

où la propriété de convergence uniforme (en ξ) a été de nouveau invoquée pour légitimer la dernière égalité. Ainsi, si la relation (4.77) est vraie pour n , elle l'est aussi pour $n + 1$; comme elle est démontrée pour $n = 1$, elle est vraie pour tout n fini, ce qui achève la démonstration par récurrence. On obtient ainsi le résultat très important : dès qu'une fonction est holomorphe (c'est-à-dire une fois dérivable), elle est *infinitement* dérivable, et on dispose maintenant en plus d'une expression²⁸ intégrale pour la dérivée d'ordre

²⁸Cette expression peut d'ailleurs visiblement être considérée comme une généralisation de la formule de Cauchy (4.62), cette dernière étant vue comme (4.77) avec $n = 0$.

n . Une autre façon d'exprimer la même propriété est de dire que toute dérivée d'une fonction holomorphe étant infiniment dérivable est holomorphe.

En bout de course on peut remarquer que les expressions (4.81) et plus généralement (4.77) des dérivées s'obtiennent directement par dérivations sous le signe intégrale : les démonstrations ci-dessus montrent que de telles opérations sont légitimes, grâce à la convergence uniforme. On aurait pu aussi invoquer le théorème énoncé p. 167, exprimé par l'équation (4.53).

La formule (4.77) permet de démontrer d'importantes inégalités, appelées inégalités de Cauchy, énonçant des bornes supérieures pour toutes les dérivées d'une fonction holomorphe. Soit M le maximum du module d'une fonction f holomorphe dans un domaine \mathcal{D} , et soit d_{\min} la plus petite distance de z à la frontière $\partial\mathcal{D}$ de \mathcal{D} (c'est-à-dire $d_{\min} = \min_{\xi \in \partial\mathcal{D}} |\xi - z|$). En notant L la longueur de $\partial\mathcal{D}$, (4.77) permet d'écrire :

$$|f^{(n)}(z)| = \frac{n!}{2\pi} \left| \int_{\partial\mathcal{D}} \frac{f(\xi)}{(\xi - z)^{n+1}} d\xi \right| \leq \frac{n!}{2\pi} \left| \int_{\partial\mathcal{D}} \frac{M}{d_{\min}^{n+1}} d\xi \right| = \frac{n!}{2\pi} \frac{ML}{d_{\min}^{n+1}}. \quad (4.85)$$

En particulier, si $f(z)$ est holomorphe (analytique) dans un disque $|z - z_0| < R$ et en choisissant ce dernier comme domaine \mathcal{D} , alors la plus petite distance est $d_{\min} = R$ et $L = 2\pi R$. Dans ces conditions, (4.85) devient :

$$\boxed{|f^{(n)}(z)| \leq \frac{n!M}{R^n} \quad (n \in \mathbb{N})}. \quad (4.86)$$

Ces inégalités permettent de démontrer l'important théorème de Liouville²⁹ :

Une fonction holomorphe pour toute valeur finie de z , et bornée, est une constante.

L'énoncé signifie que $\exists M, |f(z)| < M$ quel que soit z , et que $|f(z)|$ est borné quand $|z| \rightarrow +\infty$. En termes plus elliptiques : les seules fonctions entières bornées sont des constantes³⁰.

Soit z un complexe donné ; si $f(z)$ est bornée et entière, l'inégalité (4.86) avec $n = 1$ donne :

$$|f'(z)| \leq \frac{M}{R}, \quad (4.87)$$

z étant le centre du cercle de rayon R , et M une borne supérieure pour $|f(z')|$ quel que soit z' dans \mathbb{C} . Comme f est entière, on peut choisir R aussi grand que l'on veut, de sorte le premier membre est en fait nul : $f'(z) = 0$; comme z est quelconque, on a $f'(z) \equiv 0$.

²⁹aussi appelé *théorème de Cauchy - Liouville*, puisque Cauchy est le premier à l'avoir énoncé.

³⁰Il faut bien ajouter l'hypothèse que f est bornée. En effet, toute fonction dérivable (holomorphe) est continue ; si c'est sur un fermé, elle est donc bornée. Une fonction entière est holomorphe, donc continue, dans tout le plan \mathbb{C} , qui n'est pas fermé. Une fonction entière n'est donc pas forcément bornée.

Maintenant, par définition de la primitive d'une fonction holomorphe (ce qu'est $f'(z)$), on a :

$$f(z) - f(z_0) = \int_{z_0}^z f'(z') dz' \quad \forall z, \quad (4.88)$$

et comme $f'(z) = 0$, il vient :

$$f(z) - f(z_0) = 0 \iff f(z) = f(z_0) \quad \forall z, \quad (4.89)$$

qui montre que la fonction $f(z)$ prend la même valeur partout. En d'autres termes, une fonction entière qui n'est pas constante doit diverger *quelque part* quand $|z| \rightarrow +\infty$ (mais pas forcément quel que soit l'argument de z) ; par exemple e^z est une fonction entière (qui n'est pas constante !) qui diverge si $|z| \rightarrow +\infty$, $\Re z > 0$, mais qui tend vers zéro si $|z| \rightarrow +\infty$, $\Re z < 0$. Autre exemple, dans le même ordre d'idées³¹ : si une fonction $f(z)$ est bornée par Mz^n quand $|z|$ est assez grand, $f(z)$ est un polynôme de degré $\leq n$. Dans ce contexte, citons le *petit théorème de Picard*, loin d'être petit puisqu'il va bien plus loin : *une fonction $f(z)$ entière non constante prend toutes les valeurs complexes, sauf peut-être une* ; autrement dit³², toute fonction entière non bornée a pour image \mathbb{C} , à l'exception possible d'un point. Exemple de cette exception : la fonction exponentielle qui ne prend pas la valeur zéro dans \mathbb{C} .

Il est maintenant possible d'établir la réciproque du théorème de Cauchy, appelée théorème de Morera, qui s'énonce comme suit :

Si une fonction $f(z)$ est continue dans un domaine simplement connexe \mathcal{D} et si l'intégrale $\int_C f(z) dz$ pour tout cycle C situé dans \mathcal{D} est nulle, alors $f(z)$ est holomorphe dans ce domaine.

En effet, décomposons le cycle C en deux chemins distincts, C_1 allant d'un point A à un point B, l'autre C_2 allant de B à A. Par hypothèse, les intégrales le long de C_1 et de $-C_2$ sont égales. La fonction étant supposée continue partout dans \mathcal{D} , cette opération peut être faite pour n'importe quel cycle, d'où l'on déduit que, pour *tous* les chemins dans \mathcal{D} , l'intégrale ne dépend que des extrémités. Reprenant alors la même démarche que pour le théorème p. 165 (voir à la suite de (4.39)), où on n'a finalement utilisé que la *continuité* de $f(z)$, on voit que la relation :

$$F(z) \stackrel{\text{d'ef}}{=} \int_{z_0}^z f(z) dz \quad (4.90)$$

définit une fonction $F(z)$ possédant une dérivée $F'(z)$, de ce fait holomorphe dans \mathcal{D} . $F'(z)$, en tant que dérivée d'une fonction holomorphe, est aussi une fonction holomorphe ; comme $F' = f$, le théorème de Morera s'en trouve démontré.

³¹Une fois introduite la représentation d'une fonction par un développement en série (chapitre 5, en la circonstance une série de Taylor), cette affirmation deviendra évidente.

³²ou encore : *toute fonction entière dont l'image ne contient pas au moins deux nombres complexes est une constante.*

Ce théorème permet de montrer que si une fonction $f(z, \xi)$ est analytique en z et continue en ξ dans un domaine simplement connexe \mathcal{D} , alors la fonction $F(z)$:

$$F(z) \stackrel{\text{d\u00e9f}}{=} \int_C f(z, \xi) d\xi \quad (4.91)$$

est une fonction analytique dans \mathcal{D} , C \u00e9tant une courbe dans ce domaine. Notons d'abord que la continuit\u00e9 de $f(z, \xi)$ entra\u00eene celle de $F(z)$. Soit maintenant un cycle Γ dans \mathcal{D} , et l'int\u00e9grale I :

$$I \stackrel{\text{d\u00e9f}}{=} \int_{\Gamma} F(z) dz = \int_{\Gamma} \left(\int_C f(z, \xi) d\xi \right) dz . \quad (4.92)$$

L'hypoth\u00e8se de continuit\u00e9 sur $f(z, \xi)$ permet d'inverser l'ordre des int\u00e9grations³³, pour \u00e9crire :

$$I = \int_C \left(\int_{\Gamma} f(z, \xi) dz \right) d\xi . \quad (4.93)$$

En vertu du th\u00e9or\u00e8me de Cauchy, l'int\u00e9grale interne est nulle, d'o\u00f9 $I = 0$. Revenant \u00e0 (4.92), on en d\u00e9duit que $\int_{\Gamma} F(z) dz = 0$; $F(z)$ est donc une fonction continue dont l'int\u00e9grale sur tout cycle est nulle : d'apr\u00e8s le th\u00e9or\u00e8me de Morera, $F(z)$ est holomorphe.

4.6 Illustrations

Il s'agit d'illustrer les r\u00e9sultats importants obtenus ci-dessus \u00e0 propos de la fonction $z \rightarrow Z = z^n \equiv f(z)$ o\u00f9 $n \in \mathbb{Z}$. Ce choix n'est pas innocent, dans la perspective d'\u00e9tudier ult\u00e9rieurement la possibilit\u00e9 de repr\u00e9senter une fonction quelconque par une s\u00e9rie de puissances enti\u00e8res (positives ou n\u00e9gatives). Visiblement, selon que n est positif ou n\u00e9gatif, $f(z)$ est born\u00e9e ou non dans \mathbb{C} . Ceci justifie, pour la clart\u00e9, d'\u00e9tudier les diff\u00e9rents cas s\u00e9par\u00e9ment.

4.6.1 $n = 0, 1, 2, \dots$

Pour les valeurs enti\u00e8res positives ou nulles de n , z^n est holomorphe dans tout le plan \mathbb{C} . Par le th\u00e9or\u00e8me de Cauchy (voir (4.30)), on a donc de suite :

$$\int_C z^n dz = 0 , \quad (4.94)$$

o\u00f9 C est n'importe quel contour ferm\u00e9. Ce r\u00e9sultat s'obtient aussi d'une autre fa\u00e7on : $f(z) = z^n$ \u00e9tant holomorphe dans tout domaine simplement connexe $\{z, |z| < R\}$, sa primitive et bien d\u00e9finie : c'est $F(z) = \frac{z^{n+1}}{n+1}$. Pour un chemin quelconque reliant deux points d'affixes z_1 et z_2 , on a donc :

$$\int_{z_1}^{z_2} z^n dz = \left[\frac{z^{n+1}}{n+1} \right]_{z_1}^{z_2} = \frac{1}{n+1} (z_2^{n+1} - z_1^{n+1}) . \quad (4.95)$$

³³Il s'agit d'un th\u00e9or\u00e8me classique d'Analyse r\u00e9elle, que l'on utilise en raisonnant s\u00e9par\u00e9ment avec les parties r\u00e9elle et imaginaire.

Si on adopte la représentation polaire, $z_j = \rho_j e^{i\theta_j}$, il vient donc :

$$\int_{z_1}^{z_2} z^n dz = \frac{1}{n+1} [\rho_2^{n+1} e^{(n+1)i\theta_2} - \rho_1^{n+1} e^{(n+1)i\theta_1}] . \quad (4.96)$$

Prenons maintenant une boucle partant de z_1 et arrivant en $z_2 = z_1 = \rho_1 e^{i\theta_1}$, auquel cas l'intégrale doit être nulle ; c'est bien ce que donne le second membre de (4.96), quel que soit le contour, qu'il contienne ou non l'origine. Dans le premier cas, quand on décrit le cycle, l'argument part de la valeur θ_1 et y revient strictement à la fin du cycle. Dans ce second cas (l'origine est à l'intérieur du cycle), l'argument part de θ_1 et vaut $\theta_1 + 2\pi$ à la fin ; mais comme $e^{i(n+1)(\theta_1+2\pi)} = e^{i(n+1)\theta_1}$ puisque $e^{i(n+1)2\pi} = 1 \forall n \in \mathbb{N}$ (et même pour $n \in \mathbb{Z}$), le second membre de (4.96) est encore nul.

On peut encore raisonner comme suit, en partant d'une boucle entourant l'origine. Cette boucle peut être remplacée par un cercle de rayon ρ , paramétré suivant $z = \rho e^{i\theta}$; l'intégrale a alors l'expression explicite $\int_0^{2\pi} \rho^n e^{in\theta} d(\rho e^{i\theta})$, soit $i\rho^{n+1} \int_0^{2\pi} e^{i(n+1)\theta} d\theta$, qui est nulle quel que soit $n \in \mathbb{N}$ (et même pour $n \in \mathbb{Z} \setminus \{-1\}$). Toute autre boucle autour d'un point z_0 quelconque donne le même résultat : on pose $z = z_0 + \rho e^{i\theta}$, et on développe suivant la formule du binôme.

Illustrons maintenant la formule de Cauchy, qui s'écrit ici :

$$z^n = \frac{1}{2i\pi} \int_C \frac{\xi^n}{\xi - z} d\xi ; \quad (4.97)$$

si on veut fixer les idées, on peut avoir en tête³⁴ une fois pour toutes le contour C formé par le cercle γ_R de rayon R centré sur z .

Il s'agit donc de calculer directement le second membre de (4.97) et de vérifier qu'il vaut z^n . Pour cela, on écrit $\xi^n = [(\xi - z) + z]^n$ et on développe suivant la formule du binôme. Il vient alors :

$$\int_C \frac{\xi^n}{\xi - z} d\xi = \sum_{p=0}^n C_n^p z^{n-p} \int_C (\xi - z)^{p-1} d\xi . \quad (4.98)$$

Comme on l'a vu dans la section 4.3, le terme $p = 0$ mérite une attention particulière. En extrayant ce terme de la sommation, il vient :

$$\int_C \frac{\xi^n}{\xi - z} d\xi = z^n \int_C \frac{d\xi}{\xi - z} + \sum_{p=1}^n C_n^p z^{n-p} \int_C (\xi - z)^{p-1} d\xi \quad (4.99)$$

Toutes les intégrales de la sommation de (4.99) sont nulles³⁵, puisqu'elles sont précisément de la forme $\int_C \xi'^q d\xi'$ avec $q = p - 1 \geq 0$. Seule survit l'intégrale mise à part, venant du terme $p = 0$; d'après (4.59), elle vaut $2i\pi$. Au total, le second membre de (4.97) vaut bien $\frac{1}{2i\pi} z^n (2i\pi) = z^n$, en accord avec la formule de Cauchy.

³⁴Le peu qui a été dit sur les possibilités de déformer continûment le contour (sans faire décoller l'élastique) permet d'affirmer que les résultats énoncés avec γ_R restent vrais pour tout contour fermé entourant z .

³⁵Ceci se vérifie aussi directement, en posant $\xi = z + Re^{i\theta}$. L'intégrand est alors de la forme $(Re^{i\theta})^{p-1} iRe^{i\theta} = iR^p e^{ip\theta}$. Comme p est un entier strictement positif, l'intégrale de 0 à 2π est nulle (une telle intégrale est en fait nulle $\forall p \in \mathbb{Z}^*$).

4.6.2 $n = -1, -2, \dots$

La fonction z^n est maintenant holomorphe dans tout domaine $\{z, r < |z| < R\}$, r et R étant quelconques mais finis, domaine non simplement connexe. À nouveau, il convient pour la clarté à ce stade d'effectuer une distinction, en traitant successivement les deux cas, $n = -1$ et $n \neq -1$.

■ $n = -1$

Pour tout contour fermé qui ne contient pas l'origine en son intérieur – donc situé tout entier dans une composante simplement connexe du domaine d'holomorphie –, le théorème de Cauchy donne :

$$\int_C z^{-1} dz = 0 \quad , \quad (4.100)$$

Tant que le contour est tout entier situé dans une telle composante, la primitive³⁶ de $\frac{1}{z}$ est bien définie, c'est $\ln z$, donc :

$$\int_{z_1}^{z_2} \frac{1}{z} dz = [\ln z]_{z_1}^{z_2} = \ln z_2 - \ln z_1 \quad , \quad (4.101)$$

On sait que la fonction logarithme possède une infinité de déterminations, différant de $2ik\pi$ les unes des autres ($k \in \mathbb{Z}$) ; bien évidemment, pour calculer la variation $\ln z_2 - \ln z_1$, il faut prendre la *même* détermination pour chaque terme, de sorte que par différence, le résultat est le même quelle que soit la détermination choisie³⁷. Il en résulte :

$$\int_{z_1}^{z_2} \frac{1}{z} dz = \ln \frac{|z_2|}{|z_1|} + i\theta_{21} \quad (4.102)$$

où l'angle θ_{21} est l'angle représentant la variation de l'argument quand on se déplace continûment de z_1 à z_2 . Prenons maintenant un contour partant de z_1 et se terminant en z_2 infiniment voisin de z_1 . Si le contour choisi ne contient pas l'origine, la variation de l'argument est nulle ($\theta_{21} = 0$) et (4.100) est vérifiée. Si le cycle contient l'origine, cela a encore un sens de manipuler la primitive puisque l'on peut rester dans une composante simplement connexe (voir fig. 4.9) ; allant de z_1 à z_2 , on fait un tour (presque) complet, la variation de l'argument vaut $+2\pi - 0$ ($\theta_{21} = 2\pi$) et l'intégrale au premier membre

³⁶À un niveau élémentaire, on apprend que la primitive de z^n est $\frac{z^{n+1}}{n+1}$ pour $n \neq -1$, et $\ln z$ si $n = -1$, ce qui est indéniable. Il est toutefois intéressant de remarquer que le résultat (4.101) peut s'obtenir en partant du cas général en y faisant tendre n vers -1 dans la différence $\frac{1}{n+1}(z_2^{n+1} - z_1^{n+1})$ (et en prenant soin bien sûr de ne manipuler qu'une détermination du logarithme, choisie (arbitrairement) une fois pour toutes).

Ce "miracle" se produit tout simplement parce que la primitive de z^α est, $\forall \alpha \in \mathbb{C}$, égale à $\frac{z^{\alpha+1}}{\alpha+1}$ (comme on le voit en revenant à la définition de la fonction puissance généralisée). La limite $\alpha \rightarrow -1$ se trouve facilement en écrivant $z^{\alpha+1} = e^{(\alpha+1)\ln z}$ et en développant l'exponentielle autour de $\alpha = -1$.

³⁷Tout comme, en Physique, quand on calcule une variation d'énergie potentielle : le résultat est unique (c'est le travail de la force, qui n'a qu'une seule valeur !) et se moque de la constante additive choisie une fois pour toutes pour l'énergie potentielle.

de (4.100) vaut $i \times 2\pi$, en accord avec le résultat obtenu en (4.58) par considération d'un cycle circulaire. Noter que, ce faisant, on ne franchit pas la coupure figurée par le tortillon semi-infini à double flèche partant de l'origine (voir fig. 4.9). Tout naturellement, la primitive étant multiforme, elle ne reprend pas la même valeur une fois que l'on est arrivé en z_2 infiniment voisin de z_1 .

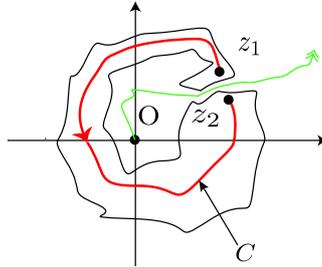


Figure 4.9: Le contour C relie deux points infiniment voisins z_1 et z_2 tout en restant dans un domaine simplement connexe où $\frac{1}{z}$ est holomorphe.

■ $n \neq -1$

À nouveau, la fonction analysée $f(z) = z^n \equiv \frac{1}{z^{|n|}}$ est holomorphe dans \mathbb{C} privé de l'origine. Pour tout contour fermé qui ne contient pas l'origine en son intérieur, on a donc toujours :

$$\int_C z^{-|n|} dz = 0, \quad (4.103)$$

ce que l'on peut retrouver aussi en calculant la variation de la primitive $F(z) = \frac{z^{n+1}}{n+1}$: c'est une fonction à une seule détermination, qui reprend donc la même valeur au départ et à l'arrivée.

Il en va encore de même si le contour C contient l'origine ; on peut à nouveau raisonner comme avec le logarithme (mais il n'y a plus de coupure), en considérant la composante simplement connexe et deux points infiniment voisins ; mais comme n est un entier $\neq -1$, la primitive $F(z)$ reprend cette fois la même valeur quand si l'argument de z a augmenté de $2\pi - 0$. Toutefois, *on ne peut pas ici invoquer le théorème de Cauchy*, qui suppose la fonction holomorphe dans le domaine ceinturé par le contour (celui-ci n'est pas réductible à un point, puisqu'il faut rester à l'extérieur d'un voisinage de l'origine, aussi petit soit-il). Ceci n'est pas non plus en contradiction avec le théorème de Morera, qui exige que la fonction soit *continue* partout dans \mathcal{D} , ce qui n'est pas le cas ici.

En rassemblant tous ces résultats, on peut finalement écrire :

$$\boxed{\int_C (z - z_0)^n dz = 2i\pi \delta_{n, -1} \quad \forall n \in \mathbb{Z}} \quad (4.104)$$

où C est un contour contenant z_0 en son intérieur ; répétons que ce résultat peut aussi être obtenu d'un coup en remplaçant le contour C par le cercle paramétré par $z = z_0 + re^{i\theta}$.

◆ *Remarques*

- Il faut bien retenir que dans les énoncés des théorèmes précédents, chaque mot pèse : holomorphe, simplement connexe et, éventuellement, quantificateur ($\forall C$, par exemple).
 1. dans l'exemple qui vient d'être donné, le théorème de Cauchy n'est pas applicable puisque le contour se situe dans un domaine *multiplement* connexe (typiquement : une couronne délimitée par deux cercles centrés à l'origine), domaine d'holomorphie de la fonction.
 2. Un autre exemple, où l'hypothèse manquante n'est pas la simple connexité, mais le fait que la fonction n'est pas holomorphe. Soit la fonction $z \rightarrow z^{*2}$, visiblement non holomorphe ; il est facile de montrer³⁸ que :

$$\frac{1}{2i\pi} \int_{\Gamma} \frac{\xi^{*2}}{\xi - z} d\xi = z^{*2} , \quad (4.105)$$

où Γ est un cercle centré sur z et de rayon quelconque. À nouveau, (4.105) a tout l'air d'être un cas particulier de la formule de Cauchy, mais il n'en est rien, puisque $f(z) \stackrel{\text{déf}}{=} z^{*2}$ n'est pas holomorphe.

- Le cas où n est quelconque (pas entier, mais réel voire complexe) est très intéressant en soi, mais exige une étude particulière : en effet, la fonction z^α étant multiforme, tous les théorèmes énoncés plus haut doivent être revisités et, éventuellement, convenablement généralisés – c'est au fond ce que l'on a fait avec le logarithme en considérant deux extrémités infiniment voisines. L'introduction des surfaces de Riemann, abordée ultérieurement (voir chapitre 5, section 5.6), est en grande partie justifiée par ce souci d'extension. Une chose est sûre : pour α réel quelconque se substituant à n , (4.96) devient :

$$\int_{z_1}^{z_2} z^\alpha dz = \frac{1}{\alpha + 1} [\rho_2^{\alpha+1} e^{i(\alpha+1)\theta_2} - \rho_1^{\alpha+1} e^{i(\alpha+1)\theta_1}] . \quad (4.106)$$

Si z_2 est infiniment proche de $z_1 = z_0$, le second membre est nul ou non, selon que l'on a tourné ou non autour de l'origine. Dans le premier cas, l'argument part d'une valeur donnée θ_0 et y revient :

$$\int_{\text{cycle ne contenant pas O}} z^\alpha dz = \frac{1}{\alpha + 1} \rho_0^{\alpha+1} [e^{i(\alpha+1)\theta_0} - e^{i(\alpha+1)\theta_0}] = 0 . \quad (4.107)$$

³⁸Il suffit de calculer directement l'intégrale en paramétrant le cercle.

Au contraire, si l'origine est à l'intérieur du contour, l'argument passe de θ_0 à $\theta_0 + 2\pi$ de sorte que l'on a :

$$\int_{\text{cycle contenant O}} z^\alpha dz = \frac{1}{\alpha + 1} \rho_0^{\alpha+1} [e^{i(\alpha+1)\theta_0} - e^{i(\alpha+1)(\theta_0+2\pi)}] =$$

$$\frac{\rho_0^{\alpha+1}}{\alpha + 1} e^{i(\alpha+1)\theta_0} [1 - e^{2i\alpha\pi}] \neq 0, \quad (4.108)$$

puisque pour α non entier, $e^{2i\alpha\pi} \neq 1$. Comme on l'a déjà vu, l'origine est un point remarquable (singulier) : de ce point de branchement part une ligne continue de singularités, infranchissable tant que le caractère holomorphe doit être préservé, c'est une coupure. \blacklozenge

Chapitre 5

Représentation des fonctions analytiques par des séries. Théorème des résidus

“Un mathématicien est une machine
à transformer le café en théorèmes.”
(Paul ERDÖS, 1913–1996)

Après un rappel et une généralisation de résultats classiques d'analyse réelle permettant la classification des singularités, un théorème capital sera démontré : le théorème des résidus.

La notion de développement de Taylor est bien connue en analyse réelle. Dans un premier temps, on va voir qu'elle s'étend aux fonctions holomorphes dans un domaine simplement connexe. Par la suite, elle sera généralisée pour montrer dans quelles conditions une fonction $f(z)$ peut être représentée par une expression du genre :

$$f(z) = \sum_{n=-\infty}^{+\infty} c_n (z - a)^n . \quad (5.1)$$

La particularité d'un tel développement (appelé *série de Laurent*, voir section 5.2) est de contenir *a priori* toutes les puissances entières du monôme $(z - a)$, positives et *négatives*. Ce résultat est d'une importance capitale, et permet notamment, quand a est une singularité de f , une classification des singularités d'une fonction ; les singularités jouent un rôle de première importance dans toute la suite. Ceci étant fait, on pourra démontrer un théorème fondamental pour les applications, le théorème des résidus (voir section 5.4).

5.1 Séries de Taylor

La série de Taylor d'une fonction s'introduit naturellement comme suit. Soit la somme géométrique¹ ($X \neq 1$) :

$$S_n(X) \stackrel{\text{déf}}{=} \sum_{p=0}^n X^p = \frac{1 - X^{n+1}}{1 - X} , \quad (5.2)$$

ce qui peut se récrire comme :

$$\frac{1}{1 - X} = 1 + X + X^2 + \dots + X^n + \frac{1}{1 - X} X^{n+1} . \quad (5.3)$$

L'expression de droite peut d'abord être vue comme un développement limité d'ordre n de la fonction $\frac{1}{1-X}$ au voisinage de $X = 0$; en outre, la fonction étant dérivable, les coefficients du développement sont en fait les nombres $\frac{1}{n!} \left(\frac{d^n}{dX^n} \frac{1}{1-X} \right)_{X=0} = 1$; le reste est bien de la forme $X^n \varepsilon(X)$ avec $\varepsilon(X) = \frac{X}{1-X} \rightarrow 0$ quand $X \rightarrow 0$. (5.3) est donc plus précisément le développement de Taylor de $\frac{1}{1-X}$ centré en $X = 0$, limité à l'ordre n .

Considérons maintenant la fraction rationnelle $\frac{1}{\xi - z}$, qui joue un rôle central dans les diverses relations et théorèmes établis au chapitre précédent. Introduisant de plus un complexe quelconque a , on a :

$$\frac{1}{\xi - z} = \frac{1}{(\xi - a) + (a - z)} = \frac{1}{\xi - a} \frac{1}{1 - \frac{z-a}{\xi-a}} , \quad (5.4)$$

En faisant la substitution $X = \frac{z-a}{\xi-a}$ dans (5.3), il vient :

$$\frac{1}{\xi - z} = \frac{1}{\xi - a} \left[1 + \left(\frac{z-a}{\xi-a} \right) + \left(\frac{z-a}{\xi-a} \right)^2 + \dots + \left(\frac{z-a}{\xi-a} \right)^n + \frac{1}{1 - \frac{z-a}{\xi-a}} \left(\frac{z-a}{\xi-a} \right)^{n+1} \right] , \quad (5.5)$$

soit :

$$\frac{1}{\xi - z} = \frac{1}{\xi - a} + \frac{z-a}{(\xi-a)^2} + \frac{(z-a)^2}{(\xi-a)^3} + \dots + \frac{(z-a)^n}{(\xi-a)^{n+1}} + \frac{1}{\xi - z} \left(\frac{z-a}{\xi-a} \right)^{n+1} . \quad (5.6)$$

Soit maintenant une fonction $f(z)$ holomorphe dans un domaine \mathcal{D} simplement connexe. Multiplions membre à membre (5.6) par $\frac{f(\xi)}{2i\pi}$ et intégrons sur un contour $C \in \mathcal{D}$ entourant a et z (ce peut être la frontière de \mathcal{D} si f y est continue). D'après la formule de Cauchy établie au chapitre 4, le premier membre de (5.6) donne $f(z)$; d'après la même formule et sa généralisation :

$$f^{(n)}(a) = \frac{n!}{2i\pi} \int_C \frac{f(\xi)}{(\xi - a)^{n+1}} d\xi \quad (n \in \mathbb{N}) , \quad (5.7)$$

¹Pour $X = 1$, on a trivialement $S_n = n + 1$. Par ailleurs, le développement classique (5.2), démontré à un niveau élémentaire pour $X \in \mathbb{R}$ en retranchant $X S_n$ à S_n , est de toute évidence également vrai pour $X \in \mathbb{C}$. Noter que (5.2) redonne le résultat bien connu : la somme des N racines N^{es} de l'unité est égale à zéro.

les différentes puissances au second membre de (5.6) donnent les dérivées successives de $f(z)$, au facteur $\frac{1}{n!}$ près. Au total, (5.6) se transforme en :

$$f(z) = f(a) + \frac{1}{1!}f'(a)(z-a) + \frac{1}{2!}f''(a)(z-a)^2 + \dots + \frac{1}{n!}f^{(n)}(a)(z-a)^n + R_n(z) , \quad (5.8)$$

où la quantité R_n , toujours appelée *reste*, a pour expression :

$$R_n(z) = \frac{(z-a)^{n+1}}{2i\pi} \int_C \frac{f(\xi)}{(\xi-z)(\xi-a)^{n+1}} d\xi . \quad (5.9)$$

Visiblement, le développement (5.8) généralise au cas d'une fonction holomorphe la notion de développement de Taylor introduit en Analyse élémentaire à propos d'une fonction réelle d'une variable réelle.

Maintenant, une question naturelle vient à l'esprit : sous quelles conditions peut-on "pousser" un tel développement à l'infini² ? Plus précisément, que faut-il pour que, dans la limite $n \rightarrow +\infty$, le reste R_n tende uniformément vers zéro ? Quand ces conditions sont réunies, alors la fonction f peut être représentée par une *série* de puissances positives du simple monôme $z - a$:

$$\boxed{f(z) = \sum_{n=0}^{+\infty} c_n(z-a)^n , \quad c_n = \frac{1}{n!}f^{(n)}(a) = \frac{1}{2i\pi} \int_C \frac{f(\xi)}{(\xi-a)^{n+1}} d\xi} \quad (5.10)$$

La réponse à cette question est fournie par le théorème suivant³ :

La fonction $f(z)$ peut être représentée par sa série de Taylor (5.10) dans tout disque ouvert de centre a dans lequel elle est holomorphe. Dans tout domaine fermé inclus dans ce disque, la série (5.10) converge uniformément.

Pour démontrer ce théorème, on introduit la variable (muette) ξ se déplaçant sur le cercle⁴ γ_R de centre a et de rayon R ; le disque correspondant est supposé tout entier contenu dans le domaine d'holomorphie de f ; soit z dans ce disque : $|z-a| < R$, soit $|z-a| = kR$ avec $0 < k < 1$ (voir fig. 5.1). Par l'inégalité triangulaire :

$$|\xi - a| \equiv |(\xi - z) + (z - a)| \leq |\xi - z| + |z - a| , \quad (5.11)$$

il vient⁵ :

$$|\xi - z| \geq |\xi - a| - |z - a| \geq R - kR = (1 - k)R > 0 . \quad (5.12)$$

Géométriquement⁶, il est évident que la plus petite valeur de $|\xi - z|$ est obtenue quand a , z et ξ sont alignés, z étant entre a et ξ .

²La même question se pose déjà dans le cas réel.

³Encore un théorème dû à Cauchy !

⁴L'équation du cercle est donc $|\xi - a| = R$.

⁵L'inégalité (5.12) est de toute façon évidente géométriquement.

⁶On assimile à nouveau un point $M(x, y)$ du plan et son affixe $z = x + iy$.

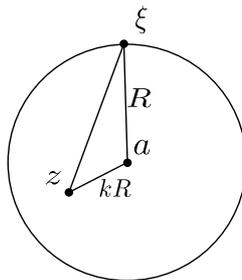


Figure 5.1: Le disque γ_R de rayon R est inclus dans le domaine d'analyticit  de $f(z)$; visiblement $\forall z, |z - a| < R, |\xi - z| \geq (1 - k)R > 0$.

Par (5.9), et M d signant le maximum du module de f dans le disque⁷ $|z - a| \leq R$, le reste R_n se majore comme suit :

$$|R_n(z)| = \left| \frac{(z - a)^{n+1}}{2i\pi} \int_{\gamma_R} \frac{f(\xi)}{(\xi - z)(\xi - a)^{n+1}} d\xi \right| \leq \frac{(kR)^{n+1}}{2\pi} \frac{M 2\pi R}{(1 - k)R R^{n+1}} = \frac{Mk^{n+1}}{1 - k} . \quad (5.13)$$

k  tant plus petit que 1, on a bien $\lim_{n \rightarrow +\infty} R_n = 0$. z  tant strictement   l'int rieur du disque, la s rie de Taylor existe bien dans tout ouvert $|z - a| < R$ inclus dans le disque ; de plus, comme la borne du reste est *ind pendante* de z , la convergence est uniforme dans tout disque ferm  de rayon kR , ce qui ach ve la d monstration du th or me. Une fonction admettant un tel d veloppement est dite *analytique*.

Il y a donc bien  quivalence entre la propri t  d'holomorphie (satisfaire les conditions n cessaires et suffisantes de Cauchy - Riemann avec les quatre d riv es premi res u_x, \dots , continues) et d'analyticit ⁸ (admettre un d veloppement en s rie enti re convergeant uniform ment dans un disque ferm ).   ce stade, on peut mesurer l'extraordinaire rigidit  d'une fonction holomorphe $f(z)$. Au d part, il s'agit "seulement" de s'assurer que $f(z)$ est d rivable (conditions de Cauchy). Comme on l'a vu ant rieurement, le fait que $f'(z)$ existe assure que *toutes* les d riv es existent  galement. Encore plus fort : f admet un d veloppement en s rie enti re dans tout disque inclus dans le domaine o  elle est holomorphe.

On sait que les s ries enti res sont des objets aux propri t s remarquables, li es

⁷Comme f est holomorphe dans le disque γ_R , elle y est born e. En effet, le principe (ou th or me) du maximum (voir p. 173) stipule qu'une fonction (non constante) holomorphe dans un domaine \mathcal{D} ne peut avoir de maximum local dans ce domaine. Autrement dit, pour tout z_0 donn , on peut trouver dans le voisinage de z_0 un point z_1 o  $|f(z_0)| < |f(z_1)|$.

Ceci peut aussi  tre compris comme une cons quence imm diate des conditions de Cauchy - Riemann ; les parties r elle et imaginaire de $f(z)$ satisfont $\Delta u = 0, \Delta v = 0$, o  Δ d signe le Laplacien   deux dimensions (pour cette raison, u et v sont dites *harmoniques*, comme on l'a vu). Il en r sulte que si l'une des d riv es secondes de u est positive, l'autre est n gative, et de m me pour v . De telles fonctions ne peuvent donc pr senter que des points-col (en selle de cheval, en *chip*). Il en r sulte que le maximum du module d'une fonction holomorphe dans \mathcal{D} survient en un (des) point(s) de la fronti re de ce domaine : ce maximum est le nombre d sign  par M dans la d monstration en cours.

⁸Selon Larousse, on peut dire *analyticit * ou *analyticit *.

à la notion de rayon de convergence : si une série entière $\sum_{n \in \mathbb{N}} c_n x^n$ converge pour $|x| = R$, elle est *uniformément* convergente pour tout x , $|x| \leq R' < R$. En particulier, on peut intégrer et dériver terme à terme⁹, les séries ainsi obtenues ayant le même disque de convergence. Dans \mathbb{C} , le rayon de convergence R devient le rayon du disque à l'intérieur duquel $\sum_{n \in \mathbb{N}} c_n z^n$ converge uniformément soit¹⁰ :

$$R = \lim_{n \rightarrow +\infty} \operatorname{Inf} |c_n|^{-\frac{1}{n}} \quad (5.14)$$

En conséquence de la convergence *uniforme*¹¹, toutes les sommes sont des fonctions continues, tout comme dans le cas réel. En fait, il existe un résultat beaucoup plus fort : à l'intérieur de son disque de convergence, une série entière $\sum_{n \in \mathbb{N}} c_n z^n \stackrel{\text{déf}}{=} f(z)$ représente une fonction analytique (holomorphe). La démonstration procède en examinant la limite $h \rightarrow 0$ de la différence $h^{-1}[f(z+h) - f(z)] - \sum_{n \in \mathbb{N}} n c_n z^{n-1}$, et en établissant qu'elle est nulle (voir [6], § 2.16 pour les détails).

Rappelons les développements de Taylor à l'infini de quelques fonctions élémentaires, qui apparaissent en fait comme la simple extension d'écriture (z au lieu de x !) des développements introduits en Analyse réelle. On a :

$$\frac{d^n}{dz^n} e^z = e^z \quad (n \in \mathbb{N}) \quad (5.15)$$

d'où il résulte que $f^{(n)}(a=0) = 1$ si $f(z) = e^z$. La série de Taylor de l'exponentielle centrée en $a=0$ est donc :

$$e^z = 1 + z + \frac{1}{2!} z^2 + \frac{1}{3!} z^3 + \dots = \sum_{n=0}^{+\infty} \frac{1}{n!} z^n \quad (5.16)$$

À partir de (5.16), en formant les bonnes combinaisons linéaires de $e^{\pm z}$, on trouve immédiatement :

$$\sinh z = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \dots = \sum_{n=0}^{+\infty} \frac{z^{2n+1}}{(2n+1)!}, \quad \cosh z = 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \dots = \sum_{n=0}^{+\infty} \frac{z^{2n}}{(2n)!} \quad (5.17)$$

Tous ces développements convergent pour z quelconque de module fini, tout comme ceux des lignes circulaires, obtenus en utilisant¹² $\cosh iz = \cos z$ et $\sinh iz = i \sin z$:

$$\sin z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots = \sum_{n=0}^{+\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}, \quad (5.18)$$

⁹Pour une série de fonctions $f(x) \stackrel{\text{déf}}{=} \sum_{n \in \mathbb{N}} c_n f_n(x)$, la série des dérivées doit être elle-même uniformément convergente pour être égale à la dérivée $f'(x)$.

¹⁰De (5.14), il résulte que la frontière du disque de convergence passe par la (ou les) singularité(s) de la fonction la(es) plus proche(s) de l'origine.

¹¹Rappelons que la convergence uniforme d'une série de fonctions (continues) $\sum_{n \in \mathbb{N}} c_n f_n(x)$ est une condition *suffisante* pour que la somme soit une fonction continue. On connaît des séries qui ne sont pas uniformément convergentes, mais dont la somme est continue (voir les exemples cités par Titchmarsh [6], § 1.3). Toutefois, pour des séries à termes positifs, il y a équivalence entre convergence uniforme et continuité de la somme.

¹²Se souvenir que $i^{2n} = (e^{i\frac{\pi}{2}})^{2n} = e^{in\pi} = (-1)^n$ et donc $i^{2n+1} = (-1)^n i$.

$$\cos z = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} + \dots = \sum_{n=0}^{+\infty} (-1)^n \frac{z^{2n}}{(2n)!} . \quad (5.19)$$

Comme ces développements convergent (uniformément) dans tout disque fermé de \mathbb{C} , on peut décaler le centre en tout z_0 fini. Par exemple :

$$e^z = \sum_{n=0}^{+\infty} \frac{1}{n!} \left(\frac{d^n}{dz^n} e^z \right)_{z_0} (z - z_0)^n = \sum_{n=0}^{+\infty} \frac{1}{n!} e^{z_0} (z - z_0)^n ; \quad (5.20)$$

à la réflexion, ceci est d'ailleurs évident puisque $e^z = e^{z_0} e^{z-z_0}$. Toutes les fonctions ci-dessus sont entières : leur série entière converge (uniformément) pour tout z fini.

Comme autre exemple, soit maintenant la fonction $\ln(1+z)$; on a :

$$\frac{d^n}{dz^n} \ln(1+z) = \frac{(-1)^{n+1}(n-1)!}{(1+z)^n} , \quad \left. \frac{d^n}{dz^n} \ln(1+z) \right|_{z=0} = (-1)^{n+1}(n-1)! . \quad (5.21)$$

La série de Taylor de $\ln(1+z)$ centrée en $z=0$ est donc :

$$\ln(1+z) = \ln 1 + \sum_{n=0}^{+\infty} \frac{1}{n!} (-1)^{n+1}(n-1)! z^n = \ln 1 + \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} z^n . \quad (5.22)$$

Le terme constant $\ln 1$ prend en compte l'aspect multiforme de la fonction logarithme (pour la détermination principale, $\ln 1 = 0$) ; la suite du développement est indépendante de la détermination choisie puisque toutes ont les mêmes dérivées, d'où ($k \in \mathbb{Z}$) :

$$\boxed{\ln(1+z) = 2ik\pi + \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} z^n \quad (|z| < 1)} \quad (5.23)$$

Cette série converge uniformément¹³ pour $|z| \leq R < 1$ (ce développement peut être retrouvé de tête en intégrant terme à terme $\frac{1}{1+z} = 1 - z + z^2 - z^3 + \dots$).

Pour la fonction $(1+z)^\alpha$, le développement de Taylor près de $z=0$ est¹⁴ :

$$(1+z)^\alpha = 1^\alpha \left[1 + \alpha z + \frac{1}{2!} \alpha(\alpha-1) z^2 + \frac{1}{3!} \alpha(\alpha-1)(\alpha-2) z^3 + \dots \right] , \quad (5.24)$$

Le choix de la détermination est reflété par la présence du 1^α en facteur ; en effet, le terme constant est $(1+z)^\alpha|_{z=0} = 1^\alpha$, le second terme implique $\frac{d}{dz}(1+z)^\alpha|_{z=0}$, soit

¹³Ceci est lié au fait que le point $z = -1$, qui annule l'argument du \ln , est un point remarquable (de branchement) et qu'il appartient au cercle de rayon unité, frontière du disque de convergence de la série entière. Noter que si $z = 1$, le développement est la série harmonique alternée, dont la somme est $\ln 2$; en revanche, pour $z = -1$, on obtient la série harmonique, qui est divergente. Une fois introduite la notion de prolongement analytique, on comprendra pourquoi la série dans (5.23) est effectivement égale à $\ln(1+z)$ sur le bord du disque à l'exception du point $z = -1$.

¹⁴Si $\alpha = N \in \mathbb{N}$, le développement s'arrête au terme en z^N : c'est celui du binôme.

En convenant que le produit de zéro facteur est égal à $(1)^\alpha$, (5.25) peut aussi s'écrire $e^{2i\alpha k\pi} \sum_{n=0}^{+\infty} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} z^n$.

$\alpha(1+z)^{\alpha-1}\big|_{z=0} = \alpha \frac{(1+z)^\alpha}{1+z}\big|_{z=0} = \alpha 1^\alpha$, et ainsi de suite. Au total, par rapport au développement “classique”, apparaît en plus 1^α en *facteur*¹⁵, ce que l’on peut écrire ($k \in \mathbb{Z}$) :

$$(1+z)^\alpha = e^{2i\alpha k\pi} \left[1 + \sum_{n=1}^{+\infty} \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} z^n \right] \quad (|z| < 1) \quad (5.25)$$

Ce développement ne converge¹⁶ uniformément que si $|z| \leq R < 1$. La fonction $(1+z)^\alpha$ est elle aussi multiforme (si $\alpha \notin \mathbb{Z}$) : la branche “principale”, vaut 1 en $z=0$ et possède la symétrie $f(z^*) = [f(z)]^*$ quand $\alpha \in \mathbb{R}$. Tout naturellement, la frontière du disque de convergence contient à chaque fois une singularité de la fonction.

Signalons que la somme $S(z)$ d’une série $\sum_{n=0}^{\infty} f_n(z)$ de fonctions analytiques dans un domaine simplement connexe \mathcal{D} , uniformément convergente dans \mathcal{D} , est une fonction analytique (Weierstrass). En effet, en raison de la convergence uniforme, d’une part la somme est une fonction continue, d’autre part, on peut intégrer terme à terme sur un contour fermé C inclus dans \mathcal{D} :

$$\int_C S(z) dz = \sum_{n=0}^{+\infty} \int_C f_n(z) dz . \quad (5.26)$$

Chaque intégrale au second membre est nulle (puisque chaque f_n est analytique dans \mathcal{D}). De $\int_C S(z) dz = 0$ et du théorème de Morera, on déduit que $S(z)$ est une fonction analytique. Par ailleurs, toute série de fonctions analytiques dans \mathcal{D} et continues sur $\bar{\mathcal{D}}$, uniformément convergente sur $\bar{\mathcal{D}}$, peut être dérivée terme à terme un nombre arbitraire de fois dans \mathcal{D} (mais pas dans $\bar{\mathcal{D}}$).

◆ *Remarques*

1. Pour démontrer la convergence dans $\bar{\mathcal{D}}$, il suffit, grâce au principe du maximum, de la prouver sur la frontière C de \mathcal{D} . En effet :

$$\max_{\{\bar{\mathcal{D}}\}} |f_{n+1}(z) + f_{n+2}(z) + \dots| = \max_{\{C\}} |f_{n+1}(z) + f_{n+2}(z) + \dots| \quad (5.27)$$

2. Pour illustrer le fait que l’on peut dériver dans \mathcal{D} mais pas (forcément) dans $\bar{\mathcal{D}}$, soit la série $S(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^2}$. Elle converge uniformément dans le disque fermé $\{z, |z| \leq 1\}$, puisque qu’elle est bornée¹⁷ par la série convergente $\sum_{n=1}^{\infty} \frac{1}{n^2}$ (qui vaut $\frac{\pi^2}{6}$). En revanche, la série obtenue par dérivation terme à terme, uniformément convergente pour $|z| < 1$ et donc y représentant la dérivée de S , $S'(z) = \sum_{n=1}^{\infty} \frac{z^{n-1}}{n}$, diverge en $z=1$ (c’est la série harmonique), affixe d’un point du cercle de rayon unité.

¹⁵Merci à Ian Jauslin pour sa question à ce sujet.

¹⁶Même remarque que dans la note 13 à propos du logarithme. Dans les deux cas, $z = -1$ est un point de branchement, sauf pour (5.25) si $\alpha \in \mathbb{N}$.

¹⁷Rappelons que cette propriété est une condition *suffisante* assurant la convergence uniforme.

3. On sait que pour une série entière $f(z) = \sum_{n \in \mathbb{N}} c_n z^n$, qui converge uniformément dans le disque de rayon R , la série des dérivées¹⁸ $\sum_{n \in \mathbb{N}^*} n c_n z^{n-1}$ a pour somme la dérivée f' de f puisqu'elle aussi converge uniformément. La série des dérivées est elle-même une série entière, dont la dérivation terme à terme produit une série dont la somme est $f''(z)$, etc. On retrouve bien ainsi qu'une fonction analytique est *infiniment* dérivable. \blacklozenge

5.2 Séries de Laurent

Le développement de Taylor d'une fonction permet de représenter cette dernière dans tout disque plongé dans le domaine d'analyticit . La propri t  essentielle pour le disque est d' tre *simplement* connexe, tout domaine d form  conviendrait tout autant pourvu qu'il le soit aussi. En fait, il s'av re aussi tr s utile de pouvoir repr senter en s rie (pas enti re !) une fonction holomorphe dans un domaine *multiplement* connexe, dont le prototype est une couronne circulaire ; c'est ainsi que s'introduisent naturellement les d veloppements dits de Laurent.

D'ailleurs, leur n cessit  s'impose au vu d'un simple exemple. Soit la fonction $z \rightarrow f(z) \stackrel{\text{d f}}{=} \frac{1}{z(R-z)}$, o  R est un r el positif ; comme $f(z)$ n'est pas d finie au point $z = 0$, on ne peut esp rer trouver un d veloppement de Taylor centr    l'origine. Toutefois, un peu de r flexion permet d' crire, si $|z| < R$:

$$\frac{1}{z(R-z)} = \frac{1}{z} \frac{1}{R-z} = \frac{1}{z} \frac{1}{R} \sum_{n=0}^{+\infty} \left(\frac{z}{R}\right)^n \equiv \frac{1}{R^2} \sum_{n=-1}^{+\infty} \left(\frac{z}{R}\right)^n ; \quad (5.28)$$

on a ainsi mis en  vidence   *la main* un d veloppement de $f(z)$ qui est parfaitement sens  si $0 < |z| < R$, ce domaine  tant pr cis ment une couronne circulaire dont le cercle int rieur a un rayon aussi petit que l'on veut, mais fini, et dont la p riph rie est le cercle de rayon R – noter que ce d veloppement contient une puissance *n gative* de z . C'est ce type de d veloppement dont on va montrer l'existence en g n ral pour toute fonction holomorphe dans une couronne donn e.

Introduisons donc la couronne circulaire K , de centre a et comprise entre les deux cercles de rayons r et R . On a :

$$r < |z - a| < R \quad \forall z \in K \quad (5.29)$$

Soit $f(z)$ une fonction holomorphe dans la couronne K et continue sur sa fronti re¹⁹, et soit un contour C' ferm  quelconque entourant $z \in K$, situ  tout entier dans

¹⁸Pour une s rie de fonctions $\phi(z) = \sum_{n \in \mathbb{N}} \gamma_n u_n(z)$ qui converge uniform ment dans un certain domaine, la s rie des d riv es $\sum_{n \in \mathbb{N}} \gamma_n u'_n(z)$ n'est  gale   la d riv e $\phi'(z)$ que si elle converge uniform ment.

¹⁹Ceci exclut  videmment les fonctions multiformes qui ne reprennent pas la m me valeur quand on a fait un tour de couronne ; c'est le cas de $\ln(z-a)$, de $(z-a)^{\alpha \notin \mathbb{N}}$, mais pas de $\ln(z-b)$ et de $(z-b)^{\alpha \notin \mathbb{N}}$ quand b est ext rieur   K .

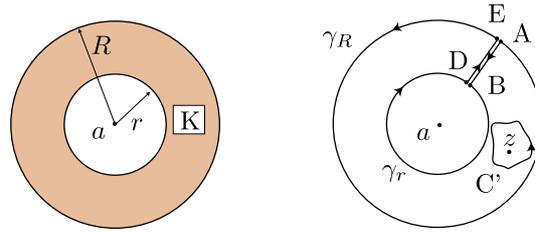


Figure 5.2: Couronne K et contour d'intégration pour l'introduction du développement de Laurent. Les deux petits segments AB et DE sont en réalité infiniment proches l'un de l'autre.

K (voir fig. 5.2). Les conditions d'applicabilité de la formule de Cauchy sont réunies et on peut écrire :

$$f(z) = \frac{1}{2i\pi} \int_{C'} \frac{f(\xi)}{\xi - z} d\xi \quad (z \in K) . \tag{5.30}$$

Maintenant, on peut déformer cette boucle C' , tout en restant dans la couronne K , pour lui donner la forme du contour $C=ABDEA$ (fig. 5.2 à droite) ; C est formé par les deux cercles (parcourus en sens contraires), reliés par deux petits segments AB et DE infiniment proches l'un de l'autre et parcourus aussi en sens opposés (voir fig. 5.2) ; comme ce contour est tout entier dans une partie simplement connexe, la formule de Cauchy s'écrit maintenant :

$$f(z) = \frac{1}{2i\pi} \int_{ABDEA} \frac{f(\xi)}{\xi - z} d\xi . \tag{5.31}$$

En désignant par γ_R et γ_r les deux cercles²⁰, cette formule s'explicite en :

$$2i\pi f(z) = \int_{\gamma_R} \frac{f(\xi)}{\xi - z} d\xi + \int_{-\gamma_r} \frac{f(\xi)}{\xi - z} d\xi + \int_{AB} \frac{f(\xi)}{\xi - z} d\xi + \int_{DE} \frac{f(\xi)}{\xi - z} d\xi ; \tag{5.32}$$

sur les deux segments infiniment proches AB et DE , la fonction f reprend les mêmes valeurs (puisque'elle n'est pas multiforme) ; comme ils sont parcourus en sens opposés, leurs deux contributions se compensent mutuellement et il reste :

$$f(z) = \frac{1}{2i\pi} \int_{\gamma_R} \frac{f(\xi)}{\xi - z} d\xi - \frac{1}{2i\pi} \int_{\gamma_r} \frac{f(\xi)}{\xi - z} d\xi \equiv f_+(z) + f_-(z) ; \tag{5.33}$$

le signe $-$ devant la deuxième intégrale vient du sens de parcours.

Pour la première intégrale $f_+(z)$ (celle sur le grand cercle), on peut refaire exactement la même gymnastique qu'en (5.6) pour la série de Taylor, puisque²¹ $|z - a| < |\xi - a|$, et écrire :

$$\frac{1}{\xi - z} = \frac{1}{\xi - a} + \frac{z - a}{(\xi - a)^2} + \frac{(z - a)^2}{(\xi - a)^3} + \dots + \frac{(z - a)^n}{(\xi - a)^{n+1}} + \dots . \tag{5.34}$$

²⁰ $\int_{-\gamma_r}$ désigne toujours l'intégrale le long du petit cercle γ_r parcouru dans le sens trigonométrique négatif.

²¹ z est dans la couronne K , alors que ξ est sur le grand cercle γ_R , qui est la périphérie externe de K .

Reportant ce développement dans l'expression intégrale de $f_+(z)$ (voir (5.33)), on trouve :

$$f_+(z) = \frac{1}{2i\pi} \int_{\gamma_R} f(\xi) \left[\frac{1}{\xi - a} + \frac{z - a}{(\xi - a)^2} + \frac{(z - a)^2}{(\xi - a)^3} + \dots + \frac{(z - a)^n}{(\xi - a)^{n+1}} + \dots \right] d\xi \quad (5.35)$$

Intégrant terme à terme (il y a toujours convergence uniforme), on met en évidence le développement :

$$f_+(z) = \sum_{n=0}^{+\infty} c_n (z - a)^n \quad (5.36)$$

avec la définition des coefficients c_n , $n \geq 0$:

$$c_n \stackrel{\text{déf}}{=} \frac{1}{2i\pi} \int_{\gamma_R} \frac{f(\xi)}{(\xi - a)^{n+1}} d\xi \quad (n = 0, 1, 2, 3, \dots) \quad (5.37)$$

La série (entière) (5.36) converge uniformément partout à l'intérieur du grand cercle.

Le développement (5.36) peut faire penser à un développement du genre Taylor au sens où seules des puissances positives (ou nulles) de $(z - a)$ apparaissent, mais il n'en est pas un. En effet, c_n ne peut pas être écrit comme $\frac{1}{n!} f^{(n)}(a)$ puisque, par hypothèse, $f(z)$ est supposée analytique dans K seulement, et peut donc fort bien ne pas l'être à l'intérieur du cercle de rayon r , en particulier en a , auquel cas $f^{(n)}(a)$ n'existe pas (répétons que z est strictement contraint à se trouver dans la couronne, soit $r < |z| < R$) ; avec la seule hypothèse que $f(z)$ est holomorphe dans la couronne, on ne saurait appliquer la formule de Cauchy pour $f^{(n)}$ à l'intérieur du petit cercle. On verra plus loin ce qui se passe quand $f(z)$ est holomorphe dans ce disque et que, notamment, a n'est pas un point singulier.

Pour la deuxième intégrale $f_-(z)$ de (5.33), on refait le même type de développement, avec la différence que maintenant $|z - a| > |\xi - a|$; pour faire apparaître une série géométrique convergente, il faut donc écrire :

$$\frac{1}{\xi - z} = -\frac{1}{z - \xi} = -\frac{1}{z - a - (\xi - a)} = -\frac{1}{z - a} \frac{1}{1 - \frac{\xi - a}{z - a}}, \quad (5.38)$$

d'où :

$$\frac{1}{\xi - z} = -\frac{1}{z - a} \left[1 + \left(\frac{\xi - a}{z - a} \right) + \left(\frac{\xi - a}{z - a} \right)^2 + \dots \right] = -\frac{1}{z - a} - \frac{\xi - a}{(z - a)^2} - \frac{(\xi - a)^2}{(z - a)^3} - \dots \quad (5.39)$$

Il en résulte :

$$f_-(z) = -\frac{1}{2i\pi} \int_{\gamma_r} f(\xi) \left[-\frac{1}{z - a} - \frac{\xi - a}{(z - a)^2} - \frac{(\xi - a)^2}{(z - a)^3} - \dots - \frac{(\xi - a)^{n-1}}{(z - a)^n} - \dots \right] d\xi \quad (5.40)$$

soit :

$$f_-(z) = \sum_{n=1}^{+\infty} c_{-n} (z - a)^{-n} \quad (5.41)$$

avec :

$$c_{-n} \stackrel{\text{déf}}{=} \frac{1}{2i\pi} \int_{\gamma_r} f(\xi)(\xi - a)^{n-1} d\xi \quad (n = 1, 2, 3, \dots) \quad (5.42)$$

La série (5.41) converge uniformément partout à l'extérieur du petit cercle.

En changeant n en $-n$ dans la sommation pour f_- , la somme $f = f_+ + f_-$ prend la forme :

$$f(z) = \sum_{n=-\infty}^{+\infty} c_n (z - a)^n, \quad (5.43)$$

avec $c_{n \geq 0}$ toujours donné par (5.37), cependant que $c_{n < 0}$ est donné par (5.42) où n est changé en $-n$:

$$c_n = \frac{1}{2i\pi} \int_{\gamma_r} f(\xi)(\xi - a)^{-n-1} d\xi \quad (n = -1, -2, -3, \dots). \quad (5.44)$$

La série (5.43) converge uniformément dans le domaine commun à l'extérieur du petit cercle et à l'intérieur du grand cercle : c'est bien la couronne K . À ce stade, on dispose de deux définitions distinctes pour les coefficients c_n : (5.37) pour $n \in \mathbb{N}$, et (5.44) pour $n \in -\mathbb{N}^*$. On va voir que ces deux relations peuvent en fait être réunies en une seule et même formule.

En effet quand on revient aux deux formules (5.37) et (5.44), on observe que les intégrales sont indépendantes de la variable z , laquelle en est totalement absente (et c'est bien normal : les c_n sont les *coefficients* du développement en série de $f(z)$). De surcroît, f étant holomorphe dans la couronne K , les cercles γ_R et γ_r peuvent être déformés continûment (tout en restant dans K , le domaine d'analyticité de f) sans changer la valeur des deux intégrales (5.37) et (5.42), indépendamment de la position de z . En particulier, on peut dans tous les cas prendre un seul et même contour, par exemple n'importe quel cercle situé dans la couronne (et centré en a). En définitive, les coefficients du développement (5.43) sont donnés par la formule unique suivante valable $\forall n \in \mathbb{Z}$:

$$c_n = \frac{1}{2i\pi} \int_{\gamma} \frac{f(\xi)}{(\xi - a)^{n+1}} d\xi \quad (n = 0, \pm 1, \pm 2, \pm 3, \dots) \quad (5.45)$$

où γ désigne un cercle de rayon compris entre r et R – ou n'importe quel cycle situé dans K homotope à ce dernier. On note que c_n a pour dimension $[f][a]^{-n}$, assurant que le produit $c_n(z - a)^n$ a la même dimension que f , comme il se doit.

Le développement (5.43) porte le nom de *développement de Laurent*. Il est remarquable – et se distingue du développement à la Taylor – au sens où il contient *a priori* toutes les puissances entières positives et négatives, et où les coefficients c_n ne sont pas, avec l'hypothèse de base sur $f(z)$, exprimables avec les dérivées $f^{(n)}(a)$. La série des puissances $n \geq 0$ ($f_+(z)$, (5.36)) est appelée *partie régulière* (ou parfois *partie entière*, par similitude avec série *entière*), celle des puissances négatives ($f_-(z)$, (5.41)) porte le nom de *partie principale*²².

²²À ne pas confondre avec la régularisation standard de l'intégrale $\int_{a < 0}^{b > 0} \frac{1}{x} dx$ appelée *partie principale de Cauchy*.

À ce stade, on a donc démontré le théorème suivant, dit de Laurent :

Dans toute couronne K , $r < |z - a| < R$, une fonction qui y est analytique (holomorphe) peut être représentée par sa série de Laurent (5.43), uniformément convergente dans tout domaine fermé appartenant à K

Notons que si $f(z)$ est en plus holomorphe dans le petit disque de rayon r , alors tous les coefficients $c_{n < 0}$ sont nuls (voir (5.44) – l’intégrand est holomorphe, étant de la forme $f(z)(z - a)^p$ avec $p \in \mathbb{N}$, et le contour est fermé) ; on retrouve alors un développement avec les seules puissances positives – plus de partie principale –, et de plus les $c_{n \geq 0}$ sont cette fois $\frac{f^{(n)}(a)}{n!}$, conformément à la généralisation de la formule de Cauchy pour les dérivées : au total, on récupère alors un développement de Taylor.

A contrario, on voit bien que c’est la partie principale – quand elle existe – qui contient l’information sur le fait que f n’est pas *a priori* holomorphe dans le petit disque : l’existence de puissances négatives montre que f n’est visiblement pas bornée pour $z = a$. La classification des *singularités* d’une fonction (section 5.3) se fera précisément sur la considération des caractéristiques de la partie principale du développement de Laurent centré sur un point singulier de cette fonction.

Bien évidemment, si on choisit un point d’affixe b situé dans la couronne K comme centre du développement, la fonction $f(z)$ possède un autre développement, en série entière cette fois, du genre $\sum_{n=0}^{+\infty} c'_n (z - b)^n$ avec c'_n donné par $\frac{1}{2i\pi} \int_{\gamma_b} \frac{f(\xi)}{(\xi - b)^{n+1}} d\xi$, γ_b étant une boucle située dans K et entourant l’affixe b . Les c'_n ne sont autres que $\frac{1}{n!} f^{(n)}(b)$, et il s’agit tout simplement du développement de Taylor d’une fonction au voisinage d’un point où elle est holomorphe.

Afin de faire le lien avec des connaissances antérieures, il est utile de noter que si la couronne K contient le cercle de rayon unité, le développement de Laurent centré à l’origine $\sum_{n \in \mathbb{Z}} c_n z^n$ de la fonction $f(z)$ n’est autre que la série de Fourier de la fonction $g(\theta) \stackrel{\text{déf}}{=} f(z = e^{i\theta})$. En effet, posant alors $\xi = e^{i\theta}$, la formule (5.45) donne :

$$c_n = \frac{1}{2i\pi} \int_0^{2\pi} \frac{f(e^{i\theta})}{(e^{i\theta})^{n+1}} d(e^{i\theta}) = \frac{1}{2\pi} \int_0^{2\pi} g(\theta) e^{-in\theta} d\theta, \quad (5.46)$$

où l’on reconnaît la formule classique donnant les coefficients de Fourier d’une fonction 2π -périodique.

Enfin, on voit que si M désigne le maximum du module de f sur la circonférence de rayon $\rho \in]r, R[$, prise comme cercle γ , alors $|\xi - a| = \rho$ et (5.45) permet d’écrire :

$$\boxed{|c_n| \leq \frac{1}{2\pi} \left| \int_{\gamma} \frac{M}{\rho^{n+1}} d\xi \right| = \frac{1}{2\pi} \frac{M}{\rho^{n+1}} 2\pi\rho = \frac{M}{\rho^n} \quad (n \in \mathbb{Z})} \quad (5.47)$$

Afin d’illustrer ce qui précède, donnons quelques exemples. Soit la fonction :

$$f(z) = \frac{1}{(z - a)(z - b)}, \quad (5.48)$$

où, pour fixer les idées²³, a et b sont supposés tous deux réels et positifs, avec $a < b$. Soit à trouver le développement de Laurent de $f(z)$ dans la couronne délimitée par les deux cercles centrés à l'origine et de rayons a et b ($a < |z| < b$), où la fonction est visiblement holomorphe. À l'inverse, $f(z)$ n'est pas holomorphe en $z = a$ et $z = b$: la frontière de la couronne contient donc des singularités de la fonction dont on cherche le développement de Laurent. On pourrait bien sûr utiliser la formule²⁴ (5.45) pour trouver les c_n , mais il est plus simple de procéder directement en recopiant le procédé utilisé ci-dessus pour développer la quantité $\frac{1}{z-\alpha}$. On écrit ainsi $\forall z, a < |z| < b$:

$$\begin{aligned} f(z) &= \frac{1}{b-a} \left(\frac{1}{z-b} - \frac{1}{z-a} \right) = \frac{1}{b-a} \left(-\frac{1}{b} \frac{1}{1-\frac{z}{b}} - \frac{1}{z} \frac{1}{1-\frac{a}{z}} \right) \\ &= \frac{1}{b-a} \left(-\frac{1}{b} \sum_{n=0}^{+\infty} \frac{z^n}{b^n} - \frac{1}{z} \sum_{n=0}^{+\infty} \frac{a^n}{z^n} \right), \end{aligned} \quad (5.49)$$

d'où le développement de Laurent cherché :

$$f(z) = \underbrace{\sum_{n=1}^{+\infty} \frac{a^{n-1}}{a-b} \frac{1}{z^n}}_{f_-(z)} + \underbrace{\sum_{n=0}^{+\infty} \frac{1}{(a-b)b^{n+1}} z^n}_{f_+(z)} \equiv \sum_{n=1}^{+\infty} c_{-n} z^{-n} + \sum_{n=0}^{+\infty} c_n z^n \quad (a < |z| < b), \quad (5.50)$$

obtenu directement sans utiliser l'expression intégrale (5.45) des c_n . On a ici :

$$c_{-n < 0} = \frac{a^{n-1}}{a-b}, \quad c_{n \geq 0} = \frac{1}{(a-b)b^{n+1}}. \quad (5.51)$$

Inversement, (5.50) permet d'écrire :

$$c_n = \frac{1}{2i\pi} \int_{\gamma} \frac{1}{\xi^{n+1}} \frac{1}{(\xi-a)(\xi-b)} d\xi = \begin{cases} \frac{a^{-n-1}}{a-b} & \text{si } n < 0 \\ \frac{1}{(a-b)b^{n+1}} & \text{si } n \geq 0 \end{cases}, \quad (5.52)$$

résultat que l'on pourra s'amuser à retrouver ultérieurement par application du théorème des résidus.

En un peu plus compliqué, soit la fonction :

$$f(z) \stackrel{\text{déf}}{=} e^{\frac{t}{2}(z-\frac{1}{z})}. \quad (5.53)$$

D'après le théorème de Laurent, cette fonction (paramétrée par $t \in \mathbb{C}$), possède un développement de Laurent convergent, à t fini, pour $0 < |z| < \infty$ (couronne avec r fini aussi petit que l'on veut, R fini aussi grand que l'on veut, où la fonction définie en (5.53) est visiblement holomorphe) ; les coefficients c_n dépendent du paramètre t , et on les note traditionnellement $J_n(t)$:

$$e^{\frac{t}{2}(z-\frac{1}{z})} = \sum_{n=-\infty}^{+\infty} J_n(t) z^n \quad (5.54)$$

²³ce détail ne change rien au fond.

²⁴En paramétrant le contour γ suivant $\xi = re^{i\theta}$.

Tout naturellement, la présence du facteur $e^{-\frac{t}{2z}}$ induit une partie principale contenant une infinité de termes. Les $J_n(t)$ sont les plus simples des illustriesses fonctions de Bessel²⁵, que l'on rencontre si souvent en Physique. D'après (5.45), on a :

$$J_n(t) = \frac{1}{2i\pi} \int_{\gamma} \frac{e^{\frac{t}{2}(\xi - \frac{1}{\xi})}}{\xi^{n+1}} d\xi, \quad (5.55)$$

où γ est n'importe quelle boucle entourant l'origine. En prenant pour γ le cercle centré à l'origine et de rayon unité²⁶, on obtient l'expression intégrale classique²⁷ :

$$J_n(t) = \frac{1}{2\pi} \int_0^{2\pi} e^{-in\phi} e^{it \sin \phi} d\phi \quad (n \in \mathbb{Z}) \quad (5.56)$$

Les $J_n(t)$ ne sont qu'une classe particulière des fonctions de Bessel²⁸ (fonctions de Bessel ordinaires d'indice entier). Notons qu'en posant $z = e^{i\theta}$, le développement (5.54) prend la forme d'une série de Fourier, celle de la fonction 2π -périodique $e^{it \sin \theta}$:

$$e^{it \sin \theta} = \sum_{n=-\infty}^{+\infty} J_n(t) e^{in\theta} \quad (5.58)$$

ainsi, comme l'exprime visiblement la formule (5.56), les coefficients de Fourier de cette fonction ne sont autres que les $J_n(t)$.

5.3 Classification des singularités d'une fonction

Une singularité est un point du plan complexe où, pour une fonction donnée $f(z)$, "il se passe quelque chose". Pour l'instant, on laisse de côté les fonctions multiformes, du genre z^α (α non entier), $\ln z, \dots$, ne considérant que des fonctions ayant une et une seule détermination. Par ailleurs, la classification ci-après porte sur les points singuliers *isolés* : on dit que z_0 est un point singulier isolé²⁹ de $f(z)$ si $f(z)$ est analytique partout dans un voisinage de z_0 à l'exception du point z_0 lui-même (et si on a trouvé un tel voisinage, il en existe autant qu'on veut). Autrement dit, $f(z)$ est analytique dans le voisinage pointé de z_0 . De telles singularités sont manifestement *discrètes*, séparées les unes des autres.

²⁵Friedrich BESSEL (1784-1846), astronome, célèbre notamment pour avoir calculé la trajectoire de la comète de Halley et les positions de plusieurs milliers d'étoiles, publiées dans son ouvrage *Fundamenta astronomiae* en 1818.

²⁶Alors $\xi = e^{i\theta}$, $d\xi = ie^{i\theta} d\theta$, $0 \leq \theta \leq 2\pi$.

²⁷On pourrait aussi intégrer de θ_0 à $\theta_0 + 2\pi$ sans changer le résultat, puisque n est entier.

²⁸Le plus souvent, on définit les fonctions de Bessel comme les solutions de l'équation différentielle :

$$y'' + \frac{1}{x} y' + \left(1 - \frac{\nu^2}{x^2}\right) y = 0 \quad (\nu \in \mathbb{C}). \quad (5.57)$$

Les J_n apparaissant ci-dessus sont les solutions correspondant à $\nu \in \mathbb{Z}$ ($\nu = n$), qui ont notamment la propriété d'être régulières (analytiques) à l'origine : n étant entier, elles possèdent un développement en série entière. Voir par exemple l'ouvrage de Schwartz [12].

²⁹Pour un exemple de singularité non isolée, voir la fonction définie en (5.84).

On a déjà rencontré ici et là cette notion. Par exemple, la fonction $f(z) = \frac{1}{z}$ est holomorphe dans $\mathbb{C} \setminus \{0\}$; de fait, lorsque $z \rightarrow 0$, le module de f diverge : f n'est pas bornée en $z = 0$, ce point est une singularité de f , clairement isolée ; dans tout domaine $0 < \varepsilon \leq |z| \leq R$, la fonction $\frac{1}{z}$ est holomorphe. Autre exemple : pour $f(z) = e^{\frac{1}{z}}$, le point $z = 0$ est aussi une singularité, bien pire si l'on peut dire, puisque le module de f n'a pas de limite quand $|z| \rightarrow 0$ (tout dépend³⁰ de l'argument de z).

Il a déjà été signalé que c'est la partie principale $f_-(z)$ du développement de Laurent centré sur la singularité isolée qui "code" l'information sur les singularités. De fait, c'est bien elle qui en permet la classification non ambiguë, en distinguant les trois cas : cette partie principale est identiquement nulle (est absente), ou bien elle comprend un nombre fini de termes, ou bien elle en contient une infinité. On peut ainsi précisément distinguer (pour les fonctions non multiformes) trois types de singularités isolées :

1. les points singuliers éliminables³¹

z_0 est un point singulier éliminable lorsque la limite de la fonction $f(z)$ existe quand $z \rightarrow z_0$. Par exemple, pour les fonctions :

$$\frac{\sin z}{z}, \quad \frac{\operatorname{Ln}(1+z)^2}{z}, \quad \frac{e^z - 1}{\pi z}, \quad \frac{1 - \cos z}{z^2}, \quad (5.59)$$

le point $z_0 = 0$ est une singularité éliminable, où les fonctions ont respectivement pour limite³² 1, 2, $\frac{1}{\pi}$ et $\frac{1}{2}$. Il peut aussi arriver que ce cas apparaisse spontanément à la suite d'une maladresse ou d'une inadvertance ; par exemple la fonction :

$$P(z) = \frac{z^n - z_0^n}{z - z_0} \quad (5.60)$$

présente à première vue une singularité en z_0 où le dénominateur s'annule... mais le numérateur aussi : le dénominateur divise exactement le numérateur et, division faite, $P(z)$ apparaît comme étant le polynôme de degré $n - 1$ égal à $\sum_{p=0}^{n-1} z_0^{n-p} z^p$, qui n'a aucune singularité (à distance finie).

Pour une singularité éliminable, on s'empresse de préciser la définition première de la fonction en ajoutant que, en ce point, la fonction est par définition, égale à sa limite : $f(z_0) \stackrel{\text{déf}}{=} f_0 \equiv \lim_{z \rightarrow z_0} f(z)$. Ainsi par exemple, partant de $\frac{\sin z}{z}$, on définira complètement la fonction comme suit :

$$f(z) \stackrel{\text{déf}}{=} \begin{cases} \frac{\sin z}{z}, & \forall z \neq 0 \\ 1, & \text{si } z = 0 \end{cases} . \quad (5.61)$$

³⁰Regarder par exemple ce qui se passe pour $|f(z)|$ quand z tend vers zéro en spiralant autour de l'origine...

³¹On dit aussi *singularités apparentes*, ou encore *singularités artificielles*.

³²Ces limites se trouvent en utilisant les développements en série entière de $\sin z$, $\operatorname{Ln}(1+z)$,... Pour le logarithme, il s'agit de la détermination telle que $\operatorname{Ln} 1 = 0$, et dont la coupure évite le point $z = 0$.

2. les pôles

z_0 est un pôle si $|f(z)| \rightarrow \infty$ quand $z \rightarrow z_0$; autrement dit, la limite du module de f existe, mais est infinie. Par exemple, l'origine est un pôle pour les fonctions $\frac{1}{z^3}$, $\frac{1}{\sin z}$. Pour cette dernière fonction, il y a en fait une infinité dénombrable de pôles, tous les nombres $z_k = k\pi$ où $k \in \mathbb{Z}$. De même, $\frac{1}{\cosh z}$ a une infinité de pôles $z_k = i(k + \frac{1}{2})\pi$, tout comme $\frac{1}{e^{az}-1}$, dont les pôles sont les nombres $2ik\frac{\pi}{a}$, $k \in \mathbb{Z}$.

3. les points singuliers essentiels

z_0 est un point singulier essentiel si la limite du module $|f(z)|$ n'existe pas quand $z \rightarrow z_0$. Par exemple, $z_0 = 0$ est une singularité essentielle pour la fonction $e^{1/z}$. En effet, si z tend vers zéro en venant de l'axe réel positif, la fonction diverge, plus précisément, elle tend vers $+\infty$. Si z vient du côté \mathbb{R}_- , la fonction tend vers zéro. Plus généralement, si z arrive vers l'origine avec un argument quelconque, la fonction oscille à toute vitesse, et ce d'autant plus vite que $|z|$ est petit ; notamment, si $-\frac{\pi}{2} < \arg z < +\frac{\pi}{2}$, les parties réelle et imaginaire oscillent de plus en plus vite entre deux exponentielles divergentes. Avec $z = re^{i\theta}$, on a $|e^{1/z}| = e^{(1/r)\cos\theta}$: le module n'a pas de limite quand $z \rightarrow 0$, signature d'une singularité essentielle³³. Ce comportement est général, d'après le théorème de Weierstrass - Sokhostski (voir p. 203).

Ces définitions s'illustrent immédiatement à propos du développement de Laurent suivant les puissances de $(z - z_0)$, c'est-à-dire centré³⁴ sur la singularité, qui s'écrit explicitement :

$$f(z) = \dots + \frac{c_{-n}}{(z - z_0)^n} + \dots + \frac{c_{-1}}{z - z_0} + c_0 + c_1(z - z_0) + \dots + c_n(z - z_0)^n + \dots \quad (5.62)$$

avec $0 < |z - z_0| < R$, où R est à préciser pour chaque fonction f considérée.

★ Pour une singularité éliminable (ou pas de singularité du tout), le développement de Laurent ne contient pas de puissances négatives (la partie principale $f_-(z)$ est identiquement nulle), puisque la limite de $f(z)$ en z_0 existe et est finie (c'est f_0 , posé égal à $f(z_0)$ quand on a pris le soin de compléter la définition de la fonction, comme noté ci-dessus, voir l'exemple (5.61)) :

$$\boxed{f(z) = c_0 + c_1(z - z_0) + \dots + c_n(z - z_0)^n + \dots \quad (z_0 \text{ est une singularité éliminable)}} \quad (5.63)$$

Par exemple, on a :

$$\frac{\sin z}{z} = 1 - \frac{z^2}{3!} + \frac{z^4}{5!} + \dots \quad (5.64)$$

³³Ce type de singularité n'est pas seulement une curiosité mathématique, mais survient fréquemment en Physique. Par exemple, la Mécanique quantique fait souvent apparaître des quantités du genre $e^{A/\hbar}$, où A est une constante (en général complexe) et où \hbar est la constante historique de Planck divisée par 2π . La limite classique correspond au cas où \hbar est très petit devant une action typique du problème – autrement dit, prendre la limite classique c'est "faire tendre \hbar vers zéro". On voit que la limite est ultra-singulière, dernier pied-de-nez de la Mécanique quantique au moment où on veut se passer de ses services...

³⁴Comparer avec l'exemple traité plus loin, (5.80).

Ceci est le développement de Laurent centré en zéro ($z_0 = 0$), et on voit qu'il n'y a pas de partie principale ($f_-(z) \equiv 0$). En définitive, il n'y a pas de différence sur ce plan entre une fonction holomorphe, et une fonction à singularité éliminable (d'ailleurs, l'achèvement de la définition de $f(z)$ comme en (5.61) la rend *de facto* holomorphe en z_0). Noter dès maintenant que dans ce cas le coefficient c_{-1} est évidemment nul ; on verra plus loin que c_{-1} joue un rôle de tout premier plan (voir (5.89)).

★ Pour un pôle, le module de f diverge, ce qui impose l'existence d'une partie principale non-identiquement nulle : alors, quand le rayon du petit disque délimitant la couronne devient de plus en plus petit, chacun des termes de la partie principale devient de plus en plus grand en module. Supposons que la première puissance négative est $(z - z_0)^{-n_0}$, $n_0 > 0$ (la partie principale contient donc au plus n_0 termes non nuls) : on dit alors que le pôle est d'ordre n_0 ; tous les coefficients c_n , $n < -n_0$, sont nuls :

$$f(z) = \frac{c_{-n_0}}{(z - z_0)^{n_0}} + \dots + \frac{c_{-1}}{z - z_0} + c_0 + c_1(z - z_0) + \dots \quad (z_0 \text{ est un pôle d'ordre } n_0)$$
(5.65)

Si z_0 est un pôle d'ordre n_0 , la fonction $(z - z_0)^{n_0} f(z)$ est holomorphe (analytique) en z_0 , où elle vaut c_{-n_0} .

Il est clair que si z_0 est un pôle pour $f(z)$, alors z_0 est un zéro pour la fonction $g(z) = \frac{1}{f(z)}$ et réciproquement. L'ordre n_0 du pôle de f est égal à l'ordre du zéro de $g(z)$. En effet, avec ces hypothèses, on a :

$$f(z) = \frac{1}{(z - z_0)^{n_0}} \psi(z) , \quad (5.66)$$

où $\psi(z)$ est holomorphe en z_0 , et alors :

$$g(z) = (z - z_0)^{n_0} \phi(z) \quad (5.67)$$

avec $\phi(z) = \frac{1}{\psi(z)}$ finie en z_0 . Ceci se voit aussi en revenant à (5.65) ; par comparaison :

$$\psi(z) = c_{-n_0} + c_{-n_0+1}(z - z_0) + c_{-n_0+2}(z - z_0)^2 + \dots + c_{-1}(z - z_0)^{n_0-1} + \dots \quad (c_{-n_0} \neq 0) ; \quad (5.68)$$

ainsi, le développement de Laurent de $\psi(z)$ n'a pas de partie principale, ce qui est le propre d'une fonction analytique en z_0 .

Traisons un exemple. Soit la fonction $f(z) = \frac{1}{z^2 + a^2}$, avec $a \in \mathbb{R}_+$ pour fixer les idées ; cette fonction a deux pôles simples en $\pm ia$ puisque le module de $f(z)$ a pour limite ∞ en ces deux points, alors que le produit $(z \mp ia)f(z)$ a une limite lorsque $z \rightarrow \pm ia$:

$$\lim_{z \rightarrow \pm ia} \left[(z \mp ia) \frac{1}{(z - ia)(z + ia)} \right] = \pm \frac{1}{2ia} . \quad (5.69)$$

Quel est son développement de Laurent centré en $z_0 = ia$? On voit de suite qu'il en existe un³⁵ dans la couronne K délimitée par les deux cercles centrés en ia et de rayons

³⁵Noter que l'autre pôle, $-ia$, n'appartient pas à la couronne K (il est sur sa périphérie).

r et R avec $0 < r < R < 2a$ (soit $0 < |z - ia| < 2a$). On écrit alors :

$$\frac{1}{(z - ia)(z + ia)} = \frac{1}{z - ia} \frac{1}{z - ia + 2ia} = \frac{1}{z - ia} \frac{1}{2ia} \frac{1}{1 + \frac{z-ia}{2ia}} = \frac{1}{2ia} \frac{1}{z - ia} \sum_{n=0}^{+\infty} (-1)^n \left(\frac{z - ia}{2ia}\right)^n, \quad (5.70)$$

soit, en séparant la seule puissance négative :

$$\frac{1}{(z - ia)(z + ia)} = \frac{1}{2ia(z - ia)} + \sum_{n=0}^{+\infty} \frac{(-1)^{n+1}(z - ia)^n}{(2ia)^{n+2}}; \quad (5.71)$$

la partie principale se réduit au terme $\frac{c_{-1}}{z-ia}$, comme il se doit pour un pôle d'ordre 1. Le coefficient c_{-1} du "bon" développement de Laurent jouera un rôle de tout premier plan dans la suite.

Autre exemple ; soit la fonction $f(z)$:

$$f(z) = \frac{e^z}{(z - 1)^2}, \quad (5.72)$$

holomorphe pour $1 < |z| < R \quad \forall R$ fini. En écrivant le numérateur sous la forme e^{z-1} , et en développant l'exponentielle en puissances de $z - 1$, on obtient d'emblée le développement de Laurent autour de $z_0 = 1$:

$$\frac{e^z}{(z - 1)^2} = \underbrace{\frac{e}{(z - 1)^2} + \frac{e}{z - 1}}_{f_-(z)} + e \underbrace{\sum_{n=0}^{+\infty} \frac{(z - 1)^n}{(n + 2)!}}_{f_+(z)} \quad (|z| > 1); \quad (5.73)$$

Le point $z_0 = 1$ est un pôle d'ordre 2, ce que l'on peut évidemment deviner en regardant l'expression de $f(z)$ (5.72), et compte tenu de l'analyse conduisant à (5.66). Dans ce cas, la formule (5.45) fournit les égalités³⁶ :

$$\frac{1}{2i\pi} \int_{\gamma} \frac{e^{\xi}}{(\xi - 1)^{n+3}} d\xi = \frac{e}{(n + 2)!} \quad (n = -2, -1, 0, 1, 2, \dots), \quad (5.74)$$

où γ est une boucle autour de l'affixe $z = 1$.

★ Enfin, pour une singularité essentielle, il existe une *infinité* de coefficients c_{-n} , $n > 0$, non-nuls, les $c_{n'}$, $n' \geq 0$ étant ce qu'ils sont (on dit parfois qu'une singularité essentielle est un pôle d'ordre *infini*) :

$$f(z) = \dots + \frac{c_{-n}}{(z - z_0)^n} + \dots + \frac{c_{-1}}{z - z_0} + c_0 + c_1(z - z_0) + \dots + c_n(z - z_0)^n + \dots$$

³⁶0! = 1.

$$\boxed{(z_0 \text{ est une singularité essentielle})} \quad (5.75)$$

et la partie principale contient de fait une *infinité* de termes non-nuls. C'est le cas pour la fonction $f(z) = e^{\alpha/z}$, dont le développement de Laurent (convergent $\forall |z| \in]0, +\infty[$) est³⁷ :

$$e^{\alpha/z} = \dots + \underbrace{\frac{\alpha^n}{n!} \frac{1}{z^n} + \dots + \frac{\alpha^2}{2!} \frac{1}{z^2} + \frac{\alpha}{z}}_{f_-(z)} + \underbrace{1}_{f_+(z)} \dots \quad (5.77)$$

Notons que la partie régulière est constante et vaut 1, le seul coefficient $c_{n \geq 0}$ non nul étant $c_0 = 1$. En revanche, il y a bien une infinité de coefficients $c_{n < 0}$ non-nuls (ici, tous sont différents de zéro), ce qui est caractéristique d'une singularité essentielle.

Si z_0 est une singularité essentielle, il n'existe pas d'entier $n_0 > 0$ tel que la fonction $(z - z_0)^{n_0} f(z)$ est holomorphe (analytique) en z_0 . Au voisinage d'une singularité essentielle, une fonction f prend *toutes* les valeurs complexes possibles (c'est l'essence du théorème dit de Weierstrass - Sokhostski), ce qui signifie que $|f(z)|$ n'a pas de limite quand $z \rightarrow z_0$. Plus précisément, si ξ est un nombre complexe quelconque, la différence $|f(z) - \xi|$ peut être rendue arbitrairement petite dans tout ε -voisinage de z_0 . En effet, supposons que $|f(z) - \xi| > \delta > 0$ pour tout z satisfaisant $0 < |z - z_0| < R$; avec ces hypothèses, $g(z) \stackrel{\text{déf}}{=} \frac{1}{f(z) - \xi}$ est holomorphe (et bornée) dans la couronne $0 < |z - z_0| < \varepsilon$. Si $\lim_{z \rightarrow z_0} g(z) \neq 0$, alors $f(z) \equiv \xi + \frac{1}{g(z)}$ est holomorphe en z_0 , contrairement aux hypothèses. De même, si $\lim_{z \rightarrow z_0} g(z) = 0$, $f(z) \equiv \xi + \frac{1}{g(z)}$ possède un pôle en z_0 , ce qui viole à nouveau les hypothèses. Ainsi, quel que soit $\delta > 0$, il existe au moins une valeur de z telle que $|f(z) - \xi| < \delta$.

À titre d'exemple, soit la fonction $f(z) \stackrel{\text{déf}}{=} e^{1/z}$, et soit à résoudre $e^{1/(re^{i\theta})} = \xi$. Posant $\xi = \rho e^{i\phi}$ ($\rho \neq 0$), il faut donc avoir $\frac{1}{r} \cos \theta = \ln \rho$ et $\frac{1}{r} \sin \theta = -\phi$ dans tout ε -voisinage de l'origine. L'équation $\tan \theta = -\frac{\phi}{\ln \rho}$ a toujours des solutions puisque pour $\theta \in [0, 2\pi]$, $\tan \theta$ prend toutes les valeurs entre $\pm\infty$. Il faut aussi $\frac{1}{r^2} = \ln^2 \rho + \theta^2$; comme θ peut être pris arbitrairement grand, cette équation a également une solution quel que soit ε contraignant r à être compris entre 0 et ε .

Le point à l'infini de \mathbb{C} peut être inclus parmi les singularités possibles d'une fonction en procédant comme suit. Une fonction $f(z)$ étant donnée, on introduit la fonction $\phi(Z) \stackrel{\text{déf}}{=} f\left(\frac{1}{z}\right)$ et on considère les différents cas possibles, en raisonnant avec son développement de Laurent centré en $z = 0$ (*a priori*, la substitution $Z = \frac{1}{z}$ n'est pas possible en $z = 0$, de sorte que $\phi(Z)$ n'est pas, toujours *a priori*, définie en ce point).

Si $\phi(Z)$ a un zéro d'ordre m en $Z = 0$, son développement de Laurent devient un simple développement de Taylor, de sorte que $\phi(Z) = \sum_{n=m}^{+\infty} c_n Z^n$; on peut alors

³⁷D'où, selon (5.45) :

$$\frac{1}{2i\pi} \int_{\gamma} \xi^{n-1} e^{\alpha/\xi} d\xi = \frac{\alpha^n}{n!} \quad (n = 0, 1, 2, \dots), \quad (5.76)$$

l'intégrale étant nulle ici si $n \in -\mathbb{N}^*$ (γ est une boucle autour de l'origine).

écrire :

$$f(z) = \sum_{n=m}^{+\infty} \frac{c_n}{z^n} . \quad (5.78)$$

En pareil cas, on dit que $f(z)$ a un zéro d'ordre m à l'infini. Au contraire, si $\phi(Z)$ a un pôle d'ordre m en $Z = 0$, le développement de Laurent est $\phi(Z) = \sum_{n=-m}^{+\infty} c_n Z^n$, et alors :

$$f(z) = c_{-m} z^m + c_{-(m-1)} z^{m-1} + \dots + \frac{c_{-1}}{z} + \sum_{n=0}^{+\infty} \frac{c_n}{z^n} , \quad (5.79)$$

et on dit que $f(z)$ a un pôle d'ordre m à l'infini (c'est par exemple un polynôme de degré m). Enfin, si $\phi(Z)$ a une singularité essentielle en $Z = 0$, on dit que $f(z)$ a une singularité essentielle à l'infini : c'est le cas pour e^z . En définitive, la classification donnée plus haut vaut aussi pour le point ∞ : c'est une singularité éliminable, un pôle ou une singularité essentielle selon que la limite $\lim_{z \rightarrow \infty} |f(z)|$ est finie, infinie ou n'existe pas. Le développement de Laurent d'une fonction centrée sur ∞ se déduit du développement centré en 0 en faisant la simple substitution $z \rightarrow \frac{1}{z}$, et se traduit donc essentiellement par l'échange des parties principale et entière f_{\pm} , le terme constant étant inclus là où il faut.

Noter que les développements (5.78) et (5.79) sont formellement de Laurent, mais ils sont centrés au point $z = 0$ qui n'est pas, dans chaque cas, une singularité³⁸ de $f(z)$.

◆ *Remarques*

1. Très souvent, la simple détection d'une singularité se fait "à vue", quand on dispose d'une expression fermée simple de la fonction. La construction explicite du développement de Laurent centré sur cette singularité permet de préciser la *nature* de cette dernière, par définition. Souvent, le savoir-faire permet d'éviter cette étape technique.
2. Clairement, pour une singularité *non isolée*, il n'y a pas de développement de Laurent centré sur cette singularité, puisqu'il n'existe de voisinage où elle serait seule à se trouver.
3. Il faut bien comprendre qu'une fonction donnée $f(z)$ n'a pas *un* développement de Laurent, mais autant qu'on veut, selon le choix du point servant de centre du développement et de la couronne. Par exemple, soit la fonction :

$$f(z) = \frac{1}{z-1} , \quad (5.80)$$

qui possède une seule singularité, un pôle d'ordre 1 en $z_0 = 1$. Un premier développement peut être écrit en prenant la couronne centrée à l'origine (qui n'est pas une

³⁸Au prix d'une légère anticipation (la définition du résidu est donnée plus loin, voir (5.91)), mais pour éviter les contre-sens, notons qu'il ne faut pas interpréter le coefficient c_{-1} comme le résidu en une singularité (voir aussi la Remarque 3 ci-dessous) ; le résidu à l'infini (voir (5.104)) est le coefficient de Z^{-1} dans le développement de Laurent de $Z^{-2}\phi(Z)$ centré en $Z = 0$.

singularité) de rayons $r > 1$ et $R > r$:

$$f(z) = \frac{1}{z} \frac{1}{1 - \frac{1}{z}} = \frac{1}{z} \sum_{n=0}^{+\infty} \frac{1}{z^n} = \sum_{n=1}^{+\infty} \frac{1}{z^n} \quad (|z| > 1) . \quad (5.81)$$

L'existence d'une infinité de termes de puissances négatives ne signifie pas que $z_0 = 0$ est une singularité essentielle. En effet, le développement de Laurent effectué ci-dessus est centré sur $z = 0$, où la fonction est bornée et dans le voisinage duquel elle est holomorphe : $z = 0$ n'est donc pas une singularité de f – et d'ailleurs (5.81) n'a ici de sens que pour $|z| > 1$: pas question d'y faire $z = 0$. Répétons que la classification ci-dessus se réfère au contraire au cas où précisément ce développement est *centré sur une singularité* (soupçonnée). Ainsi, le développement de Laurent centré sur l'unique singularité $z = 1$ existe et n'est autre que ...

$$f(z) = \frac{1}{z-1} \quad (!!!) \quad (5.82)$$

Tous les coefficients c_n sont nuls, sauf c_{-1} , qui vaut 1. De la même façon, un autre développement dans la couronne $0 < |z| < 1$ est :

$$f(z) = -\frac{1}{1-z} = -\sum_{n=0}^{+\infty} z^n \quad (|z| < 1) , \quad (5.83)$$

et ne doit pas faire conclure, au motif que toutes les puissances sont positives, que $f(z)$ n'a pas de pôle : il y en a bien un ($z_0 = 1$), en dehors du domaine de convergence uniforme de la série entière (5.83) (tout disque inclus dans le cercle de rayon unité)³⁹. Visiblement le second membre de (5.83) est en fait un développement de Taylor, puisqu'il n'y a pas de partie principale. Dans le même ordre d'idées, revenons à l'exemple donné en (5.48) et (5.49) : la fonction f définie en (5.48) a deux pôles a et b d'ordre 1, alors que le développement centré sur l'origine, (5.50), contient *toutes* les puissances z^n avec $n \in \mathbb{Z}$, mais n'est pas valide en $z = a$ ou b .

4. Très souvent en Physique, on utilise des techniques conduisant naturellement à l'écriture des quantités d'intérêt sous la forme d'un développement (c'est tentant quand un petit paramètre *sans dimension* a pu être mis en évidence). Dans ces conditions⁴⁰, l'identification des singularités de la fonction cherchée peut être périlleuse si l'on ne précise pas avec soin le domaine où le développement a un sens, qui implique de trancher la question de la convergence uniforme (pour cela, il n'est pas toujours nécessaire de connaître explicitement la forme du terme général).

Raisonnement en sens contraire – disposer d'un développement obtenu d'une façon ou d'une autre et lui faire dire des choses inconsidérément – peut conduire à des âneries, d'où la difficulté pour le physicien, quand il ne dispose que d'un développement

³⁹D'ailleurs, le fait que les développements (5.81) et (5.83) divergent pour $z=1$ laisse suspecter l'existence d'au moins une singularité sur le cercle unité.

⁴⁰L'idéal est évidemment de savoir resommer le développement pour identifier la fonction – ce qui est bien rare en pratique ! Souvent, un peu de savoir-faire allié à des considérations physiques, permet de localiser les singularités ou, à tout le moins, de savoir où il n'y en a pas.

qu'il ne sait pas resommer ou, pire, quand il ne connaît que des *premiers* termes d'un développement qu'il suppose exister sans en avoir en général d'autre preuve qu'une intime conviction... Les exemples ci-dessus, s'ils gardent un caractère académique dans la mesure où toutes les séries se resomment trivialement, permettent néanmoins de mettre le doigt sur la subtilité et inciter à la prudence⁴¹. En toute hypothèse, toute information *a priori* sur les propriétés analytiques de la fonction cherchée est précieuse et doit être exploitée à fond pour réduire les risques d'erreur⁴². ♦

Pour en terminer avec la classification des singularités, donnons un exemple de singularité *non-isolée*, qui y échappe par nature. La fonction :

$$f(z) = \frac{1}{e^{1/z} + 1} \quad (5.84)$$

a une infinité de pôles⁴³ puisque le dénominateur s'annule linéairement en $(z - z_k)$ pour $e^{1/z} = -1 = e^{(2k+1)i\pi}$, $k \in \mathbb{Z}$, soit pour la suite de nombres :

$$z_k = -i \frac{1}{(2k+1)\pi} . \quad (5.85)$$

Cette suite tend vers zéro quand $k \rightarrow +\infty$: l'origine est un point limite (point d'accumulation) des zéros du dénominateur qui font diverger le module de f : l'origine est bien une singularité, mais elle n'est pas isolée – dans tout disque autour de l'origine on trouve un nombre *infini* (dénombrable) de singularités (l'origine est un point-limite). Certains auteurs (Whittaker et Watson [7] par exemple) définissent (pour une fonction monovaluée⁴⁴) une singularité essentielle comme une singularité qui n'est ni éliminable, ni un pôle (lequel est, par définition, une singularité isolée) ; avec cette convention, le point d'accumulation ci-dessus est classé parmi les singularités essentielles.

Dans un genre différent, la coupure relative à une fonction multiforme est en fait une ligne *continue* de singularités⁴⁵ – lesquelles sortent également de la classification en cours : un point de branchement d'une fonction multiforme n'est manifestement pas une singularité isolée.

Selon la nature des singularités, il s'avère utile de distinguer deux classes⁴⁶ de fonctions analytiques parmi d'autres :

⁴¹On verra par la suite (chapitre 7) comment s'introduit la notion de développement asymptotique, elle-même liée au problème dit de la "*sommation des séries divergentes*". De tels développements sont d'une extrême utilité, en Physique et ailleurs.

⁴²Par exemple, on verra dans la suite comment le Principe de causalité donne aux fonctions de réponse des propriétés d'analyticit e bien pr ecises ( etre analytiques dans un certain demi-plan de \mathbb{C}) : m eme si on est incapable de calculer exactement ces fonctions de r eponse, on connaît d'avance certaines de leurs propri et es.

⁴³Chaque z_k est bien une singularit e *isol ee*, et c'est un p ole d'ordre 1.

⁴⁴c'est- a-dire non multiforme.

⁴⁵Par exemple, pour la branche du logarithme dont la coupure est le demi-axe \mathbb{R}_- , $\ln z$ a un saut  egal  a $2i\pi$ pour les deux valeurs $x \pm i0$, *quel que soit* $x < 0$. Une coupure est une ligne dense o u la fonction est discontinue – tout comme l'aimantation spontan ee d'un ferromagn etique de part et d'autre du segment d elimit e par $T = 0$ et la temp erature de Curie T_C .

⁴⁶ A nouveau, r ep etons que l'on ne consid ere toujours pas le cas de fonctions multiformes.

- les fonctions *entières*, qui n'ont aucune singularité à distance finie⁴⁷. Toute fonction entière peut être représentée par un développement en série entière :

$$f(z) = \sum_{n=0}^{+\infty} c_n z^n \quad (5.86)$$

qui converge dans le plan ouvert. e^z est une fonction entière⁴⁸, tout comme les fonctions trigonométriques (circulaires ou hyperboliques). En définitive, une fonction entière est la somme d'une série entière dont le rayon de convergence est aussi grand que l'on veut.

Une fonction entière peut être considérée comme un polynôme de degré infini ; ceci est acceptable au vu de la série entière 5.86, mais est encore plus compréhensible quand on sait qu'une fonction entière peut être représentée par un produit infini⁴⁹ (voir section 6.8.2).

- les fonctions *méromorphes*, dont les seules singularités sont des pôles, éventuellement à l'infini. Il en résulte que dans tout domaine *borné*, une fonction méromorphe ne peut avoir qu'un nombre *fini* de pôles : dans le cas contraire, il existerait une suite de pôles convergeant vers une singularité qui ne saurait être une singularité isolée. En revanche, dans tout le plan, une fonction méromorphe peut avoir une infinité de pôles : c'est le cas par exemple de $\frac{1}{\sin z}$. La somme, le produit et le quotient de deux fonctions méromorphes sont des fonctions méromorphes. Une fonction entière qui a des pôles à l'infini est une fonction méromorphe : l'intersection de ces deux ensembles de fonctions n'est pas vide.

Les fonctions méromorphes les plus simples sont les fractions rationnelles. À l'inverse, on peut considérer que la notion de fonction méromorphe généralise celle de fraction rationnelle ; cette généralisation prend tout son sens quand on démontre qu'une telle fonction peut être représentée en *série* d'éléments rationnels simples (voir sous-section 6.8.1).

5.4 Théorème des résidus

Ce théorème est probablement le plus connu de tous ceux relatifs aux fonctions d'une variable complexe : sans aucun doute, on s'en souvient encore quand on a oublié tous les

⁴⁷Selon le théorème de Liouville (voir chapitre 4), une fonction bornée qui n'a aucune singularité dans le plan fermé $\bar{\mathbb{C}}$ est constante. Une fonction ne varie que si elle a des singularités quelque part, éventuellement à l'infini, $|z| = \infty$. Pour une fonction entière, l'infini du plan complexe, c'est un peu comme le nuage de Oort pour les comètes : sans lui, il ne se passerait rien.

⁴⁸... mais n'est visiblement pas constante ! e^z diverge si z tend vers l'infini le long de l'axe réel positif : il y a bien une singularité quelque part, dont la nature sera précisée ultérieurement (e^z a une singularité essentielle à l'infini, tout comme $e^{1/z}$ a une singularité essentielle à l'origine).

⁴⁹Par exemple :

$$\sin z = z \prod_{n \in \mathbb{N}^*} \left(1 - \frac{z^2}{n^2 \pi^2} \right) \quad (5.87)$$

autres. Cette notoriété est justifiée par l'immensité de ses applications dans toutes les disciplines où un peu de mathématique est nécessaire. En Physique, notamment, il joue un rôle de tout premier plan, et vient même illustrer de façon spectaculaire certains grands principes physiques, comme le principe de causalité : on verra comment la nécessité exprimée par ce principe (les effets sont postérieurs aux causes) impose aux fonctions de réponse d'un système, *via* ce théorème⁵⁰, de posséder des propriétés d'analyticit e bien pr ecises.

5.4.1 D efinitions et recettes de calcul des r esidus

Parmi tous les coefficients du d eveloppement de Laurent centr e en z_0 d'une fonction holomorphe dans la couronne $0 < r < |z - z_0| < R$, l'un d'entre eux joue un r ole de tout premier plan : c'est c_{-1} . Pour prendre conscience de cette pr e eminence, soit une boucle γ_{z_0} tout enti ere dans la couronne et calculons $\int_{\gamma_{z_0}} f(z) dz$. Comme la s erie converge uniform ement, on a :

$$\int_{\gamma_{z_0}} f(z) dz = \sum_{n=-\infty}^{+\infty} c_n \int_{\gamma_{z_0}} (z - z_0)^n dz \quad (5.88)$$

Toutes les int egrales sur les puissances positives ou nulle sont nulles puisque la fonction $(z - z_0)^{n \in \mathbb{N}}$ est holomorphe dans \mathbb{C} . En ce qui concerne les puissances n egatives, mais $n \neq -1$, on a vu  egalement qu'elles donnent z ero puisque la primitive $\frac{1}{1+n} z^{n+1}$ n'est pas multiforme et reprend la m eme valeur au d epart et  a l'arriv ee ; enfin, quant au terme $n = -1$, on sait que son int egrale vaut $2i\pi$. Au total⁵¹ :

$$\int_{\gamma_{z_0}} (z - z_0)^n dz = \begin{cases} 0 & \text{si } n \neq -1 \\ 2i\pi & \text{si } n = -1 \end{cases}, \quad (5.89)$$

d'o u :

$$\boxed{\int_{\gamma_{z_0}} f(z) dz = 2i\pi c_{-1}} \quad (5.90)$$

ce qui montre bien la *singularit e* du coefficient c_{-1} compar e  a tous les autres. Ce caract ere exceptionnel justifie qu'on lui donne un nom pour le distinguer : on l'appelle *r esidu* de f en z_0 ; ainsi, par d efinition :

$$\boxed{\text{Res}(f, z_0) \stackrel{\text{d ef}}{=} c_{-1} = \frac{1}{2i\pi} \int_{\gamma_{z_0}} f(z) dz} \quad (5.91)$$

o u γ_{z_0} est une circonf erence $|z - z_0| = \rho$ suffisamment petite pour n'inclure aucune autre singularit e⁵², et parcourue une fois dans le sens positif, laissant donc toujours le

⁵⁰En r ealit e, c'est un peu l'histoire de la poule et de l' o uf...

⁵¹Rappelons que ce r esultat vient du fait que $\int_{\gamma} z^n dz = \int_0^{2\pi} (re^{i\theta})^n d(re^{i\theta})$, γ  tant une boucle autour de l'origine d eform ee en un cercle de rayon r . Le r esultat est $2i\pi r^{n+1} \delta_{n, -1} = 2i\pi r^0 \delta_{n, -1} = 2i\pi \delta_{n, -1}$, o u $\delta_{n, n'}$ est le symbole de Kronecker.

⁵²d'o u l'importance de distinguer entre singularit e isol ee - le cas ici - et singularit e non isol ee.

point z_0 à gauche. Par le théorème de Cauchy, quand toutes ces conditions sont satisfaites, l'intégrale ne dépend pas du rayon ρ de la circonférence décrite, laquelle pourrait d'ailleurs être remplacée par un contour fermé de forme quelconque (mais ne contenant toujours que la seule singularité z_0). On retiendra ce fait assez remarquable : l'intégration $\frac{1}{2i\pi} \int_{\gamma_{z_0}} f(z) dz$ filtre le développement de Laurent pour ne retenir que le coefficient du terme en $(z - z_0)^{-1}$. Noter que le résidu est nul si la fonction est holomorphe en z_0 (pas de singularité ou une singularité éliminable) ou si le terme $\propto \frac{1}{z-z_0}$ est absent de la partie principale : la nullité du résidu ne signifie pas que la fonction n'a pas de singularité en z_0 !

En pratique, plusieurs procédés permettent de calculer commodément le résidu d'une fonction en un pôle. Si z_0 est un pôle d'ordre 1, alors, par définition d'une telle singularité :

$$f(z) = \frac{c_{-1}}{z - z_0} + c_0 + c_1(z - z_0) + \dots \quad (5.92)$$

et une façon rapide d'avoir $c_{-1} = \text{Res}(f, z_0)$ est de prendre tout simplement la limite $\lim_{z \rightarrow z_0} [(z - z_0)f(z)]$; par (5.92), il vient :

$$\lim_{z \rightarrow z_0} [(z - z_0)f(z)] = \lim_{z \rightarrow z_0} [c_{-1} + c_0(z - z_0) + c_1(z - z_0)^2 + \dots] = c_{-1} ; \quad (5.93)$$

d'où la recette :

$$\boxed{\text{Res}(f, z_0) = \lim_{z \rightarrow z_0} [(z - z_0)f(z)] \quad (\text{pôle d'ordre 1})} \quad (5.94)$$

Pour un pôle d'ordre n , on a, toujours par définition :

$$f(z) = \frac{c_{-n}}{(z - z_0)^n} + \dots + \frac{c_{-1}}{z - z_0} + c_0 + c_1(z - z_0) + \dots \quad (5.95)$$

et la méthode précédente ne marche plus (en formant le produit $(z - z_0)f(z)$, le second membre diverge à la limite). En revanche, l'astuce suivante fonctionne ; on multiplie d'abord membre à membre (5.95) par $(z - z_0)^n$, ce qui donne :

$$(z - z_0)^n f(z) = c_{-n} + \dots + c_{-2}(z - z_0)^{n-2} + c_{-1}(z - z_0)^{n-1} + c_0(z - z_0)^n + c_1(z - z_0)^{n+1} + \dots \quad (5.96)$$

Ensuite, on dérive $n - 1$ fois par rapport à z , ce qui fait disparaître les termes contenant $c_{-n}, c_{-n+1}, \dots, c_{-2}$:

$$\frac{d^{n-1}}{dz^{n-1}} [(z - z_0)^n f(z)] = (n - 1)! c_{-1} + \frac{n!}{1!} c_0(z - z_0) + \frac{(n + 1)!}{2!} c_1(z - z_0)^2 + \dots \quad (5.97)$$

Il suffit maintenant de prendre la limite $z \rightarrow z_0$ pour que seul subsiste le terme c_{-1} au second membre. D'où la formule générale :

$$\boxed{\text{Res}(f, z_0) \equiv c_{-1} = \frac{1}{(n - 1)!} \lim_{z \rightarrow z_0} \frac{d^{n-1}}{dz^{n-1}} [(z - z_0)^n f(z)] \quad (\text{pôle d'ordre } n)} \quad (5.98)$$

Clairement, (5.94) est un cas particulier de (5.98) (il suffit d'y faire $n = 1$).

Une autre formule utile en pratique est la suivante. Quand $f(z) = \frac{\phi(z)}{\psi(z)}$ où $\psi(z)$ a un zéro simple en z_0 ($\psi(z_0) = 0$) et où $\phi(z)$ est holomorphe en z_0 , z_0 est un pôle simple de $f(z)$; d'après (5.94), on a :

$$\text{Res}(f, z_0) = \lim_{z \rightarrow z_0} \left[(z - z_0) \frac{\phi(z)}{\psi(z)} \right], \quad (5.99)$$

ce qui s'écrit aussi :

$$\lim_{z \rightarrow z_0} \left[(z - z_0) \frac{\phi(z)}{\psi(z) - \psi(z_0)} \right] = \lim_{z \rightarrow z_0} \left[\frac{\phi(z)}{\frac{\psi(z) - \psi(z_0)}{z - z_0}} \right] = \frac{\phi(z_0)}{\psi'(z_0)}. \quad (5.100)$$

Bien évidemment, si $\phi(z_0) = 0$, le résidu est nul, ce qui signifie tout simplement ici que z_0 est en fait une singularité éliminable⁵³. En partant du résultat (5.98), il est possible d'écrire des formules analogues à (5.100) pour un pôle d'ordre supérieur, donnant le résidu en fonction des dérivées de la fonction ; par exemple, avec $f(z) = \frac{1}{\psi(z)}$ où $\psi(z)$ a un zéro d'ordre 2 en z_0 , $f(z)$ y a un pôle double et on trouve :

$$\boxed{\text{Res}(f, z_0) = \lim_{z \rightarrow z_0} \frac{d}{dz} \left[(z - z_0)^2 \frac{1}{\psi(z)} \right] = -\frac{2\psi^{(3)}(z_0)}{3[\psi^{(2)}(z_0)]^2}} \quad (5.101)$$

Enfin, il arrive aussi que le développement de Laurent se trouve à vue, auquel cas le calcul du résidu est tout fait (voir par exemple la fonction définie en (5.48), ou $e^{\frac{z}{z}}$). Dans tous les cas, il est utile en pratique, pour une vérification rapide, de se souvenir que c_{-1} est homogène à $[z] \times [f]$.

On définit aussi le résidu à l'infini d'une fonction par :

$$\boxed{\text{Res}[f(z), \infty] = \frac{1}{2i\pi} \int_{-C_\infty} f(z) dz} \quad (5.102)$$

où $-C_\infty$ est un contour décrit dans le sens *négatif* de façon que le voisinage du point à l'infini soit à gauche (tout comme pour le résidu en un point à distance finie) et suffisamment grand pour ne laisser à gauche aucune singularité de $f(z)$ de module fini. En pratique, on peut calculer ce résidu en paramétrant le contour par $z = Re^{i\theta}$ et en prenant R suffisamment grand. Si la notion de résidu à l'infini est un peu troublante, on peut toujours se ramener au calcul d'un résidu à l'origine. En effet, soit une fonction $f(z)$; par définition, on a :

$$\text{Res}[f(z), \infty] = -\frac{1}{2i\pi} \int_{C_\infty} f(z) dz ; \quad (5.103)$$

⁵³ $\phi(z)$ étant supposée holomorphe en z_0 , son développement de Laurent est de la forme

$$c'_0 + c'_1(z - z_0) + \dots$$

Dans ces conditions, le résidu de $f(z)$ est égal à $\frac{c_0}{\psi'(z_0)}$; si $c_0 = 0$, donnant $\phi(z_0) = 0$, alors le point z_0 est de fait une singularité apparente.

posons $Z = \frac{1}{z}$, $\phi(Z) \stackrel{\text{déf}}{=} f(z = \frac{1}{Z})$, il vient⁵⁴ :

$$\text{Res}[f(z), \infty] = -\frac{1}{2i\pi} \int_{C_0} \phi(Z) \frac{-dZ}{Z^2} = \frac{1}{2i\pi} \int_{C_0} \frac{1}{Z^2} \phi(Z) dZ = \text{Res}[Z^{-2}\phi(Z), 0] . \quad (5.104)$$

5.4.2 Démonstration du théorème des résidus

Pour démontrer le théorème des résidus dans le cas le plus simple, considérons une fonction $f(z)$ holomorphe dans un domaine simplement connexe \mathcal{D} , à l'exception d'un seul point singulier isolé z_0 . $f(z)$ est donc analytique dans toute couronne dont la frontière extérieure est celle de \mathcal{D} , et dont l'autre est une boucle isolant le point singulier z_0 . Soit l'intégrale de $f(z)$ sur un contour fermé Γ quelconque, inclus dans \mathcal{D} et ne contenant pas z_0 (donc appartenant à la couronne). D'après le théorème de Cauchy, on a :

$$\int_{\Gamma} f(z) dz = 0 . \quad (5.105)$$

Maintenant, déformons le contour *ad libitum*, $\Gamma \rightarrow \Gamma' \rightarrow \Gamma''$ (voir fig. 5.3) tout en lui faisant éviter le point z_0 – en vertu du théorème de Cauchy, la valeur de l'intégrale ne change pas et est toujours égale à zéro :

$$\int_{\Gamma} f(z) dz = \int_{\Gamma'} f(z) dz = \dots = 0 . \quad (5.106)$$

En particulier, on peut en venir au contour Γ'' , formé d'une petite circonférence γ_{z_0} au-

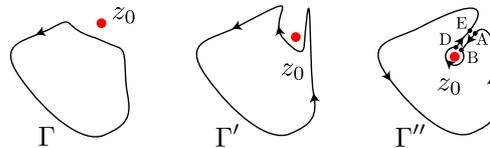


Figure 5.3: Le contour Γ et ses avatars évitant soigneusement la singularité z_0 .

tour de z_0 , des deux petits segments de droite⁵⁵ AB et DE infiniment proches (parcourus en sens opposés) et d'une ligne C "refermant le tout". Les contributions des deux petits segments se compensent, et il reste⁵⁶ :

$$\int_C f(z) dz + \int_{-\gamma_{z_0}} f(z) dz = 0 \iff \int_C f(z) dz = \int_{\gamma_{z_0}} f(z) dz . \quad (5.107)$$

⁵⁴Si z décrit le contour à l'infini dans le sens négatif, Z^{-1} décrit dans le sens positif un cercle de rayon infiniment petit autour de l'origine.

⁵⁵qui pourraient d'ailleurs être tout autant deux petits arcs de forme quelconque mais, pour la démonstration, toujours infiniment proches l'un de l'autre.

⁵⁶Les deux points A et E sont en fait confondus, de sorte que le "grand" contour C est bien fermé, l'adjonction ou la suppression du point A (ou E) ne changeant rien à la valeur de l'intégrale de ligne puisque f est continue.

Par la définition du résidu (5.91), cette dernière équation se lit :

$$\int_C f(z) dz = 2i\pi \operatorname{Res}(f, z_0) . \quad (5.108)$$

Ainsi, l'intégrale sur le contour C est égale à l'intégrale sur la (petite) circonférence γ_{z_0} .

Le point remarquable est que le calcul de l'intégrale sur C se ramène au calcul de quantités *locales* : le résidu en z_0 , valeur de l'intégrale le long d'une circonférence (ou clairement le long de toute boucle) éventuellement infiniment petite entourant z_0 . On retrouve à nouveau l'invariance de l'intégrale quand le contour est déformé en restant dans le domaine d'holomorphic : en étirant la petite circonférence, on peut la superposer à C et inversement ; ce résultat majeur est illustré sur la fig. 5.4. Notons que les deux boucles C et γ sont parcourues dans le *même* sens, et que si le sens de parcours est inversé, alors toutes les intégrales changent de signe.

Ce résultat se généralise immédiatement au cas d'un nombre fini, N , de singularités z_k , $k = 1, 2, \dots, N$: pour chacune d'entre elles, on se livre aux mêmes déformations du contour que pour z_0 ci-dessus. Dès lors, on peut énoncer précisément le théorème des résidus (Cauchy, 1825) :

Soit $f(z)$ une fonction holomorphe sur la frontière C d'un domaine simplement connexe \mathcal{D} et partout holomorphe dans ce domaine sauf peut-être en un nombre fini de points z_k , $k = 1, 2, \dots, N$. Alors :

$$\int_C f(z) dz = 2i\pi \sum_{k=1}^N \operatorname{Res}(f, z_k) , \quad (5.109)$$

où la frontière C est décrite de sorte que le domaine \mathcal{D} se trouve à gauche dans le sens de parcours

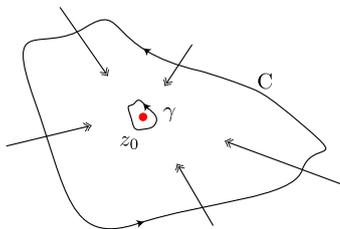


Figure 5.4: Illustration du théorème des résidus dans le cas d'une seule singularité z_0 . On peut étirer le petit cercle γ pour le superposer à la grande boucle C , ou inversement, sans changer la valeur de l'intégrale.

On verra que ce théorème est d'une extrême utilité en pratique, et quelques exemples d'application immédiate seront données au chapitre 6. Quoi qu'il en soit, il

exprime d'une nouvelle façon ce fait majeur : s'agissant des fonctions holomorphes, la forme précise du contour d'intégration importe peu (dans le cas d'une seule singularité z_0 , l'intégrale sur le grand contour C est égale à l'intégrale sur tout petit cercle⁵⁷ γ_{z_0}). On peut le déformer à souhait à deux conditions : ne pas le faire décoller du plan et éviter les singularités, ce qui, pour un mouvement dans le plan, revient à les contourner. Rappelons une fois encore l'image utile déjà mentionnée : le plan \mathbb{C} est une planche en bois, les singularités sont des clous à demi-enfoncés, le contour est un élastique dont les extrémités sont fixées par deux punaises, mais qui peut être déformé sans perdre le contact avec la planche, les clous se chargeant de lui faire obstacle pour contourner les singularités. Il est clair que pour pouvoir jouer à ce petit jeu, l'identification exhaustive des singularités est un préalable absolu à tout calcul.

D'un autre côté, rien d'autre n'est requis en ce qui concerne la nature des singularités (pôles, singularités essentielles, ...). Qu'il s'agisse d'un point singulier éliminable (alors le résidu est évidemment nul), d'un pôle (quel que soit son ordre) ou d'une singularité essentielle, le théorème est vrai⁵⁸. Une façon de bien s'en convaincre est de revenir au développement de Laurent centré sur la singularité et de reconsidérer l'égalité (5.89).

À titre d'illustration, considérons la fonction $f(z) = z^{n \in \mathbb{Z}} e^{1/z}$, qui a une singularité essentielle en $z = 0$, quel que soit l'entier n . Son développement de Laurent centré à l'origine s'obtient sans calcul :

$$f(z) = z^n \sum_{p \in \mathbb{N}} \frac{1}{p!} \left(\frac{1}{z}\right)^p = \sum_{p \in \mathbb{N}} \frac{1}{p!} z^{p-n} = \sum_{m=-\infty}^n \frac{1}{(n-m)!} z^m, \quad (5.110)$$

sur lequel on voit que le coefficient c_{-1} de $\frac{1}{z}$ est égal à $\frac{1}{(n+1)!}$ si $n \geq -1$, à zéro si $n < -1$, donc que $\text{Res}(f, 0) = \frac{1}{(n+1)!}$ ou 0. Le théorème des résidus permet donc d'écrire :

$$I_n \stackrel{\text{dét}}{=} \int_C z^n e^{1/z} dz = \frac{2i\pi}{(n+1)!} \quad (n \geq -1), \quad (5.111)$$

et $I_n = 0$ si $n < -1$, C étant n'importe quel contour fermé entourant l'origine et parcouru dans le sens positif. À titre d'exercice, ce résultat peut être retrouvé directement de diverses façons ; par exemple, prenons pour C le cercle γ de rayon R centré à l'origine, sur lequel $z = Re^{i\theta}$, θ variant sur un intervalle d'amplitude 2π , de 0 à 2π par exemple :

$$I_n = \int_0^{2\pi} (Re^{i\theta})^n e^{\frac{1}{R}e^{-i\theta}} d(Re^{i\theta}). \quad (5.112)$$

En développant en série l'exponentielle $e^{\frac{1}{R}e^{-i\theta}}$ et en s'appuyant sur la convergence uniforme, on obtient :

$$I_n = i \sum_{q \in \mathbb{N}} \frac{R^{n-q+1}}{q!} \int_0^{2\pi} e^{i(n-q+1)\theta} d\theta. \quad (5.113)$$

⁵⁷ou toute petite boucle.

⁵⁸Dans le cas des points de branchement, le théorème n'est plus applicable, évidemment, mais les règles du jeu de l'élastique restent en vigueur : défense de franchir la coupure, tant que... (voir note 59)

En utilisant maintenant $\int_0^{2\pi} e^{im\theta} d\theta = 2\pi\delta_{m,0}$:

$$I_n = i \sum_{q \in \mathbb{N}} \frac{R^{n-q+1}}{q!} \delta_{q, n+1} = i \times 2\pi \frac{1}{(n+1)!} \quad (5.114)$$

si $n+1 \geq 0$, $I_n = 0$ autrement, en accord avec (5.111) donné directement par le théorème des résidus.

Le théorème des résidus entraîne que pour une fonction ayant un nombre fini, N , de singularités z_n toutes de module fini, la somme de tous les résidus, y compris le résidu à l'infini, est nulle. En effet, l'intégrale le long d'un contour parcouru dans le sens positif et englobant les N singularités est égale à $2i\pi \sum_{n=1}^N \text{Res}(f, z_n)$; mais l'intégrale en question n'est autre, par définition, que $-2i\pi \text{Res}(f, \infty)$, d'où :

$$\boxed{\text{Res}(f, \infty) + \sum_{n=1}^N \text{Res}(f, z_n) = 0} \quad (5.115)$$

Cette relation est utile en pratique : si le contour d'une intégrale à calculer englobe précisément (presque) toutes les singularités finies, mieux vaut calculer d'emblée le résidu à l'infini, plutôt que de déterminer tous les résidus et d'en faire la somme.

◆ *Remarques*

1. Le théorème des résidus s'applique exclusivement dans le cas de singularités *isolées*, puisque d'un bout à l'autre on s'appuie sur le théorème de Laurent. Prenons par exemple le cas d'une fonction ayant un point de branchement en z_0 ; de ce point part une ligne continue infranchissable par l'élastique⁵⁹. Partant de n'importe quelle boucle laissant la coupure à l'extérieur, on peut évidemment la déformer et la faire s'approcher aussi près que l'on veut de la coupure, mais elle ne devra jamais sauter par-dessus. Si la coupure va jusqu'à l'infini, il restera toujours deux demi-“droites” (de part et d'autre de la coupure) où, précisément, la fonction prend des valeurs distinctes. Si la coupure est un segment de ligne, on peut certes envisager un contour d'intégration qui le contient en son intérieur, mais on ne pourra jamais faire s'effondrer le contour comme on l'a fait sur la fig. 5.4 (voir fig. 5.5)
2. Comme on le verra par la suite, il arrive souvent que des portions du contour d'intégration se situent à l'infini (par exemple pour la sommation de séries, voir chapitre 6). L'application du Théorème des résidus se fait alors en deux temps :
 - (a) on raisonne avec un contour d'extension *finie*, et on applique le Théorème des résidus. Dans un tel contour, et s'agissant de singularités *isolées*, il ne peut y en avoir qu'un nombre fini, d'où l'application sans souci du théorème.

⁵⁹... tant que l'on ne connaît que le premier feuillet de Riemann, voir plus loin.

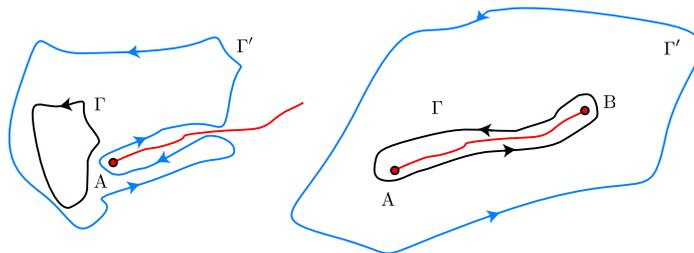


Figure 5.5: Illustration, en présence d’une coupure (demi-ligne allant de A à l’infini, ou segment de ligne AB), des déformations possibles du contour d’intégration ($\Gamma \rightarrow \Gamma' \rightarrow \dots$) qui ne changent la valeur de l’intégrale. La coupure interdit de réduire le contour à une boucle aussi petite que l’on veut.

- (b) On effectue le passage à la limite du contour infini, en examinant éventuellement les questions de convergence, et en restant attentif à la possibilité de singularités à l’infini (voir chapitre 6). Cette procédure constitue de fait, au cas par cas, l’extension du théorème des résidus à un nombre infini *dénombrable* de singularités.

Par exemple, l’intégrale $\int_{-\infty}^{+\infty} \frac{1}{\cosh x} dx = \pi$ se calcule par résidus⁶⁰ en fermant par un demi-cercle de rayon R et prenant la limite $R \rightarrow +\infty$; dans cette limite, la contribution du demi-cercle est nulle et la somme de résidus devient une *série* dont la somme est égale à π . \blacklozenge

5.5 Prolongement analytique

La notion de prolongement analytique est simple à appréhender. Soit une fonction holomorphe $f(z)$ définie dans un domaine \mathcal{D}_1 du plan complexe par une certaine relation d’égalité (expression, série, intégrale, ...) dénuée d’ambiguïté. On peut alors se poser la question suivante : dans quelle mesure, la fonction $f(z)$ ainsi définie (avec les contraintes inhérentes à l’expression première de f) existe-t-elle en fait dans un domaine \mathcal{D}_2 distinct de \mathcal{D}_1 , et englobant tout ou partie⁶¹ de \mathcal{D}_1 ? Autrement dit, dans quelle mesure peut-on *étendre* le domaine de définition de la fonction au-delà du domaine apparaissant “naturellement” lors de la première définition ?

L’intérêt de cette question n’est pas purement d’ordre mathématique. Il est fréquent en Physique de devoir étendre le domaine “naturel” d’existence de fonctions représentant certaines grandeurs, ou paramétrées par des quantités dont il est nécessaire

⁶⁰On sait aussi la calculer par des moyens élémentaires, en la prenant comme $2 \int_0^{+\infty} \frac{2e^{-x}}{1-e^{-2x}} dx$, en développant en série entière par rapport à e^{-2x} , obtenant alors la série entière de $4\text{Arctg } 1 = \pi$.

Le calcul par résidus introduit une somme partielle de série non convergente au sens usuel, qui doit être régularisée en suivant la prescription d’Euler.

⁶¹autrement, la question est sans intérêt.

de les généraliser à des valeurs un peu absconses *a priori*. Par exemple, on définira au chapitre 7 la fonction Γ d'Euler, $\Gamma(z)$, dont les valeurs aux points entiers ne sont rien d'autre que celles de la factorielle définie à un niveau élémentaire ($\Gamma(n+1) = n!$, $n \in \mathbb{N}$). La Mécanique statistique a parfois besoin des caractéristiques géométriques (volume, surface, ...) de la sphère dans \mathbb{R}^D (avec $D \sim$ le nombre d'Avogadro \mathcal{N}) – et c'est justement la fonction $\Gamma(z)$ que l'on retrouve⁶² ; avec des nombres aussi faramineux, il est prévisible que l'on aura besoin, physiquement, de généraliser la notion de factorielle à des valeurs non-entières. En plus élaboré, la description quantique d'un état instable (atome à l'état excité, particule éphémère, résonance au sens large, ...) se fait très commodément par prolongement analytique dans le domaine des énergies complexes de la bonne quantité (le propagateur ou l'un de ses avatars) ; on trouve alors tout naturellement que la durée de vie τ est (à des détails près) essentiellement donnée par l'inverse de la partie imaginaire⁶³ de l'"énergie" E , $\tau \simeq \frac{\hbar}{|\Im E|}$ (un état à durée de vie infiniment longue correspond à une "énergie" réelle). Les mêmes arguments, appliqués à l'énergie libre (de Helmholtz par exemple), permettent de trouver la durée de vie d'un état métastable (un liquide surfondu) ; inversement, on passe du propagateur $U(t, t_0)$ de la Mécanique quantique à une fonction de partition $Z(\beta)$ de la Mécanique statistique en "passant en temps imaginaire" ($t \rightarrow -i\hbar\beta$). Un dernier exemple, célèbre, est l'étude des phénomènes critiques, où l'on est conduit à considérer des dimensions spatiales D non-entières, parfois négatives, quand elles ne sont pas tout simplement ... complexes ; le succès des "développements en $4-\varepsilon$ " pour le calcul des exposants critiques est la preuve de l'utilité d'un prolongement analytique par rapport à la dimension spatiale. Si ceci est très contre-intuitif, il est aisé d'en comprendre la contingente nécessité : connaître la valeur numérique d'une certaine fonction f pour la dimensionnalité $D = 3$ est une information bien pauvre comparée à la connaissance de la fonction $f(D)$ dans le voisinage de $D = 3$, grand ou petit, voisinage qui a parfaitement le droit d'être un *disque* du plan complexe centré sur le point réel $D = 3$. Dit encore plus schématiquement : connaître la valeur d'une fonction en un point ne dit rien sur ce que vaut sa dérivée !

Pour préciser les choses sur le plan conceptuel, donnons des exemples précis n'impliquant que des fonctions élémentaires. Soit la fonction $f_1(z)$ définie comme étant égale à la somme de la série géométrique :

$$f_1(z) \stackrel{\text{déf}}{=} \sum_{n \in \mathbb{N}} z^n . \quad (5.116)$$

On sait bien que cette série ne converge que si $|z| < 1$ – elle n'est autre qu'une certaine série de Taylor ; autrement dit, $f_1(z)$, en tant que fonction de z , n'est définie par la relation (5.116) qu'à l'intérieur du disque unité centré à l'origine. Une écriture plus précise (et plus correcte !) est donc :

$$f_1(z) = \sum_{n \in \mathbb{N}} z^n \quad \forall z, |z| < 1 . \quad (5.117)$$

⁶²D'où l'intérêt de la formule de Stirling, grâce à laquelle on peut aisément démontrer un résultat majeur : l'entropie est une grandeur extensive.

⁶³Les détails ont parfois leur importance : ils sont ici responsables d'un déclin *algébrique* (par opposition à *exponentiel*) aux temps longs ; l'émergence de lois-puissances ($\propto t^{-\alpha}$) est le signe d'une relaxation *lente* dans la phase ultime.

Soit maintenant la fonction $f_2(z)$:

$$f_2(z) \stackrel{\text{déf}}{=} \frac{1}{1-z}, \quad (5.118)$$

qui est définie partout dans \mathbb{C} sauf au point $z = 1$. Quelle est la relation entre les deux fonctions f_1 et f_2 ? On sait bien que la somme de la série (5.116) est égale à $(1-z)^{-1}$, autrement dit la somme de la série (5.116) – qui n'existe que pour $|z| < 1$ – est égale partout où elle est définie à la fonction f_2 qui existe, elle, dans un domaine beaucoup plus grand de \mathbb{C} . On peut dire que la fonction $f_2(z)$ est le *prolongement* de la fonction $f_1(z)$ apparaissant initialement en tant que somme de la série géométrique (5.116) quand elle converge, évidemment. Ce prolongement est sensé dans et au-delà du disque unité (sauf en $z = 1$) ; comme $f_2(z)$ est une fonction analytique, sauf en $z = 1$, on dit qu'elle est le prolongement *analytique* de $f_1(z)$.

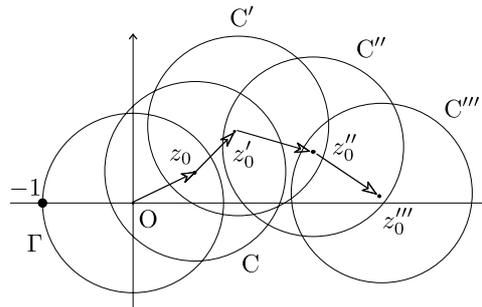


Figure 5.6: Illustration des prolongements analytiques successifs de $f_1(z)$, (5.116), en procédant de proche en proche par des disques de rayon unité (le premier disque C est centré en z_0 et correspond à (5.123)). Ceci permet de comprendre que la fonction peut bien être prolongée dans tout le plan et que son expression compacte n'est autre que $\frac{1}{1-z}$, holomorphe dans $\mathbb{C} \setminus \{1\}$.

D'ailleurs, on peut aussi prolonger directement la fonction définie par la série (5.116) – en feignant d'ignorer que l'on sait la resommer et qu'elle vaut $\frac{1}{1-z}$ – en trouvant son développement de Laurent autour de tout point z_0 judicieusement choisi, et donnant lieu à un domaine de convergence recouvrant au moins partiellement⁶⁴ le domaine initial $|z| < 1$. Récrivons (5.116) sous la forme :

$$f_1(z) = \sum_{n=0}^{+\infty} (z - z_0 + z_0)^n = \sum_{n=0}^{+\infty} \sum_{p=0}^n C_n^p z_0^{n-p} (z - z_0)^p, \quad (5.119)$$

z_0 étant n'importe où dans le disque Γ de rayon 1 centré en O . Dans le domaine commun aux deux disques $|z| < 1$ et $|z - z_0| < 1$, toutes les séries rencontrées convergent absolument (et uniformément) et on peut échanger les ordres de sommation :

$$f_1(z) = \sum_{p=0}^{+\infty} (z - z_0)^p \sum_{n=p}^{+\infty} C_n^p z_0^{n-p} = \sum_{p=0}^{+\infty} (z - z_0)^p \sum_{n=0}^{+\infty} C_{n+p}^p z_0^n. \quad (5.120)$$

⁶⁴La nécessité d'un tel recouvrement sera précisée plus loin (voir l'énoncé au dessus de (5.129)).

En utilisant⁶⁵ :

$$\frac{1}{(1-z)^{p+1}} = \sum_{n=0}^{+\infty} C_{n+p}^p z^n \implies \sum_{n=0}^{+\infty} C_{n+p}^p z_0^n = \frac{1}{(1-z_0)^{p+1}}, \quad (5.122)$$

(5.120) devient :

$$f_1(z) = \sum_{p=0}^{+\infty} \frac{1}{(1-z_0)^{p+1}} (z-z_0)^p, \quad (5.123)$$

qui est le développement de Laurent⁶⁶ de f_1 centré cette fois sur z_0 (sans surprise, la partie principale est nulle). La série ainsi obtenue converge dans le disque $|z-z_0| < |1-z_0|$, qui étend le domaine de définition de f_1 : ici encore, il s'agit d'un prolongement analytique. La procédure peut d'ailleurs être poursuivie, élargissant de proche en proche le domaine de définition de f_1 (voir fig. 5.6). On vérifiera que toutes les expressions obtenues, une fois resommées (ici on sait resommer !) donnent toutes $\frac{1}{1-z}$, comme il se doit.

Comme second exemple, soit l'intégrale⁶⁷ :

$$f_1(z) \stackrel{\text{déf}}{=} \int_0^{+\infty} e^{-zt} dt. \quad (5.124)$$

t étant toujours positif, cette intégrale n'existe⁶⁸ que pour $\Re z > 0$ (sinon l'intégrand diverge exponentiellement à l'infini). Une écriture plus correcte pour définir la fonction f_1 dans (5.124) est donc :

$$f_1(z) = \int_0^{+\infty} e^{-zt} dt \quad \forall z, \Re z > 0. \quad (5.125)$$

Autrement dit, l'égalité (5.124) ne définit la fonction $f_1(z)$ que dans le demi-plan de droite ouvert (nombres complexes à partie réelle *strictement* positive). Cette condition étant satisfaite, une intégration élémentaire donne $f_1(z) = \frac{1}{z}$. Au total :

$$f_1(z) = \frac{1}{z} \quad \forall z, \Re z > 0. \quad (5.126)$$

Soit maintenant la fonction :

$$f_2(z) \stackrel{\text{déf}}{=} \frac{1}{z} \quad \forall z, z \neq 0. \quad (5.127)$$

⁶⁵L'égalité (5.122) est le développement de Taylor en $x = 0$ de $(1-x)^{-(p+1)}$, construit en remarquant que :

$$\frac{d^n}{dz^n} \frac{1}{(1-z)^{p+1}} = \frac{(p+1)(p+2)\dots(p+n)}{(1-z)^{p+1}} = \frac{(n+p)!}{p!(1-z)^{p+1}} = n! C_{n+p}^p \frac{1}{(1-z)^{p+1}}; \quad (5.121)$$

la série entière (5.121) converge (uniformément) $\forall z, |z-1| > 1$.

⁶⁶On peut aussi trouver ce développement en utilisant la formule (5.45), et prenant pour $f(\xi)$ la série (5.116).

⁶⁷Comme on le verra, $f_1(z)$ est la transformée de Laplace de la fonction qui vaut partout 1 sur le demi-axe réel positif.

⁶⁸Avec $z = x + iy$, on a $|f_1(z)| \leq \int_0^{+\infty} |e^{-xt} e^{iyt}| dt = \int_0^{+\infty} e^{-xt} dt$; cette dernière intégrale existe $\forall x \in \mathbb{R}_+$.

La fonction f_2 , définie partout sauf en $z = 0$, réalise le prolongement de f_1 bien au-delà du demi-plan de droite – en fait dans tout le plan excepté le point $z = 0$. De la même façon, on dit que $f_2(z)$ est le prolongement analytique de f_1 dans $\mathbb{C} \setminus \{0\}$.

Donnons maintenant un autre argument permettant de se convaincre que l'intégrale (5.124) définit bien une fonction prolongeable au-delà du demi-plan de droite. L'expression (5.124) peut être considérée comme une intégrale sur la variable *complexe* t sur le chemin C formé du demi-axe réel positif. Grâce au théorème de Cauchy, ce chemin peut être déformé en maintenant ses extrémités fixes ; en particulier, l'intégrale ne change pas si C est déformé en une demi-droite inclinée par rapport à l'axe Ox et rejoignant le point à l'infini sur \mathbb{R}_+ par un arc de cercle lui-même situé à l'infini. z étant fixé (et donc son argument aussi), la contribution de l'arc à l'infini est nulle si $|\text{Arg}zt| < \frac{\pi}{2}$, et l'intégrale de départ est égale à l'intégrale le long de la demi-droite inclinée (qui converge aussi puisque $\Re(zt) > 0$). Ainsi, quand l'argument de z augmente et va dépasser $\frac{\pi}{2}$ pour sortir de la région où la définition (5.125) a un sens, il suffit d'incliner simultanément la droite d'intégration au-dessous de l'axe Ox pour ne pas violer la dernière inégalité et récupérer *in fine* une intégrale convergente. Modifier directement le contour d'intégration est un procédé usuel pour prolonger analytiquement à vue une intégrale. On en rencontrera d'autres exemples dans la suite.

Ces deux exemples montrent bien qu'une relation donnée, valide dans un certain domaine, produit l'expression d'une fonction qui peut en réalité avoir un sens (être définie) dans un domaine beaucoup plus vaste. Par définition, on dira qu'une fonction $f_2(z)$ holomorphe dans \mathcal{D}_2 est le prolongement analytique d'une fonction $f_1(z)$ holomorphe dans \mathcal{D}_1 ssi il existe une fonction analytique $f(z)$ égale à $f_i(z)$ si $z \in \mathcal{D}_i$:

$$f(z) = \begin{cases} f_1(z) & \forall z \in \mathcal{D}_1 \\ f_2(z) & \forall z \in \mathcal{D}_2 \end{cases} . \quad (5.128)$$

Cela étant, le théorème suivant énonce les conditions permettant d'affirmer qu'une fonction f_2 est en effet le *prolongement analytique* d'une fonction f_1 :

Soit deux domaines \mathcal{D}_1 et \mathcal{D}_2 dont les frontières possèdent un arc commun Γ et soit deux fonctions f_1 et f_2 analytiques respectivement dans \mathcal{D}_1 et \mathcal{D}_2 . Si ces deux fonctions sont continues sur $\mathcal{D}_1 \cup \Gamma$ et $\mathcal{D}_2 \cup \Gamma$, et si elles prennent les mêmes valeurs en tout point de l'arc commun Γ , alors f_2 (resp. f_1) est le prolongement analytique de f_1 (resp. f_2) dans le domaine \mathcal{D}_2 (resp. \mathcal{D}_1).

En effet, soit la fonction précisément définie comme :

$$f(z) \stackrel{\text{d\u00e9f}}{=} \begin{cases} f_1(z) & \forall z \in \mathcal{D}_1 \\ f_1(z) = f_2(z) & \forall z \in \Gamma \\ f_2(z) & \forall z \in \mathcal{D}_2 \end{cases} \quad (5.129)$$

qui est une fonction continue sur $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$. On va montrer que son intégrale le long de tout contour fermé situé dans \mathcal{D} est nulle, ce qui, compte tenu du théorème de Morera, permet de conclure que f est analytique dans \mathcal{D} .

Il suffit de considérer deux contours C_1 et C_2 parcourus dans le même sens (positif par exemple), respectivement situés dans \mathcal{D}_1 et \mathcal{D}_2 et ayant un arc longeant de l'intérieur de chaque domaine l'arc commun Γ . Comme chaque fonction f_j est analytique dans son \mathcal{D}_j , on a :

$$\int_{C_j} f_j(z) dz = 0 \quad (j = 1, 2) . \quad (5.130)$$

Il en résulte notamment :

$$\int_{C_1} f_1(z) dz + \int_{C_2} f_2(z) dz = 0 \iff \int_{C_1 \cup C_2} f(z) dz = 0 . \quad (5.131)$$

Les deux contributions le long de l'arc commun sont manifestement de signes opposés, quel que soit le sens de parcours. Il en résulte que la contribution de l'arc commun disparaît du total et que seule subsiste l'intégrale le long du contour C formé de la réunion des C_j d'où Γ a été effacé :

$$\int_C f(z) dz = 0 . \quad (5.132)$$

L'intégrale de $f(z)$ le long du contour fermé C inclus dans \mathcal{D} est donc nulle. D'après le théorème de Morera (voir p.178), f est donc une fonction analytique dans \mathcal{D} . Ceci montre que le prolongement par *continuité* d'une fonction est une fonction analytique : non seulement il étend la définition de la fonction de départ à un domaine plus vaste, mais encore le résultat est une fonction douée d'analyticit⁶⁹.

Quand les deux domaines \mathcal{D}_j se recouvrent, et que les fonctions sont égales entre elles dans la région commune, on dit souvent que l'une réalise le prolongement analytique *immédiat* de l'autre. Revenons aux deux exemples ci-dessus, où on a su construire d'instinct le prolongement analytique, ce qui est une façon de se convaincre qu'il est en soi possible. Clairement, les conditions d'application théorème du prolongement analytique sont satisfaites, puisque les deux domaines ont à chaque fois bien plus qu'un arc de frontière en commun, le plus petit étant complètement inclus dans l'autre. Il s'agit donc bien ici de prolongement analytique immédiat.

Fondamentalement, la possibilité du prolongement résulte du fait que sur la frontière *fermée* de définition de f_1 , la fonction prolongée ne possède qu'un nombre dénombrable de singularités (pour les exemples ci-dessus, il n'y en a qu'une : $z = 1$ pour la série géométrique, la frontière étant le cercle de rayon unité, $z = 0$ pour l'intégrale de Laplace (5.124), la frontière étant l'axe imaginaire). Ce qui veut dire qu'un piéton peut toujours sortir de \mathcal{D}_1 en contournant toute singularité qu'il rencontrerait sur son chemin, ou encore que l'ensemble des singularités est plein de vide, particulièrement lacunaire (comme la matière ordinaire). L'image que l'on peut retenir : il doit être possible de se *faufiler* entre les singularités et sortir de \mathcal{D}_1 .

⁶⁹Notons que ce théorème n'est pas vrai pour des fonctions réelles : soit deux fonctions dérivables sur deux intervalles $[a, b]$ et $[b, c]$, égales au point commun b . Leur réunion n'est pas forcément dérivable (il peut y avoir un point anguleux en b).

De fait, le prolongement analytique n'est pas toujours possible. Par exemple⁷⁰, soit la fonction :

$$f_1(z) \stackrel{\text{d\u00e9f}}{=} \sum_{n=1}^{+\infty} z^{2^n} . \quad (5.133)$$

Cette s\u00e9rie converge $\forall z$, $|z| < 1$, mais diverge visiblement pour $z = 1$ (et m\u00eame effroyablement si $|z| > 1$!), r\u00e9v\u00e9lant ainsi l'existence d'une singularit\u00e9 en $z = 1$. Par ailleurs, on a trivialement :

$$f_1(z) = z^2 + z^4 + z^8 + z^{16} \dots = z^2 + [(z^2)^2 + (z^2)^4 + (z^2)^8 \dots] , \quad (5.134)$$

de sorte que la relation fonctionnelle suivante est vraie :

$$f_1(z) = z^2 + f_1(z^2) . \quad (5.135)$$

Ceci montre que $f_1(z)$ a aussi⁷¹ des singularit\u00e9s en $z^2 = 1$, soit $z = \pm 1$. En it\u00e9rant une fois la relation fonctionnelle (5.135), on a aussi :

$$f_1(z) = z^2 + [(z^2)^2 + f_1((z^2)^2)] = z^2 + z^4 + f_1(z^4) ; \quad (5.136)$$

ceci montre que $f_1(z)$ a aussi des singularit\u00e9s en $z^4 = 1$, soit en $e^{ik\frac{\pi}{2}}$, $k = 1, 2, 3, 4$ (on retrouve les singularit\u00e9s d\u00e9j\u00e0 d\u00e9tect\u00e9es), et ainsi de suite. D'une fa\u00e7on g\u00e9n\u00e9rale, la forme it\u00e9r\u00e9e :

$$f_1(z) = z^2 + z^4 + z^8 + \dots + z^{2^n} + f_1(z^{2^n}) , \quad (5.137)$$

r\u00e9v\u00e8le des singularit\u00e9s en $z^{2^n} = 1$, c'est-\u00e0-dire aux sommets du polygone r\u00e9gulier \u00e0 2^n c\u00f4t\u00e9s inscrit dans le cercle de rayon unit\u00e9. Au total, l'ensemble des singularit\u00e9s de $f_1(z)$ est $\{z_{k_n n} \stackrel{\text{d\u00e9f}}{=} e^{ik_n \frac{\pi}{2^n - 1}}\}$, $k_n = 1, 2, \dots, 2^n$, $n \in \mathbb{N}^*$, toutes situ\u00e9es sur le cercle unit\u00e9. La *densit\u00e9* de ces points fait du cercle un rempart infranchissable⁷² r\u00e9v\u00e9lant l'impossibilit\u00e9 de prolonger analytiquement f_1 au-del\u00e0 de ce cercle. On dit alors que le cercle est la fronti\u00e8re *essentielle* de $f_1(z)$, que le disque est le *domaine d'existence* de f_1 , qui est qualifi\u00e9e de fonction analytique *compl\u00e8te*.

Plus g\u00e9n\u00e9ralement, une fois que l'on effectu\u00e9 tous les prolongements possibles d'une certaine fonction, celle-ci est de m\u00eame qualifi\u00e9e de *compl\u00e8te*. Par exemple, soit \u00e0 trouver toutes les d\u00e9terminations de la fonction logarithme ; l'id\u00e9e est, dans la premi\u00e8re \u00e9tape, d'utiliser $(\ln z)' = \frac{1}{z}$, c'est-\u00e0-dire d'int\u00e9grer la fonction $\frac{1}{z}$ en utilisant la convergence uniforme de son d\u00e9veloppement en s\u00e9rie enti\u00e8re. On \u00e9crit :

$$\frac{1}{z} = \frac{1}{1 + (z - 1)} = \sum_{n \in \mathbb{N}} (-1)^n (z - 1)^n , \quad (5.138)$$

⁷⁰Un autre exemple : la fonction $\sum_{n \in \mathbb{N}} z^{n!}$.

⁷¹ce qui est imm\u00e9diat, $f_1(Z)$ ayant une singularit\u00e9 en $Z = 1$ (il suffit de prendre $Z = z^2$).

⁷²"The attempt to carry out the process of continuation will therefore be frustrated by the existence of this unbroken front of singularities, beyond which it is impossible to pass." ([7], \u00a5 5.501).

L'aspect *infranchissable* peut se percevoir intuitivement en notant que le "nombre" de singularit\u00e9s sur le cercle de circonf\u00e9rence 2π est \u00e9gal \u00e0 $\lim_{N \rightarrow +\infty} 2^N = 2^{\aleph_0}$ o\u00f9 \aleph_0 est le cardinal de \mathbb{N} . Le cardinal de \mathbb{R} \u00e9tant par d\u00e9finition not\u00e9 \aleph_1 , l'*hypoth\u00e8se du continu* de la th\u00e9orie des ensembles est \u00e9quivalente \u00e0 l'\u00e9galit\u00e9 $2^{\aleph_0} = \aleph_1$, d'o\u00f9 la densit\u00e9 de la fronti\u00e8re circulaire impossible \u00e0 franchir (voir par exemple [33])...

la série convergeant uniformément $\forall z, |z - 1| < 1$, c'est-à-dire à l'intérieur du disque de rayon 1 centré en 1. Intégrant terme à terme du point 1 jusque vers n'importe quel z situé dans ce disque, on obtient :

$$f_0(z) \stackrel{\text{déf}}{=} \int_1^z \frac{1}{\xi} d\xi = \sum_{n \in \mathbb{N}^*} (-1)^{n-1} \frac{(z-1)^n}{n} \quad (|z-1| < 1) ; \quad (5.139)$$

la fonction $f_0(z)$ coïncide avec la détermination du logarithme qui prend la valeur 0 au point $z = 1$: $f_0(z) = \ln z \stackrel{\text{déf}}{=} f(z)$, $\ln 1 = 0 = f(1)$.

Le même travail peut être effectué en partant de n'importe quel point du disque unité d'affixe z_1 , par exemple avec $z_1 = e^{i\pi/4}$, définissant une certaine fonction $f_1(z)$:

$$f_1(z) \stackrel{\text{déf}}{=} f(z_1) + \sum_{n \in \mathbb{N}^*} \frac{1}{n!} f_1^{(n)}(z_1) (z - z_1)^n , \quad (5.140)$$

avec $f_1(z_1) = \int_1^{z_1} \frac{1}{\xi} d\xi = \int_0^{\pi/4} \frac{1}{e^{i\theta}} i e^{i\theta} d\theta = \frac{i\pi}{4}$ et $f_1^{(n)}(z_1) = \frac{(-1)^{n-1} (n-1)!}{z_1^n}$, d'où le développement :

$$f_1(z) = \frac{i\pi}{4} + \sum_{n \in \mathbb{N}^*} (-1)^{n-1} \frac{(z - e^{i\pi/4})^n}{n e^{in\pi/4}} \quad (|z - e^{i\pi/4}| < 1) ; \quad (5.141)$$

Plus généralement, le développement de Taylor centré au point $z_k = e^{ik\pi/4}$ permet de définir la fonction :

$$f_k(z) \stackrel{\text{déf}}{=} \frac{ik\pi}{4} + \sum_{n \in \mathbb{N}^*} (-1)^{n-1} \frac{(z - e^{ik\pi/4})^n}{n e^{ink\pi/4}} \quad (|z - e^{ik\pi/4}| < 1) . \quad (5.142)$$

Pour $k=8$, on se retrouve intégralement dans le disque de départ avec le développement :

$$f_8(z) \stackrel{\text{déf}}{=} 2i\pi + \sum_{n \in \mathbb{N}^*} (-1)^{n-1} \frac{(z-1)^n}{n} \quad (|z-1| < 1) , \quad (5.143)$$

qui diffère de la fonction $f_0(z)$ définie en (5.139) par la constante additive $2i\pi$. Ainsi, le prolongement analytique continu de proche en proche ne reproduit pas forcément la même fonction. On retrouve ici l'augmentation de la partie imaginaire de $\ln z$ à chaque tour de cadran : clairement, en répétant ce processus à l'infini (dans un sens ou dans un autre), on peut de fait engendrer toutes les branches de la fonction logarithme ; visiblement, le choix d'un décalage de $\frac{\pi}{4}$ est juste pour l'exemple : n'importe quel angle $\frac{\pi}{q}$ conviendrait tout autant, redonnant pour $f_{2q}(z)$ la fonction de départ agmentée de $2i\pi$ et développée dans le disque de départ. Le même procédé peut être répété avec la fonction $f(z) = z^{1/2}$, $f(1) = 1$.

5.6 Fonctions multiformes. Coupures. Notion de surface de Riemann

Les fonctions multiformes les plus simples ont déjà été rencontrées : $\ln z$, z^α , $\alpha \notin \mathbb{Z}$. La caractéristique d'une fonction multiforme est de ne pas reprendre forcément la même

valeur quand on décrit certains chemins continus fermés dans le plan. Soit à nouveau la branche de fonction $z^{\frac{1}{2}}$ qui vaut 1 en $z = 1$, soit $\sqrt{|z|} e^{i\theta/2}$, $0 \leq \theta < 2\pi$. Effectuons un tour complet en suivant n'importe quelle courbe partant de $z = 1$ et y revenant⁷³ *en contournant l'origine*. Au départ, la branche ainsi définie vaut 1, à l'arrivée elle vaut $(e^{2i\pi})^{\frac{1}{2}} = e^{i\pi} = -1$; le même phénomène se produit pour l'autre branche de la fonction racine carrée. Il en va de même pour le logarithme, quelle que soit la branche choisie : quand z décrit une courbe fermée autour de l'origine, son argument varie de 2π et la branche choisie, quelle qu'elle soit, diffère⁷⁴ de $\pm 2i\pi$ de la valeur de départ. Le parcours d'un tel contour fermé induit donc pour une fonction multiforme une sorte de discontinuité, à l'instar d'une fonction d'une variable réelle présentant en un point d'abscisse x_0 un saut de première espèce : selon que l'on arrive à gauche ou à droite de x_0 , la limite de la fonction n'est pas la même ; un point sur \mathbb{R} est de dimension égale à zéro. Pour une fonction multiforme, ceci se produit en tout point de la coupure préalablement définie ; dans $\mathbb{C} \sim \mathbb{R}^2$, espace bidimensionnel, toute ligne a pour dimension un : en ce sens, une coupure dans \mathbb{C} est l'extension bidimensionnelle d'un point de discontinuité dans \mathbb{R} .

Il en irait de même si le point de départ (et le point d'arrivée) était n'importe quel point du plan d'argument quelconque, pas forcément un point de l'axe réel positif, pourvu que l'on fasse un circuit fermé autour de O : ce fait montre bien que le problème n'est pas le demi-axe réel, mais bel et bien l'origine O . Un tel point bizarre (c'est en fait un point singulier d'une nouvelle sorte) est un *point de branchement* : c'est de ce point que part une demi-droite⁷⁵ infranchissable tant que l'on s'interdit de provoquer des discontinuités dans la fonction considérée. On peut voir une coupure comme un ensemble dense de singularités⁷⁶ : où que l'on soit sur la coupure, la fonction a un saut en deux points infiniment voisins situés de part et d'autre de celle-ci.

Le choix de la coupure est *a priori* conventionnel⁷⁷, mais, une fois fait, il faut s'y tenir absolument pour éviter les ambiguïtés, qui conduisent tôt ou tard à des erreurs ; c'est seulement à ce moment que la branche considérée est parfaitement définie. Revenons à la fonction $z^{\frac{1}{2}}$; une première détermination peut être définie comme :

$$f_1^{(1)}(z = re^{i\theta}) = \sqrt{r} e^{i\theta/2} , \quad 0 \leq \theta < 2\pi , \quad (5.144)$$

⁷³Si le chemin suivi ne contient pas l'origine en son intérieur, l'argument de z part de zéro, augmente, puis diminue et revient à zéro à l'arrivée : alors la branche choisie reprend la même valeur.

⁷⁴Le signe dépend du sens de parcours, positif ou négatif.

⁷⁵En réalité l'aspect *demi-droite* est tout à fait secondaire : on pourrait prendre comme coupure n'importe quelle "demi-courbe" partant du point de branchement (un *spagnetti* semi-infini fait tout autant l'affaire). Tout ceci n'est qu'affaire de commodité et autant faire simple en prenant une (demi)-ligne rectiligne !

⁷⁶On note que la coupure peut toujours être contournée, au contraire de la frontière *essentielle* d'une fonction analytique complète.

⁷⁷Cette affirmation est correcte en-dehors de tout contexte. Il arrive au contraire que la coupure soit déterminée par la nature du problème considéré. Il est fréquent, en Physique, qu'un grand principe (celui de causalité, . . .), ou une symétrie (par rapport au renversement du temps, . . .), impose d'emblée certaines relations de conjugaison complexe aux fonctions représentant des grandeurs physiques. Supposons que la contrainte physique oblige une certaine fonction à satisfaire $f(z^*) = [f(z)]^*$, et que $f(z)$ soit précisément $z^{\frac{1}{2}}$: alors, il n'y a pas de doute, c'est la branche définie en (5.145) qui s'impose d'elle-même.

où \sqrt{r} désigne la racine carrée positive “ordinaire” d’un nombre réel positif ou nul ($\sqrt{9} = 3$, etc). Sur l’axe réel positif, $z = x \in \mathbb{R}_+$ et $f_1^{(1)}(x) = \sqrt{x}$. Sur l’axe réel négatif, on a $f_1^{(1)}(re^{i\pi}) = \sqrt{r}e^{i\pi/2} = i\sqrt{r}$. Enfin, juste au-dessous de l’axe réel positif, $f_1^{(1)}(re^{i(2\pi-0)}) = \sqrt{r}e^{i(\pi-0)} = -\sqrt{r}$. On retrouve bien le fait crucial qu’ayant fait un tour continu complet autour de l’origine, la fonction ne reprend pas la même valeur : partant de \sqrt{r} , elle vaut $-\sqrt{r}$ quand on revient au “même” point de l’axe réel (ce n’est pas tout à fait le même : le point d’arrivée est juste au-dessous du point de départ). En tout cas, quel que soit $r > 0$, la fonction $z^{\frac{1}{2}}$ saute de $+\sqrt{r}$ (bord supérieur de la coupure) à $-\sqrt{r}$ (bord inférieur) ; tout au long de la coupure, la fonction a un saut d’amplitude $2\sqrt{r}$, ce qui qualifie la coupure en tant que ligne continue de singularités. À la branche $f_1^{(1)}$ est tout naturellement associée la seconde branche de la racine carrée complexe⁷⁸ $f_1^{(2)}(z = re^{i\theta}) = \sqrt{r}e^{i\theta/2+i\pi} = -f_1^{(1)}(z = re^{i\theta})$.

On peut tout autant définir une autre branche, par exemple :

$$f_2^{(1)}(z = re^{i\theta}) = \sqrt{r}e^{i\theta/2}, \quad -\pi < \theta < \pi ; \quad (5.145)$$

Une telle fonction est maintenant continue à la traversée de l’axe réel, et le saut se produit de part et d’autre du demi-axe \mathbb{R}_- , qui constitue la coupure de la fonction ainsi proprement définie. Noter que cette fonction satisfait la relation de conjugaison $f_2^{(1)}(z^*) = [f_2^{(1)}(z)]^*$, que ne satisfait pas $f_1^{(1)}(z)$ (voir note 77).

Quoi qu’il en soit, la coupure étant choisie une fois pour toutes, la branche de la fonction $z^{\frac{1}{2}}$ considérée est clairement analytique en-dehors de la coupure et tout ce qui a été démontré pour les fonctions analytiques reste donc valable. En particulier, l’intégrale sur une ligne évitant la coupure ne dépend que des extrémités, en aucune façon du chemin effectivement suivi pour les relier, l’intégrale sur un cycle où la fonction est continue⁷⁹ est nulle, etc. Ceci est évidemment vrai pour n’importe quelle fonction multiforme.

L’obstacle présenté par la coupure est infranchissable tant que l’on en reste au plan complexe ordinaire considéré jusqu’ici. Pour cette raison, on ne peut pas décrire des cycles entourant le point de branchement sans buter sur une discontinuité de la fonction : en l’état actuel des choses, le théorème de Cauchy ne vaut que pour les contours évitant la coupure, et ses applications (notamment le théorème des résidus) s’en trouvent limitées à certains circuits dans le plan.

L’astuce géniale de Riemann consiste à *feuilleter* ce plan en l’imaginant comme étant en réalité le collage de feuillets plans, en nombre à préciser dans chaque cas (c’est-à-dire pour chaque fonction multiforme), de telle sorte que l’on soit autorisé à franchir la coupure à la condition expresse de *changer de feuillet* et de prolonger par continuité la fonction dans le nouveau feuillet. Pour pouvoir se livrer à cette gymnastique, il suffit en définitive de séparer les feuillets les uns des autres en les maintenant collés exclusivement le long de la coupure, qui constitue l’endroit de passage obligé afin de préserver la continuité de la fonction. L’empilement de plans ainsi constitué forme la *surface de Riemann*

⁷⁸C’est la généralisation à \mathbb{C} du fait que l’équation $x^2 = a > 0$ a deux racines $\pm\sqrt{a}$.

⁷⁹Ce qui exclut une boucle dont les points d’arrivée et de départ, infiniment proches, sont de part et d’autre de la coupure.

de la fonction multiforme considérée ; sur sa surface de Riemann, une fonction multiforme peut être traitée comme une fonction analytique “ordinaire”, et le théorème des résidus retrouve sa pleine et entière utilité – dans la mesure où il est possible d’effectuer un contour fermé partant d’un certain feuillet, passant dans les autres et revenant au point de départ du feuillet de départ en tournant toujours dans le même sens⁸⁰.

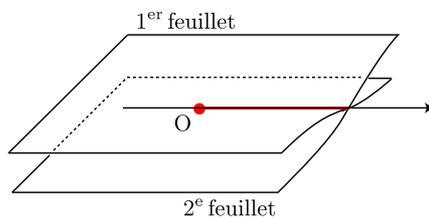


Figure 5.7: Surface de Riemann de la détermination de $z^{\frac{1}{2}}$ valant $+1$ en $z = 1 + i0$.

À titre d’exemple, reprenons la fonction $z^{\frac{1}{2}}$ et sa détermination (5.144), dont la coupure est le demi-axe \mathbb{R}_+ . Les feuillets de sa surface sont donc collés les uns aux autres par la coupure, qui est la ligne de passage obligé d’un feuillet à l’autre. Le premier feuillet correspond tout naturellement aux valeurs de θ comprises entre 0 et 2π et, quand θ dépasse cette dernière valeur, le point d’affixe z glisse dans le deuxième feuillet (on dit souvent “descend dans le deuxième feuillet”, mais *monter* ou *descendre* est purement conventionnel), tandis que son argument augmente par continuité au-delà de 2π : entre un point d’argument dans le premier feuillet $2\pi - 0$ et son voisin dans le deuxième feuillet d’argument $2\pi + 0$, la fonction n’a plus de saut. Ainsi, grâce au feuilletage, la fonction devient *uniforme* et, toute autre singularité éventuelle mise à part, devient une fonction analytique. Il devient alors licite de considérer des contours allant d’un feuillet à l’autre et pour lesquels les règles de l’intégration démontrées pour les fonctions “uniformes” retrouvent leur pleine et entière validité.

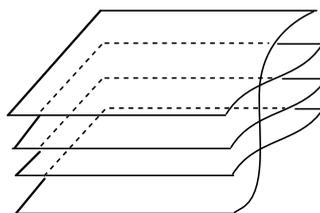


Figure 5.8: La surface de Riemann de la détermination de $z^{1/n}$ valant 1 en $z = 1$ possède n feuillets (la figure est faite avec $n = 4$).

⁸⁰Ceci n’est pas toujours possible, par exemple pour la fonction $\ln z$ dont la surface de Riemann a une structure hélicoïdale (“vis sans fin”).

Dans le deuxième feuillet, l'angle varie de 2π à 4π , atteignant cette valeur en revenant sur la coupure après avoir contourné l'origine une seconde fois. À ce moment, la détermination (5.144) vaut $e^{i\frac{4\pi}{2}} = e^{2i\pi} = 1 \equiv (e^{i\frac{\theta}{2}})_{\theta=0}$ et reprend donc la même valeur qu'au départ ($\theta = 0$) : sans créer de discontinuité, on peut alors remonter dans le premier feuillet. La surface de Riemann de (5.144) est donc formée de deux feuillets reliés par \mathbb{R}_+ . On peut se la représenter comme deux plans parallèles pincés le long de cet axe (voir fig. 5.7). Ainsi, le long d'un cycle partant dans le premier feuillet, glissant dans le second quelque part sur \mathbb{R}_+ pour revenir dans le premier après deux tours complets, l'intégrale de $z^{-\frac{1}{2}}$ est nulle : la primitive est $2z^{\frac{1}{2}}$, et reprend la même valeur au départ et à l'arrivée. Pour une fonction multiforme ayant des singularités, le théorème des résidus fonctionne et "ramasse" les résidus de toutes les singularités présentes dans les différents feuillets⁸¹.

Avec les mêmes arguments, on montre que la surface de Riemann de la fonction $z^{\frac{1}{n}}$, $n \in \mathbb{N}$ contient n feuillets ; on descend du feuillet k au feuillet $k + 1$ au droit de \mathbb{R}_+ , remontant du n^{e} au premier au même endroit.

Autre exemple : la fonction $(z^2 - a^2)^{\frac{1}{2}}$, $a \in \mathbb{R}_+$. Il est facile de voir qu'il y a deux points de branchement⁸² en $z = \pm a$, d'où partent deux coupures. Si l'une file⁸³ de $+a$ vers $+\infty \in \mathbb{R}$, l'autre de $-a$ vers $+\infty \in \mathbb{R}$, on constate qu'en fait ces deux coupures s'annihilent mutuellement à droite de a , puisqu'alors les deux monômes $(z \pm a)^{\frac{1}{2}}$ gagnent chacun un signe – après un tour complet autour de l'origine. Ici, la coupure est donc le segment $[-a, +a]$ de l'axe réel ; de part et d'autre de ce segment la fonction prend deux valeurs opposées. La surface de Riemann est donc simplement constituée de deux plans parallèles pincés le long du segment $[-a, +a]$. À nouveau, la(es) coupures peuvent être choisies autrement : on peut tout autant prendre⁸⁴ les deux demi-droites $] -\infty, -a]$ et $[+a, +\infty[$, auquel cas il n'y a pas une mais deux coupures (ou, si on préfère, la seule et unique coupure n'est apparemment pas connexe⁸⁵). Cette fois, la fonction est continue à la traversée du segment de $] -a, +a[$ de l'axe réel et a des valeurs opposées de part et d'autre de chaque demi-droite $] -\infty, -a]$, $[+a, +\infty[$. Avec ce choix, la surface de Riemann est formée de deux plans pincés le long de ces deux demi-droites.

Une surface de Riemann peut contenir une infinité de feuillets : c'est le cas de la fonction logarithme (voir fig. 5.9), puisque sa partie imaginaire augmente de 2π à chaque

⁸¹Une fonction peut ne pas avoir de singularité dans le premier feuillet, et en avoir dans un autre. Par exemple, supposons que la quantité $z + z^{\frac{1}{2}}$ apparaisse au dénominateur d'une certaine fonction. Les singularités de cette fonction sont alors données par l'équation $z + z^{\frac{1}{2}} = 0$, dont une solution est $z = 0$. Avec la branche (5.144), les autres solutions sont données par $\sqrt{r}e^{i\frac{\theta}{2}} + 1 = 0$, soit $\sqrt{r} \cos \frac{\theta}{2} + 1 = 0$ et $\sqrt{r} \sin \frac{\theta}{2} = 0$; avec $0 \leq \theta < 2\pi$, il n'y en a pas. En revanche, si θ dépasse 2π (on plonge dans le deuxième feuillet), il y a une solution $r = 1$, $\theta = 2\pi$.

Ce scénario est typique dans la description d'un système instable en Mécanique quantique, un atome à l'état excité par exemple – voir section 5.7.

⁸²là où s'annule la quantité sous le radical.

⁸³Ce qui revient à prendre pour $(z \pm a)^{\frac{1}{2}}$ la même détermination qu'en (5.144) pour $z^{\frac{1}{2}}$.

⁸⁴Ce qui revient à considérer la branche de $(z - a)^{\frac{1}{2}}$ dont la coupure est le demi-axe $[+a, +\infty[$, et pour $(z + a)^{\frac{1}{2}}$ la coupure suivant $[-\infty, -a]$.

⁸⁵Elle l'est en fait par le point à l'infini.

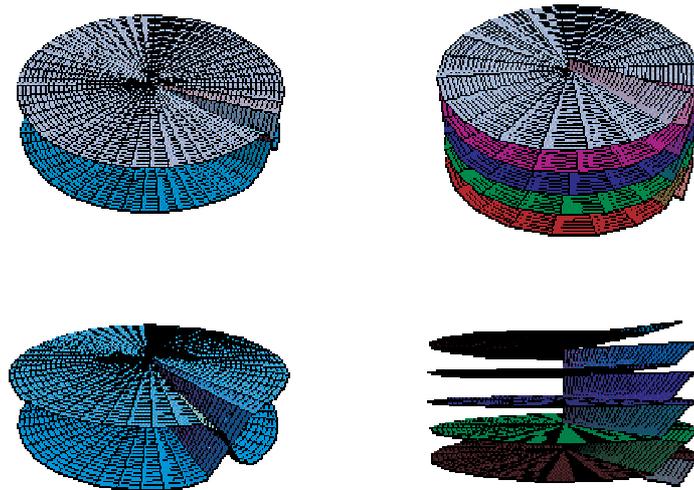


Figure 5.9: Surfaces de Riemann de $z^{\frac{1}{2}}$ (à gauche) et de $\ln z$ (à droite), vues sous deux angles.

tour décrit dans le sens positif : une fois le manège commencé, la fonction \ln ne revient *jamaïs* à sa valeur initiale si on tourne toujours dans le même sens. Si la coupure est prise le long de \mathbb{R}_+ , sa surface de Riemann est constituée d'une infinité de feuillets, sur lesquels on ne fait que descendre (ou de monter) quand on passe de l'un au suivant et si l'argument ne cesse de croître ou de décroître. Bien évidemment, il n'est pas interdit d'inverser le sens de rotation le long du chemin (*i.e.* de faire demi-tour), après avoir été pêcher une singularité (c'est même en pratique l'un des intérêts principaux de cette gymnastique).

5.7 Exemples physiques de prolongement analytique : émission spontanée, métastabilité

C'est justement cet exercice qui permet de calculer proprement les durées de vie en Mécanique quantique (ou en Mécanique statistique). Tout système situé dans un état d'énergie élevée a une tendance naturelle à transiter tôt ou tard vers un état d'énergie inférieure. Il peut s'agir d'un noyau instable (radioactivité), ou d'un atome excité, voire d'un système thermodynamique dans un état métastable (vapeur surchauffée ou liquide surfondu). Pour fixer les idées, il peut s'agir d'un atome porté à l'état excité qui, au bout d'un temps *aléatoire*, finit par émettre un photon par émission spontanée, c'est-à-dire en l'absence de tout champ (plus précisément lorsque les *moyennes* (quantiques) du champ

électromagnétique sont nulles⁸⁶). Pour un tel système instable, c'est l'existence d'un *continuum* d'énergie pour le système "nu" qui donne une coupure au propagateur. En prolongeant analytiquement celui-ci dans le deuxième feuillet, on voit que ce prolongement y possède des singularités (usuellement des pôles), qui sont précisément les énergies complexes évoquées plus haut. C'est le Principe de causalité qui impose à ces "énergies" d'avoir une partie imaginaire négative, que l'on appelle alors plutôt résonances (l'inverse de la largeur de la résonance est précisément la durée de vie de l'état excité). En regardant le premier feuillet en transparence, on voit ces énergies en-dessous de l'axe réel ; dans la situation de couplage faible, les résonances sont situées à proximité des énergies (réelles) du système "nu", le décalage de la partie réelle constituant le *déplacement de Lamb* (*Lamb shift*).

Par la suite (voir chapitre 7, section 7.7), on rencontrera un autre exemple de prolongement analytique, constituant un scénario typique pour faire apparaître une partie imaginaire dans une fonction de partition Z (donc dans une énergie libre F), permettant ainsi de rendre compte de la durée de vie finie d'un état métastable (par exemple la nucléation de gouttes de liquide dans une vapeur surchauffée).

⁸⁶Ce sont les fluctuations (quantiques) du champ dans l'état vide qui provoquent l'émission spontanée.

Chapitre 6

Applications élémentaires du théorème des résidus

“Everything should be made as simple as possible,
but not simpler.”

(Albert EINSTEIN, 1879–1955)

<i>Il s'agit de présenter les applications les plus élémentaires du théorème des résidus (calcul d'intégrales, de la somme de séries...)</i>
--

Avant de présenter les applications les plus courantes en Physique du théorème des résidus, il convient d'établir les lemmes de Jordan, dont l'usage est tellement systématique qu'avec un peu de savoir-faire et d'habitude, on en vient à un usage mentionné pour la forme quand il n'est pas implicite. Ceci fait, les quelques exemples traités ensuite montreront que le théorème des résidus constitue un outil de calcul d'une puissance extraordinaire, et d'une rare élégance.

6.1 Lemmes de Jordan

Le théorème de Cauchy permet de déformer tout contour d'intégration dans le plan \mathbb{C} , à condition de prendre en compte les singularités de l'intégrand ; par ailleurs, le théorème des résidus permet de calculer une intégrale portant sur un contour *fermé* (cycle). En pratique, une intégrale implique souvent un chemin d'intégration ouvert (la droite réelle, la demi-droite réelle, ...). La méthode consiste à compléter ce dernier chemin afin de définir un cycle et d'appliquer le théorème des résidus (ce que l'on appelle *refermer le contour*). L'idéal est bien sûr que les morceaux ainsi ajoutés donnent une contribution

nulle ; alors, l'intégrale sur le contour (calculée par résidus) donne directement l'intégrale initiale sans calculs additionnels.

Les lemmes de Jordan sont précisément des petits théorèmes d'usage systématique quand il faut compléter ou modifier un contour d'intégration. Le premier lemme s'énonce comme suit.

Lemme 1

Soit une fonction $f(z)$ continue sur le secteur défini par $z = Re^{i\theta}$, $R > 0$, $0 \leq \theta_1 \leq \theta \leq \theta_2 \leq 2\pi$, telle que $\lim_{|z| \rightarrow \infty} |zf(z)| = 0$. Alors :

$$\boxed{\lim_{R \rightarrow \infty} \int_{C_R} f(z) dz = 0} \quad (6.1)$$

où C_R est l'arc de cercle de rayon R compris entre les deux angles θ_1 et θ_2 .

L'intégrale se met immédiatement sous forme d'une intégration réelle $\int_{\theta_1}^{\theta_2} f(Re^{i\theta}) iRe^{i\theta} d\theta$. Le module de l'intégrale est borné supérieurement par :

$$\int_{\theta_1}^{\theta_2} |f(Re^{i\theta})| R d\theta \equiv \int_{\theta_1}^{\theta_2} |zf(z)| d\theta. \quad (6.2)$$

Par hypothèse, la fonction étant continue, l'intégrand tend vers zéro : $\forall \varepsilon > 0, \exists R_\varepsilon$ tel que $\forall |z| > R_\varepsilon, \text{Max}|zf(z)| < \varepsilon$; l'intervalle d'intégration est fini, donc l'intégrale est nulle à la limite. On peut dire les choses autrement, en introduisant $M \stackrel{\text{d\'ef}}{=} \max_{C_R} |f(z)|$; on a alors :

$$\left| \int_{\theta_1}^{\theta_2} f(z) dz \right| \leq R(\theta_2 - \theta_1)M ; \quad (6.3)$$

si $\lim_{R \rightarrow +\infty} RM = 0$, l'intégrale sur l'arc de cercle est nulle à la limite ¹.

Il existe une autre version de ce lemme, pour un contour en arc de cercle dont le rayon r tend vers zéro :

Lemme 2

Soit une fonction $f(z)$ continue sur le secteur défini par $z = re^{i\theta}$, $r > 0$, $0 \leq \theta_1 \leq \theta \leq \theta_2 \leq 2\pi$, telle que $\lim_{|z| \rightarrow 0} |zf(z)| = 0$. Alors :

$$\boxed{\lim_{r \rightarrow 0} \int_{C_r} f(z) dz = 0} \quad (6.4)$$

¹Noter qu'il s'agit d'une condition suffisante : l'intégrale peut être nulle sans que cette condition soit satisfaite.

où C_r est l'arc de cercle de rayon r compris entre les deux angles θ_1 et θ_2

La démonstration est la même que pour (6.1), et considère un petit arc de cercle de rayon r délimité par les angles θ_1 et θ_2 .

Lemme 3

Un autre lemme très utile concerne les intégrands contenant un facteur de Fourier e^{iz} . Il s'énonce comme suit :

Soit une fonction $f(z)$ continue sur le secteur défini par $z = Re^{i\theta}$, $R > 0$, $0 \leq \theta_1 \leq \theta \leq \theta_2 \leq \pi$, telle que $\lim_{|z| \rightarrow \infty} |f(z)| = 0$. Alors :

$$\boxed{\lim_{R \rightarrow \infty} \int_{C_R} e^{iz} f(z) dz = 0} \quad (6.5)$$

où C_R est l'arc de cercle de rayon R compris entre les deux angles θ_1 et θ_2 .

Dans le cas où $\theta_1 > 0$ et $\theta_2 < \pi$, la démonstration est immédiate, puisque l'intégrand est borné en module par $e^{-R \sin \theta_{\min}} |f(z)|$ où θ_{\min} est $\min(\theta_1, \pi - \theta_2)$, qui tend exponentiellement vers zéro quand $R \rightarrow \infty$. La démonstration est moins facile pour le secteur semi-circulaire, très fréquent en pratique, où $\theta_1 = 0$ et $\theta_2 = \pi$.

Soit l'intégrale sur le demi-cercle de rayon R situé dans le demi-plan supérieur, où $z = Re^{i\theta}$, $0 \leq \theta \leq \pi$:

$$I = \int_0^\pi e^{iz} f(z) iRe^{i\theta} d\theta = \int_0^\pi e^{iR(\cos \theta + i \sin \theta)} f(z) iRe^{i\theta} d\theta . \quad (6.6)$$

On décompose l'intégrale en deux, posant $\theta' = \pi - \theta$ dans la seconde. Prenant les modules², il vient :

$$|I| \leq \int_0^\pi e^{-R \sin \theta} |f(z)| R d\theta = R \int_0^{\frac{\pi}{2}} e^{-R \sin \theta} |f(Re^{i\theta})| d\theta + R \int_0^{\frac{\pi}{2}} e^{-R \sin \theta'} |f(Re^{i(\pi - \theta')})| d\theta' \equiv I_1 + I_2 . \quad (6.7)$$

Pour $0 \leq \theta \leq \frac{\pi}{2}$, on a $\sin \theta \geq \frac{2}{\pi} \theta$, comme on le voit en faisant un dessin. La première intégrale au second membre de (6.7) a donc la borne supérieure :

$$I_1 \leq R \int_0^{\frac{\pi}{2}} e^{-\frac{2}{\pi} R \theta} |f(Re^{i\theta})| d\theta . \quad (6.8)$$

²et en se souvenant que le module d'une somme est inférieur ou égal à la somme des modules.

$|f|$ tendant uniformément vers zéro quand $R \rightarrow +\infty$, $\forall \varepsilon > 0$, $\exists R_\varepsilon$ tel que $\forall |z|=R > R_\varepsilon$, $\text{Max}|f| < \varepsilon$. D'où :

$$\forall R > R_\varepsilon : I_1 \leq \varepsilon R \int_0^{\frac{\pi}{2}} e^{-\frac{2}{\pi}R\theta} d\theta = \varepsilon R \frac{\pi}{2R} (1 - e^{-R}) = \frac{\pi}{2} \varepsilon (1 - e^{-R}) \leq \frac{\pi}{2} \varepsilon . \quad (6.9)$$

Ainsi, $\forall R > R_\varepsilon$, $|f| < \varepsilon$, la première intégrale au second membre de (6.7) est bornée par $\frac{\pi}{2}\varepsilon$; pour la deuxième intégrale I_2 de (6.7), l'argument est exactement le même. Au total $|\int_{C_R} e^{iz} f(z) dz| \leq \pi\varepsilon$, qui est aussi petit que l'on veut.

Attention !

Par la suite, on rencontrera des intégrales de ce type avec non pas e^{iz} mais e^{ikz} , $k \in \mathbb{R}$. Alors, si $k > 0$, le demi-cercle le long duquel l'intégrale tend vers zéro si $R \rightarrow +\infty$ est comme ci-dessus le demi-cercle supérieur, mais si $k < 0$, c'est l'intégrale sur le demi-cercle *inférieur* qui tend vers zéro, puisque $ikz = ik(x + iy) = ikx - ky$; quand $k < 0$, il faut bien prendre $y < 0$ pour que le demi-cercle donne une contribution tendant vers zéro quand R tend vers l'infini.

6.2 Calcul d'intégrales définies

Il s'agit de montrer comment le calcul d'une intégrale sur un intervalle réel d'une fonction d'une variable réelle³ $\int_a^b f(x)dx$ peut se ramener au calcul d'une certaine intégrale complexe $\int_C \tilde{f}(z)dz$, de laquelle le théorème des résidus donne rapidement la valeur.

Il y a *a priori* deux éléments à préciser : quel est le contour C ? Quelle est la fonction \tilde{f} ? Souvent, mais pas toujours, on a $\tilde{f} = f$; il faut parfois faire preuve d'ingéniosité et travailler par essai et par erreur ; un peu d'entraînement permet de faire vite le tour de presque toutes les astuces permettant de trouver le bon contour C et la bonne fonction \tilde{f} . On trouvera ci-dessous (section 6.3) quelques exemples où \tilde{f} est franchement différente de f .

6.2.1 Fractions rationnelles

Soit l'intégrale :

$$\int_a^b \frac{P(x)}{Q(x)} dx \quad (6.10)$$

où P et Q sont des polynômes. À un niveau élémentaire, on apprend à calculer ce type d'intégrale en effectuant la division des polynômes et/ou en décomposant en éléments

³La fonction elle-même n'est pas forcément réelle mais peut tout à fait être à valeurs complexes – par exemple, ce peut être l'intégrand apparaissant dans la transformation de Fourier d'une fonction, elle-même réelle ou complexe.

simples. On va voir que, dans le cas où les bornes a et b sont $\pm\infty$, le théorème des résidus donne le résultat de façon quasi immédiate, en tout cas moins laborieuse. Plutôt que de donner des résultats formels, traitons quelques exemples d'intégrales du genre (6.10) avec $a = -\infty$, $b = +\infty$. Alors, pour que l'intégrale soit bien définie, il faut que le rapport $\frac{P}{Q}$ tende plus vite que $|x|^{-1}$ à l'infini, soit : degré de $P <$ degré de $Q - 1$ (alors, $\lim_{|z| \rightarrow +\infty} |z \frac{P(z)}{Q(z)}| = 0$ et le 1^{er} lemme de Jordan est applicable si besoin est). En outre, Q est supposé n'avoir aucun zéro réel ; dans le cas contraire⁴, tout zéro simple donne une intégrale impropre, que l'on peut régulariser en tant que partie principale de Cauchy (voir section 6.4) ; si le zéro est multiple, d'autres régularisations doivent être envisagées.

Exemple 1

Soit à calculer l'intégrale⁵ :

$$I \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} \frac{1}{x^4 + 1} dx . \quad (6.11)$$

Pour appliquer le théorème des résidus, il faut fermer le contour. Partant de la droite réelle figurant dans (6.11), il n'est pas forcément facile de voir comment s'y prendre. Commençons par considérer l'intégrale (6.11) avec des bornes finies $\pm R$, sans avoir d'état d'âme sur le fait que la limite pour en arriver à (6.11) doit en toute rigueur être prise en faisant tendre *indépendamment* vers l'infini les deux bornes⁶. Le point de départ est donc :

$$\int_{-R}^{+R} \frac{1}{x^4 + 1} dx . \quad (6.12)$$

La façon la plus simple pour former un contour fermé est de relier les deux extrémités

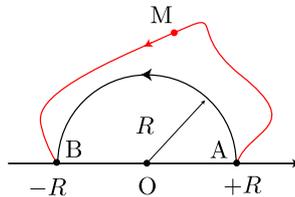


Figure 6.1: Contour C d'intégration pour le calcul de l'intégrale réelle (6.12) en passant par (6.13).

du segment $[-R, +R]$ par un demi-cercle centré à l'origine et de rayon R . Par en-haut

⁴On suppose que $P(x)$ et $Q(x)$ n'ont aucun zéro réel en commun. Dans le cas contraire, il suffit d'affiner l'analyse.

⁵Cette intégrale peut aussi se calculer par des moyens élémentaires (décomposition en éléments simples et utilisation de primitives connues).

⁶Dans le cas présent, il n'y a aucune difficulté puisque l'intégrand se comporte comme $|x|^{-4}$ à l'infini.

ou par en bas ? *A priori*, aucun choix ne s'impose à l'esprit, et commençons par refermer par en-haut. Enfin, aucune raison ne semble devoir changer la fonction f en une autre fonction \tilde{f} . On essaie donc de trouver I en considérant la fonction $f(z) = \frac{1}{z^4+1}$ – qui n'est autre que la fonction de départ calculée pour un complexe z quelconque – et son intégrale :

$$\tilde{I} \stackrel{\text{d\'ef}}{=} \int_C \frac{1}{z^4+1} dz \quad (6.13)$$

où C est le contour indiqué sur la fig.6.1. Tout ce que l'on sait permet d'affirmer que le choix d'un *demi-cercle* est fait uniquement pour la commodité du calcul (que d'ailleurs on ne fera pas : il suffira d'invoquer le 1^{er} lemme de Jordan) : on pourrait tout autant refermer par n'importe quelle ligne du genre AMB, cela ne changerait pas la valeur de l'intégrale. Le choix du demi-cercle étant fait, celle-ci, pour R fini, est la somme des deux intégrales :

$$\tilde{I}_R = \int_{-R}^{+R} \frac{1}{x^4+1} dx + \int_0^\pi \frac{1}{(Re^{i\theta})^4+1} d(Re^{i\theta}) \quad (6.14)$$

La première intégrale, dans la limite $R \rightarrow +\infty$, est celle que l'on cherche ; le théorème des résidus va vite donner la valeur de \tilde{I}_R , soit de la somme des deux intégrales de (6.14). Il faut donc être en mesure de calculer l'intégrale sur le demi-cercle, le cas "idéal" étant celui où elle vaut zéro quand $R \rightarrow +\infty$. C'est le cas ici, par le lemme de Jordan (6.1) : la fonction $f(z) = \frac{1}{z^4+1}$ est bien telle $|zf(z)| \rightarrow 0$ quand $|z| \rightarrow +\infty$, de sorte que la deuxième intégrale dans (6.14) est nulle à la limite $R \rightarrow +\infty$ (ceci saute d'ailleurs aux yeux sur la forme explicite apparaissant dans (6.14)). Il reste donc :

$$\lim_{R \rightarrow +\infty} \tilde{I}_R = \int_{-\infty}^{+\infty} \frac{1}{x^4+1} dx, \quad (6.15)$$

qui est bien l'intégrale cherchée (6.11). Maintenant, \tilde{I}_R , de par sa définition (6.13), se calcule rapidement par résidus, quel que soit R assez grand⁷ : il suffit de localiser, d'identifier les singularités de $\tilde{f} = \frac{1}{z^4+1}$ et de trouver les résidus. Le dénominateur est un polynôme du quatrième degré, il y a donc des pôles, qui sont les zéros de z^4+1 , c'est-à-dire les quatre racines de l'équation $z^4+1=0$, soit $z_k = e^{i(2k+1)\pi/4}$, $k=0, 1, 2, 3$, affixes des quatre sommets d'un carré parallèle aux axes, de côté $\sqrt{2}$. Seuls les deux pôles z_0 et z_1 ont une partie imaginaire positive et se trouvent de ce fait dans le contour choisi (demi-cercle dans le demi-plan *supérieur*). Le dénominateur s'écrit $\prod_{k=0}^3 (z-z_k)$; le résidu en z_0 est donc simplement $\lim_{z \rightarrow z_0} \frac{1}{(z-z_1)(z-z_2)(z-z_3)}$, soit $\frac{1}{(z_0-z_1)(z_0-z_2)(z_0-z_3)}$, etc. ; on trouve successivement⁸ :

$$\text{Res}\left(\frac{1}{z^4+1}, z_0\right) = \frac{1}{i\sqrt{2}(2+2i)}, \quad \text{Res}\left(\frac{1}{z^4+1}, z_1\right) = \frac{1}{i\sqrt{2}(2-2i)}. \quad (6.17)$$

⁷Il suffit que le demi-cercle contienne tous les pôles de $f(z)$, soit $R > 1$: dès lors, il ne se passe plus rien lorsque R augmente, sauf que la contribution de la périphérie du grand demi-cercle tend vers zéro quand $R \rightarrow +\infty$.

⁸Les résidus se calculent aussi par :

$$\text{Res}\left(\frac{1}{z^4+1}, z_k\right) = \frac{1}{(z^4+1)'_{z_k}} = \frac{1}{4z_k^3} = -\frac{z_k}{4} = -\frac{1}{4}e^{i(2k+1)\pi/4}. \quad (6.16)$$

d'où, dès que $R > 1$:

$$\tilde{I}_R \equiv \int_{-R}^{+R} \frac{1}{x^4 + 1} dx + \int_0^\pi \frac{1}{(\operatorname{Re}^{i\theta})^4 + 1} d(\operatorname{Re}^{i\theta}) = 2i\pi \left[\operatorname{Res}\left(\frac{1}{z^4 + 1}, z_0\right) + \operatorname{Res}\left(\frac{1}{z^4 + 1}, z_1\right) \right] . \quad (6.18)$$

Compte tenu de (6.17), il vient :

$$\int_{-R}^{+R} \frac{1}{x^4 + 1} dx + \int_0^\pi \frac{1}{(\operatorname{Re}^{i\theta})^4 + 1} d(\operatorname{Re}^{i\theta}) = 2i\pi \left[\frac{1}{i\sqrt{2}(2 + 2i)} + \frac{1}{i\sqrt{2}(2 - 2i)} \right] = \frac{\pi}{\sqrt{2}} , \quad (6.19)$$

d'où à la limite R infini :

$$\boxed{\int_{-\infty}^{+\infty} \frac{1}{x^4 + 1} dx = \frac{\pi}{\sqrt{2}}} \quad (6.20)$$

Noter que l'on peut tout autant refermer le contour par le demi-cercle situé dans le demi-plan inférieur, le résultat fourni par le 1^{er} lemme de Jordan ne dépendant pas ici du signe de $\Im z$. Le parcours s'effectuant alors dans le sens négatif, il ne faut pas oublier le signe global, et prendre tous les résidus relatifs aux pôles ayant une partie imaginaire négative ; on a alors :

$$\int_{-\infty}^{+\infty} \frac{1}{x^4 + 1} dx = (-2i\pi) [\operatorname{Res}\left(\frac{1}{z^4 + 1}, z_2\right) + \operatorname{Res}\left(\frac{1}{z^4 + 1}, z_3\right)] . \quad (6.21)$$

La somme des résidus est trouvée égale à $\frac{i}{\sqrt{2}}$, de sorte que tous calculs faits, on retrouve le même résultat (heureusement !) qu'en (6.20). D'ailleurs, en comparant les sommes des deux résidus dans un calcul et dans l'autre, on observe que :

$$\operatorname{Res}\left(\frac{1}{z^4 + 1}, z_0\right) + \operatorname{Res}\left(\frac{1}{z^4 + 1}, z_1\right) + \operatorname{Res}\left(\frac{1}{z^4 + 1}, z_2\right) + \operatorname{Res}\left(\frac{1}{z^4 + 1}, z_3\right) = 0 . \quad (6.22)$$

Il n'est pas difficile de comprendre le sens d'une telle relation. En effet, considérons un grand contour *faisant le tour du plan complexe* à l'infini, par exemple un cercle C_R de rayon R dont on va prendre la limite $R = \infty$. Par le théorème des résidus, on a :

$$\lim_{R \rightarrow +\infty} \int_{C_R} f(z) dz = 2i\pi \sum_{k=0}^3 \operatorname{Res}[f(z), z_k] , \quad (6.23)$$

où la somme court cette fois sur *tous* les pôles (à la limite, tous les pôles sont dans le contour). D'un autre côté, l'intégrale au premier membre de (6.23) est nulle à la limite, comme on le voit en appliquant le lemme de Jordan aux deux demi-cercles, d'où (6.22).

L'intégrale au premier membre de (6.23) n'est autre que celle apparaissant dans la définition du *résidu à l'infini* d'une fonction $f(z)$ (voir (5.102)). Avec cette définition, l'équation (6.23) se relit comme suit :

$$\operatorname{Res}[f(z), \infty] + \sum_{k=0}^3 \operatorname{Res}[f(z), z_k] = 0 , \quad \forall z_k, |z_k| < \infty . \quad (6.24)$$

Ce résultat est tout à fait général, comme on l'a vu : pour toute fonction méromorphe, la somme de tous les résidus (y compris le résidu à l'infini) est toujours nulle :

$$\sum_k \operatorname{Res}[f(z), z_k] = 0, \quad \forall z_k, |z_k| \leq \infty. \quad (6.25)$$

En effet, il suffit de partir du contour à l'infini et de le déformer en tirant des brins en direction de chaque pôle pour vérifier géométriquement que l'égalité est vraie. Ceci permet de réaliser que si l'on cherche la somme de presque tous les résidus aux pôles à distance finie, il est souvent moins laborieux de calculer les résidus des pôles délaissés et du pôle à l'infini, d'en faire la somme et de changer le signe pour trouver la somme d'intérêt. Pour une fonction telle que $|zf(z)| \rightarrow 0$ uniformément par rapport à l'argument de z , le résidu à l'infini est nul (c'est le cas pour $\frac{1}{z^4+1}$).

Exemple 2

Soit à calculer l'intégrale⁹ :

$$I_n \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{1}{x^{2n}+1} dx, \quad n \in \mathbb{N}^* \quad (6.26)$$

Afin d'utiliser ultérieurement un grand demi-cercle, on utilise la parité de l'intégrand pour récrire :

$$I_n = \frac{1}{2} \int_{-\infty}^{+\infty} \frac{1}{x^{2n}+1} dx. \quad (6.27)$$

On procède comme pour (6.11), en introduisant provisoirement le même contour que sur la fig.6.1 : le segment $[-R, +R]$ de l'axe réel refermé par le grand demi-cercle supérieur (l'autre demi-cercle donnerait le même résultat), dont la contribution tend vers zéro quand R tend vers l'infini puisque l'intégrand se comporte comme $\frac{1}{R^{2n}}$. Il vient donc, après passage à la limite :

$$I_n + 0 = \frac{1}{2} 2i\pi \sum_{k, \Im z_k > 0} \operatorname{Res} \left(\frac{1}{z^{2n}+1}, z_k \right), \quad (6.28)$$

où la somme ne porte que sur les résidus associés aux pôles ayant une partie imaginaire strictement positive. Les z_k sont les zéros du dénominateur, donnant chacun un pôle d'ordre un :

$$z_k^{2n} + 1 = 0 \iff z_k = e^{i\frac{(2k+1)\pi}{2n}}, \quad k = 0, 1, 2, \dots, 2n-1. \quad (6.29)$$

Le résidu vaut $\frac{1}{2nz_k^{2n-1}} = \frac{z_k}{2nz_k^{2n}} = -\frac{1}{2n} e^{i\frac{(2k+1)\pi}{2n}}$; les résidus des pôles z_k à partie imaginaire positive correspondent aux valeurs de k telles que $0 < \frac{2k+1}{2n}\pi < \pi$, soit

⁹ n est ici supposé entier, afin de commencer par traiter des intégrales n'impliquant pas de fonctions multiformes. L'intégrale (6.26) avec n non-entier est considérée plus loin, voir (6.148).

$k = 0, 1, \dots, n-1$. La \sum_k dans (6.28) fait alors intervenir une progression géométrique de raison $e^{i\frac{\pi}{n}}$ qui se somme facilement, et on trouve finalement¹⁰ :

$$I_n = \frac{\pi}{2n \sin \frac{\pi}{2n}} \quad (n \in \mathbb{N}^*) \quad (6.30)$$

On note que si $n \gg 1$, on a $I_n \simeq 1 + \frac{\pi^2}{24n^2}$, un résultat qui ne saute pas aux yeux au vu de la définition intégrale (6.26). (6.30) donne $2I_{n=2} = \frac{\pi}{\sqrt{2}}$, en conformité avec (6.20). À titre d'exercice, montrer (sans calcul) que l'intégrale $I_n(a) \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{1}{x^{2n} + a^{2n}} dx$ ($a \in \mathbb{R}$) est égale à $\frac{1}{|a|^{2n-1}} I_n$ (voir aussi l'exemple (6.35) p. 237 et les commentaires de la note 11).

Exemple 3

Soit à calculer l'intégrale :

$$I_1 \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{x^6}{(x^4 + 1)^2} dx . \quad (6.31)$$

On utilise le même contour que précédemment, en remarquant que la parité de l'intégrand permet d'écrire :

$$I_1 = \frac{1}{2} \int_{-\infty}^{+\infty} \frac{x^6}{(x^4 + 1)^2} dx . \quad (6.32)$$

Le lemme de Jordan s'applique puisque quand R est grand, l'intégrand se comporte comme $R^6/(R^4)^2 = R^{-2}$: l'intégrale sur le demi-cercle vaut zéro à la limite R infini. On retrouve les mêmes pôles que dans l'exemple 1 ci-dessus, mais ils sont ici d'ordre deux. Après calcul, on trouve :

$$I_1 = \frac{3\pi}{8\sqrt{2}} \quad (6.33)$$

C'est un bon exemple pour montrer comment, par un changement d'échelle (*scaling* en anglais), ce résultat permet de trouver la valeur d'une intégrale plus générale, et aussi de rencontrer un prolongement analytique¹¹. Soit en effet :

$$I(a) \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{x^6}{(x^4 + a^4)^2} dx \quad (a \in \mathbb{R}_+) . \quad (6.35)$$

¹⁰Pour $n = 1$, I_n peut s'obtenir par des moyens élémentaires, qui donnent immédiatement $I_1 = \frac{\pi}{2}$: c'est bien ce que dit aussi (6.30).

¹¹La restriction $a \in \mathbb{R}_+$ est essentielle pour pouvoir obtenir le résultat cherché à partir de (6.36), puisque 1 est un nombre positif. D'ailleurs, le résultat (6.37) est visiblement faux pour $a < 0$ par exemple, puisque l'intégrale de départ est encore dans ce cas une quantité visiblement positive. Ceci étant précisé, rien n'empêche de prolonger analytiquement l'expression (6.37) à condition de l'écrire précisément :

$$I(a) = \frac{3\pi\sqrt{2}}{16(a^4)^{1/4}} \quad (a \in \mathbb{C}, -\frac{\pi}{4} < \text{Arg } a < +\frac{\pi}{4}) , \quad (6.34)$$

la branche de $z^{1/4}$ étant celle qui prend des valeurs positives pour $z \in \mathbb{R}_+$, dont la coupure est le demi-axe réel négatif. La restriction sur $\text{Arg } a$ se comprend comme suit : la branche $z^{1/4}$ considérée est holomorphe dans le plan coupé du demi-axe réel négatif, donc $-\pi < \text{Arg } z < +\pi$. a^4 devant avoir son

On voit que $I(a = 1) = I_1$; en posant $x = ax'$, il vient (a étant positif) :

$$I(a) = \int_0^{+\infty} \frac{a^6 x'^6}{a^8 (x'^4 + 1)^2} a dx' = \frac{1}{a} \int_0^{+\infty} \frac{x^6}{(x^4 + 1)^2} dx = \frac{1}{a} I_1, \quad (6.36)$$

d'où :

$$I(a) = \frac{3\pi}{8a\sqrt{2}} \quad (a > 0). \quad (6.37)$$

Le *scaling* $I(a) \propto \frac{1}{a}$ était d'emblée prévisible : il suffit d'imaginer que a a une dimension physique (une longueur, par exemple), auquel cas x est aussi une longueur ; l'analyse dimensionnelle sur l'expression de départ (6.36) montre immédiatement que $[I] = L^{-1}$, de sorte que, a étant la seule longueur disponible, on a forcément $I(a) = \text{nombre} \times \frac{1}{a}$. Noter enfin que l'intégrale (6.31) peut se calculer en prenant un contour en secteur circulaire revenant à l'origine suivant un angle de $\frac{2\pi}{4} = \frac{\pi}{2}$.

Exemple 4

Comme exemple impliquant des pôles d'ordre n , soit l'intégrale :

$$I_n(a) \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{1}{(x^2 + a^2)^n} dx \quad (a > 0, n = 1, 2, 3, \dots). \quad (6.38)$$

Pour $n = 1$, on trivialement $I_1 = \frac{\pi}{2a}$. Ceci étant, la dépendance en a de I_n est visiblement très simple : l'homogénéité montre que $I(a)$ est homogène à a^{-2n+1} ; on peut se débarrasser de a en posant $x = ax'$; en outre, l'intégrand étant pair :

$$I_n(a) = \frac{1}{2a^{2n-1}} \int_{-\infty}^{+\infty} \frac{1}{(x'^2 + 1)^n} dx'. \quad (6.39)$$

La suite du calcul procède comme précédemment, à l'aide d'un contour comme sur la fig. 6.1 (mais on pourrait ici tout autant refermer par le bas, sans oublier le signe devant $2i\pi$!). La fonction à intégrer a deux pôles d'ordre n en $z = \pm i$, mais seul l'un d'entre eux est à considérer, $+i$ si on referme en haut, $-i$ dans le cas contraire. Si on ferme en haut, il faut donc calculer le résidu au pôle d'ordre n $z = i$; la formule générale s'écrit :

$$\begin{aligned} \text{Res} \left(\frac{1}{(z^2 + 1)^n}, z = +i \right) &= \frac{1}{(n-1)!} \left[\frac{d^{n-1}}{dz^{n-1}} \left((z-i)^n \frac{1}{(z^2 + 1)^n} \right) \right]_{z=i} = \\ &= \frac{1}{(n-1)!} \left[\frac{d^{n-1}}{dz^{n-1}} \frac{1}{(z+i)^n} \right]_{z=i}. \end{aligned} \quad (6.40)$$

argument dans cet intervalle, il faut que celui de a soit dans $\pm \frac{\pi}{4}$. À titre de vérification, on voit que si $\text{Arg } a = \pm \frac{\pi}{4}$, a^4 est négatif, et l'intégrale (6.35) est singulière. Enfin, connaître $I(a)$ dans ce secteur en a définit bien l'ensemble de toutes les valeurs de cette intégrale (où elle a un sens) puisque y restreindre a donne bien à a^4 toutes les valeurs complexes possibles (hormis les valeurs négatives de a^4 pour lesquelles, répétons-le, l'intégrale n'est pas bien définie).

La dérivée multiple se calcule facilement, et on trouve :

$$\operatorname{Res} \left(\frac{1}{(z^2 + 1)^n}, z = +i \right) = \frac{(-1)^{n-1} n(n+1)(n+2) \dots (2n-2)}{(n-1)! (2i)^{2n-1}}, \quad (6.41)$$

si $n \geq 2$; si $n = 1$:

$$\operatorname{Res} \left(\frac{1}{z^2 + 1}, z = +i \right) = \frac{1}{2i}. \quad (6.42)$$

Au total, l'intégrale $I_n(a)$ vaut :

$$I_n(a) = \frac{\pi}{a^{2n-1}} \frac{n(n+1)(n+2) \dots (2n-2)}{2^{2n-1}(n-1)!} \quad (n \geq 2), \quad (6.43)$$

et peut se mettre sous la forme :

$$I_n(a) = \frac{\pi}{a^{2n-1}} \frac{1.3.5 \dots (2n-3)}{2^n(n-1)!} \quad (n \geq 2), \quad I_1 = \frac{\pi}{2a} \quad (6.44)$$

Ce résultat s'exprime aussi à l'aide de la factorielle, donc à l'aide de la fonction Γ d'Euler¹² (voir chapitre 7). On peut montrer que quand n est un réel ν :

$$I_\nu(a) = \frac{\sqrt{\pi}}{2a^{2\nu-1}} \frac{\Gamma(\nu - \frac{1}{2})}{\Gamma(\nu)}, \quad \forall a > 0, \forall \nu > \frac{1}{2} \quad (6.45)$$

La restriction $\nu > \frac{1}{2}$ se comprendra bien après étude de la fonction d'Euler $\Gamma(z)$, qui a un pôle en $z = 0$: revenant à l'intégrale de départ (6.38), on voit qu'elle diverge si $2n \leftrightarrow \nu$ tend formellement vers $\frac{1}{2}$.

6.2.2 Intégrales d'une fonction rationnelle de lignes trigonométriques

Ce sont des intégrales du genre :

$$\int_0^{2\pi} R(\cos \theta, \sin \theta) d\theta \quad (6.46)$$

où $R(X, Y)$ est une fonction rationnelle (ne contenant que des puissances entières) de lignes trigonométriques circulaires. L'astuce consiste à remarquer que sur le cercle centré à l'origine et de rayon unité, on a $z = e^{i\theta}$, θ variant de 0 à 2π (par exemple) quand l'affixe de z parcourt une fois et une seule le cercle dans le sens positif ; sur le cercle, on

¹²Encore un "prolongement analytique" des entiers $n \geq 1$ vers les réels $\nu > \frac{1}{2}, \dots$ et en fait vers les complexes $\Re \nu > \frac{1}{2}$. Essayer de comprendre pourquoi l'expression (6.45) a un point singulier pour $\nu = \frac{1}{2}$.

Comme $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, l'expression (6.45) implique la combinaison $\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ avec $p = \nu - \frac{1}{2}$, $q = \frac{1}{2}$. Par définition, cette combinaison remarquable est une autre fonction spéciale importante, notée usuellement $B(p, q)$ (voir également le chapitre 7).

a $dz = ie^{i\theta}d\theta$, soit $d\theta = i^{-1}e^{-i\theta}dz = \frac{dz}{iz}$. En utilisant les formules d'Euler, une intégrale telle que (6.46) s'écrit donc :

$$\int_{\text{cercle unité}} R\left(\frac{1}{2}\left(z + \frac{1}{z}\right), \frac{1}{2i}\left(z - \frac{1}{z}\right)\right) \frac{dz}{iz} \equiv \int_{\text{cercle unité}} \mathcal{R}(z) dz, \quad (6.47)$$

où \mathcal{R} est visiblement une fraction rationnelle en z .

Par exemple, soit à calculer :

$$I_n(a) \stackrel{\text{déf}}{=} \int_0^{2\pi} \frac{\cos n\theta}{a + \cos \theta} d\theta \quad (n \in \mathbb{N}; a \in \mathbb{R}, a > 1) \quad (6.48)$$

(par la formule de Moivre, $\cos n\theta$ est bien une expression rationnelle en $\cos \theta$). La restriction $a \in]1, +\infty[$ assure que l'intégrale est bien définie. En posant $z = e^{i\theta}$, il vient¹³ :

$$I_n(a) = \int_{\text{cercle unité}} \frac{\frac{1}{2}\left(z^n + \frac{1}{z^n}\right)}{a + \frac{1}{2}\left(z + \frac{1}{z}\right)} \frac{dz}{iz}, \quad (6.49)$$

soit :

$$I_n(a) = i^{-1} \int_{\text{cercle unité}} \frac{z^{2n} + 1}{z^n(z^2 + 2az + 1)} dz. \quad (6.50)$$

Il y a un pôle d'ordre n à l'origine, ce qui est une bonne occasion d'appliquer (astucieusement) la formule donnant le résidu en un tel pôle¹⁴. En outre, les deux racines (simples, produit des racines égal à 1) du trinôme au dénominateur donnent chacune un pôle d'ordre un, $z_{\pm} = -a \pm \sqrt{a^2 - 1} \in \mathbb{R}$. z_+ est effectivement de module inférieur à 1, donc dans le cercle (comme $z_-z_+ = 1$, z_- est en-dehors du cercle). D'où :

$$I_n(a) = 2i\pi \frac{1}{i} (\text{Res au pôle d'ordre } n \text{ en } 0 + \text{Res au pôle d'ordre } 1 \text{ en } z_+). \quad (6.52)$$

Le calcul donne (utiliser $z_- = \frac{1}{z_+}$) :

$$\text{Résidu au pôle d'ordre } 1 \text{ en } z_+ = \frac{z_+^n + z_-^n}{z_+ - z_-}, \quad (6.53)$$

$$\text{Résidu au pôle d'ordre } n \text{ en } 0 = \frac{z_+^n - z_-^n}{z_+ - z_-}. \quad (6.54)$$

Au total, il vient :

$$I_n(a) = \frac{2\pi}{\sqrt{a^2 - 1}} (\sqrt{a^2 - 1} - a)^n \quad (n \in \mathbb{N}; a \in \mathbb{R}, |a| > 1) \quad (6.55)$$

¹³Noter que si n n'était pas entier, l'intégrand ne serait pas rationnel. Par ailleurs, de toute évidence $I_{-n} = I_n$.

¹⁴qui s'écrit ici :

$$c_{-1} = \frac{1}{(n-1)!} \lim_{z \rightarrow 0} \frac{d^{n-1}}{dz^{n-1}} \frac{z^{2n} + 1}{z^2 + 2az + 1}. \quad (6.51)$$

On calcule la dérivée par la formule de Newton donnant la dérivée n^e d'un produit en considérant $(z^{2n} + 1) \times \frac{1}{z^2 + 2az + 1}$; quand $z = 0$, un seul terme subsiste, que l'on trouve facilement en décomposant $\frac{1}{z^2 + 2az + 1}$ en éléments simples.

Cette fonction de a se prolonge analytiquement par continuité, en notant que l'intégrale (6.48) est réelle si $a \in \mathbb{R}$ et $|a| > 1$, et bornée¹⁵ si $a \in \mathbb{C}$, $\Im a \neq 0$; en outre, sur la définition (6.48), on voit que $I(a^*) = I^*(a)$. Les deux points de branchement étant ± 1 , la coupure est le segment $[-1, +1]$ de l'axe réel.

Par ailleurs, sur la définition (6.48), on voit que $I_{-n} = I_n$, une symétrie qui peut s'interpréter comme suit. Compte tenu du fait que $\sin \theta$ est impair sur $[0, 2\pi]$ de part et d'autre de π , I_n donné par (6.48) peut aussi s'écrire :

$$I_n(a) = \int_0^{2\pi} \frac{e^{-in\theta}}{a + \cos \theta} d\theta \quad (n \in \mathbb{N}; a \in \mathbb{R}, a > 1) . \quad (6.56)$$

Autrement dit, $\frac{1}{2\pi}I_n$ est la composante de Fourier de la fonction (paire) 2π -périodique $\frac{1}{a + \cos \theta}$. On en déduit l'égalité :

$$\frac{1}{a + \cos \theta} = \sum_{n=-\infty}^{+\infty} \frac{I_n(a)}{2\pi} e^{in\theta} . \quad (6.57)$$

Afin de simplifier les écritures, on pose $a = \cosh \phi$ ($\Re \phi > 0$ si $a \in \mathbb{C}$, $\phi > 0$ si $a > 1$), de sorte que :

$$z_+ = -e^{-\phi} , \quad z_- = -e^{+\phi} , \quad I_n(a) = 2\pi \frac{(-1)^n e^{-|n|\phi}}{\sinh \phi} ; \quad (6.58)$$

alors (6.57) s'écrit :

$$\frac{1}{\cosh \phi + \cos \theta} = \frac{1}{\sinh \phi} \sum_{n=-\infty}^{+\infty} e^{-|n|\phi} e^{in(\theta+\pi)} , \quad (6.59)$$

soit :

$$\frac{1}{\cosh \phi - \cos \theta} = \frac{1}{\sinh \phi} \sum_{n=-\infty}^{+\infty} e^{-|n|\phi} e^{in\theta} . \quad (6.60)$$

La série au second membre peut d'ailleurs se sommer facilement, et on retrouve bien ce qu'il faut.

◆ *Remarque*

L'évocation des séries de Fourier qui vient d'être faite permet de revoir le lien déjà mentionné (voir (5.46) p.196) entre composantes de Fourier d'une fonction périodique et coefficients de Laurent d'une fonction qui lui est associée. Soit

¹⁵Dès que a a une partie imaginaire non-nulle, le dénominateur dans l'intégrand de (6.48) ne peut jamais s'annuler : l'intégrale est de ce fait toujours définie $\forall a$, $\Im a \neq 0$. En revanche, si $a \in \mathbb{R}$, $I(a)$ n'est pas définie sur $[-1, +1]$.

$f(\theta)$ une fonction 2π -périodique munie des propriétés voulues pour admettre le développement de Fourier :

$$f(\theta) = \sum_{n=-\infty}^{+\infty} f_n e^{in\theta} , \quad f_n = \frac{1}{2\pi} \int_0^{2\pi} e^{-in\theta} f(\theta) d\theta . \quad (6.61)$$

Définissons maintenant la fonction $F(z = e^{i\theta}) = f(\theta)$. La relation donnant les composantes f_n est alors :

$$f_n = \frac{1}{2\pi} \int_C z^{-n} F(z) \frac{dz}{iz} = \frac{1}{2i\pi} \int_C \frac{F(z)}{z^{n+1}} dz , \quad (6.62)$$

où C est le cercle de rayon unité centré à l'origine. (6.62) montre que f_n n'est autre que le coefficient c_n du développement de Laurent de $F(z)$ centré en O . Cette identification est en fait évidente sur la définition même des deux fonctions :

$$f(\theta) = \sum_{n=-\infty}^{+\infty} f_n e^{in\theta} = \sum_{n=-\infty}^{+\infty} f_n z^n = F(z) ; \quad (6.63)$$

si F a le développement de Laurent $\sum_{n=-\infty}^{+\infty} c_n z^n$, l'égalité de droite dit immédiatement que $c_n = f_n$. \blacklozenge

6.2.3 Intégrales de Fourier

Ce sont des intégrales du genre :

$$\int_{-\infty}^{+\infty} e^{ikx} f(x) dx \quad (k \in \mathbb{R}) . \quad (6.64)$$

Pour évaluer ces intégrales, l'idée est de refermer par un demi-cercle comme sur la fig. 6.1 et d'invoquer le lemme 3 de Jordan, (6.5), mais attention : le signe de k conditionne le "bon" demi-cercle, celui qui donne une contribution nulle dans la limite $R \rightarrow +\infty$. En effet, sur le demi-cercle supérieur $|e^{ikRe^{i\theta}}| = e^{-kR \sin \theta}$, qui ne tend vers zéro que si $0 < k \sin \theta$: c'est bien le cas pour ce demi-cercle supérieur quand $k > 0$. Quand $k < 0$, il faut visiblement refermer par le demi-cercle *inférieur* pour que l'intégrale supplémentaire, apparaissant pour avoir un contour fermé, donne une contribution nulle¹⁶.

Soit par exemple à calculer l'intégrale :

$$I(k) \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} \frac{e^{ikx}}{x^2 + a^2} dx \quad (k \in \mathbb{R}, a \in \mathbb{R}_+) , \quad (6.65)$$

¹⁶On peut aussi refermer avec le "mauvais" demi-cercle – qui donne une contribution finie, évidemment –, mais il faut alors calculer l'intégrale correspondante. Si cette façon de faire est particulièrement malhabile, elle donne évidemment le même résultat, moyennant un calcul beaucoup plus laborieux.

qui apparaît très souvent en Physique¹⁷. On introduit donc $f(z) = \frac{e^{ikz}}{z^2+a^2}$, une fonction qui a deux pôles simples en $\pm ia$, et on referme en haut pour $k > 0$, d'où l'intégrale de contour :

$$\tilde{I}_R = \int_{-R}^{+R} \frac{e^{ikx}}{x^2+a^2} dx + \int_0^\pi \frac{e^{ikRe^{i\theta}}}{(Re^{i\theta})^2+a^2} d(Re^{i\theta}) . \quad (6.66)$$

La deuxième intégrale est nulle à la limite $R \rightarrow +\infty$, d'où :

$$\lim_{R \rightarrow +\infty} \tilde{I}_R = I(k) . \quad (6.67)$$

D'un autre côté, \tilde{I}_R s'obtient immédiatement par résidus ; seul le pôle en $z = ia$ est dans le contour, et le résidu vaut $\lim_{z \rightarrow ia} \left[(z-ia) \frac{e^{ikz}}{z^2+a^2} \right] = \frac{e^{(ik)(ia)}}{2ia}$; dès que $R > a$, on a donc :

$$\tilde{I}_R = 2i\pi \frac{e^{(ik)(ia)}}{2ia} = \frac{\pi}{a} e^{-ka} \quad (k > 0, R > a) ; \quad (6.68)$$

utilisant alors (6.67), on obtient l'intégrale cherchée :

$$I(k) = \frac{\pi}{a} e^{-ka} \quad (k > 0) . \quad (6.69)$$

Pour trouver $I(k)$ quand $k < 0$, il suffit de noter que le changement de variable $x' = -x$ dans (6.65) donne immédiatement $I(-k) = I(k)$, de sorte que :

$$I(k) = \frac{\pi}{a} e^{-|k|a} \quad \forall k \in \mathbb{R} . \quad (6.70)$$

Toutefois, cela vaut la peine de retrouver ceci par résidus. Quand k est négatif, on ferme par le demi-cercle inférieur ; le contour étant parcouru dans le sens négatif, il apparaît un signe $-$ global, que l'on traduit par le facteur $(-2i\pi)$ dans l'application du théorème des résidus. Le pôle pertinent est maintenant en $-ia$, le résidu vaut $\frac{e^{(ik)(-ia)}}{-2ia}$. On trouve donc :

$$I(k) = (-2i\pi) \frac{e^{(ik)(-ia)}}{-2ia} = \frac{\pi}{a} e^{ka} \quad (k < 0) . \quad (6.71)$$

En comparant (6.69) et (6.71), on retrouve bien (6.70), dont le graphe est une courbe en toile de tente. Enfin, notons que la valeur de l'intégrale en $k = 0$ est obtenu en prenant la primitive, et donne $I(k = 0) = \frac{\pi}{a}$. Au total :

$$\boxed{I(k) = \frac{\pi}{a} e^{-|k|a} \quad (\forall k \in \mathbb{R})} \quad (6.72)$$

Il est instructif de remarquer que $I(k)$, (6.70), n'est pas dérivable en $k = 0$. Ceci vient du fait que l'intégrand de (6.65) ne décroît pas assez vite quand $x \rightarrow +\infty$ (on comprendra pourquoi au chapitre 8 consacré à l'analyse de Fourier). On doit déjà retenir que, pour une intégrale de Fourier, ce qui se passe près de $k = 0$ est conditionné par le

¹⁷C'est la transformée de Fourier d'une lorentzienne.

comportement du facteur de e^{ikx} dans l'intégrand quand $|x| \rightarrow +\infty$ (et réciproquement) ; la décroissance *lente* de $\frac{1}{x^2+a^2}$ à *l'infini* pour x induit une singularité à *l'origine* pour $I(k)$. k et x sont en quelque sorte des variables¹⁸ *duales* (on dit aussi *conjuguées*). D'ailleurs, si on fabrique l'intégrale $I^{(1)}$ en prenant la dérivée en k de l'intégrand, et en précisant la définition concernant les bornes infinies :

$$I^{(1)}(k) \stackrel{\text{déf}}{=} \lim_{R \rightarrow +\infty} \int_{-R}^{+R} \frac{ix e^{ikx}}{x^2 + a^2} dx \quad (k \in \mathbb{R}, a \in \mathbb{R}_+) , \quad (6.73)$$

un calcul de résidus montre que :

$$I^{(1)}(k) = -\pi \operatorname{sgn} k e^{-|k|a} , \quad (6.74)$$

où $\operatorname{sgn} k = +1$ (resp. -1) si $k > 0$ (resp. $k < 0$). $I^{(1)}(k)$ est effectivement égal à $I'(k)$, sauf en $k = 0$: selon (6.73), $I^{(1)}(k = 0)$ est visiblement nulle par symétrie, alors que la dérivée $I'(k)$ n'est pas définie en $k = 0$. C'est un exemple où la limite de l'intégrale n'est pas l'intégrale de la limite.

On note enfin que $I^{(1)}(k)$ a une singularité plus violente que $I(k)$: c'est parce que l'intégrand de (6.73) décroît à l'infini encore plus lentement que celui de (6.65) (comme d'habitude, la dérivation aggrave les singularités quand il y en a).

On s'y prend de la même façon pour calculer :

$$J_a(k) \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} \frac{e^{ikx}}{x - ia} dx \quad (a \in \mathbb{R}_+, k \in \mathbb{R}^*) . \quad (6.75)$$

Le calcul par résidus donne $J_a(k) = 2i\pi e^{-ka}$ si $k > 0$, $J_a(k) = 0$ si $k < 0$. Pour $k = 0$, on convient de *définir* la valeur de la fonction $J(k)$ comme le nombre obtenu en faisant $k = 0$ *sous* l'intégrale ; on connaît alors une primitive de l'intégrand (c'est $\ln(x - ia)$), et le calcul direct donne $i\pi$ (variation de cette primitive entre les deux bornes $\pm\infty - ia$). En définitive :

$$J_a(k) = \begin{cases} 2i\pi e^{-ka} & \text{si } k > 0 \\ i\pi & \text{si } k = 0 \\ 0 & \text{si } k < 0 \end{cases} . \quad (6.76)$$

[?] Combien vaut $J_a(k)$ si $a \in \mathbb{R}_-$? Faire le lien entre la somme $J_a(k) + J_{-a}(k)$ et $I(k)$ définie en (6.65).

La même méthode fonctionne encore dans le cas suivant. Soit à calculer l'intégrale :

$$I \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} \frac{x^3 \sin x}{x^4 + 5x^2 + 4} dx . \quad (6.77)$$

En examinant les propriétés de parité de l'intégrand, on voit que I est aussi égale à¹⁹ :

$$I = \int_{-\infty}^{+\infty} \frac{x^3 e^{ix}}{x^4 + 5x^2 + 4} dx ; \quad (6.78)$$

¹⁸Ces considérations sont le fondement technique des relations d'incertitude de Heisenberg.

¹⁹ I est donc bel et bien une intégrale de Fourier.

(on a ajouté un intégrand impair dont l'intégrale entre $\pm\infty$ est nulle). On trouve facilement les zéros du dénominateur (racines d'une équation bicarrée), qui sont tous simples : tous les pôles sont donc d'ordre un. Fermant avec le demi-cercle dans le demi-plan supérieur, et utilisant le lemme 3 de Jordan pour les intégrands de Fourier, on trouve :

$$I = \frac{\pi}{3e^2}(4 - e) . \quad (6.79)$$

[?] Combien vaut $\int_{-\infty}^{+\infty} \frac{x^3 e^{ikx}}{x^4 + 5k^2 x^2 + 4k^4} dx$? Cette intégrale dépend-elle de k ?

6.2.4 Exemples divers

Exemple 1

Comme premier exemple, soit à calculer l'intégrale²⁰ :

$$I(k) \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{\sin kx}{x} dx , \quad (6.80)$$

comprise comme étant la limite $\lim_{R \rightarrow +\infty} \int_0^R \frac{\sin kx}{x} dx$ (l'intégrale converge simplement) ; en utilisant la parité de l'intégrand, on a :

$$I(k) = \frac{1}{2} \lim_{R \rightarrow +\infty} \int_{-R}^{+R} \frac{\sin kx}{x} dx . \quad (6.81)$$

Noter que cette intégrale est ainsi bien définie, le point $x = 0$ étant en fait une singularité artificielle de l'intégrand. En outre, la définition (6.80) dit clairement que la fonction $I(k)$ est impaire en k : $I(-k) = -I(k)$; en posant $kx = X$, on a (si $k > 0$) :

$$I(k) = \frac{1}{2} \lim_{R \rightarrow +\infty} \int_{-R/k}^{+R/k} \frac{\sin X}{X} dX , \quad (6.82)$$

ce qui peut faire croire que I ne dépend pas de k , du moins à première vue ; en fait, si $k < 0$, le même changement de variable donne $\frac{1}{2} \int_{R/|k|}^{-R/|k|} \frac{\sin X}{X} dX$, de sorte que²¹ :

$$I(k) = \frac{1}{2} \operatorname{sgn} k \int_{-\infty}^{+\infty} \frac{\sin X}{X} dX = \frac{1}{2} \operatorname{sgn} k \Im \int_{-\infty}^{+\infty} \frac{e^{iX}}{X} dX ; \quad (6.83)$$

d'ailleurs, la définition (6.80) montre clairement que $I(k)$ est une fonction impaire de k . Pour calculer cette dernière intégrale, on choisit un contour formé des segments $[-R, -r]$

²⁰Selon l'usage courant, l'intégrale $I(k)$ est associée aux noms de Laplace ou de Dirichlet ; elle a cependant été calculée pour la première fois par Euler en 1781.

²¹Noter que seule compte la partie imaginaire, tant que l'on ne s'intéresse qu'à l'intégrale $I(k)$, qui est bien définie. En revanche, la partie réelle de $\int_{-\infty}^{+\infty} \frac{e^{iX}}{X} dX$ n'a de sens qu'en tant que partie principale de Cauchy (et alors elle vaut zéro).

et $[r, R]$ de l'axe réels, reliés par un petit demi-cercle c_r de centre O et de rayon r , et un grand demi-cercle C_R centré en O et de rayon R , tous deux situés dans le demi-plan supérieur²² ; en vertu du théorème de Cauchy, l'intégrale le long de ce contour est nulle. Par le 3^e lemme de Jordan, l'intégrale le long du grand demi-cercle donne zéro dans la limite $R \rightarrow +\infty$. Il reste donc :

$$\int_{-R}^{-r} \frac{e^{iX}}{X} dX + \int_{\pi}^0 \frac{e^{ire^{i\theta}}}{re^{i\theta}} d(re^{i\theta}) + \int_r^R \frac{e^{iX}}{X} dX = 0 . \quad (6.84)$$

L'intégrale du milieu est $-i \int_0^{\pi} e^{ire^{i\theta}} d\theta$; comme le développement en série entière de e^z converge uniformément dans le plan ouvert, cette intégrale est égale à

$$-i \sum_{n \in \mathbb{N}} \frac{(ir)^n}{n!} \int_0^{\pi} e^{in\theta} d\theta ;$$

dans la limite $r \rightarrow 0$, seul subsiste²³ le terme $n = 0$ de la sommation, donnant au total $-i\pi$. Une fois prises les deux limites $r \rightarrow 0$ et $R \rightarrow +\infty$, la partie imaginaire de (6.84) s'écrit :

$$\Im \left(\int_{-\infty}^{+\infty} \frac{e^{iX}}{X} dX - i\pi \right) = 0 , \quad (6.85)$$

d'où finalement (utilisant (6.83)) :

$$\boxed{\int_0^{+\infty} \frac{\sin kx}{x} dx = \operatorname{sgn} k \frac{\pi}{2}} \quad (6.86)$$

En Physique, on rencontre ici et là l'intégrale :

$$A(k) = \int_{-\infty}^{+\infty} \frac{\sin kx}{x} \phi(x) dx , \quad (6.87)$$

où, si $k > 0$ est l'inverse d'une longueur (un nombre d'onde par exemple), x est une longueur. La présence du sinus entraîne que les valeurs de x qui comptent vraiment sont celles autour de 0 pour lesquelles le produit kx est au plus voisin de l'unité : $k|x| \lesssim 1$. Dès lors, et si $\phi(x)$ est une fonction lentement variable à l'échelle $\Delta x \stackrel{\text{déf}}{=} k^{-1}$ ($k\Delta x \gg 1$, ce que l'on traduit sommairement par $k \rightarrow +\infty$), on peut écrire :

$$A(k) \simeq \phi(0) \int_{-\infty}^{+\infty} \frac{\sin kx}{x} dx = \phi(0) \times \pi \quad (k > 0) . \quad (6.88)$$

Ce résultat intuitif se formalise comme suit. Dire que ϕ varie lentement à l'échelle k^{-1} , c'est, physiquement parlant, s'autoriser à "passer à la limite formelle $k \rightarrow +\infty$ ", d'où

²²Il est vivement conseillé de reprendre ce calcul en choisissant le petit demi-cercle dans le demi-plan inférieur pour voir ce qui se passe.

²³Noter que, indépendamment de la valeur de r , toutes les intégrales avec n pair non-nul sont nulles.

l'égalité utile²⁴ :

$$\lim_{k \rightarrow +\infty} \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\sin kx}{x} \phi(x) dx = \phi(0) , \quad (6.90)$$

valable pour toute "bonne" fonction $\phi(x)$ lentement variable. Cette égalité est d'usage fréquent en Physique (par exemple en Mécanique Quantique), et est l'une des pierres angulaires de la Théorie des distributions²⁵.

[?] On rencontre aussi souvent $\int_{-\infty}^{+\infty} \frac{\sin^2 x}{x^2} dx$: montrer que cette intégrale est égale à π . En déduire que si $\delta_T(\omega - \omega_0) \stackrel{\text{d\'ef}}{=} \frac{1}{\omega - \omega_0} \sin[\frac{1}{2}(\omega - \omega_0)T]$, et si $\omega_0 T \gg 1$, alors :

$$[\delta_T(\omega - \omega_0)]^2 \simeq \frac{\pi}{2} \delta(\omega - \omega_0) , \quad (6.92)$$

où $\delta(x)$ désigne la "fonction" de Dirac. Une telle relation est utilisée notamment en Mécanique quantique, où elle permet de définir une probabilité de transition indépendante du temps pour un processus réel (où l'énergie est conservée).

Exemple 2

Donnons un exemple qui utilise une astuce revenant assez souvent en pratique sous une forme ou une autre. Soit à calculer l'intégrale (due à Euler) :

$$I(a) \stackrel{\text{d\'ef}}{=} \int_{-\infty}^{+\infty} \frac{e^{ax}}{e^x + 1} dx \quad (0 < \Re a < 1) ; \quad (6.93)$$

cette intégrale intervient fréquemment en Mécanique statistique quand il est question de bosons. Il est facile de voir que cette intégrale converge aux deux infinis (et indépendamment) pourvu que $0 < \Re a < 1$, d'où la précision ci-dessus à droite ; la relation (6.93) définit donc une fonction $I(a)$ analytique dans la bande ouverte délimitée par les deux droites verticales d'abscisses 0 et 1. Notons que l'intégrale peut se calculer par des moyens élémentaires : on la coupe en deux, de $-\infty$ à 0 et de 0 à $+\infty$, on pose $x' = -x$ dans l'un des deux morceaux, et on développe en série entière $(1 + e^{-x})^{-1}$ ($x > 0$) ; il reste alors à effectuer des intégrations élémentaires²⁶.

²⁴qui s'écrit aussi :

$$\lim_{k \rightarrow +\infty} \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\sin kx}{x} \phi(x) dx = \lim_{k \rightarrow +\infty} \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\sin X}{X} \phi(X/k) dX = \phi(0) . \quad (6.89)$$

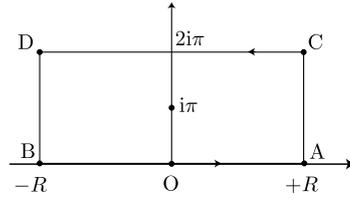
²⁵Tout comme :

$$\boxed{\lim_{n \rightarrow +\infty} \frac{n}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-n^2 x^2 / 2} \phi(x) dx = \phi(0)} \quad (6.91)$$

(voir chapitre 2).

²⁶On trouve ainsi :

$$I(a) = \sum_{n \in \mathbb{Z}} \frac{(-1)^n}{n + a} . \quad (6.94)$$

Figure 6.2: Contour C d'intégration pour le calcul de l'intégrale réelle (6.93).

Tout naturellement, on introduit la fonction $f(z)$:

$$f(z) = \frac{e^{az}}{e^z + 1}, \quad (6.95)$$

et l'astuce consiste à utiliser le fait que si z augmente de $2i\pi$, f est simplement multipliée par $e^{2i\pi a}$:

$$f(z + 2i\pi) = \frac{e^{a(z+2i\pi)}}{e^{z+2i\pi} + 1} = \frac{e^{2i\pi a} e^z}{e^z + 1} = e^{2i\pi a} f(z). \quad (6.96)$$

Ceci suggère d'introduire le contour rectangulaire BACDB (parcouru dans le sens positif) dont les grands côtés sont le segment BA $[-R, +R]$ sur l'axe réel, et le segment CD translaté verticalement de $+2i\pi$ (fig.6.2). Les petits côtés sont les deux segments verticaux AC et DB allant de $\pm R$ à $\pm R + 2i\pi$. À l'intérieur de ce contour, il y a un pôle d'ordre un en $z = i\pi$; le résidu correspondant est²⁷ :

$$\lim_{z \rightarrow i\pi \mp \varepsilon} (z - i\pi) \frac{e^{a(i\pi + \varepsilon)}}{e^{i\pi + \varepsilon} + 1} = \lim_{\varepsilon \rightarrow 0} \varepsilon \frac{e^{a(i\pi + \varepsilon)}}{-e^\varepsilon + 1} = -e^{i\pi a}. \quad (6.97)$$

D'où :

$$\int_{\text{BACDB}} \frac{e^{az}}{e^z + 1} dz = (2i\pi)(-e^{i\pi a}). \quad (6.98)$$

Par ailleurs, grâce à la propriété exprimée en (6.96), on a :

$$\int_{\text{BACDB}} f(z) dz = (1 - e^{2i\pi a}) \int_{\text{BA}} f(z) dz + \int_{\text{AC} \cup \text{DB}} f(z) dz. \quad (6.99)$$

Montrons maintenant en détail que les deux segments verticaux AC et DB donnent une contribution nulle à la limite R infini. Le long du segment vertical AC, $z = R + iy$, $0 \leq y \leq 2\pi$, donc e^{R+iy} part de e^R en $y = 0$, vaut $-e^R$ pour $y = \pi$ et revient à sa valeur de départ pour $y = 2i\pi$. En conséquence, le module du dénominateur varie de $|1 + e^R|$ à $|1 - e^R|$ sur ce segment. Pour $R > 0$, la borne inférieure du module du dénominateur est donc $|e^R - 1|$. En ce qui concerne le numérateur, et posant²⁸ $a = \alpha e^{i\phi}$, $\alpha > 0$, $-\frac{\pi}{2} < \phi < +\frac{\pi}{2}$, on a :

$$|e^{a(R+iy)}| = |e^{\alpha R e^{i\phi}} e^{i\alpha y e^{i\phi}}| = e^{\alpha R \cos \phi} e^{-\alpha y \sin \phi} \leq e^{\alpha R \cos \phi}. \quad (6.100)$$

²⁷Le résidu se calcule aussi par $\left(\frac{e^{az}}{(e^z+1)'}\right)_{z=e^{i\pi}} = \frac{e^{i\pi a}}{e^{i\pi}} = -e^{i\pi a}$.

²⁸Comme a est dans la bande $0 < \Re a < 1$, l'argument de a est borné entre $\pm \frac{\pi}{2}$.

En rassemblant les majorations, il vient :

$$\left| \frac{e^{a(R+iy)}}{e^{R+iy} + 1} \right| \leq e^{\alpha R \cos \phi} \frac{1}{e^R - 1} = \frac{e^{-(1-\alpha \cos \phi)R}}{1 - e^{-R}} . \quad (6.101)$$

Comme $\alpha \cos \phi \equiv \Re a < 1$, le module de la contribution du segment AC est borné par une quantité tendant vers zéro exponentiellement quand $R \rightarrow +\infty$. *Mutatis mutandis*, le même argument vaut pour le segment DB. En conséquence :

$$\lim_{R \rightarrow +\infty} \int_{\text{BACDB}} f(z) dz = (1 - e^{2i\pi a}) \lim_{R \rightarrow +\infty} \int_{\text{BA}} f(z) dz , \quad (6.102)$$

de sorte que, selon (6.98) :

$$(1 - e^{2i\pi a}) \lim_{R \rightarrow +\infty} \int_{-R}^{+R} \frac{e^{az}}{e^z + 1} dz = (2i\pi)(-e^{i\pi a}) , \quad (6.103)$$

soit :

$$(1 - e^{2i\pi a}) I = -2i\pi e^{i\pi a} \iff I = \frac{\pi}{\sin \pi a} . \quad (6.104)$$

Au total :

$$\boxed{\int_{-\infty}^{+\infty} \frac{e^{ax}}{e^x + 1} dx = \frac{\pi}{\sin \pi a} \quad (0 < \Re a < 1)} \quad (6.105)$$

Sur l'expression finale on vérifie que $I(a)$ diverge si $\Re a = 0$ ou 1 , ce que l'on sait depuis le début ; il ne faut jamais négliger ce genre de vérification²⁹. D'un autre côté, le second membre de (6.105) est défini partout dans le plan $\mathbb{C} \setminus \mathbb{Z}$; les deux fonctions $I(a)$ et $I_1(a) = \frac{\pi}{\sin \pi a}$ coïncident dans la bande $0 < \Re a < 1$: la fonction $I_1(a)$ est bien le prolongement analytique³⁰ de $I(a)$ à l'extérieur de la bande.

? En déduire l'égalité :

$$\boxed{\frac{1}{\sin \pi z} = \frac{1}{\pi} \sum_{n \in \mathbb{Z}} \frac{(-1)^n}{z - n}} \quad (6.106)$$

qui est un exemple de développement de Mittag - Leffler (développement d'une fonction méromorphe sur ses pôles – voir sous-section 6.8.1).

²⁹qui n'en est une que si l'intégrale de départ est absolument convergente, ce qui est le cas ici. En revanche, une intégrale *semi-convergente* peut présenter des discontinuités (l'intégrale de la limite peut être distincte de la limite de l'intégrale).

³⁰D'après le théorème du prolongement analytique vu antérieurement, il suffit que les fonctions coïncident sur la frontière commune des domaines.

Exemple 3

Par des moyens analogues, on montre que^{31,32} :

$$I(a, \alpha) \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{\cos ax}{\cosh x + \cosh \alpha} dx = \frac{\pi \sin \alpha a}{\sinh \pi a \sinh \alpha} \quad (\alpha \in \mathbb{R}_+, -1 < \Im a < +1) . \quad (6.107)$$

Sur le résultat final, on observe que $I(a, \alpha)$, considérée comme une fonction de α seul, a une singularité éliminable en $\alpha = 0$: en effet, l'intégrale de départ est bien définie en ce point³³. De la même façon, considéré comme une fonction de a , le membre de droite a aussi une singularité éliminable en $a = 0$, en conformité avec le fait que l'intégrale $I(a, \alpha)$ est bien définie en $a = 0$. Par ailleurs, le résultat (6.107) se prolonge analytiquement et définit une fonction méromorphe de α ou de a ayant des pôles en $\alpha_k = ik\pi$ et en $a_l = il$ (k et l entiers non-nuls).

Exemple 4

D'une façon plus générale, si $f(z)$ est une fonction méromorphe de période $i\omega$ ayant des pôles z_k dans la bande $0 < \Im z < \omega$ et bornée par $e^{-p\Re z}$ quand $\Re z \rightarrow \pm\infty$ ($p > 0$). Alors :

$$\int_{-\infty}^{+\infty} e^{ax} f(x) dx = \frac{2i\pi}{1 - e^{ia\omega}} \sum_k \text{Res}[e^{az} f(z), z_k] \quad (|\Re a| < p) . \quad (6.109)$$

Par exemple :

$$\int_{-\infty}^{+\infty} \frac{e^{ax} - e^{bx}}{1 - e^x} dx = \pi(\cot \pi a - \cot \pi b) \quad (0 < \Re a, \Re b < 1) . \quad (6.110)$$

³¹Noter l'existence des deux pôles simples pertinents : $z_{\pm} = \pm\alpha + i\pi$ dans le contour rectangulaire.

³²Petite suggestion à l'usage des fanatiques du calcul formel sur machine qui croient qu'à notre époque moderne il n'est plus utile d'apprendre l'intégration : soumettre cette intégrale à un logiciel de calcul formel, et observer la réponse (!?) de la machine... (pour un autre canular pas drôle, voir la note 46).

Un an plus tard : on n'arrête pas le progrès : la version $N + 1$ du même logiciel sait maintenant se tirer d'affaire pour l'intégrale ci-dessus ; pour l'autre, faut-il attendre la version $N + q$?

³³D'où, en sous-produit, les égalités :

$$\int_0^{+\infty} \frac{\cos ax}{\cosh x + 1} dx = \frac{\pi a}{\sinh \pi a} , \quad \int_0^{+\infty} \frac{1}{\cosh x + \cosh \alpha} dx = \frac{\alpha}{\sinh \alpha} , \quad (6.108)$$

et d'autres encore (prendre a et α imaginaires purs, etc).

Exemple 5

L'intégrale gaussienne³⁴ apparaît très fréquemment. Soit à calculer :

$$I(a, k) \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} e^{-ax^2+ikx} dx . \quad (6.111)$$

Pour l'instant, on suppose $a \in \mathbb{R}_+$ et $k \in \mathbb{R}$ (noter qu'avec ces hypothèses, $I(a, k) \in \mathbb{R}$ puisque la partie imaginaire de l'intégrand est impaire en x). Le calcul peut se conduire de plusieurs façons équivalentes ; l'une d'entre elles est la suivante.

L'argument de l'exponentielle se transforme comme suit³⁵ :

$$ax^2 - ikx = a\left(x^2 - i\frac{kx}{a}\right) = a\left[\left(x - i\frac{k}{2a}\right)^2 + \frac{k^2}{4a^2}\right] , \quad (6.112)$$

de sorte que :

$$I(a, k) = e^{-\frac{k^2}{4a}} \int_{-\infty}^{+\infty} e^{-a\left(x - i\frac{k}{2a}\right)^2} dx . \quad (6.113)$$

Posons maintenant $z = x - i\frac{k}{2a}$, d'où $dz = dx$; maintenant, le chemin suivi par la nouvelle variable d'intégration est la droite D parallèle à Ox d'ordonnée $-\frac{k}{2a}$, parcourue de gauche à droite :

$$\int_{-\infty}^{+\infty} e^{-a\left(x - i\frac{k}{2a}\right)^2} dx = \int_D e^{-az^2} dz . \quad (6.114)$$

En vertu du théorème de Cauchy, cette dernière intégrale a la même valeur si on déforme D en rabattant cette droite sur l'axe Ox , tout en maintenant ses extrémités aux points $\pm\infty - i\frac{k}{2a}$. En fait, les intégrales le long de ces deux "petits" segments verticaux sont nulles³⁶, de sorte que finalement :

$$\int_D e^{-az^2} dz = \int_{\mathbb{R}} e^{-az^2} dz \equiv \int_{-\infty}^{+\infty} e^{-ax^2} dx . \quad (6.115)$$

On verra³⁷ au chapitre 7 comment calculer cette dernière intégrale, qui vaut $\sqrt{\frac{\pi}{a}}$. En définitive, selon (6.113) :

$$I(a, k) = \sqrt{\frac{\pi}{a}} e^{-\frac{k^2}{4a}} \quad (a \in \mathbb{R}_+, k \in \mathbb{R}) . \quad (6.116)$$

³⁴du nom de Carl Friedrich GAUSS (1777-1855). "L'œuvre du mathématicien allemand Carl Friedrich Gauss est un monument d'une ampleur et d'une richesse sans égale" (*Encyclopædia Universalis*). Essayer de calculer en quelques secondes la somme des 100 premiers nombres entiers, ce que fit le petit Gauss un jour où le maître d'école, lassé du chahut, avait cru trouver la paix quelques minutes, sinon plus, en posant cette question à titre de *pensum*.

³⁵Cette opération élémentaire est souvent désignée par "compléter le carré".

³⁶Le vérifier par un passage à la limite $R \rightarrow +\infty$ après introduction des segments reliant $\pm R$ à $\pm R - i\frac{k}{2a}$.

³⁷Cette intégrale peut aussi se calculer par des moyens élémentaires, grâce à l'astuce suivante, bien connue. Soit $I_2 = \int_{\mathbb{R}^2} e^{-a(x^2+y^2)} dx dy$, qui n'est autre que le carré de $\int_{\mathbb{R}} e^{-ax^2} dx$. Pour trouver I_2 , on passe en polaires, de sorte que $I_2 = \int_0^{+\infty} \int_0^{2\pi} e^{-ar^2} r dr d\theta$. L'intégration en θ donne juste un facteur 2π , d'où $I_2 = 2\pi \int_0^{+\infty} e^{-ar^2} r dr = 2\pi \int_0^{+\infty} e^{-aX} (\frac{1}{2} dX) = \frac{\pi}{a}$, et le résultat annoncé.

Sans aucune ambiguïté, le symbole $\sqrt{}$ désigne ici la racine carrée ordinaire ($\sqrt{a^2} = |a|$).

Jusqu'ici, on a supposé $a > 0$. Posons-nous maintenant la question du prolongement analytique en a de la fonction $I(a, k)$, en faisant quitter à a l'axe réel à partir d'une certaine abscisse positive, et en considérant d'emblée que l'intégrale (6.111) est une intégrale dans le plan complexe sur le chemin formé par l'axe réel. On voit tout de suite que ce prolongement est possible, par les arguments du même type de ceux déjà développés dans le chapitre 5 à propos de l'intégrale $\int_0^{+\infty} e^{-zt} dt$: quand l'argument de a devient positif et croît au point de se rapprocher de la valeur³⁸ "dangereuse" $\frac{\pi}{2}$, il suffit d'incliner la droite d'intégration (alors la variable d'intégration x prend des valeurs complexes z) en la faisant tourner dans le sens négatif, de sorte qu'à l'infini la partie réelle du produit az^2 reste positive. a parcourt tout le plan quand on restreint son argument à l'intervalle $[-\pi, +\pi]$, et la droite d'intégration pour z doit se trouver dans le secteur défini par les angles tels que $|\arg(az^2)| < \frac{\pi}{2}$, soit $|\arg z| < \frac{3\pi}{4}$. Cela étant, et par le théorème du prolongement analytique, l'expression (6.116) conserve sa validité – à la condition expresse toutefois de préciser la définition de la fonction multiforme notée pour l'instant " \sqrt{a} " (mais ce qui vient d'être dit permet d'anticiper que c'est la branche dont la coupure est le demi-axe réel négatif).

Pour savoir quelle est la bonne branche, il suffit de noter que, tant que k est réel, la relation de définition (6.111) montre que l'on a :

$$I(a^*, k) = [I(a, k)]^* . \quad (6.117)$$

La branche de " \sqrt{a} " à retenir est donc celle qui possède cette propriété de symétrie (et prend des valeurs réelles sur l'axe réel) ; comme on l'a vu au chapitre 4, c'est la branche dont la coupure est le demi-axe réel négatif. On peut aussi donner un argument de continuité : quand à $a > 0$ on ajoute une petite partie imaginaire, l'intégrale acquiert visiblement une (petite) partie imaginaire de même signe. Si donc a oscille un peu verticalement de part et d'autre d'une abscisse positive, $\Im I$ en fait autant et n'a pas de discontinuité (ce qui serait le cas si la coupure était le long de \mathbb{R}_+). Enfin, le prolongement à k complexe se fait suivant les mêmes arguments (le terme en k dans l'exponentielle n'est d'ailleurs pas critique, le terme en z^2 l'emportant toujours sur celui en z).

Ainsi, il est maintenant possible d'écrire :

$$\int_{-\infty}^{+\infty} e^{-ax^2+ikx} dx = \left(\frac{\pi}{a}\right)^{\frac{1}{2}} e^{-\frac{k^2}{4a}} , \quad -\pi < \operatorname{Arg} a < +\pi, k \in \mathbb{C} \quad (6.118)$$

où $z^{\frac{1}{2}}$ est à comprendre comme il vient d'être précisé. En guise de sous-produits, prenant $a = -i$ et³⁹ $k = 0$, on a :

$$\int_{-\infty}^{+\infty} e^{ix^2} dx = \sqrt{\pi} i^{1/2} = \sqrt{\frac{\pi}{2}}(1+i) , \quad (6.119)$$

³⁸Cette valeur est dangereuse : avec x réel et $\arg a = \frac{\pi}{2}$, la partie réelle de ax^2 s'annule (elle était auparavant positive).

³⁹La limite $k \rightarrow 0$ est visiblement uniforme puisque e^{ikx} est de module unité pour $k \in \mathbb{R}$.

d'où, séparant les parties réelles et imaginaires, les intégrales dites de Fresnel⁴⁰ :

$$\int_0^{+\infty} \cos x^2 dx = \int_0^{+\infty} \sin x^2 dx = \frac{1}{2} \sqrt{\frac{\pi}{2}} . \quad (6.120)$$

Exemple 6

Soit enfin à calculer :

$$I \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{x}{\sinh \pi x} dx . \quad (6.121)$$

Cette intégrale peut se trouver par des moyens élémentaires en écrivant $\frac{1}{\sinh \pi x} = \frac{2e^{-\pi x}}{1-e^{-2\pi x}}$, puis en développant en série géométrique $\frac{1}{1-e^{-2\pi x}}$ et en utilisant $\int_0^{+\infty} xe^{-\lambda x} dx = \frac{1}{\lambda^2}$ ($\lambda > 0$) ; on trouve ainsi⁴¹ :

$$I = 2 \sum_{n=1}^{+\infty} \frac{1}{(2n+1)^2 \pi^2} = \frac{2}{\pi^2} \left(\sum_{n=1}^{+\infty} \frac{1}{n^2} - \sum_{n=1}^{+\infty} \frac{1}{(2n)^2} \right) = \frac{2}{\pi^2} \left(1 - \frac{1}{4} \right) \zeta(2) = \frac{1}{4} . \quad (6.123)$$

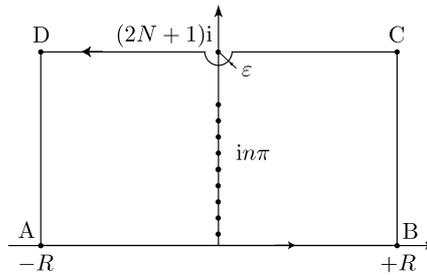


Figure 6.3: Contour C d'intégration pour le calcul de l'intégrale réelle (6.121).

⁴⁰La convergence de ces deux intégrales n'a rien d'évident *a priori* ; elle peut être prouvée en posant $x^2 = X$ et en faisant une intégration par parties qui met en évidence une intégrale visiblement convergente. Intuitivement : ces intégrales existent grâce au fait que plus x est grand, plus l'argument des lignes trigonométriques oscille vite. Par contraste, $\int_{-\infty}^{+\infty} \cos x dx$ n'existe pas au sens usuel : l'argument du cosinus n'oscille pas assez vite – toutefois, cette dernière intégrale fournit une représentation de la “fonction” de Dirac : $\delta(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ikx} dx$.

Noter que les intégrales de Fresnel peuvent aussi se calculer en considérant le contour formé du segment $[0, R]$ de l'axe réel, suivi d'un huitième de cercle de rayon R et revenant à l'origine suivant le segment incliné de $\frac{\pi}{4}$ (et en prenant bien sûr la limite $R \rightarrow +\infty$).

⁴¹ $\zeta(2) = \frac{\pi^2}{6}$, voir (6.192) p. 270. Par la même méthode, montrer que :

$$I_p \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{x^p}{\sinh \pi x} dx = \frac{2p!}{\pi^{p+1}} \left(1 - \frac{1}{2^{p+1}} \right) \zeta(p+1) \equiv \frac{2}{\pi^{p+1}} \left(1 - \frac{1}{2^{p+1}} \right) \Gamma(p+1) \zeta(p+1) \quad (6.122)$$

où $\zeta(z)$ est la fonction zeta de Riemann (voir une première définition en (6.193), et chapitre 7, section 7.3) et Γ la fonction d'Euler (voir chapitre 7).

Le calcul de I peut se faire aussi par résidus, partant de $2I = \int_{-\infty}^{+\infty} \frac{z}{\sin \pi z} dz$, puis en considérant l'intégrale $I_{N,R}$ associée au contour de la figure 6.3, formé en refermant par une droite horizontale d'ordonnée $2N+1$, $N \in \mathbb{N}^*$, contournant le pôle en $(2N+1)i\pi$, par en-bas par exemple. La fonction $\frac{z}{\sin \pi z}$ a des pôles simples en $z_n = in$, $n \in \mathbb{Z}^*$, le résidu valant $\frac{in}{\pi \cosh in\pi}$. On voit facilement que la contribution des deux côtés verticaux tend exponentiellement vite vers zéro quand ils partent à l'infini. En ce qui concerne le trajet CD contournant le pôle en $(2N+1)i\pi$, il s'écrit :

$$\int_{+R}^{+\varepsilon} \frac{x + (2N+1)i}{\sinh \pi(x + (2N+1)i)} dz + \int_{-\varepsilon}^{-R} \frac{x + (2N+1)i}{\sinh \pi(x + (2N+1)i)} dz + I_\varepsilon \quad (6.124)$$

où $I_\varepsilon \stackrel{\text{d'ef}}{=} \int_0^{-\pi} \frac{\varepsilon e^{i\theta} + (2N+1)i}{\sinh \pi[\varepsilon e^{i\theta} + (2N+1)i]} d(\varepsilon e^{i\theta})$. Avec $\sinh \pi[x + (2N+1)i] = -\sinh \pi x$, et une fois prises les limites $R \rightarrow +\infty$ et $\varepsilon \rightarrow 0$, les deux premières intégrales reconstituent la partie principale de Cauchy $\mathcal{P} \int_{-\infty}^{+\infty} \frac{z}{\sinh \pi z} dz$, mais comme l'intégrand est fini en $x = 0$, c'est simplement $2I_{N,+\infty}$; l'intégrale I_ε sur le petit demi-cercle donne $-(2N+1)$ dans la limite $\varepsilon \rightarrow 0$. L'application du théorème des résidus donne alors, $z = 0$ étant une singularité éliminable :

$$4I_{N,+\infty} - (2N+1) = 2i\pi \sum_{n=1}^{2N} \text{Res}\left(\frac{z}{\sinh \pi z}, z = in\right) = 2i\pi \sum_{n=1}^{2N} \frac{in}{\pi(-1)^n}, \quad (6.125)$$

d'où :

$$4I \equiv \lim_{N \rightarrow +\infty} 4I_{N,+\infty} = \lim_{N \rightarrow +\infty} \left((2N+1) - 2 \sum_{n=1}^{2N} (-1)^n n \right) \equiv \lim_{N \rightarrow +\infty} ((2N+1) - 2N) = 1, \quad (6.126)$$

et finalement⁴² :

$$\boxed{\int_0^{+\infty} \frac{x}{\sinh \pi x} dx = \frac{1}{4}} \quad (6.127)$$

6.3 Calcul d'intégrales de fonctions multiformes

Avec des fonctions multiformes, la première opération est de définir proprement la(es) coupure(s). Ceci étant fait, la règle de l'élastique est toujours de vigueur : outre des possibles singularités isolées, la coupure est un mur infranchissable dans toute déformation du contour, ou par toute ligne le complétant. Pour illustrer ceci, le mieux est de traiter quelques exemples, où on ne quittera pas le premier feuillet de Riemann. Bien sûr, les méthodes employées s'étendent aux chemins empruntant différents feuillets.

⁴²On peut donner un sens à la série des résidus, en régularisant cette série divergente suivant la méthode d'Abel, c'est-à-dire en la prenant comme $\lim_{x \rightarrow 1^-} \sum_{n \in \mathbb{N}^*} (-1)^{n-1} n x^{n-1}$, soit $\lim_{x \rightarrow 1^-} \frac{d}{dx} \frac{-1}{1+x} = \frac{1}{4}$.

Le procédé d'Abel permet de aussi donner un sens à l'égalité $1 - 2 + 3 - 4 + \dots = \frac{1}{4}$; pour la curiosité : tracer le graphe de $f(N) \stackrel{\text{d'ef}}{=} \sum_{n=1}^N (-1)^{n-1} n$.

Exemple 1

Soit à calculer l'intégrale⁴³ :

$$I \stackrel{\text{d\'ef}}{=} \int_0^{+\infty} \frac{1}{1+x^3} dx \equiv \int_0^{+\infty} f(x) dx . \quad (6.128)$$

L'intégrand de (6.128) n'a pas de parité déterminée, et on ne voit pas l'utilité d'étendre l'intégration à tout l'axe réel et de refermer par un grand cercle d'un côté ou de l'autre. L'astuce consiste ici à introduire une fonction $\tilde{f}(z)$ nettement différente de $f(z)$ qui est tout juste le prolongement analytique immédiat de f au champ complexe.

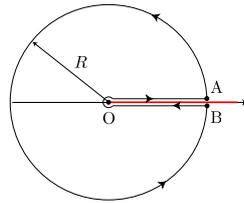


Figure 6.4: Contour utilisé pour l'intégrale (6.129). La coupure est la demi-droite en gras. On fait bien sûr tendre vers l'infini le rayon du grand cercle coupé, et vers zéro celui du petit cercle coupé autour de l'origine.

Considérons l'intégrale :

$$\tilde{I} = \int_C \tilde{f}(z) dz , \quad \tilde{f}(z) = \frac{\ln z}{1+z^3} , \quad (6.129)$$

où $\ln z$ est la branche dont la coupure est le demi-axe \mathbb{R}_+ , et où C est le contour de la figure 6.4. Sur le bord supérieur de la coupure ($z = x + i0$, $x > 0$), la branche de $\ln z$ ainsi définie vaut tout simplement $\ln x \in \mathbb{R}$; partant d'un tel point et faisant faire un tour à z autour de O , on se retrouve sur le bord inférieur de la coupure ($z = x - i0$) et la même branche ainsi suivie par continuité vaut maintenant $\ln x + 2i\pi$. Par le 1^{er} lemme de Jordan, le grand "cercle"⁴⁴ ne contribue pas à la limite. Quant au petit cercle autour de l'origine, on montre facilement que sa contribution tend vers zéro comme $|z| \ln |z|$ (c'est aussi ce que dit le lemme 2). Il en résulte :

$$\tilde{I} = \int_{\varepsilon=0_+}^{R=\infty} \frac{\ln x}{1+x^3} dx + \int_{R=\infty}^{\varepsilon=0_+} \frac{\ln x + 2i\pi}{1+x^3} dx = -2i\pi \int_0^{+\infty} \frac{1}{1+x^3} dx \equiv -2i\pi I , \quad (6.130)$$

d'où la relation précise entre I (que l'on cherche) et \tilde{I} (que l'on va calculer par résidus). En effet, $\tilde{I} = 2i\pi \Sigma \text{Res}$; dans le contour, il y a trois pôles d'ordre un, les trois racines de

⁴³Même remarque que dans la note 5.

⁴⁴Les guillemets sont là pour rappeler que le cercle est coupé d'un point.

$z^3 + 1 = 0$. Le résidu en $z_k = e^{\frac{i}{3}(2k+1)\pi}$ est $\frac{\ln z_k}{3z_k^2} = -\frac{1}{3}z_k \ln z_k$:

$$\tilde{I} = 2i\pi \sum_{k=0,1,2} (-1)^k \frac{1}{3} e^{\frac{i}{3}(2k+1)\pi} \ln e^{\frac{i}{3}(2k+1)\pi} = \frac{2\pi^2}{9} \sum_{k=0,1,2} (2k+1) e^{\frac{i}{3}(2k+1)\pi} = \frac{2\pi^2}{9} [e^{i\frac{\pi}{3}} - 3 + 5e^{-i\frac{\pi}{3}}] \quad (6.131)$$

soit $\tilde{I} = -\frac{4\pi^2 i}{3\sqrt{3}}$ et, selon (6.130), l'intégrale cherchée :

$$\int_0^{+\infty} \frac{1}{1+x^3} dx = \frac{2\pi}{3\sqrt{3}}. \quad (6.132)$$

☐ Calculer l'intégrale (6.128) en utilisant un secteur d'angle $\frac{2\pi}{3}$.

☐ Montrer que $\int_0^{+\infty} \frac{1}{x^6+1} dx = \frac{\pi}{3}$ (on pourra aussi utiliser un secteur d'angle $\frac{\pi}{3}$).

Exemple 2

Soit à calculer l'intégrale :

$$I \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{\phi(x)}{\ln^2 x + \pi^2} dx, \quad (6.133)$$

où $\phi(x)$ est une fonction supposée paire, méromorphe, analytique en $z = 0$ et sur tout le demi-axe \mathbb{R}_- , et telle que $\lim_{|z| \rightarrow \infty} |z\phi(z)| \rightarrow 0$. L'astuce consiste ici à prendre la branche du logarithme dont la coupure est le demi-axe réel négatif, et à considérer :

$$\tilde{I} = \int_C \tilde{f}(z) dz, \quad \tilde{f}(z) = \frac{\phi(z)}{\ln z}, \quad (6.134)$$

où C est le contour de la fig. 6.5.

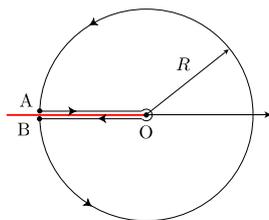


Figure 6.5: Contour utilisé pour l'intégrale (6.134). La coupure est la demi-droite en gras. On fait bien sûr tendre vers l'infini le rayon du grand cercle coupé, et vers zéro celui du petit cercle autour de O.

Sur le bord supérieur de la coupure, $z = -x + i0$ où $x \in \mathbb{R}_+$ et $\ln z = \ln |-x| + i\pi$; sur le bord inférieur $z = -x - i0$ et $\ln z = \ln |-x| - i\pi$. En décomposant le contour en morceaux et en ignorant les contributions du grand “cercle” à l’infini et du petit “cercle” autour de O qui ne comptent pas à la fin⁴⁵, il vient :

$$\tilde{I} = \int_{+\infty}^{0+} \frac{\phi(-x)}{\ln x + i\pi} d(-x) + \int_{0+}^{+\infty} \frac{\phi(-x)}{\ln x - i\pi} d(-x) = \int_{0+}^{+\infty} \frac{-2i\pi \phi(x)}{\ln^2 x + \pi^2} dx = -2i\pi I . \quad (6.135)$$

Par ailleurs, \tilde{I} se trouve par résidus en considérant les pôles de $\frac{\phi(z)}{\ln z}$, c’est-à-dire les pôles z_k de $\phi(z)$ et le pôle en $z = 1$ de $\frac{1}{\ln z}$, dont le résidu est $\lim_{z \rightarrow 1} \frac{z-1}{\ln z} = 1$. Si ϕ est analytique en $z = 1$, on a donc :

$$\tilde{I} = 2i\pi \left[\sum_k \operatorname{Res}\left[\frac{\phi(z)}{\ln z}, z_k\right] + 1 \times \phi(1) \right] , \quad (6.136)$$

d’où par (6.135) l’intégrale cherchée :

$$I = - \sum_{z_k \text{ pôle de } \phi} \operatorname{Res}\left[\frac{\phi(z)}{\ln z}, z_k\right] - \phi(1) . \quad (6.137)$$

À titre d’exemple, soit $\phi(x) = \frac{1}{x^2+a^2}$; le calcul ci-dessus exige que a^2 ne soit pas négatif (on a supposé que $\phi(x)$ était analytique sur \mathbb{R}_-), soit $\operatorname{Arg} a \neq \pm \frac{\pi}{2}$. Les résidus de $\frac{\phi(z)}{\ln z}$ aux pôles $\pm ia$ de $\phi(z)$ sont $\pm \frac{1}{2ia} \frac{1}{\ln(\pm ia)}$, la branche du \ln à considérer étant évidemment celle qui sert de définition à \tilde{I} ; (6.137) donne⁴⁶ :

$$I(a) \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{1}{x^2+a^2} \frac{1}{\ln^2 x + \pi^2} dx = -\frac{1}{2ia} \left(\frac{1}{\ln(ia)} - \frac{1}{\ln(-ia)} \right) - \frac{1}{1+a^2} , \quad (6.138)$$

avec $|a| > 0$, $|\operatorname{Arg} a| < \frac{\pi}{2}$. Ce résultat est vrai tant que $\pm ia$ ne rejoint pas la coupure, d’où la restriction ci-dessus à droite (et d’ailleurs $\ln(\pm ia)$ qui figure dans (6.138) correspond à la branche de $\ln z$ définie pour $|\operatorname{Arg} z| < \pi$). Cela étant, ia et $-ia$ ont le même module

⁴⁵Le petit cercle de rayon r autour de O donne une contribution en $\frac{r}{\ln r} \rightarrow 0$ quand $r \rightarrow 0$

⁴⁶Un certain logiciel de calcul formel déclare “*The integral is not convergent*” quand on lui soumet naïvement cette intégrale (!?). Décidément... (revenir à la note 32). Dans un cas la machine ne sait pas faire, dans l’autre, elle donne une réponse grossièrement fausse.

$|a|$ mais leurs arguments diffèrent de π , donc⁴⁷ $\ln(ia) - \ln(-ia) = i\pi$, d'où :

$$I(a) = \frac{1}{2a} \frac{\pi}{\ln(ia) \ln(-ia)} - \frac{1}{1+a^2}, \quad |\operatorname{Arg} a| < \frac{\pi}{2} \quad (6.141)$$

où la branche du logarithme est toujours la même. Dans le cas où $a \in \mathbb{R}_+$, ceci se simplifie en⁴⁸ :

$$I(a) = \frac{\pi}{2a} \frac{1}{\ln^2 a + \frac{\pi^2}{4}} - \frac{1}{1+a^2}, \quad a \in \mathbb{R}_+ . \quad (6.142)$$

Notons que si $\phi(x)$ a des pôles sur la coupure, il faut les contourner par deux petits demi-cercles et prendre convenablement la limite du rayon nul. Par exemple, on trouve ainsi :

$$\int_0^{+\infty} \frac{\ln x}{x^2-1} dx = \frac{\pi^2}{4}, \quad (6.143)$$

en employant le contour de la figure 6.4 modifié par deux petits demi-cercles autour du point $z = 1$ (et dont la contribution est ici nulle⁴⁹).

[?] Montrer que $\int_0^{+\infty} \frac{\ln x}{(x+1)\sqrt{x}} dx = 0$ et que⁵⁰ $\int_0^{+\infty} \frac{\ln x}{(x^2+a^2)} dx = \frac{\pi}{2a} \ln a$ ($\Re a > 0$).

[?] Montrer que pour $-1 < \Re \alpha < 2$, $\int_0^1 x^\alpha (1-x)^{1-\alpha} dx = \frac{\pi \alpha (1-\alpha)}{2 \sin \pi \alpha}$.

Exemple 3

Soit maintenant à calculer l'intégrale :

$$I \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{x^\alpha}{P(x)} dx, \quad (6.144)$$

⁴⁷Bien noter que cette égalité n'est vraie que pour $|\operatorname{Arg} a| < \frac{\pi}{2}$; si par exemple $\operatorname{Arg} a$ dépasse $\frac{\pi}{2}$, on a $\ln(ia) - \ln(-ia) = -i\pi$ (il faut toujours garder la même branche du logarithme), et le premier terme dans (6.141) devient $\frac{1}{2a} \frac{-\pi}{\ln(ia) \ln(-ia)}$. Quand a tend vers le demi-axe \mathbb{R}_- (soit $\operatorname{Arg} a \rightarrow +\pi_-$), ce terme vaut $\frac{-\pi}{2(-a)} \frac{1}{\ln^2 |a| + \frac{\pi^2}{4}}$. Par comparaison avec (6.142), on constate bien que $I(-a) = I(a)$, comme le montre la définition (6.138). En définitive :

$$I(a) = \begin{cases} +\frac{1}{2a} \frac{\pi}{\ln(ia) \ln(-ia)} - \frac{1}{1+a^2} & \text{si } 0 \leq |\operatorname{Arg} a| < \frac{\pi}{2} \\ -\frac{1}{2a} \frac{\pi}{\ln(ia) \ln(-ia)} - \frac{1}{1+a^2} & \text{si } \frac{\pi}{2} < |\operatorname{Arg} a| \leq \pi \end{cases}, \quad (6.139)$$

et si $a \in \mathbb{R}$:

$$I(a) = \frac{\pi}{2|a|} \frac{1}{\ln^2 |a| + \frac{\pi^2}{4}} - \frac{1}{1+a^2}, \quad \forall a \in \mathbb{R}. \quad (6.140)$$

⁴⁸Comment se comporte $I(a)$ lorsque $a \rightarrow +\infty$?

⁴⁹Il s'avère ainsi qu'il n'est pas nécessaire de définir l'intégrale du premier membre comme la partie principale de Cauchy (pour un rappel de cette définition, voir section 6.4, (6.157)), puisque en $x = 1$, l'intégrand a une limite finie : si les petits demi-cercles ne donnent rien à la limite du rayon nul, c'est parce que $z = 1$ est une singularité apparente. Dit autrement : vue simplement comme une intégrale réelle, (6.143) implique un intégrand qui ne diverge pas en $x = 1$.

Rendons justice au logiciel de calcul formel : il sait se débrouiller avec l'intégrale (6.143).

⁵⁰Poser $x^2 = X$.

où $P(x)$ est un polynôme de degré n fini en $x = 0$, n'ayant aucun zéro réel positif⁵¹, et où $-1 < \alpha < n - 1$. Par les lemmes de Jordan, l'intégrale le long d'un grand cercle de rayon R tend vers zéro si $R \rightarrow +\infty$, tout comme celle le long d'un cercle centré en O dont le rayon tend vers zéro. En prenant le même contour que celui de la figure 6.4, on définit une intégrale \tilde{I} calculable par résidus, et qui est égale à $(1 - e^{2i\pi\alpha})I$. Par le théorème des résidus, on a :

$$\tilde{I} = 2i\pi \sum \operatorname{Res}\left[\frac{z^\alpha}{P(z)}, z_k\right]. \quad (6.145)$$

En particulier, si les n zéros de $P(z)$ sont tous des zéros d'ordre 1, tous les pôles de $\frac{1}{P(z)}$ sont d'ordre un et il vient finalement :

$$I = \frac{2i\pi}{1 - e^{2i\pi\alpha}} \sum_{k=1}^n \frac{z_k^\alpha}{P'(z_k)}. \quad (6.146)$$

C'est ainsi que l'on établit :

$$\int_0^{+\infty} \frac{x^\alpha}{x^n + 1} dx = \frac{\pi}{n} \frac{1}{\sin \frac{\alpha+1}{n}\pi}, \quad (\alpha > -1, n - \alpha > 1). \quad (6.147)$$

L'expression au second membre diverge pour $\alpha \rightarrow -1$ et pour $\alpha \rightarrow n - 1$, comme il se doit⁵².

Exemple 4

Soit enfin à calculer l'intégrale :

$$I(\nu) \stackrel{\text{d'ef}}{=} \int_0^{+\infty} \frac{1}{1 + x^{2\nu}} dx, \quad (6.148)$$

qui ne converge que si $\nu > \frac{1}{2}$. En posant $x^{2\nu} = X$, et en choisissant la branche $x = X^{\frac{1}{2\nu}}$ d'argument nul si $X \in \mathbb{R}_+$, $I(\nu)$ s'écrit :

$$I(\nu) = \frac{1}{2\nu} \int_0^{+\infty} \frac{X^\alpha}{1 + X} dX, \quad \alpha = \frac{1}{2\nu} - 1, -1 < \alpha < 0; \quad (6.149)$$

utilisant (6.147) avec $n = 1$, on en déduit :

$$I(\nu) = \frac{\pi}{2\nu} \frac{1}{\sin \frac{\pi}{2\nu}}. \quad (6.150)$$

Il est intéressant de rapprocher ceci du résultat (6.30), p. 237 : on voit que (6.150) le généralise au cas non-entier (c'est aussi, pour $\nu = \frac{3}{2}$, l'intégrale (6.128)). On ne peut

⁵¹Si c'était le cas, l'intégrale serait impropre et relèverait d'une régularisation, voir section 6.4.

⁵²? Que se passe-t-il pour $\alpha = -1$? Par ailleurs, quelle est l'expression de l'intégrale si $\alpha \in \mathbb{C}$, avec $-1 < \Re\alpha < n - 1$?

pas dire trop vite qu'il s'agit d'un prolongement analytique au sens strict, car la quantité I_n définie en (6.26) n'est pas une *fonction analytique* de la variable n : telle qu'elle est définie, n est d'emblée supposé entier⁵³ – on alors a notamment utilisé ce fait en invoquant la parité de l'intégrand pour introduire l'intégrale entre $\pm\infty$. La généralisation (le "prolongement") d'une quantité ponctuelle vers une fonction continue est fréquente ; par exemple, c'est ainsi que la fonction $\Gamma(z)$ d'Euler⁵⁴ généralise la factorielle : on verra alors que $\Gamma(z = n \in \mathbb{N}^*) = (n - 1)!$. Toutefois, qu'une fonction $f(z)$ prenne pour des valeurs discrètes z_n des valeurs coïncidant avec une séquence f_n préalablement rencontrée est bien une sorte de généralisation, mais il ne s'agit nullement d'un prolongement au sens strict puisque l'identification se fait dans le sens continu \rightarrow discret : il serait bien aventureux de prétendre, sans autre information, que la fonction $f(z)$ égale à $(-1)^n$ pour $z = n\pi$ est la fonction $\cos z$!

Ce point étant précisé, il convient d'ajouter que ce type de généralisation est fréquent en Physique. Il arrive que l'on soit contraint d'introduire des fonctions $f(x)$, alors que le problème physique n'est concerné que par des valeurs données de la variable, entières positives par exemple : x peut être la dimension D de l'espace physique, soit usuellement $x = 1, 2, 3$. Le calcul direct de $f(n)$ peut se révéler difficile, voire impossible ; c'est ce qui conduit à considérer des espaces de dimension non-entières, voire négatives (!), et pourquoi pas complexes (!!), à calculer $f(x \in \mathbb{R})$, ou $f(z \in \mathbb{C})$, afin de trouver la quantité pertinente en faisant, en fin de calcul, le choix physique $x = n$. Un exemple célèbre est celui des phénomènes critiques où, pour traiter des problèmes en dimension d'espace $D = 3$, on part de la dimension $D = 4$ (où l'on connaît certains résultats) et on fait des développements en $\varepsilon = 4 - D$, espérant pouvoir, en fin de compte, poser $\varepsilon = 1$ pour retrouver le problème physique.

Un autre exemple : soit à déterminer une certaine quantité qui se présente sous la forme⁵⁵ $F = \ln Z$. On peut trouver F comme :

$$F = \lim_{\nu \rightarrow 0} \frac{Z^\nu - 1}{\nu} . \quad (6.151)$$

Cette relation, triviale telle qu'elle est, un peu absconse même, est à la base d'une méthode puissante utilisée en Mécanique statistique, introduite dans les années 1970 par Parisi (*Méthode des répliques*).

Exemple 5

Terminons par la description semi-quantitative d'un exemple apparaissant fréquemment en Mécanique quantique, notamment lorsqu'il s'agit de décrire un état instable (atome à l'état excité, particule éphémère, quasi-particule en phase condensée...). En anticipant sur la suite du cours, soit la transformée de Fourier du propagateur $U(t) \stackrel{\text{def}}{=} e^{\frac{1}{i\hbar}Ht}$, qui est essentiellement⁵⁶ $G(z) = \frac{1}{z-H}$ (on l'appelle aussi *résolvante*, *fonction de Green*,...).

⁵³ \mathbb{N} n'est pas un *domaine*.

⁵⁴voir chapitre 7.

⁵⁵On devine qu'il est question (à des facteurs près) d'énergie libre, et de fonction de partition.

⁵⁶On devrait plus précisément écrire $G(z) = \frac{1}{z\mathbf{1}-H}$ où $\mathbf{1}$ note l'opérateur identité.

La caractéristique d'un système instable est de posséder un spectre *continu* d'énergie⁵⁷, à partir d'une valeur E_{fond} (l'énergie de l'état fondamental), prise comme zéro d'énergie. Dans ces conditions, l'allure typique de l'élément de matrice de $G(z)$, $G_e(z)$, décrivant la dynamique de l'état instable est la suivante :

$$G_e(z) = \frac{1}{z - E_e - \Sigma(z)} , \quad (6.152)$$

où $E_e > 0$ est l'énergie du système isolé (donc stable dans une description élémentaire) et où la fonction $\Sigma(z)$, souvent appelée *self-energy*, est de la forme :

$$\Sigma(z) \stackrel{\text{def}}{=} \int_0^{+\infty} \frac{|V(E)|^2}{z - E} dE ; \quad (6.153)$$

$V(E)$ est un élément de matrice traduisant le couplage entre le système "nu" (stable) et le "champ" qui le rend instable, précisément⁵⁸. La forme même de $\Sigma(z)$ montre que $\Im z$ et $\Im \Sigma$ sont de signes contraires quel que soit $z \in \mathbb{C}$.

Si $\Re z < 0$, le dénominateur de l'intégrand ne peut pas s'annuler : $\Sigma(z)$ est donc certainement holomorphe dans le demi-plan de gauche, et prend d'ailleurs des valeurs réelles (négatives) si $z \in \mathbb{R}_-$ (à la traversée par z du demi-axe réel négatif, de haut en bas par exemple, $\Im \Sigma(z)$ s'annule continûment, passant des valeurs négatives à des valeurs positives).

Si la partie réelle $\Re z$ est positive, $\Sigma(z)$ est bien définie tant que $\Im z \neq 0$, puisque le dénominateur $z - E$ ne s'annule pas ; en revanche, la partie imaginaire de $\Sigma(z)$ *ne tend pas vers zéro*⁵⁹ quand z tend vers l'axe \mathbb{R}_+ (tout en étant toujours d'un signe opposé à celui de $\Im z$). Il en résulte $\Im \Sigma(x + i0) = -\Im \Sigma(x - i0)$, entraînant que $\Sigma(x + i0) \neq \Sigma(x - i0) \neq 0$ si $x \in \mathbb{R}_+$, révélant que $\Sigma(z)$ a une coupure sur le demi-axe réel $[0, +\infty[$ des énergies et est donc une fonction multiforme. Au total, la propriété des signes de z et $\Sigma(z)$ interdit au dénominateur de $G(z)$ de s'annuler dans le premier feuillet de Riemann. En définitive, $G(z)$ n'a aucun pôle tant que l'on reste dans le premier feuillet.

L'inversion de Fourier peut se faire par résidus et notamment en glissant dans le deuxième feuillet, où le prolongement analytique continu de $G(z)$ possède en général des pôles z_e , situés dans le demi-plan inférieur (heureusement !), et en général pas très loin de l'énergie "nue" E_e (le couplage atome - champ est piloté par la constante de structure fine et ses différentes puissances, tous des nombres petits devant 1). En posant $z_e = \tilde{E}_e - i\hbar\Gamma_e$, on voit que le résidu correspondant donne dans $U(t)$ un terme du genre

⁵⁷Très schématiquement : quand le spectre d'énergie est discret, toutes les expressions donnant l'évolution temporelle des valeurs moyennes sont *de facto* des sommes discrètes (séries de Fourier ou leur généralisation), impliquant des fréquences de Bohr isolées les unes des autres ; suivant la distribution précise de ces fréquences (commensurables ou non), la variation en temps est périodique (oscillateur harmonique) ou quasi-périodique.

Au contraire, si les fréquences sont distribuées *continûment*, les sommes discrètes du genre Fourier sont remplacées par des *intégrales*, elles aussi du genre Fourier, donnant une évolution temporelle *irréversible*.

⁵⁸On suppose que $|V(E)|^2$ est une "bonne" fonction qui, notamment, décroît à l'infini plus vite que $E^{-\alpha}$, $\alpha > 0$.

⁵⁹Avec $E_0 \equiv \Re z > 0$, on a $G(z = E_0 \pm i0) = \lim_{\varepsilon \rightarrow 0} \int_0^{+\infty} \frac{|V(E)|^2}{E_0 \pm i\varepsilon - E} dE = \mathcal{P} \int_0^{+\infty} \frac{|V(E)|^2}{E_0 - E} dE \mp i\pi |V(E_0)|^2$.

$e^{\frac{1}{i\hbar}\tilde{E}_e t} e^{-\Gamma_e t}$. \tilde{E}_e s'interprète comme l'énergie du système nu modifiée par couplage avec le champ (on dit "renormalisée" et on parle de *déplacement de Lamb (Lamb shift)*) ; quant à Γ_e^{-1} , c'est essentiellement la durée de vie de l'état excité, si on s'en tient à la phase initiale du déclin, où l'exponentielle domine les autres contributions à $U(t)$. Ces dernières finissent tôt ou tard par l'emporter sur l'exponentielle : aux grands temps, le déclin est régi par des lois en puissance $\propto t^{-\mu}$; d'une façon générale, l'exposant μ est lié à la masse de la particule émise – pour l'atome, c'est un photon (masse nulle).

6.4 Calcul d'intégrales impropres

D'une façon générale (et quelque peu vague), on appelle intégrale impropre une intégrale dont la définition exige quelques précautions ou quelques précisions. C'est le cas lorsque l'une ou les borne(s) est (sont) infinie(s). Un autre cas, très fréquent en Physique, est celui où le chemin d'intégration passe par une singularité de l'intégrand. Par exemple, on rencontre souvent des intégrales du type :

$$\int_a^b \frac{f(x)}{x - x_0} dx , \quad (6.154)$$

où x_0 est un réel compris entre a et b . Clairement, si $f(x_0) \neq 0$, l'expression (6.154) n'a pas de sens (elle diverge logarithmiquement) et il faut lui en donner un ; c'est typiquement ce que l'on appelle une *régularisation*.

Les deux types de difficultés peuvent d'ailleurs se présenter simultanément, quand d'une part une(les) borne(s) est(sont) infinie(s), et que d'autre part le chemin d'intégration rencontre une singularité. Afin de raisonner dans un cas concret (et important en Physique) et ne pas accumuler les difficultés, soit l'intégrale :

$$I(x_0) = \int_{-\infty}^{+\infty} \frac{f(x)}{x - x_0} dx \quad (f(x_0) \neq 0) , \quad (6.155)$$

où les bornes sont d'emblée prises infinies mais où la fonction f est supposée posséder toutes les propriétés pour que la convergence à l'infini de l'intégrale soit assurée sans devoir faire d'acrobaties (par exemple f est supposée décroître assez vite pour que chacun des morceaux à l'infini tende vers zéro indépendamment de l'autre) ; on suppose en outre que $f(z)$ n'a aucune singularité sur l'axe réel. On peut d'ailleurs sans perte de généralité supposer $x_0 = 0$; considérons donc :

$$I \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} \frac{f(x)}{x} dx . \quad (6.156)$$

Pour la simplicité du raisonnement, la fonction f est en outre supposée n'avoir aucune singularité sur \mathbb{R} (en particulier elle est supposée continue en $x = 0$ où elle vaut $f(0) \neq 0$). Dans ces conditions, le seul ennui est la présence de la singularité de l'intégrand en $x = 0$. Si on exclut le petit segment $[-\varepsilon, +\varepsilon]$ et que l'on estime le comportement de l'intégrale

amputée \tilde{I} , on voit tout de suite que $\tilde{I} \sim \ln(\varepsilon/\varepsilon')$, de sorte que si on fait tendre vers zéro indépendamment les deux quantités positives ε et ε' , l'intégrale a une divergence de type logarithmique, et peut prendre n'importe quelle valeur réelle selon la façon dont sont prises les deux limites. En revanche, si on prend $\varepsilon = \varepsilon'$, on ampute l'intégrale de deux segments symétriques⁶⁰ de part et d'autre de l'origine, les aires algébriques se compensent et la limite de l'intégrale tronquée existe, aussi petit que soit ε . À la limite $\varepsilon = 0_+$ ceci est très exactement la définition d'une régularisation particulière appelée *partie principale de Cauchy*, et ici notée \mathcal{P} ou \mathcal{f} , la petite barre évoquant que l'on a "enlevé un point" – et pas n'importe comment :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{f(x)}{x} dx \stackrel{\text{d'ef}}{=} \lim_{\varepsilon \rightarrow 0_+} \left[\int_{-\infty}^{-\varepsilon} \frac{f(x)}{x} dx + \int_{+\varepsilon}^{+\infty} \frac{f(x)}{x} dx \right] \equiv \mathcal{f}_{-\infty}^{+\infty} \frac{f(x)}{x} dx . \quad (6.157)$$

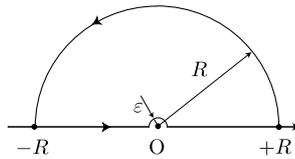


Figure 6.6: Contour utilisé pour illustrer la définition (6.157) de la partie principale.

Comment se traduit une telle opération en terme d'une intégrale dans le plan complexe ? Pour fixer les idées, supposons que :

$$\lim_{|z| \rightarrow \infty} |f(z)| = 0 \quad \forall \text{Arg} z \in [0, 2\pi] ; \quad (6.158)$$

auquel cas le premier lemme de Jordan est applicable puisque $|z \frac{f(z)}{z}| \rightarrow 0$ si $|z| \rightarrow +\infty$. Cela étant, on peut calculer par résidus en refermant le contour par un grand demi-cercle, soit par en-haut soit pas en-bas, demi-cercle dont la contribution est nulle à la limite. Par exemple, considérons l'intégrale sur le contour dessiné sur la figure 6.6 ; elle s'écrit, désignant par ... la contribution du grand demi-cercle supposée nulle à la limite R infini :

$$\left(\int_{-R}^{-\varepsilon} + \int_{\varepsilon}^{+R} \right) \frac{f(x)}{x} dx + \int_{\pi}^0 d(\varepsilon e^{i\theta}) \frac{f(\varepsilon e^{i\theta})}{\varepsilon e^{i\theta}} + \dots . \quad (6.159)$$

et vaut $2i\pi \sum_k \text{Res} \left[\frac{f(z)}{z}, z_k \right]$ où les z_k désignent les pôles de $f(z)$ ayant une partie imaginaire strictement positive⁶¹. Après simplification dans l'intégrale sur le petit demi-cercle autour de l'origine, le théorème des résidus s'écrit :

$$\left(\int_{-R}^{-\varepsilon} + \int_{\varepsilon}^{+R} \right) \frac{f(x)}{x} dx + i \int_{\pi}^0 d\theta f(\varepsilon e^{i\theta}) + \dots = 2i\pi \sum_{k, \Im z_k > 0} \text{Res} \left[\frac{1}{z} f(z), z_k \right] . \quad (6.160)$$

⁶⁰À condition que ε soit suffisamment petit, seule importe la valeur de f en $x = 0$, et les deux aires sont bien égales en valeur absolue, quelle que soit l'éventuelle asymétrie de $f(x)$ dans un plus grand voisinage de l'origine.

⁶¹Rappelons que, par hypothèse, $f(z)$ n'a pas de singularité sur l'axe réel.

Comme $f(x)$ est continue en $x = 0$, dans la limite $\varepsilon \rightarrow 0$, $R \rightarrow +\infty$, la définition (6.157) permet d'écrire :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{f(x)}{x} dx + i f(0) \int_{\pi}^0 d\theta = 2i\pi \sum_k \operatorname{Res} \left[\frac{1}{z} f(z), z_k \right] \quad (\Im z_k > 0) . \quad (6.161)$$

soit :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{f(x)}{x} dx + i(-\pi)f(0) = 2i\pi \sum_k \operatorname{Res} \left[\frac{1}{z} f(z), z_k \right] \quad (\Im z_k > 0) . \quad (6.162)$$

Au second membre n'apparaît pas le résidu en $z = 0$ de $\frac{1}{z}f(z)$ puisque le pôle correspondant est à l'extérieur du contour. (6.162) s'écrit en définitive :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{f(x)}{x} dx = i\pi f(0) + 2i\pi \sum_k \operatorname{Res} \left[\frac{1}{z} f(z), z_k \right] \quad (\Im z_k > 0) . \quad (6.163)$$

Il n'y a pas de doute que le premier terme au second membre, $i\pi f(0)$, se lit comme $2i\pi \frac{1}{2} \operatorname{Res} \left[\frac{1}{z} f(z), z = 0 \right]$. En définitive, (6.163) constitue une sorte de généralisation du théorème des résidus dans le cas où l'un des pôles est *sur* le contour : une fois régularisée à la Cauchy, l'intégrale proposée se calcule comme d'habitude, sauf qu'il faut prendre seulement la *moitié* des résidus des pôles rencontrés sur le contour.

Bien sûr, on peut refaire le calcul en prenant les deux demi-cercles (le grand et le petit) dans le demi-plan inférieur ; on trouve alors :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{f(x)}{x} dx + i(+\pi)f(0) = -2i\pi \sum_k \operatorname{Res} \left[\frac{1}{z} f(z), z_k \right] \quad (\Im z_k < 0) ; \quad (6.164)$$

à nouveau, le résidu en $z = 0$ est absent de la sommation au second membre. L'équation (6.164) se lit encore :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{f(x)}{x} dx = -i\pi f(0) - 2i\pi \sum_k \operatorname{Res} \left[\frac{1}{z} f(z), z_k \right] \quad (\Im z_k < 0) \quad (6.165)$$

et peut encore être interprétée comme une généralisation du théorème des résidus :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{f(x)}{x} dx = -i\pi f(0) - 2i\pi \sum_k \operatorname{Res} \left[\frac{1}{z} f(z), z_k \right] \quad (\Im z_k < 0) \quad (6.166)$$

le premier terme n'étant autre que $-2i\pi \frac{1}{2} \operatorname{Res} \left[\frac{1}{z} f(z), z = 0 \right]$.

On peut enfin prendre le grand demi-cercle en haut et le petit demi-cercle en bas ; on a alors :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{f(x)}{x} dx + i(+\pi)f(0) = 2i\pi \sum_k \operatorname{Res} \left[\frac{1}{z} f(z), z_k \right] \quad (\Im z_k > 0 \text{ et } z_k = 0) ; \quad (6.167)$$

le résidu supplémentaire donne la contribution $2i\pi f(0)$, mais se recombine avec le terme $+i(+\pi)f(0)$ du premier membre : au total, les deux égalités (6.162) et (6.167) sont identiques.

En rapprochant les deux expressions (6.163) et (6.166) de $\mathcal{P} \int_{-\infty}^{+\infty} \frac{1}{x} f(x) dx$, on obtient :

$$f(0) + \sum_{k, \Im z_k \neq 0} \operatorname{Res} \left[\frac{1}{z} f(z), z_k \right] = 0, \quad (6.168)$$

où la somme porte sur *tous* les résidus, que les pôles soient en haut ou en bas ; le terme isolé $f(0)$ est le résidu au pôle en $z = 0$ de $\frac{1}{z} f(z)$. Le résidu à l'infini étant ici nul (en vertu de (6.158) et du Lemme 1 de Jordan), on retrouve le résultat général, affirmant que la somme de *tous* les résidus est nulle. C'est aussi pour cette raison que le résultat ne dépend pas du fait que l'on ferme le contour par en-haut ou par en-bas (comme on l'a vu, ceci n'est pas vrai pour une intégrale de Fourier, pour laquelle on n'a pas⁶² la propriété (6.158)).

En translatant tout de x_0 , on déduit de (6.162) la relation plus générale⁶³ :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{f(x)}{x - x_0} dx = i\pi f(x_0) + 2i\pi \sum_k \operatorname{Res} \left[\frac{1}{z - z_0} f(z), z_k \right] \quad (\Im z_k > 0), \quad (6.169)$$

ou son équivalente :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{f(x)}{x - x_0} dx = -i\pi f(x_0) - 2i\pi \sum_k \operatorname{Res} \left[\frac{1}{z} f(z), z_k \right] \quad (\Im z_k < 0). \quad (6.170)$$

Les deux relations (6.169) et (6.170) constituent donc une réécriture de la régularisation de Cauchy à l'aide du théorème des résidus, avec les deux hypothèses : $f(z)$ est analytique dans un domaine incluant l'axe réel et $|f(z)| \rightarrow 0$ si $|z| \rightarrow +\infty$ quel que soit l'argument de z .

Une fois précisé le lien entre la partie principale au sens de Cauchy et le théorème des résidus, il est possible d'établir des relations remarquables, dites de Kramers - Kronig, dont l'universalité en Physique tient au fait qu'elles reposent fondamentalement sur un principe intangible, le Principe de causalité.

Supposons maintenant d'une part que $|f(z)| \rightarrow 0$ si $|z| \rightarrow +\infty$ dans le demi-secteur $0 \leq \operatorname{Arg} z \leq \pi$, d'autre part que $f(z)$ est analytique dans tout le plan supérieur $\Im z \geq 0$ - on verra que cette propriété est une conséquence du Principe de causalité⁶⁴.

⁶²La présence d'un terme du genre e^{ikz} entraîne que le module de l'intégrale diverge exponentiellement dans un demi-secteur circulaire fixé par le signe de la partie réelle de k .

⁶³Bien sûr, c'est maintenant le point x_0 que l'on "retire" de l'intégrale, en extrayant le segment $[x_0 - \varepsilon, x_0 + \varepsilon]$.

Noter aussi que, pour la partie principale, c'est la moitié du résidu qui compte ($i\pi f(0) = \frac{1}{2}(2i\pi f(0))$), comme si rencontrer la singularité sur le contour c'était laisser un demi-pôle à gauche, un demi-pôle à droite !

⁶⁴moyennant une convention précise pour la transformation de Fourier. Avec la convention opposée, l'analyticité est obtenue dans le demi-plan *inférieur*.

On peut refaire le même calcul que ci-dessus en prenant le demi-cercle dans le demi-plan supérieur ; comme, par hypothèse, $f(z)$ n'a pas de pôles $\Im z_k > 0$, (6.169) se simplifie en⁶⁵ :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{f(x)}{x - x_0} dx = i\pi f(x_0) . \quad (6.171)$$

Cette relation, vraie pour toute fonction $f(z)$ analytique dans le demi-plan supérieur, et telle que $\lim_{|z| \rightarrow +\infty} |f(z)| = 0$ avec $0 \leq \text{Arg } z \leq \pi$, établit une dépendance très importante et très forte entre la partie principale et les propriétés analytiques de $f(z)$ (dans le cas discuté ici, l'absence de singularités dans le demi-plan supérieur). Comme la plupart des relations déduites du théorème des résidus (c'est-à-dire plus fondamentalement du théorème de Cauchy), elle montre une fois de plus que les valeurs d'une fonction analytique sur une ligne (propriétés *étendues* de la fonction) sont intimement liées à des valeurs de cette fonction en un (ou plusieurs) point(s) (propriétés *locales*).

À la réflexion, (6.171) est presque évidente, et peut s'exprimer dans le langage des mots : compte tenu des hypothèses sur $f(z)$, le produit $\phi(z) \stackrel{\text{déf}}{=} \frac{f(z)}{z - x_0}$ est analytique dans $\mathbb{C}_{x_0}^{(+)} \stackrel{\text{déf}}{=} \{z : \Im z \geq 0, |z - x_0| > 0\}$. Le théorème de Cauchy dit que l'intégrale de $\phi(z)$ sur toute boucle située dans $\mathbb{C}_{x_0}^{(+)}$ est nulle. Par déformation continue, une telle boucle peut être superposée au contour de la fig.6.6, l'intégrale étant toujours nulle, ce qui se traduit par $\mathcal{P} \int_{-\infty}^{+\infty} \frac{f(x)}{x - x_0} dx - i\pi f(x_0) = 0$.

Posant $f(z) = f_1(z) + if_2(z)$, et séparant les parties réelle et imaginaire dans (6.171), on obtient :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{f_1(x)}{x - x_0} dx = -\pi f_2(x_0) , \quad \mathcal{P} \int_{-\infty}^{+\infty} \frac{f_2(x)}{x - x_0} dx = \pi f_1(x_0) . \quad (6.172)$$

En Physique, de telles relations⁶⁶ portent le nom de *Relations de Kramers - Kronig*. Elles jouent un rôle essentiel pour analyser la réponse d'un système, décrite par sa *susceptibilité* $\chi(\omega)$, et permettent ainsi de remonter à la dynamique interne de ce système. Comme on le verra par la suite (voir chapitre 10), l'absence de singularités dans un demi-plan pour la susceptibilité est effectivement la traduction du Principe de causalité.

? Combien vaut $\mathcal{P} \int_{-\infty}^{+\infty} \frac{e^{i\frac{\pi x}{2}}}{x^2 - 1} dx$?

⁶⁵Le résultat (6.171) peut à nouveau se lire comme une généralisation du théorème des résidus où, comme le pôle est *sur* le contour, il faut prendre seulement – quand on régularise au sens de Cauchy – la *moitié* du résidu ($i\pi$ au lieu de $2i\pi$ au second membre).

Par ailleurs, cette relation peut sembler étrange à première vue si l'on imagine que $f(z)$ prend des valeurs réelles sur tout l'axe réel : si c'était le cas, le premier membre de (6.171) serait réel, alors que le second serait imaginaire pur. En fait, il y a incompatibilité entre les hypothèses faites sur f pour arriver à (6.171), et imaginer que, en plus, f prend des valeurs réelles sur tout l'axe réel : (6.171) dit aussi qu'une fonction analytique ayant les propriétés de f ne *peut* pas être à valeurs réelles sur tout l'axe réel.

⁶⁶Le plus souvent, elles apparaissent pour la susceptibilité χ en fonction de la pulsation ω (transformée de Fourier de la réponse en temps du système) ; en transcrivant les notations, on a :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{\chi(\omega)}{\omega - \omega_0} d\omega = i\pi\chi(\omega_0) \quad (\chi(\omega) \text{ analytique dans le demi-plan supérieur}) . \quad (6.173)$$

6.5 Calcul de la somme de séries

On applique parfois le théorème des résidus “à l’envers”, au sens où voulant calculer une certaine expression, on arrive à identifier celle-ci avec le second membre du théorème des résidus appliqué à une certaine fonction dont on sait calculer *directement* l’intégrale. Un exemple d’application de cette méthode est le calcul de certaines sommes, finies ou infinies, puisque le théorème implique une *somme* (de résidus).

S’agissant d’abord d’une somme d’un nombre infini de termes (série), la toute première chose est d’examiner la convergence. Une fois celle-ci établie, il est évidemment très utile de savoir calculer effectivement la somme $S \stackrel{\text{d'éf}}{=} \lim_{N \rightarrow \infty} \sum_{n=1}^N u_n$; le terme général peut dépendre d’un paramètre (ou plus généralement être une certaine fonction $u_n(x)$) ; ici encore, le théorème des résidus est d’une grande utilité.

La méthode consiste à trouver une fonction ayant une infinité de singularités paramétrables par l’entier n et à dénicher le bon contour de façon à faire apparaître la série S au second membre du théorème des résidus quand on l’applique à l’intégrale ainsi construite. Autrement dit, il faut remplir l’équation suivante :

$$\int_{\text{contour}} f(z) dz = 2i\pi (S + \text{autre chose que l'on sait calculer}) . \quad (6.174)$$

Pour l’instant, on n’est guère avancé : il faut évidemment savoir calculer l’intégrale au premier membre (quand on aura dit ce qu’elle est), l’idéal étant qu’elle soit nulle...

Tout naturellement, on explore cette voie en commençant par introduire une fonction $f(z)$ qui est simplement l’expression de u_n en y remplaçant l’entier n par la variable complexe z :

$$f(z = n) \stackrel{\text{d'éf}}{=} u_n . \quad (6.175)$$

Il faut ensuite trouver une fonction $\phi(z)$ ayant des singularités (des pôles simples, par exemple) réparties comme \mathbb{N} (ou \mathbb{Z} , en jouant éventuellement avec la parité de $f(n)$), afin de construire à partir de résidus la somme S au second membre de (6.174). Une fonction simple ayant une infinité de pôles simples indexables par $n \in \mathbb{Z}$ est la fonction $\frac{1}{\sin z}$ (tous les pôles sont alors en $z_n = n\pi$ et sont d’ordre un). En prenant plutôt $\frac{1}{\sin \pi z}$, les pôles sont strictement les entiers relatifs. Le résidu correspondant est :

$$\left(\frac{1}{(\sin \pi z)'} \right)_{z=n\pi} = \frac{1}{\pi \cos n\pi} = \frac{(-1)^n}{\pi} , \quad (6.176)$$

ou, ce qui revient au même :

$$\lim_{z \rightarrow n\pi} (z - n\pi) \frac{1}{\sin \pi z} = \lim_{z \rightarrow n\pi} (z - n\pi) \frac{1}{\pi(z - n\pi) \cos n\pi} = \frac{(-1)^n}{\pi} . \quad (6.177)$$

Le signe alternant $(-1)^n$ n’est pas *a priori* souhaitable pour une série du genre⁶⁷ $\sum_n u_n$, et le facteur $\frac{1}{\pi}$ cosmétiquement désagréable ; comme $\cos n\pi = (-1)^n$, on en vient ainsi à

⁶⁷mais est au contraire ce qu’il faut pour une série alternée $\sum_n (-1)^n u_n$ (voir plus loin p. 273).

considérer la fonction $\pi \cot \pi z$, qui a les mêmes pôles que $\frac{1}{\sin \pi z}$ et avec un résidu égal à $+1$ quel que soit n :

$$\lim_{z \rightarrow n\pi} (z - n\pi) \frac{\pi \cos \pi z}{\sin \pi z} = \lim_{z \rightarrow n\pi} (z - n\pi) \frac{\pi \cos \pi z}{\pi(z - n\pi) \cos n\pi} = \frac{\pi(-1)^n}{\pi(-1)^n} = 1 . \quad (6.178)$$

En d'autres termes, le $\cos n\pi$ donne lui aussi une alternance de signe, qui vient compenser celle du résidu de $\frac{1}{\sin \pi z}$. Le résultat important ici s'exprime donc comme suit :

$$\boxed{\text{Res}(\pi \cot \pi z, n) = \frac{1}{2i\pi} \int_{\gamma_n} \pi \cot \pi z \, dz = 1} \quad (6.179)$$

où γ_n est un contour fermé (un petit cercle par exemple) entourant le point d'abscisse entière $n \in \mathbb{Z}$ sur l'axe réel, à l'exclusion de tout autre entier.

Soit maintenant à calculer une série du genre :

$$S = \sum_{n=0}^{+\infty} f(n) , \quad (6.180)$$

où f est une fonction paire, $f(-n) = f(n)$, n'ayant aucune singularité⁶⁸ sur \mathbb{Z} , et supposée méromorphe pour simplifier⁶⁹. On a :

$$S = \frac{1}{2}f(0) + \frac{1}{2} \sum_{n=-\infty}^{+\infty} f(n) \iff \sum_{n=-\infty}^{+\infty} f(n) = 2S - f(0) , \quad (6.181)$$

Considérons maintenant des grands contours C_R centrés sur l'origine (carré de côté R , cercle de rayon R, \dots) et soit l'intégrale :

$$I = \int_{C_R} \pi \cot \pi z f(z) \, dz , \quad (6.182)$$

où $f(z)$ est donc le "prolongement analytique" immédiat de la fonction servant de terme général de la série et initialement définie sur les entiers⁷⁰. Plus R est grand, plus grand est le nombre de valeurs entières de n situées dans le contour. Par ailleurs, supposons que $\lim_{|z| \rightarrow +\infty} |zf(z)| = 0 \forall \text{Arg} z \in [0, 2\pi]$: alors l'intégrale sur le grand contour est nulle

⁶⁸Dans le cas contraire, le terme général de la série ne serait pas défini partout sur \mathbb{Z} .

⁶⁹Si f a des coupures, il convient de faire un calcul spécifique, dont la recette n'est pas aussi simple mais qui relève des mêmes idées.

⁷⁰Pour ne pas compliquer la discussion de façon inessentielle, on suppose que $f(z)$ est analytique aux pôles de $\cot \pi z$. Dans le cas contraire, le calcul se fait de la même façon, mais en prenant en compte une telle particularité. À titre d'exercice, on cherchera la somme de la série $\sum_{n=1}^{+\infty} \frac{1}{n^2(n^2+a^2)}$ et on montrera qu'elle vaut $\frac{\pi^2}{6a^2} - \frac{\pi}{2a^3} \cot \pi a + \frac{1}{2a^4}$ (cette somme se trouve aussi à l'aide de (6.191), en écrivant $\frac{1}{n^2(n^2+a^2)} = \frac{1}{a^2} \left(\frac{1}{n^2} - \frac{1}{n^2+a^2} \right)$.

à la limite⁷¹ (encore un avatar du lemme de Jordan). Ainsi, après passage à la limite R infini, (6.182) devient :

$$0 = \int_{C_\infty} \pi \cot \pi z f(z) dz ; \quad (6.184)$$

mais par ailleurs, l'application du théorème des résidus donne⁷² :

$$\int_{C_\infty} \pi \cot \pi z f(z) dz = 2i\pi \sum_{\alpha} \text{Res}[\pi \cot \pi z f(z), z_\alpha] , \quad (6.185)$$

où la somme court sur tous les résidus du produit $\pi \cot \pi z f(z)$. D'après (6.184), il vient⁷³ :

$$0 = \sum_{\alpha} \text{Res}[\pi \cot \pi z f(z), z_\alpha] . \quad (6.186)$$

Par hypothèse, la fonction $f(z)$ a des pôles z_k qui, après la limite $R \rightarrow +\infty$, se trouvent forcément tous à l'intérieur du grand contour. En séparant les contributions venant des résidus des pôles de $\pi \cot \pi z$ et celles liées aux pôles z_k de f (toujours dans l'hypothèse où il n'y a pas de coïncidences, $z_k \neq n\pi \forall k, n$), (6.186) donne :

$$0 = \sum_{n \in \mathbb{Z}} f(n) \times 1 + \sum_{z_k \text{ pôle de } f} \pi \cot \pi z_k \text{Res}[f(z), z_k] , \quad (6.187)$$

où le 1 en facteur de $f(n)$ est le résidu propre à $\pi \cot \pi z$. Alors, selon (6.181) :

$$0 = 2i\pi [2S - f(0)] + 2i\pi \sum_{z_k \text{ pôle de } f} \pi \cot \pi z_k \text{Res}[f(z), z_k] , \quad (6.188)$$

et finalement la somme S cherchée :

$$S = \frac{1}{2}f(0) - \frac{1}{2} \sum_{z_k \text{ pôle de } f} \pi \cot \pi z_k \text{Res}[f(z), z_k] . \quad (6.189)$$

Le calcul de S se réduit donc au calcul des résidus de f ; en pratique, la méthode est efficace quand f a un *petit* nombre de pôles⁷⁴.

⁷¹En pratique, c'est essentiellement dans ce cas que cette méthode astucieuse est utile !

Soit un grand carré dont les côtés, parallèles à Ox et Oy , coupent ces axes aux points $\pm(N + \frac{1}{2})$ et $\pm i(N + \frac{1}{2})$ ($N > 0$). Sur le côté vertical de droite, $z = N + \frac{1}{2} + iy$, d'où $\cot \pi z = \cot(\frac{\pi}{2} + i\pi y) = -i \tanh \pi y$, et donc $|\cot \pi z| \leq 1$. Pour le côté horizontal supérieur, on utilise :

$$|\cot \pi(x + iy)| = \left| \frac{e^{i\pi(x+iy)} + e^{-i\pi(x+iy)}}{e^{i\pi(x+iy)} - e^{-i\pi(x+iy)}} \right| \leq \frac{e^{-\pi y} + e^{\pi y}}{|e^{i\pi(x+iy)} - e^{-i\pi(x+iy)}|} \leq \frac{e^{-\pi y} + e^{\pi y}}{|e^{-\pi y} - e^{\pi y}|} , \quad (6.183)$$

la dernière égalité venant de $|a - b| \geq ||a| - |b||$, d'où finalement $|\cot \pi(x + iy)| \leq \coth \pi y$. Sur le côté horizontal supérieur, $z = x + i(N + \frac{1}{2})$, d'où $|\cot \pi z| \leq \coth(N + \frac{1}{2})\pi$; N_0 étant un entier quelconque, et puisque $\coth X$, $X \in \mathbb{R}$, est une fonction décroissante, $|\cot \pi z| \leq \coth(N + \frac{1}{2})\pi \leq \coth(N_0 + \frac{1}{2})\pi \forall N \geq N_0$. Il en va de même pour les deux autres côtés du carré : au total, la fonction $\cot \pi z$ est bornée en module sur le carré de côté $2N + 1$; il suffit donc bien d'avoir $\lim_{|z| \rightarrow \infty} |zf(z)| = 0$ pour invoquer le lemme de Jordan.

⁷²Maintenant le contour infini contient *tous* les pôles de $\pi \cot \pi z$.

⁷³On retrouve le fait que, y compris le résidu à l'infini, la somme de tous les résidus d'une fonction méromorphe est nulle.

⁷⁴... mais peut aussi servir pour démontrer l'égalité de deux séries, quand f a une infinité de pôles.

D'ailleurs, le recours à un grand contour étiré à l'infini n'est pas obligatoire, le calcul de la série pouvant s'effectuer différemment. En effet, la relation évidente (voir (6.179)) :

$$\sum_{n=-\infty}^{+\infty} f(n) = \sum_{n=-\infty}^{+\infty} f(n) \times 1 = \sum_{n \in \mathbb{Z}} \text{Res}[f(z) \pi \cot \pi z, z = n] \quad (6.190)$$

peut se lire autrement, en disant que, par définition du résidu, la série de gauche est égale à la somme des intégrales sur des petits contours fermés (pourquoi pas circulaires) renfermant chacun un et un seul pôle de $\pi \cot \pi z$. Par des déformations continues successives (voir fig. 6.7), le contour devient deux droites situées l'une juste en-dessous, l'autre juste au dessus de l'axe réel. Chacune des intégrales correspondantes se calcule en refermant par un grand demi-cercle situé du même côté qu'elle de l'axe réel et en invoquant à nouveau le 1^{er} lemme de Jordan⁷⁵.

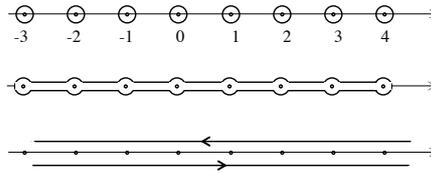


Figure 6.7: Modifications par continuité du contour pour le calcul d'une série du genre (6.180).

D'une façon ou d'une autre, c'est ainsi que l'on peut établir les égalités suivantes :

$$\sum_{n \in \mathbb{Z}} \frac{1}{n^2 + a^2} = \frac{\pi}{a} \coth \pi a \iff \sum_{n=1}^{+\infty} \frac{1}{n^2 + a^2} = \frac{\pi}{2a} \coth \pi a - \frac{1}{2a^2}, \quad (6.191)$$

et bien d'autres du même genre (ces expressions sont valides tant que $a \neq i \times \text{entier}$). On en déduit, par exemple⁷⁶ :

$$\zeta(2) \stackrel{\text{d\'ef}}{=} \sum_{n \in \mathbb{N}^*} \frac{1}{n^2} = \frac{\pi^2}{6}, \quad \zeta(4) \stackrel{\text{d\'ef}}{=} \sum_{n \in \mathbb{N}^*} \frac{1}{n^4} = \frac{\pi^4}{90}. \quad (6.192)$$

La fonction $\zeta(z)$:

$$\zeta(z) \stackrel{\text{d\'ef}}{=} \sum_{n \in \mathbb{N}^*} n^{-z} \quad (\Re z > 1) \quad (6.193)$$

⁷⁵La fonction $\pi \cot \pi z$ est aussi bornée sur l'arc de cercle rejoignant les deux extrémités de chaque droite. En effet, limitant pour l'instant la droite au segment $[-R + i\alpha, R + i\alpha]$ ($\alpha > 0$), sur l'arc de cercle on a $|\cot \pi z| \leq \cot(\pi R \sin \theta_R)$ où $\theta_R > 0$ est le plus petit argument de z sur l'arc de cercle. Dans la limite $R \rightarrow +\infty$, $R \sin \theta_R \rightarrow \alpha$ et donc $|\cot \pi z| \leq \cot \pi \alpha < +\infty$. Noter que α est tout nombre strictement positif, en outre borné supérieurement par la plus petite partie imaginaire des singularités de f .

⁷⁶ces résultats se trouvent aussi par la considération des bonnes séries de Fourier. Par exemple, $\zeta(2)$ s'obtient en sommant les coefficients de Fourier de la fonction $f(x) = x^2$, $-\pi \leq x \leq +\pi$ et périodisée sur tout \mathbb{R} , après avoir choisi $x = \pi$.

s'appelle *fonction de Riemann*, et joue un rôle fondamental en Mathématiques⁷⁷ (notamment en Théorie des nombres, plus particulièrement celle des nombres premiers). On rencontre souvent les sommes $\zeta(z)$ en Physique : $\zeta(4)$ apparaît dans le calcul selon Planck de la constante phénoménologique de Wien, $\zeta(D)$ dans la condensation de Bose en D dimensions, dans le modèle d'Ising à D quelconque, traité à la manière de Bethe-Peierls... Notons que l'égalité (6.191) se lit aussi :

$$\cot z = \frac{1}{z} + \sum_{n \in \mathbb{N}^*} \left(\frac{1}{z - n\pi} + \frac{1}{z + n\pi} \right) = \frac{1}{z} + 2z \sum_{n \in \mathbb{N}^*} \frac{1}{z^2 - n^2\pi^2} . \quad (6.194)$$

et constitue le *développement de Mittag-Leffler*⁷⁸ de la fonction $\cot z$, un exemple de développement d'une fonction méromorphe sur ses pôles ((6.106) est un autre exemple de développement de ce type) – pour une généralisation de cette écriture, voir sous-section 6.7.

? De (6.194), déduire le développement de $\sin z$ en produit infini :

$$\sin z = z \prod_{n \in \mathbb{N}^*} \left(1 - \frac{z^2}{n^2\pi^2} \right) , \quad (6.195)$$

que l'on retrouvera plus tard (voir (6.260)). Par ailleurs, comment utiliser le résultat (6.195) pour trouver $\zeta(2)$? Enfin, en examinant le résidu de $\frac{z}{\sinh \pi z}$ en $z = n_0$, montrer que :

$$\prod_{n \in \mathbb{N}^*, n \neq n_0} \frac{n^2}{n^2 - n_0^2} = (-1)^{n_0+1} 2 . \quad (6.196)$$

Ce type de résultat sera généralisé dans la section 6.8.2

À titre de dernier exemple, citons l'une des relations extraordinaires dues à Ramanujan⁷⁹ :

$$S_3 \stackrel{\text{déf}}{=} \sum_{n=1}^{+\infty} \frac{1}{n^7} \coth n\pi = \frac{19\pi^7}{56700} = 1.012\,091 \dots . \quad (6.199)$$

⁷⁷À propos de cette fonction, il existe une fameuse conjecture, due précisément à Riemann, selon laquelle tous les zéros z_k de $\zeta(z)$ compris dans la bande $0 \leq z \leq 1$ sont en fait sur la demi-droite verticale d'abscisse $\frac{1}{2}$ (par ailleurs, on peut montrer que la fonction ζ est nulle pour toutes les valeurs paires et négatives : $\zeta(-2n) = 0 \forall n \in \mathbb{N}^*$, dites *zéros triviaux*). À l'heure actuelle, cette conjecture n'est pas démontrée, mais n'a toujours pas été prise en défaut malgré d'énormes calculs rendus possibles par des ordinateurs de plus en plus puissants... Voir par exemple l'ouvrage de Du Sautoy [34]. On reviendra sur $\zeta(z)$ dans le chapitre 7, section 7.3.

⁷⁸Gösta MITTAG-LEFFLER (1846-1927), grand mathématicien suédois et *mentor* de Sonya KOWA-LEVSKAYA, a aussi laissé son nom dans la petite histoire des Mathématiques. Une légende persistante affirme que, brillant, mondain et plutôt bel homme, il aurait eu une "affaire" avec l'une des dames de cœur d'Alfred NOBEL, Sophie HESS (qui n'était pas mal non plus). Certains vont même jusqu'à affirmer que c'est la raison pour laquelle il n'y a pas de Prix Nobel de Mathématiques, mais la relation de cause à effet est plus que douteuse.

⁷⁹Mathématicien indien de génie, d'abord autodidacte... qui ne connaissait pas le théorème des résidus et qui pourtant... Sur le site <http://fr.wikipedia.org/wiki>, on peut lire :

Nous lui devons la bagatelle de six mille théorèmes en théorie des nombres. Ramanujan était un homme hors du commun ; il avait beaucoup d'imagination et jouissait d'une intuition prodigieuse. Il raisonnait très vite en se fondant sur des résultats qu'il considérait comme évidents, et donnait souvent ses théorèmes sans preuve mathématique. Ces

On peut calculer S_3 en intégrant sur le grand carré à l'infini la fonction $\phi(z)$ définie comme :

$$\phi(z) \stackrel{\text{déf}}{=} \frac{1}{z^7} \pi \cot \pi z \coth \pi z .$$

Le lemme de Jordan dit que l'intégrale est nulle ; outre les pôles de $\pi \cot \pi z$ en $z = n \in \mathbb{Z}$, il y a tous les pôles de $\coth \pi z$ en $z = in\pi$, $n \in \mathbb{Z}$. Le point $z = 0$ est visiblement un pôle d'ordre 9, et doit faire l'objet d'un traitement à part. Le résidu de $\phi(z)$ en $z = in\pi$ ($n \neq 0$) est $\frac{1}{(in)^7} \cot in\pi = \frac{1}{n^7} \coth n\pi$: la somme sur tous ces pôles reconstruit $2S$, aussi donné par la somme sur tous les pôles de $\pi \cot \pi z$. Comme la somme de tous les résidus est nulle (une autre façon de traduire le fait qu'il en est de même pour l'intégrale), il vient :

$$2S + 2S + \text{Res} \left[\frac{1}{z^7} \pi \cot \pi z \coth \pi z, z = 0 \right] = 0 , \quad (6.200)$$

d'où :

$$S = -\frac{1}{4} \text{Res} \left[\frac{1}{z^7} \pi \cot \pi z \coth \pi z, z = 0 \right] . \quad (6.201)$$

Il reste à trouver le développement à l'ordre 6 de $\pi \cot \pi z \coth \pi z$; à l'aide d'une machine, on trouve :

$$\pi \cot \pi z \coth \pi z = \frac{1}{\pi z^2} - \frac{7\pi^3}{45} z^2 - \frac{19\pi^7}{14175} z^6 + \mathcal{O}(z^{10}) , \quad (6.202)$$

d'où $S = \left(-\frac{1}{4}\right) \left(-\frac{19\pi^7}{14175}\right)$ et le résultat (6.199). Plus généralement, la considération⁸⁰ des sommes S_p :

$$S_p \stackrel{\text{déf}}{=} \sum_{n=1}^{+\infty} \frac{1}{n^{2p+1}} \coth n\pi \quad (6.203)$$

permet de comprendre pourquoi la partie régulière du développement de Laurent de $\pi \cot \pi z \coth \pi z$ ne contient que les puissances z^{4k+2} ($k \in \mathbb{N}$), une propriété qui ne saute

théorèmes demandèrent un siècle pour être tous démontrés.

Voir également le livre de Hardy [35], grand mathématicien lui-même, qui "donna à Ramanujan la note 100 alors qu'il n'attribuait que 80 à Hilbert, 30 à Littlewood, et 25 à lui-même" (mathematiques.fauriel.org/bio-ramanujan.pdf). Pour la réflexion (et l'anecdote) : Hardy écrit qu'il a beaucoup peiné pour établir l'égalité suivante, également due à Ramanujan :

$$\int_0^{+\infty} \frac{1}{(1+x^2)(1+r^2x^2)(1+r^4x^2)\dots} dx = \frac{\pi}{2(1+r+r^3+r^6+r^{10}+\dots)} . \quad (6.197)$$

Une autre formule de Ramanujan, sans doute encore plus extraordinaire, relie entre eux les deux nombres fondamentaux e et π via une série et une fraction continue :

$$1 + \frac{1}{1.3} + \frac{1}{1.3.5} + \frac{1}{1.3.5.7} + \dots + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \dots}}}}} = \sqrt{\frac{e\pi}{2}} \quad (6.198)$$

⁸⁰Faire cet exercice !

pas aux yeux au vu de l'expression de cette fonction⁸¹.

Pour calculer la somme d'une série alternée, impliquant à nouveau une fonction paire $f(-n) = f(n)$, du genre :

$$S = \sum_{n=1}^{+\infty} (-1)^n f(n) = \frac{1}{2} \sum_{n \in \mathbb{Z}, n \neq 0} (-1)^n f(n) , \tag{6.205}$$

et puisque $\text{Res}\left[\frac{\pi}{\sin \pi z}, z = n\right] = (-1)^n$, il suffit de considérer des intégrales du type :

$$I = \int_C \frac{\pi}{\sin \pi z} f(z) dz , \tag{6.206}$$

où le contour C , indiqué sur la fig. 6.8 en haut, prend en compte le fait que l'intégrand possède un pôle "de trop" en $z = 0$. Cette intégrale vaut :

$$I = 2i\pi \sum_{n \in \mathbb{Z}^*} \text{Res}\left[\frac{\pi f(z)}{\sin \pi z}, n\right] , \tag{6.207}$$

puisque les deux "lacets" peuvent être déformés en une série de petits cercles autour de chaque pôle, parcourus dans le sens positif, d'où $I = 4i\pi S$.

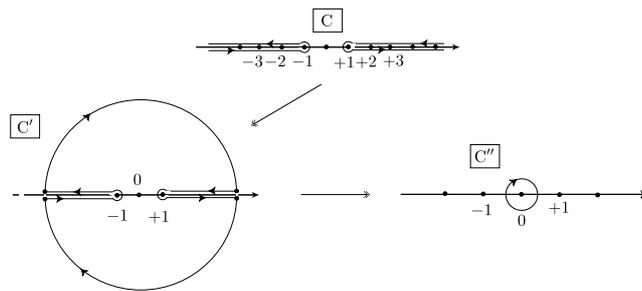


Figure 6.8: Fermeture et déformation du contour pour le calcul d'une série du genre (6.205).

On peut ensuite rabouter les deux lacets, les contributions des deux grands demi-cercles ajoutés étant nulles : I est donc aussi égale à l'intégrale le long du contour C' en bas à gauche de la fig. 6.8 ; mais cette dernière intégrale est égale à $-2i\pi \sum_{z_k} \text{Res}\left[\frac{\pi f(z)}{\sin \pi z}, z_k\right]$

⁸¹Sachant que (!!!) :

$$\pi \cot \pi z \coth \pi z = \frac{1}{\pi z^2} - \frac{7\pi^3}{45} z^2 - \frac{19\pi^7}{14175} z^6 - \frac{2906\pi^{11}}{212837625} z^{10} - \frac{13687\pi^{15}}{97692469875} z^{14} - \frac{15417074\pi^{19}}{10719306257034375} z^{18} + \mathcal{O}(z^{22}) , \tag{6.204}$$

combien vaut S_9 ? Pourquoi peut-on dire, sans *aucun* calcul, que $S_9 \simeq 1$?

où les z_k sont les pôles de $\frac{\pi}{\sin \pi z} f(z)$ situés à l'intérieur de C' (contour qui contient notamment le pôle en $z = 0$ venant de $\frac{1}{\sin \pi z}$, sans préjuger d'une singularité de $f(z)$ en ce point). D'où finalement, pour la somme (6.205) :

$$S = -\frac{1}{2} \sum_{z_k} \operatorname{Res} \left[\frac{\pi f(z)}{\sin \pi z}, z_k \right], \quad (6.208)$$

le signe venant du sens de parcours sur C' . De cette façon, on établit⁸² :

$$\sum_{n=1}^{+\infty} \frac{(-1)^n}{n^2 + a^2} = \frac{\pi}{2a \sinh \pi a} - \frac{1}{2a^2} \quad (a \notin i\mathbb{Z}). \quad (6.209)$$

Dans le cas où $f(z)$ n'a aucune singularité dans les demi-plans supérieur et inférieur, on peut faire collapser presque complètement le contour, le réduisant à un petit cercle C'' autour de l'origine parcouru dans le sens *négalif* (fig. 6.8 en bas à droite) ; il vient alors :

$$S = \sum_{n=1}^{+\infty} (-1)^n f(n) = \frac{1}{2} \frac{-1}{2i\pi} \int_{C''} \frac{\pi}{\sin \pi z} f(z) dz = -\frac{1}{2} \operatorname{Res} \left[\frac{\pi}{\sin \pi z} f(z), 0 \right]. \quad (6.210)$$

C'est ainsi que, prenant $f(z) = \frac{1}{z^4}$, on peut démontrer l'égalité :

$$\boxed{\sum_{n=1}^{+\infty} \frac{(-1)^n}{n^4} = -\frac{7\pi^4}{720}} \quad (6.211)$$

après avoir calculé le résidu de $\frac{\pi}{z^4 \sin \pi z}$ à l'origine. On peut le trouver par la formule générale donnant le résidu en un pôle d'ordre 5, mais aussi, ayant effectué le développement limité de $\frac{1}{z^{-1} \sin \pi z}$, comme le coefficient de $\frac{1}{z}$ de l'expression :

$$\frac{1}{z^5} \left[1 + \left(\frac{\pi^2 z^2}{3!} - \frac{\pi^4 z^4}{5!} \right) + \left(\frac{\pi^2 z^2}{3!} \right)^2 \right], \quad (6.212)$$

soit $+\frac{7\pi^4}{360}$, d'où (6.211) en vertu de (6.210).

6.6 Calcul de sommes finies

Le théorème des résidus, à nouveau appliqué "à l'envers", est aussi très utile pour trouver l'expression compacte de certaines somme *finies*. Un exemple est celui du calcul des sommes de Gauss, qui sont certaines sommes de puissances des racines N^{es} de l'unité. Posant précisément $\omega = e^{2i\pi/N}$, une certaine somme de Gauss est par exemple :

$$\boxed{G_N \stackrel{\text{d'éf}}{=} \sum_{n=0}^{N-1} \omega^{n^2} \equiv \sum_{n=0}^{N-1} e^{2i\pi(n^2/N)}} \quad (6.213)$$

⁸²égalité où l'on peut vérifier que les deux membres ont bien le même comportement près de $a = 0$.

Afin de simplifier le calcul ultérieur par résidus, il est utile de transformer quelque peu l'expression de G_N . Notons d'abord que $\omega^{(\frac{N}{2}-n)^2} = (\omega^N)^{\frac{N}{4}-n} \omega^{n^2} = \omega^{n^2}$ puisque $\omega^N = 1$. Ceci permet de récrire l'expression de G_N comme suit, selon la parité de N :

$$G_N = \begin{cases} 2 + 2 \sum_{n=1}^{\frac{N}{2}-1} e^{2i\pi(n^2/N)} & \text{si } N \text{ est pair} \\ 1 + 2 \sum_{n=1}^{\frac{N-1}{2}} e^{2i\pi(n^2/N)} & \text{si } N \text{ est impair} \end{cases} . \quad (6.214)$$

Dans tous les cas, G_N s'exprime à l'aide d'une somme sur des valeurs entières de l'indice, il faut à nouveau construire une fonction $f(z)$ ayant des pôles aux points entiers de l'axe réel. La nécessité pressentie de devoir en cours de route calculer une certaine intégrale (puisque la somme à trouver est *finie*) suggère le choix le plus simple, à savoir $e^{2i\pi z} - 1$. Par ailleurs, la forme du terme sommé dans (6.213) suggère fortement d'introduire le facteur $e^{2i\pi(z^2/N)}$. De proche en proche, on en vient ainsi à définir la fonction :

$$f(z) \stackrel{\text{déf}}{=} \frac{e^{2i\pi(z^2/N)}}{e^{2i\pi z} - 1} . \quad (6.215)$$

Par construction, cette fonction a des pôles en $z = n \in \mathbb{Z}$, et le résidu vaut $\frac{e^{2i\pi(n^2/N)}}{2i\pi}$.

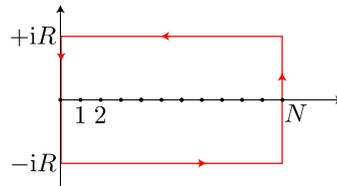


Figure 6.9: Contour utilisé pour calculer l'intégrale (6.216) donnant la somme de Gauss (6.213).

Il faut donc trouver un contour englobant un certain segment de l'axe réel ; on considère C_1 ou C_2 selon la parité de N (figure 6.9), l'abscisse précise du côté vertical étant choisie pour des raisons qui seront plus claires par la suite. Dans le cas où N est pair, le théorème des résidus donne :

$$I_2 \stackrel{\text{déf}}{=} \int_{C_2} \frac{e^{2i\pi(z^2/N)}}{e^{2i\pi z} - 1} dz = 2i\pi \left(\frac{1}{2} \frac{1}{2i\pi} + \sum_{n=1}^{\frac{N}{2}-1} \frac{e^{2i\pi(n^2/N)}}{2i\pi} + \frac{1}{2} \frac{e^{2i\pi(\frac{N}{2})^2/N}}{2i\pi} \right) ; \quad (6.216)$$

les facteurs $\frac{1}{2}$ prennent en compte le fait que les deux pôles $n = 0$ et $n = \frac{N}{2}$ sont *sur* le contour, l'intégrale étant implicitement régularisée à la *Cauchy* (voir section 6.4). De plus, N étant pair, $\frac{N}{2}$ est entier et le numérateur du terme de droite vaut 1 ; on a ainsi (voir (6.214)) :

$$I_2 = 1 + \sum_{n=1}^{\frac{N}{2}-1} e^{2i\pi(n^2/N)} \equiv \frac{1}{2} G_N . \quad (6.217)$$

Ainsi, l'intégrale I_2 est très simplement reliée à la somme G_N , d'où l'idée de la calculer *directement*. On commence par prendre la limite $R \rightarrow +\infty$, ce qui annule la contribution des côtés horizontaux du rectangle. En effet, sur CD, l'intégrand est :

$$\frac{e^{2i\pi(x+iR)^2/N}}{e^{2i\pi(x+iR)} - 1} = \frac{e^{2i\pi(x^2-R^2)/N} e^{-4\pi Rx/N}}{e^{2i\pi x} e^{-2\pi R} - 1} \underset{R \rightarrow \infty}{\sim} -e^{-4\pi Rx/N} ; \quad (6.218)$$

sur le côté AB, l'intégrand est :

$$\frac{e^{2i\pi(x-iR)^2/N}}{e^{2i\pi(x-iR)} - 1} = \frac{e^{2i\pi(x^2-R^2)/N} e^{4\pi Rx/N}}{e^{2i\pi x} e^{2\pi R} - 1} \underset{R \rightarrow \infty}{\sim} e^{-2\pi R(1-2x/N)} . \quad (6.219)$$

Comme dans tous les cas $x < \frac{N}{2}$, ce terme tend aussi vers zéro quand $R \rightarrow \infty$ - c'est ici qu'apparaît la raison du choix précis des abscisses des côtés verticaux du rectangle. Une fois cette limite prise, l'intégrale I_2 est :

$$I_2 = \int_{-\infty}^{+\infty} \left[\frac{e^{2i\pi(\frac{N}{2}+iy)^2/N}}{e^{2i\pi(\frac{N}{2}+iy)} - 1} - \frac{e^{2i\pi(iy)^2/N}}{e^{2i\pi(iy)} - 1} \right] idy = i \int_{-\infty}^{+\infty} \frac{i^N e^{-2\pi y} - 1}{e^{-2\pi y} - 1} e^{-i\frac{2\pi}{N}y^2} dy . \quad (6.220)$$

où on a utilisé le fait que N est ici pair, et l'intégrale étant à prendre comme une partie principale de Cauchy. Quand N est de la forme $4k$ (k entier), l'intégrale est gaussienne et on a $I_2 = i(\frac{N}{2i})^{1/2}$, d'où $G_N = (1+i)\sqrt{N}$. Si N est de la forme $4k+2$, l'intégrale est nulle comme on le voit en la décomposant en $\int_{-\infty}^0 + \int_0^{+\infty}$ et en effectuant le changement de variable $y = -y' \rightarrow -y$ dans la deuxième intégrale.

Le calcul pour N impair se conduit de la même façon. Le théorème des résidus avec C_1 donne (seul le pôle en zéro est sur le contour) :

$$I_1 \stackrel{\text{d'ef}}{=} \int_{C_1} \frac{e^{2i\pi(z^2/N)}}{e^{2i\pi z} - 1} dz = 2i\pi \left(\frac{1}{2} \frac{1}{2i\pi} + \sum_{n=1}^{\frac{N-1}{2}} \frac{e^{2i\pi(n^2/N)}}{2i\pi} \right) , \quad (6.221)$$

d'où, selon (6.214) :

$$I_1 = \frac{1}{2} + \sum_{n=1}^{\frac{N-1}{2}} e^{2i\pi(n^2/N)} \equiv \frac{1}{2} G_N . \quad (6.222)$$

Se tournant alors vers le calcul direct de I_1 , on voit à nouveau que les contributions des côtés horizontaux sont nulles dans la limite $R \rightarrow \infty$, et il reste (avec cette fois N impair) :

$$I_1 = \int_{-\infty}^{+\infty} \left[\frac{e^{2i\pi(\frac{N}{2}+iy)^2/N}}{e^{2i\pi(\frac{N}{2}+iy)} - 1} - \frac{e^{2i\pi(iy)^2/N}}{e^{2i\pi(iy)} - 1} \right] idy = -i \int_{-\infty}^{+\infty} \left[\frac{i^N e^{-2\pi y}}{e^{-2\pi y} + 1} + \frac{1}{e^{-2\pi y} - 1} \right] e^{-i\frac{2\pi}{N}y^2} dy , \quad (6.223)$$

la deuxième intégrale étant à prendre comme une partie principale de Cauchy. On a, coupant à nouveau l'intégrale en deux et posant $y = -y' \rightarrow -y$:

$$\int_{-\infty}^{+\infty} \frac{e^{-2\pi y}}{e^{-2\pi y} + 1} e^{-i\frac{2\pi}{N}y^2} dy = \int_0^{+\infty} \left[\frac{i^N e^{-2\pi y}}{e^{-2\pi y} + 1} + \frac{i^N e^{2\pi y}}{e^{2\pi y} + 1} + \frac{1}{e^{-2\pi y} - 1} + \frac{1}{e^{2\pi y} - 1} \right] e^{-i\frac{2\pi}{N}y^2} dy ; \quad (6.224)$$

la somme des deux premiers du crochet de droite vaut i^N , la somme des deux suivants vaut -1 , d'où $I_1 = -i(i^N - 1)\frac{1}{2}\left(\frac{N}{2i}\right)^{1/2} = \frac{1}{2\sqrt{2}}i^{1/2}(1 - i^N)\sqrt{N}$, soit $G_N = \frac{1}{2}(1+i)(1 - i^N)\sqrt{N}$. Pour N de la forme $4k + 1$, $G_N = \sqrt{N}$, pour $N = 4k + 3$, $G_N = i\sqrt{N}$.

Toutes les expressions de G_N peuvent se rassembler $\forall N$, dans la formule suivante :

$$G_N = \frac{1}{2}(1+i)(1 - i^{-N})\sqrt{N} , \quad (6.225)$$

soit explicitement :

$$G_N = \begin{cases} (1+i)\sqrt{N} & \text{si } N = 0 \text{ (4)} \\ \sqrt{N} & \text{si } N = 1 \text{ (4)} \\ 0 & \text{si } N = 2 \text{ (4)} \\ i\sqrt{N} & \text{si } N = 3 \text{ (4)} \end{cases} \quad (6.226)$$

La fonction $\frac{1}{\sqrt{N}}G_N$ est donc 4-périodique : $\frac{1}{\sqrt{N+4}}G_{N+4} = \frac{1}{\sqrt{N}}G_N$, soit :

$$G_{N+4} = \sqrt{1 + \frac{4}{N}} G_N , \quad (6.227)$$

une propriété qui ne saute pas aux yeux sur la définition (6.213). À l'inverse, la récurrence (6.227) posée *a priori* ne doit pas faire tomber dans le piège consistant à affirmer (trop) rapidement que la limite $\lim_{N \rightarrow +\infty} G_N$ existe ...

6.7 Quelques compléments sur les fonctions méromorphes

Il s'agit ici de donner quelques théorèmes et résultats parmi bien d'autres, le choix étant dicté par leur pertinence et leur utilité en Physique.

6.7.1 Formule intégrale de Poisson

Soit $f(z = re^{i\theta}) = u(r, \theta) + iv(r, \theta)$ une fonction holomorphe dans le disque $|z| < R$, noté \mathcal{D}_R , et continue sur sa frontière (le cercle C_R). La formule intégrale de Poisson donne une expression de chaque partie (réelle ou imaginaire) en fonction d'elle-même à la manière de la formule de Cauchy (4.62). Pour l'obtenir, il ne suffit pas de séparer parties réelle et imaginaire dans cette dernière (on trouverait chaque u ou v comme une intégrale portant sur u et v).

$f(z)$ a un développement en série entière $\sum_{n \in \mathbb{N}} a_n z^n$ dans \mathcal{D}_R . Sans nuire à la généralité du résultat, on peut supposer tous les a_n réels. En effet, dans le cas contraire et posant $a_n = \alpha_n + i\beta_n$, on a :

$$f(z) = \sum_{n \in \mathbb{N}} \alpha_n z^n + i \sum_{n \in \mathbb{N}} \beta_n z^n \stackrel{\text{déf}}{=} f_1(z) + if_2(z) , \quad (6.228)$$

d'où $\Re f = \Re f_1 - \Im f_2$, $\Im f = \Im f_1 + \Re f_2$; comme $|\alpha_n| \leq |a_n|$ et $|\beta_n| \leq |a_n|$, f_1 et f_2 sont aussi analytiques dans le disque \mathcal{D}_R . L'intégration étant une opération linéaire, le résultat quand $a_n \in \mathbb{C}$ résulte immédiatement du cas où $a_n \in \mathbb{R}$.

Soit $\xi = Re^{i\phi}$ un point de C_R ; la formule de Cauchy (4.62) s'écrit :

$$u(r, \theta) + iv(r, \theta) = \frac{1}{2i\pi} \int_0^{2\pi} \frac{u(R, \phi) + iv(R, \phi)}{Re^{i\phi} - re^{i\theta}} d(Re^{i\phi}) . \quad (6.229)$$

Avec $0 < r < R$, le point $\frac{R^2}{z}$ est à l'extérieur de \mathcal{D}_R et on a aussi $\frac{1}{2i\pi} \int_{C_R} \frac{f(\xi)}{\xi - R^2/z} d\xi = 0$, soit :

$$\frac{1}{2i\pi} \int_0^{2\pi} \frac{u(R, \phi) + iv(R, \phi)}{Re^{i\phi} - (R^2/r)e^{-i\theta}} d(Re^{i\phi}) = 0 , \quad (6.230)$$

ou encore :

$$\frac{1}{2\pi} \int_0^{2\pi} \frac{[u(R, \phi) + iv(R, \phi)]re^{i\theta}}{re^{i\theta} - Re^{-i\phi}} d\phi = 0 . \quad (6.231)$$

Les a_n étant réels, $f(\xi^*) = (f(\xi))^*$: $u(R, \phi)$ est donc paire en ϕ et $v(R, \phi)$ est impaire ; posant $\phi = -\phi'$ dans (6.231), il vient :

$$0 = \frac{1}{2\pi} \int_0^{2\pi} \frac{[u(R, \phi') - iv(R, \phi')]re^{i\theta}}{Re^{i\phi'} - re^{i\theta}} d\phi' ; \quad (6.232)$$

additionnant maintenant (6.229) et (6.232) membre à membre, il vient :

$$u(r, \theta) + iv(r, \theta) = \frac{1}{2\pi} \int_0^{2\pi} \left[\frac{Re^{i\phi} + re^{i\theta}}{Re^{i\phi} - re^{i\theta}} u(R, \phi) + iv(R, \phi) \right] d\phi ; \quad (6.233)$$

prenant la partie réelle, on obtient :

$$\boxed{u(r, \theta) = \frac{1}{2\pi} \int_0^{2\pi} \frac{R^2 - r^2}{R^2 - 2rR \cos(\theta - \phi) + r^2} u(R, \phi) d\phi} \quad (6.234)$$

Retranchant (6.229) et (6.232) membre à membre, et prenant la partie imaginaire, on voit que la même relation vaut aussi pour la partie imaginaire :

$$\boxed{v(r, \theta) = \frac{1}{2\pi} \int_0^{2\pi} \frac{R^2 - r^2}{R^2 - 2rR \cos(\theta - \phi) + r^2} v(R, \phi) d\phi} \quad (6.235)$$

Ces relations sont vraies en général pour des fonctions harmoniques satisfaisant l'équation de Laplace $\sum_{n=1}^N \partial_{x_n}^2 f = 0$, et ont donc leur équivalent dans \mathbb{R}^N .

6.7.2 Pôles et zéros d'une fonction méromorphe

Un résultat souvent utile en pratique est le suivant. Soit $f(z)$ une fonction holomorphe à l'intérieur d'un contour fermé C à l'exception d'un nombre (fini) de pôles, continue et ne s'annulant pas sur C ; alors :

$$\boxed{\frac{1}{2i\pi} \int_C \frac{f'(z)}{f(z)} dz = N - P} \quad (6.236)$$

où N et P sont respectivement les nombres de zéros et de pôles à l'intérieur de C , chacun d'entre eux étant compté un nombre de fois égal à son ordre. La démonstration procède comme suit. D'après le théorème des résidus, on a :

$$\frac{1}{2i\pi} \int_C \frac{f'(z)}{f(z)} dz = \sum_k \operatorname{Res} \left[\frac{f'(z)}{f(z)}, z_k \right] \quad (6.237)$$

Soit z_0 un zéro d'ordre m_0 de $f(z)$; dans le voisinage de z_0 , on a $f(z) = (z - z_0)^{m_0} g(z)$ où $g(z)$ est analytique et $g(z_0) \neq 0$, d'où $f'(z) = m_0(z - z_0)^{m_0-1} g(z) + (z - z_0)^{m_0} g'(z)$, et le rapport :

$$\frac{f'(z)}{f(z)} = \frac{m_0}{z - z_0} + \frac{g'(z)}{g(z)} . \quad (6.238)$$

Le second terme est analytique en $z = z_0$, le premier a un pôle simple en $z = z_0$, et le résidu correspondant vaut m_0 ; en raisonnant de même pour tous les zéros, on voit que la contribution des zéros de $f(z)$ au second membre du théorème des résidus est la somme des m_0 , soit la somme des multiplicités de tous les zéros de $f(z)$. Soit maintenant z_p un pôle d'ordre p_0 de $f(z)$; on peut alors écrire dans le voisinage de z_p : $f(z) = \frac{h(z)}{(z - z_p)^{p_0}}$ où $h(z)$ est analytique en z_p . Il en résulte :

$$\frac{f'(z)}{f(z)} = -\frac{p_0}{z - z_p} + \frac{h'(z)}{h(z)} . \quad (6.239)$$

Le même argument que ci-dessus montre que seul le premier terme a un résidu non nul, égal à $-p_0$; tous les pôles donnent donc une contribution totale égale à l'opposé de la somme de leurs multiplicités⁸³. Au total, on obtient le résultat (6.236).

Si $f(z)$ est analytique l'intérieur de C , on obtient simplement :

$$\frac{1}{2i\pi} \int_C \frac{f'(z)}{f(z)} dz = N . \quad (6.240)$$

Comme $\frac{f'(z)}{f(z)} = \frac{d}{dz} \ln z$, ceci s'écrit aussi $\frac{1}{2i\pi} \Delta_C [\ln f(z)] = N$, où $\Delta_C [\dots]$ est la variation sur le contour, soit $\Delta_C [\ln |f(z)| + i \arg f(z)]$; $\ln |f(z)|$ reprenant la même valeur au départ et à l'arrivée, on obtient :

$$\boxed{\Delta_C [\arg f(z)] = 2\pi N} \quad (6.241)$$

Clairement, le choix de la détermination du logarithme est indifférent.

Il est maintenant possible de démontrer l'important théorème de Rouché, qui joue un rôle de premier plan pour localiser les zéros d'une fonction dans certains domaines de \mathbb{C} . Il s'énonce :

*Soit $f(z)$ et $g(z)$ deux fonctions holomorphes dans et sur un contour C .
Si $|g(z)| < |f(z)|$ sur C , alors $f(z)$ et $f(z) + g(z)$ ont le même nombre de zéros à l'intérieur de C .*

⁸³En définitive, pour passer des zéros aux pôles, il suffit de changer le signe de la multiplicité, $m_0 \rightarrow -p_0$.

Notons d'abord que les hypothèses assurent que ni $f(z)$ ni $f(z) + g(z)$ n'ont de zéro sur le contour C : $|f(z)|$ est *strictement* borné par $|g(z)|$ qui est positif ou nul ; par ailleurs, $|f(z) + g(z)|$ s'annule si $g(z) = -f(z)$, mais ceci n'est pas possible puisque $|g(z)| < |f(z)|$. Maintenant, soit N_f et N_{fg} le nombre de zéros de f et $f + g$; d'après (6.241) :

$$\Delta_C[\arg f(z)] = 2\pi N_f \quad , \quad \Delta_C[\arg(f(z) + g(z))] = 2\pi N_{fg} \quad . \quad (6.242)$$

En notant que $f + g = f(1 + \frac{g}{f})$ et avec $\arg(z_1 z_2) = \arg z_1 + \arg z_2$, la deuxième relation s'écrit :

$$\Delta_C[\arg f(z) + \Delta_C\left[\arg\left(1 + \frac{g(z)}{f(z)}\right)\right]] = 2\pi N_{fg} \quad . \quad (6.243)$$

Établir l'égalité $N_{fg} = N_f$ revient à montrer que $\Delta_C[\arg(1 + \frac{g}{f})] = 0$. Comme $|g| < |f|$, le point $Z = 1 + \frac{g}{f}$ est à l'intérieur du cercle de \mathbb{C} centré en $(1, 0)$ et de rayon 1 ; en conséquence, l'argument de Z est borné par $\pm\frac{\pi}{2}$. Quand z revient à son point de départ, son argument a varié de 0 ou de 2π selon que le contour ne contient pas ou contient l'origine. Z aussi revient à son point de départ, mais son argument ne peut varier de 2π puisqu'il est compris entre $\pm\frac{\pi}{2}$: sa variation ne peut donc qu'être nulle, ce qui établit le théorème de Rouché.

À titre d'exemple, soit l'équation :

$$E(z) \stackrel{\text{déf}}{=} z^4 + z^3 + 4z^2 + 2z + 3 = 0 \quad , \quad (6.244)$$

dont on cherche à localiser les racines dans \mathbb{C} . Il n'existe pas de racines réelles. En effet, il n'y a visiblement pas de racines positives ; posant $z = -x$, l'équation s'écrit $x^4 - x^3 + 4x^2 - 2x + 3 = 0$. Pour $0 < x < 1$, les trois premiers termes donnent une quantité positive, tout comme les deux derniers ; pour $x > 1$, les deux premiers termes donnent $x^3(x - 1) > 0$, le reste est aussi positif. Posant maintenant $z = iy$, l'équation devient $y^4 - iy^3 - 4y^2 + 2iy + 3 = 0$: on voit que les deux équations $y^4 - 4y^2 + 3 = 0$ et $-y^3 + 2y = 0$ n'ont pas de racines communes, ce qui signifie que (6.244) n'a pas non plus de racines imaginaires pures.

Soit maintenant le contour formé du segment $[0, R]$ de l'axe réel, du quart de cercle de rayon R , et de l'intervalle allant de iR à l'origine sur l'axe imaginaire, la question étant de trouver la variation de l'argument du premier membre de (6.244) quand z décrit cette boucle. Le long de $[0, R]$, l'argument est constant et vaut zéro ; après le quart de tour, une fois parvenu en haut de l'axe imaginaire, l'argument de $E(z)$ vaut $\arg[(Re^{i\pi/2})^4(1 + \mathcal{O}(1/R))] \simeq 2\pi$. Quand $z = iy$ redescend l'axe imaginaire, l'argument de $E(z)$ est $\arctg\frac{-y^3 + 2y}{y^4 - 4y^2 + 3}$. L'étude de cette fraction rationnelle montre que $\arg(E(z))$ varie de 0 à $-\frac{\pi}{2}$ quand y varie de $+\infty$ à $\sqrt{3}_+$, puis décroît de π quand y passe de $\sqrt{3}_-$ à 1_+ , et décroît enfin de $\frac{\pi}{2}$ quand y passe de 1_- à 0_+ . Globalement, l'argument de $E(z)$ décroît de 2π entre $+i\infty$ et $+i0$, et au total, $\Delta[\arg(E(z))] = 0$ le long du contour. Par (6.240), on en déduit que le polynôme $E(z)$ n'a pas de zéros dans le premier quadrant. Comme il est à coefficients réels, la conclusion se tire d'elle-même : les 4 zéros sont dans les deuxième et troisième quadrants, deux à deux complexes conjugués.

Le résultat exprimé par (6.240) permet de démontrer simplement le théorème fondamental de l'algèbre : un polynôme $P(z)$ de degré d a d zéros, chaque zéro étant compté un nombre de fois égal à son ordre. En effet, soit $P(z) \stackrel{\text{déf}}{=} \sum_{n=0}^d a_n z^n$; on a :

$$\frac{P'(z)}{P(z)} = \frac{da_d z^{d-1} + \dots}{a_d z^d + \dots} = \frac{d}{z} + \mathcal{O}(|z|^{-2}) , \quad (6.245)$$

la deuxième égalité supposant $|z| \gg 1$. Si C est un grand cercle de rayon R , il vient :

$$\int_C \frac{P'(z)}{P(z)} dz = \int_C \left(\frac{d}{z} + \mathcal{O}(R^{-2}) \right) dz ; \quad (6.246)$$

la deuxième intégrale est $\sim \mathcal{O}(R^{-1})$ et tend vers zéro dans la limite $R \rightarrow +\infty$. L'intégrale restante vaut $2i\pi d$; suivant (6.240), elle vaut aussi $2i\pi N$, d'où le théorème $N = d$.

6.8 Représentation des fonctions méromorphes et des fonctions entières

Les fonctions méromorphes et les fonctions entières constituent deux grandes familles de fonctions pour lesquelles on sait trouver des expressions dont la forme même illustre bien leurs propriétés. Pour les premières, il existe un développement en éléments rationnels simples, attestant du fait que les fonctions méromorphes peuvent être considérées comme constituant une généralisation des fractions rationnelles. Ce développement est un développement sur les pôles de la fonction, et s'appelle aussi *développement de Mittag-Leffler*.

Quant à elles, les fonctions entières peuvent s'écrire sous la forme d'un produit infini, montrant par là qu'elles peuvent être vues comme une généralisation des polynômes.

6.8.1 Décomposition des fonctions méromorphes en série d'éléments simples

Pour commencer, donnons l'idée centrale de ce genre de développement. Soit une fraction rationnelle $\frac{P(z)}{Q(z)}$, où P et Q sont des polynômes et avec $n_P \equiv$ degré de P , inférieur à n_Q degré de Q . On peut toujours l'exprimer sous la forme d'une somme d'éléments simples du genre $\frac{c_k}{(z-z_k)^{n_k}}$; dans le cas où $Q(z)$ n'a que des zéros simples, on a ainsi :

$$\frac{P(z)}{Q(z)} = \sum_{k=1}^{n_Q} \frac{c_k}{z - z_k} . \quad (6.247)$$

En intégrant $\frac{P}{Q}$ sur une petite boucle ne contenant que l'un des pôles, z_k , on voit que c_k n'est autre que le résidu de $\frac{P(z)}{Q(z)}$ en z_k . En définitive, le développement en éléments

simples n'est autre que le développement de la fraction rationnelle sur ses pôles. C'est ceci que l'on peut généraliser à une fonction méromorphe quelconque.

Soit une fonction méromorphe $f(z)$ dont, pour simplifier, on supposera que tous les pôles sont *simples* (la généralisation à des pôles d'ordre supérieur à 1 est sans difficulté), et indexés de sorte que $0 < |z_1| \leq |z_2| \leq |z_3| \dots$; les résidus correspondants sont notés r_1, r_2, r_3, \dots . Considérons l'intégrale $I(z) \stackrel{\text{def}}{=} \frac{1}{2i\pi} \int_{C_N} \frac{f(\xi)}{\xi(\xi-z)} d\xi$, où C_N est un cercle⁸⁴ de rayon R_N contenant tous les z_k , $1 \leq k \leq N$, et avec z à l'intérieur de C_N . Par le théorème des résidus, on a :

$$I(z) = \sum_{k=1}^N \frac{r_k}{z_k(z_k - z)} - \frac{f(0)}{z} + \frac{f(z)}{z} . \quad (6.248)$$

D'un autre côté, on a :

$$|I(z)| \leq \frac{2\pi R_N}{2\pi R_N |R_N - z|} \text{Max}_{C_N} |f(\xi)| . \quad (6.249)$$

Supposons maintenant que $\frac{1}{R_N} \text{Max}_{C_N} |f(\xi)| \rightarrow 0$ quand $R_N \rightarrow \infty$; alors l'intégrale tend vers zéro dans cette limite, et on en déduit :

$$f(z) = f(0) + z \sum_{k=1}^{\infty} \frac{r_k}{z_k(z - z_k)} , \quad (6.250)$$

soit :

$$f(z) = f(0) + \sum_{k=1}^{\infty} r_k \left(\frac{1}{z - z_k} + \frac{1}{z_k} \right) \quad \forall z \neq z_k \quad (6.251)$$

La démonstration montre aussi que la série converge uniformément à l'intérieur de tout contour laissant tous les pôles à l'extérieur. À titre d'exemple, soit $f(z) \stackrel{\text{def}}{=}} \frac{1}{\sin z} - \frac{1}{z}$, qui a des pôles simples en $z_k = k\pi$, $k \in \mathbb{Z}^*$ (la singularité en $z=0$ est éliminable). Le résidu en $k\pi$ s'obtient par $\lim_{z \rightarrow k\pi} [(z - k\pi) (\frac{1}{\sin z} - \frac{1}{z})] = (-1)^k$. Il n'est pas difficile de voir que $\frac{1}{\sin z}$ est borné sur le cercle de rayon R_N , de sorte que le développement (6.251) existe et s'écrit :

$$\frac{1}{\sin z} - \frac{1}{z} = \sum_{k \in \mathbb{Z}^*} (-1)^k \left(\frac{1}{z - k\pi} + \frac{1}{k\pi} \right) , \quad (6.252)$$

soit :

$$\frac{1}{\sin z} = \frac{1}{z} + 2z \sum_{k=1}^{+\infty} \frac{(-1)^{k+1}}{k^2 \pi^2 - z^2} \quad (6.253)$$

On trouve de même :

$$\tan z = 2z \sum_{k=1}^{+\infty} \frac{1}{(k + \frac{1}{2})^2 \pi^2 - z^2} \quad (6.254)$$

⁸⁴Clairement, on peut remplacer le cercle par un contour fermé quelconque Γ_N contenant les mêmes pôles; alors, le rayon R_N dans l'argument qui suit est remplacé par la plus petite distance \tilde{R}_N entre un point de Γ_N et l'origine, et $2\pi R_N$ par la longueur L_N du contour Γ_N .

Le développement de $\cot z$ a déjà été rencontré, et obtenu par un moyen détourné (voir (6.194)).

6.8.2 Décomposition des fonctions entières en produits infinis

Une fonction entière est une fonction holomorphe pour tout z de module fini, et peut être considérée, de ce point de vue, comme la généralisation d'un polynôme. Cette propriété saute aux yeux quand la fonction peut être écrite sous la forme d'un produit infini.

Soit une fonction entière $f(z)$, dont les zéros z_k sont supposés simples pour la seule commodité du raisonnement. Au voisinage de l'un quelconque d'entre eux, z_k , on a $f(z) = (z - z_k)g(z)$ où $g(z)$ est holomorphe et telle que $g(z_k) \neq 0$. Par dérivation il vient $f'(z) = g(z) + (z - z_k)g'(z)$ puis :

$$\frac{f'(z)}{f(z)} = \frac{1}{z - z_k} + \frac{g'(z)}{g(z)}, \quad (6.255)$$

le second terme étant analytique en z_k ; la fonction $\frac{f'(z)}{f(z)}$ a donc un pôle simple en z_k avec le résidu 1. Supposons maintenant que $\frac{f'(z)}{f(z)}$ a les propriétés requises pour avoir un développement de Mittag - Leffler (voir (6.251)) ; avec cette hypothèse, les zéros de $f(z)$ sont les pôles de $\frac{f'}{f}$, avec $r_k = 1$ quel que soit k d'après (6.255) :

$$\frac{f'(z)}{f(z)} = \frac{f'(0)}{f(0)} + \sum_{k=1}^{\infty} \left(\frac{1}{z - z_k} + \frac{1}{z_k} \right). \quad (6.256)$$

Intégrons maintenant membre à membre sur un chemin C partant de l'origine pour arriver en z et évitant évidemment tout pôle ; il vient :

$$\ln f(z) - \ln f(0) = z \frac{f'(0)}{f(0)} + \sum_{k=1}^{\infty} \left[\ln(z - z_k) - \ln(-z_k) + \frac{z}{z_k} \right], \quad (6.257)$$

toutes les fonctions \ln étant à prendre suivant *une* certaine détermination, soit encore :

$$\ln \frac{f(z)}{f(0)} = z \frac{f'(0)}{f(0)} + \sum_{k=1}^{\infty} \left[\ln \frac{z - z_k}{-z_k} + \frac{z}{z_k} \right], \quad (6.258)$$

Prenant les exponentielles (ce qui efface d'ailleurs toute absence de précision sur la branche du logarithme choisie), on obtient :

$$\boxed{f(z) = f(0) e^{z \frac{f'(0)}{f(0)}} \prod_{k=1}^{\infty} \left(1 - \frac{z}{z_k} \right) e^{\frac{z}{z_k}}} \quad (6.259)$$

Soit par exemple $f(z) = \frac{1}{z} \sin z$; dans ce cas, $\frac{f'(z)}{f(z)} = \cot z - \frac{1}{z}$ a un développement de Mittag - Leffler et des pôles simples aux points $z_k = k\pi$. De plus $f(0) = 1$, $f'(0) = 0$, de sorte que (6.259) donne ici $\frac{\sin z}{z} = \prod_{k \in \mathbb{Z}^*} \left(1 - \frac{z}{k\pi}\right) e^{\frac{z}{k\pi}}$, soit :

$$\boxed{\sin z = z \prod_{k=1}^{\infty} \left(1 - \frac{z^2}{k^2 \pi^2}\right)} \quad (6.260)$$

De la même façon, on a :

$$\boxed{\cos z = \prod_{k=0}^{\infty} \left(1 - \frac{z^2}{\left(k + \frac{1}{2}\right)^2 \pi^2}\right)} \quad (6.261)$$

Les développements de $\sinh z$ et $\cosh z$ s'obtiennent en substituant iz à z , comme d'habitude. Un autre exemple, un peu plus ésotérique, est fourni par l'expression historique de la fonction d'Euler $\Gamma(z)$ (voir chapitre 7, section 7.1.1, éq. (7.1)).

Chapitre 7

Quelques applications de la théorie des fonctions d'une variable complexe

“La mathématique est l'art de donner le même nom
à des choses différentes.”

(Henri POINCARÉ, 1854–1912)

*Il s'agit d'utiliser et d'illustrer les concepts précédemment introduits
à propos d'objets ou de méthodes d'usage très courant en Physique.*

Les applications de la théorie des fonctions d'une variable complexe sont innombrables. Il s'agit ici de présenter quelques applications importantes en Physique des questions exposées précédemment, permettant d'introduire des fonctions spéciales apparaissant fréquemment, et dont certaines – la fonction de Riemann par exemple – constituent des points de convergence parfois surprenants entre Mathématiques et Physique à propos de questions fondamentales – en l'occurrence, la distribution des nombres premiers. Il s'agit aussi de présenter des outils d'usage courant (méthode du col, développements asymptotiques, sommation des séries divergentes, ...) autorisant l'analyse fine de problèmes “insolubles” attaqués sous l'angle des méthodes perturbatives.

7.1 Les fonctions d'Euler $\Gamma(z)$ et $B(p, q)$

7.1.1 La fonction Gamma $\Gamma(z)$

Afin d'en venir graduellement et intuitivement à la fonction $\Gamma(z)$ d'Euler¹, considérons l'intégrale suivante :

$$\Gamma_n \stackrel{\text{déf}}{=} \int_0^{+\infty} t^{n-1} e^{-t} dt, \quad (7.2)$$

où n est un entier supérieur ou égal à 1. Cette intégrale, qui existe de toute évidence, se calcule par les moyens élémentaires en effectuant d'abord une intégration par parties :

$$\int_0^{+\infty} t^{n-1} e^{-t} dt = [-t^{n-1} e^{-t}]_0^{+\infty} + (n-1) \int_0^{+\infty} t^{n-2} e^{-t} dt; \quad (7.3)$$

le terme tout intégré vaut δ_{n1} , l'intégrale de droite n'est autre que Γ_{n-1} , d'où la relation de récurrence entre les Γ_n :

$$\Gamma_n = (n-1)\Gamma_{n-1} \quad \text{si } n > 1, \quad \Gamma_1 = 1. \quad (7.4)$$

Il en résulte immédiatement que, pour $n \in \mathbb{N}^*$:

$$\boxed{\Gamma_n = (n-1)(n-2)\dots 2.1 = (n-1)! \quad (n \in \mathbb{N}^*)} \quad (7.5)$$

Noter que la convention habituelle $0! = 1$ se trouve *ipso facto* justifiée par $\int_0^{+\infty} e^{-t} dt = 1$.

Il est clair que cette expression simple et explicite de Γ_n est liée au fait que le nombre n est un entier positif. D'un autre côté, revenant à l'intégrale au second membre de (7.2), rien n'interdit de remplacer cet entier par un certain réel x , ou même un complexe z , quitte à examiner de plus près par la suite le domaine de définition de la fonction ainsi obtenue. On en vient ainsi tout naturellement à définir la fonction $\Gamma(z)$ par l'égalité² :

$$\boxed{\Gamma(z) \stackrel{\text{déf}}{=} \int_0^{+\infty} t^{z-1} e^{-t} dt \quad (\Re z > 0)} \quad (7.6)$$

où l'intégration se fait toujours sur la variable réelle t . Pour le coup, avec $z \in \mathbb{C}$, aucune expression n'existe pour $\Gamma(z)$ en terme de fonctions élémentaires. La précision $\Re z > 0$ est essentielle et est expliquée ci-dessous.

¹Aussi appelée fonction d'Euler de deuxième espèce. La fonction d'Euler de première espèce est la fonction $B(p, q)$ introduite plus loin (voir sous-section 7.1.2)).

Historiquement, Euler a introduit $\Gamma(z)$ à partir du produit infini :

$$\Gamma z = \frac{1}{z} \prod_{n=1}^{+\infty} \left[\left(1 + \frac{1}{n}\right)^z \left(1 + \frac{z}{n}\right)^{-1} \right]; \quad (7.1)$$

cette expression montre clairement que $\Gamma(z)$ a des pôles d'ordre 1 pour $z = 0, -1, -2, \dots$, un résultat qui sera établi autrement plus loin (voir (7.31)).

²On peut ainsi dire que $\Gamma(z)$ est la transformée de Mellin de e^{-t} .

La définition intégrale (7.6) de $\Gamma(z)$ est sans doute la plus naturelle, étant entendu qu'il existe d'autres écritures équivalentes, rencontrées ici et là dans la suite. Elle montre déjà, compte tenu de (7.5), que $\Gamma(z = n \in \mathbb{N}^*) \equiv \Gamma_n = (n-1)!$, soit :

$$\boxed{\Gamma(n+1) = n! \quad (n \in \mathbb{N})} \quad (7.7)$$

Dans le contexte de l'intégration des fonctions d'une variable complexe, (7.6) s'écrit aussi :

$$\Gamma(z) = \int_{\mathbb{R}_+} t^{z-1} e^{-t} dt ; \quad (7.8)$$

$\Gamma(z)$ est ainsi l'intégrale le long du chemin \mathbb{R}_+ du plan complexe de la branche de la fonction $f(t) = t^{z-1} e^{-t}$ qui prend des valeurs réelles pour $z \in \mathbb{R}_+$. De fait, l'une ou l'autre de ces écritures définit proprement $\Gamma(z)$ dans tout le demi-plan de droite. En effet, avec $z = x + iy$ et pour $t \in \mathbb{R}_+$:

$$|t^{z-1} e^{-t}| = |t^{x+iy-1}| e^{-t} = t^{x-1} |e^{iy \ln t}| e^{-t} = t^{x-1} e^{-t} , \quad (7.9)$$

de sorte que :

$$|\Gamma(z)| = \left| \int_0^{+\infty} t^{z-1} e^{-t} dt \right| \leq \int_0^{+\infty} |t^{z-1} e^{-t}| dt = \int_0^{+\infty} t^{x-1} e^{-t} dt . \quad (7.10)$$

L'intégrand est sommable à l'origine pourvu que $\Re z \equiv x > 0$. Considéré comme une fonction de la variable complexe t , l'intégrand contient certes la fonction multiforme t^{z-1} , dont l'origine est un point de branchement ; toutefois rien n'interdit de faire arriver l'intervalle d'intégration en t jusqu'à ce point³, tant que $\Re z > 0$. Quant à la limite supérieure, elle ne pose pas de problème, l'exponentielle l'emportant toujours sur t^{x-1} quel que soit z fini fixé. On peut ainsi affirmer que, quand $\Re z > 0$, $|\Gamma(z)|$ est borné par une intégrale convergente : (7.6), ou (7.8), définit une fonction majorable dans tout le demi-plan de droite.

$\Gamma(z)$ est une fonction continue pour $\Re z > 0$, la fonction $f(z, t) \stackrel{\text{def}}{=} t^{z-1} e^{-t}$ étant séparément continue en t et en z ; sa dérivée $\partial_z f$ l'étant aussi, la dérivation sous le signe somme (voir (4.53), et aussi [12]) donne⁴ :

$$\boxed{\Gamma'(z) = \int_0^{+\infty} (\ln t) t^{z-1} e^{-t} dt} \quad (7.12)$$

³Quand $z \in \mathbb{R}$ tend vers zéro, l'intégrale (7.10) diverge, la majoration (7.10) devient vide de sens, et on ne peut plus rien dire. On verra par la suite que $\Gamma(z) \propto \frac{1}{z}$ quand $|z| \rightarrow 0$: $\Gamma(z)$ est donc encore définie pour $\Re z = 0$ (et bien au-delà, comme on le verra) à condition que $\Im z \neq 0$, mais au stade actuel on n'en sait rien.

⁴On a aussi :

$$\Gamma''(z) = \int_0^{+\infty} (\ln t)^2 t^{z-1} e^{-t} dt , \quad (7.11)$$

et ainsi de suite.

Ceci se voit aussi directement en formant la dérivée :

$$\Gamma'(z) = \lim_{h \rightarrow 0} \frac{1}{h} \int_0^{+\infty} t^{z-1} (t^h - 1) e^{-t} dt ; \quad (7.13)$$

la parenthèse est $e^{h \ln t} - 1 = h \ln t + h\eta(h)$ où $\lim_{h \rightarrow 0} \eta(h) = 0$, compte tenu de la continuité de la dérivée. Il reste ainsi :

$$\Gamma'(z) = \lim_{h \rightarrow 0} \int_0^{+\infty} t^{z-1} [\ln t + \eta(h)] e^{-t} dt = \int_0^{+\infty} t^{z-1} \ln t e^{-t} dt . \quad (7.14)$$

Ces expressions, définies $\forall z, \Re z > 0$, montrent que $\Gamma(z)$ a une dérivée partout à droite. Au total, à ce stade, on a démontré que l'intégrale (7.6) définit une fonction qui est holomorphe $\forall z, \Re z > 0$; les singularités de $\Gamma(z)$ se trouvent donc toutes dans le demi-plan $\Re z \leq 0$.

Une fois $\Gamma(z)$ ainsi définie, on peut démontrer aisément la relation fonctionnelle remarquable suivante⁵ :

$$\boxed{\Gamma(z+1) = z\Gamma(z) \quad (\Re z > 0)} \quad (7.15)$$

dont pour l'instant on pourra seulement affirmer qu'elle est vraie⁶ $\forall z, \Re z > 0$, puisqu'elle va être démontrée à partir de la définition intégrale (7.6) – d'où la précision (provisoire) à droite. Partant de celle-ci, écrite pour $z+1$, une intégration par parties donne :

$$\Gamma(z+1) = \int_0^{+\infty} t^{(z+1)-1} e^{-t} dt = \left| -t^z e^{-t} \right|_0^{+\infty} - z \int_0^{+\infty} t^{z-1} (-e^{-t}) dt \quad (\Re z > 0) . \quad (7.16)$$

Le terme tout intégré est toujours nul tant que $\Re z > 0$ puisqu'il vaut $e^{z \ln t} e^{-t}$: le premier facteur est nul à la limite $t = 0$, en $t = +\infty$, c'est le tour du second. L'intégrale de droite n'est autre que $-\Gamma(z)$, d'où la relation (7.15), qui généralise au cas non-entier la relation de récurrence (7.4) écrite pour $n+1$: $\Gamma(n+1) \equiv \Gamma_{n+1} = n\Gamma_n \equiv n\Gamma(n)$. De par sa définition (7.6), pour les valeurs entières de sa variable, Γ n'est autre que la factorielle⁷, dont on sait qu'il s'agit d'une fonction définie sur \mathbb{N} croissant à une vitesse vertigineuse.

Montrons maintenant comment la fonction $\Gamma(z)$ peut être prolongée analytiquement dans le demi-plan de gauche ($\Re z \leq 0$), en-dehors bien sûr de singularités à identifier ultérieurement. Pour cela, considérons toujours la branche de t^{z-1} qui vaut 1 quand $t = 1$ et $(e^{2i\pi})^{(z-1)} = e^{2i\pi z}$ quand t a décrit un tour autour de O et revient infiniment près de 1 par en-dessous : la coupure s'étend de l'origine à $+\infty$, c'est le demi-axe réel positif \mathbb{R}_+ . Soit la fonction $F(z)$ définie comme :

$$F(z) \stackrel{\text{d\'ef}}{=} \int_C t^{z-1} e^{-t} dt , \quad (7.17)$$

où C est le contour représenté sur la fig. 7.1 à gauche (un lacet faisant le tour l'origine).



Figure 7.1: Contours utilisés pour le prolongement analytique de la fonction $\Gamma(z)$ (voir (7.17)).

Pour pouvoir exploiter cette définition, il convient de préciser les propriétés de la fonction $F(z)$. Le voisinage de l'origine ne soulève aucune question, parce que le "rayon" du circuit autour de l'origine est ce qu'il est, et ne tend pas vers zéro (il reste fini). La seule question est donc de savoir si l'intégrale converge à l'infini. Pour élucider ce point, rabattons chaque brin juste au-dessus et juste au-dessous de l'axe réel au-delà d'une certaine abscisse fixe t_0 (la variable d'intégration t devenant alors réelle, voir fig. 7.1 à droite), ce qui permet de majorer comme suit. Si z est fini, alors $\exists M$ tel que $|z| < M$, donc $|x = \Re z| < M$; comme $|x - 1| \leq |x| + 1$, il en résulte que pour⁸ $t > t_0 > 1$, $|e^{(z-1)\ln t}| = e^{|x-1|\ln t} < e^{(M+1)\ln t} = t^{M+1}$; l'intégrale $\int_{t_0}^{+\infty} t^{M+1} e^{-t} dt$ étant convergente $\forall M$ fixé, l'intégrale impropre associée à chaque brin courant à l'infini converge uniformément en z . En conséquence ([9], théorème 4 du § 16) la fonction $F(z)$ est holomorphe quel que soit z de module fini : c'est une fonction entière⁹. De surcroît, une intégration par parties montre immédiatement que $F(z)$ satisfait aussi la relation fonctionnelle :

$$\boxed{F(z+1) = zF(z)} \quad (7.18)$$

Enfin, on note que $F(z \in \mathbb{N}^*) = 0$: en pareil cas, l'intégrand n'est pas multiforme, prend les mêmes valeurs sur les deux bords, de sorte que l'intégrale donnant $F(z)$ est nulle¹⁰.

On va maintenant montrer que, moyennant une déformation par continuité du contour initial choisi pour définir $F(z)$ selon (7.17), il est possible d'établir une relation simple entre $F(z)$ et la fonction $\Gamma(z)$, relation vraie dans tout le demi-plan de droite. Dès lors, on aura amplement satisfait la nécessité requise par le théorème du prolongement analytique : celui-ci exige que les deux expressions coïncident sur l'arc de frontière commune, alors que l'on aura montré la coïncidence $\forall z, \Re z > 0$. L'égalité étant acquise, mais $F(z)$ étant comme on vient de le voir définie dans tout le plan, le prolongement analytique sera effectué.

⁵Attention, cette relation n'est pas *caractéristique* de la fonction Γ : toute fonction $\Phi(z) = f(z)\Gamma(z)$ où la fonction f est 1-périodique, $f(z+1) = f(z)$, satisfait également la relation $\Phi(z+1) = z\Phi(z)$ (par exemple $\Phi(z) = e^{2i\pi z}\Gamma(z)$, ou $\Phi(z) = F(z)$ définie en (7.17)).

⁶En temps utile, on verra que cette relation se prolonge dans tout le plan complexe, partout où Γ est analytique, c'est-à-dire dans $\mathbb{C} \setminus (-\mathbb{N})$.

⁷Se souvenir du décalage $\Gamma(n+1) = n!$.

⁸Se souvenir que pour $t > t_0$, t est réel.

⁹Rappelons qu'une fonction est entière ssi elle ne possède aucune singularité à distance finie (c'est le cas de e^z). Une telle fonction possède au plus des singularités pour $|z| = \infty$; quand elle n'en possède même pas à l'infini, c'est la fonction constante (Théorème de Liouville).

¹⁰On peut aussi dire ceci : quand $z \in \mathbb{N}^*$, il n'y a plus de coupure et l'intégrand est une fonction entière ; on peut donc déformer le contour en le tirant tout entier à l'infini à droite, ce qui donne

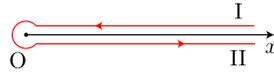


Figure 7.2: Déformation du contour de la fig. 7.1 pour établir la relation (7.24) entre $\Gamma(z)$ et $F(z)$.

Pour établir cette relation entre les deux fonctions $F(z)$ et $\Gamma(z)$ dans le demi-plan de droite $x = \Re z > 0$, et sachant que $F(z)$ est holomorphe partout, déformons le contour initial pour le transformer en deux demi-droites sur les bords supérieur (I) et inférieur (II) de la coupure de t^{z-1} et un petit cercle γ_r de rayon r autour de l'origine (voir fig. 7.2), on peut écrire explicitement :

$$F(z) = \int_I t^{z-1} e^{-t} dt + \int_{\gamma_r} t^{z-1} e^{-t} dt + \int_{II} t^{z-1} e^{-t} dt . \quad (7.19)$$

Sur le bord supérieur de la coupure on a simplement $(t^{z-1})_I = |t|^{z-1} = t^{z-1}$, l'argument de t valant 0 ; quand t arrive sur le bord inférieur après contournement de l'origine, l'argument de t a augmenté de 2π : $(t^{z-1})_{II} = (te^{2i\pi})^{z-1} \equiv |t|^{z-1} e^{2i\pi(z-1)} = |t|^{z-1} e^{2i\pi z}$, d'où $(t^{z-1})_{II} = e^{2i\pi z} (t^{z-1})_I$. Prenant en compte les sens de parcours des demi-droites I et II, il vient (pour le contour déformé de la fig. 7.2) :

$$F(z) = (e^{2i\pi z} - 1) \int_r^{+\infty} t^{z-1} e^{-t} dt + \int_0^{2\pi} r^{z-1} e^{i(z-1)\theta} e^{-re^{i\theta}} d(ir e^{i\theta}) . \quad (7.20)$$

r étant pour l'instant fixé, l'intégrale (impropre) de r à ∞ converge uniformément $\forall z$ fini ; en effet, le module de l'intégrand est :

$$|t^{z-1}| e^{-t} = |e^{(x-1+iy) \ln t}| e^{-t} = t^{x-1} e^{-t} ; \quad (7.21)$$

pour tout z fini ($|z| < M$), cette quantité est majorée par $t^{M-1} e^{-t}$, borne indépendante de z .

Prenons la limite $r \rightarrow 0$; le second terme du deuxième membre de (7.20) est $ir^z \int_0^{2\pi} e^{iz\theta} e^{-re^{i\theta}} d\theta$ et tend vers zéro¹¹ (il est en module $\propto r^x$ où $x = \Re z > 0$). Dans cette même limite, la première intégrale au second membre de (7.20) n'est autre que celle apparaissant dans (7.6), c'est donc $\Gamma(z)$, d'où la relation entre F et Γ , vraie dans tout le demi-plan $\Re z > 0$:

$$F(z) = (e^{2i\pi z} - 1) \Gamma(z) \quad \forall z, \Re z > 0 . \quad (7.22)$$

Cette égalité permet d'abord de préciser le comportement de $F(z)$ quand z est au voisinage d'un entier positif (on sait déjà que $F(n \in \mathbb{N}) = 0$, voir note 10 p. 289). En effet, $\Gamma(z)$ est analytique à droite, et admet donc un développement en série entière (de Taylor,

visiblement une intégrale nulle. Le comportement de $F(z)$ pour z dans le voisinage d'un entier positif est donné en (7.23).

¹¹Ceci n'est vrai que pour $\Re z > 0$, mais répétons que c'est justement dans le seul plan de droite que l'on a besoin de la relation précise entre F et Γ (pour l'instant, Γ n'est définie qu'à droite !).

pas de puissances négatives) ; posant $\xi = z - n$ ($n \in \mathbb{N}^*$), (7.22) donne :

$$F(z = n + \xi) = (e^{2i\pi(n+\xi)} - 1) [\Gamma(n) + \mathcal{O}(\xi)] = (e^{2i\pi\xi} - 1) [\Gamma(n) + \mathcal{O}(\xi)] = 2i\pi\Gamma(n)\xi + \mathcal{O}(\xi^2) ; \quad (7.23)$$

ainsi, dans le voisinage d'un entier positif n , $F(z)$ s'annule comme la différence $z - n$. En outre, et c'est le point important, (7.22) lue à l'envers donne :

$$\Gamma(z) = \frac{F(z)}{e^{2i\pi z} - 1} \quad \forall z, \Re z > 0 ; \quad (7.24)$$

cette relation est vraie partout à droite¹², mais comme F est en réalité définie dans tout le plan par (7.17), la relation (7.24) prolonge Γ au-delà du demi-plan de droite. En définitive, l'égalité :

$$\boxed{\Gamma(z) = \frac{1}{e^{2i\pi z} - 1} \int_C t^{z-1} e^{-t} dt \quad \forall z \in \mathbb{C}} \quad (7.25)$$

constitue une autre représentation intégrale de la fonction $\Gamma(z)$ affranchie de la restriction au demi-plan $\Re z > 0$ qu'imposait la définition première (7.8). Les deux fonctions définies par (7.8) et (7.25) sont une seule et même fonction $\Gamma(z)$ dans le demi-plan de droite, mais la seconde est également définie pour $\Re z \leq 0$: les conditions du théorème du prolongement analytique sont donc largement remplies. La fonction $\Gamma(z)$ se trouve ainsi définie dans \mathbb{C} , à l'exception de singularités ne pouvant se trouver que dans le demi-plan $\Re z \leq 0$, et dont on va maintenant préciser la nature. Notons aussi que la relation fonctionnelle (7.15), alors établie avec la restriction $\Re z > 0$, se prolonge automatiquement ; par (7.18) et (7.24), on peut maintenant écrire :

$$\boxed{\Gamma(z+1) = z\Gamma(z) \quad \forall z} \quad (7.26)$$

$F(z)$ est une fonction entière (voir la discussion après (7.17)) ; l'expression (7.24), rapport de deux fonctions entières, est donc analytique partout dans le plan sauf *a priori* aux zéros du dénominateur, qui sont tous les entiers positifs ou négatifs. Pour les valeurs entières strictement positives, il s'agit visiblement de singularités apparentes (éliminables) puisque l'on sait depuis le début que $\Gamma(z)$ est holomorphe à droite ; les seules singularités de $\Gamma(z)$ sont donc tous les entiers négatifs ou nuls. Pour en préciser complètement la nature¹³, itérons la relation fonctionnelle (7.15) écrite avec $z \rightarrow z + n$, on tire :

$$\Gamma(z + n + 1) = (z + n)\Gamma(z + n) = (z + n)(z + n - 1)\Gamma(z + n - 1) = \dots ; \quad (7.27)$$

en répétant ces opérations, on arrive à :

$$\Gamma(z + n + 1) = (z + n)(z + n - 1)\dots z\Gamma(z) , \quad (7.28)$$

¹² $\Gamma(z)$ est analytique $\forall z, \Re z > 0$. Pour les valeurs entières *positives* de z , le numérateur et le dénominateur de l'expression (7.24) s'annulent ensemble, les zéros du dénominateur étant pour $\Re z > 0$ des singularités *éliminables*.

¹³Après tout, $F(z)$ pourrait aussi s'annuler pour les entiers négatifs, ouvrant la possibilité que ceux-ci soient aussi des singularités éliminables.

soit :

$$\Gamma(z) = \frac{\Gamma(z+n+1)}{(z+n)(z+n-1)(z+n-2)\dots z} . \quad (7.29)$$

Dans le voisinage de 1, Γ a un développement de Taylor ; avec $|z+n| < \varepsilon$ ($n \in \mathbb{N}$), on a $\Gamma(1+z+n) = \Gamma(1) + c_1(z+n) + \dots$ où $c_1 = \Gamma'(1)$; par ailleurs, le dénominateur est $(z+n)[(-1)(-2)\dots(-n) + \mathcal{O}(z+n)]$, d'où :

$$z \simeq -n : \Gamma(z) = \frac{\Gamma(1)}{(z+n)(-1)(-2)\dots(-n)} [1 + \mathcal{O}((z+n))] , \quad (7.30)$$

c'est-à-dire :

$$\boxed{z \simeq -n : \Gamma(z) = \frac{(-1)^n}{n!} \frac{1}{z+n} + \mathcal{O}((z+n)^0)} \quad (7.31)$$

Ceci n'est rien d'autre que le début du développement de Laurent autour de $-n$, montrant que *tout* entier négatif ou nul est bel et bien une singularité, qu'il s'agit d'un pôle d'ordre un, et que le résidu est :

$$\boxed{c_{-1, -n} \stackrel{\text{d\'ef}}{=} \text{Res} [\Gamma(z), z = -n] = \frac{(-1)^n}{n!}} \quad (7.32)$$

(à titre mnémotechnique, on pourra retenir que le résidu de Γ coïncide avec le coefficient du développement en série de e^{-z}).

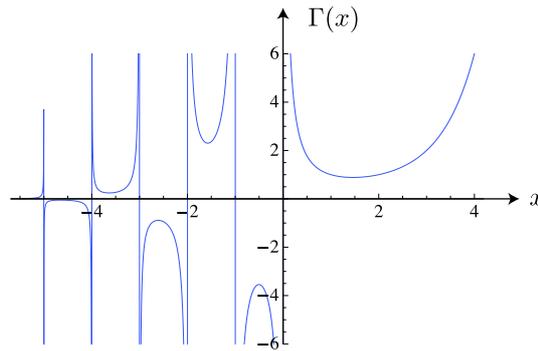


Figure 7.3: Variations de $\Gamma(z)$ pour z réel.

Le graphe de la fonction $\Gamma(z = x \in \mathbb{R})$ est tracé sur la figure 7.3. Son allure peut se comprendre qualitativement. Pour $x > 0$, la dérivée seconde $\Gamma''(x)$ est positive (voir (7.11)) ; $\Gamma(x)$ est donc une fonction convexe, qui est soit monotone, soit décroissante puis croissante. Comme $\Gamma(1) = \Gamma(2) = 1$, c'est le deuxième cas qui est réalisé. Pour $x < 0$, il est facile de montrer ([12] p. 353) que le produit $\Gamma(x)\Gamma''(x)$ est positif¹⁴ : $\Gamma(x)$

¹⁴En utilisant (7.12) et l'inégalité de Schwarz, on peut montrer plus précisément que :

$$\boxed{\Gamma''(x)\Gamma(x) - \Gamma'(x)^2 \geq 0} \quad (7.33)$$

est convexe quand elle est positive, concave quand elle est négative. Par ailleurs, $\Gamma(x)$ diverge en $x = -n$ ($n \in \mathbb{N}$), et se comporte comme $\frac{(-1)^n}{n!} \frac{1}{x+n}$ près de ces valeurs (voir (7.32)). Enfin, on établira plus tard que $\Gamma(z)$ ne s'annule pas dans \mathbb{C} (voir p. 297).

La fig. 7.4 donne une idée de $\Gamma(z)$ pour z complexe.

◆ *Remarque*

Au chapitre précédent, lors de la sommation de séries (et à propos de la fonction cot z), on a rencontré un exemple d'application du théorème de Mittag - Leffler¹⁵, permettant de développer sur ses pôles une fonction méromorphe satisfaisant la condition $\lim_{|z| \rightarrow +\infty} |f(z)| \rightarrow 0$, uniformément quel que soit l'argument de z ; un tel développement n'existe pas pour la fonction $\Gamma(z)$, qui ne possède pas la propriété requise.

□ ? Montrer que $\sum_{n \in \mathbb{N}} \frac{(-1)^n}{n!} \frac{1}{z+n} = \int_1^{+\infty} t^{z-1} e^{-t} dt$ (une autre façon de se convaincre que le premier membre, qui évoque un développement de Mittag-Leffler, n'est pas égal à $\Gamma(z)$). ◆

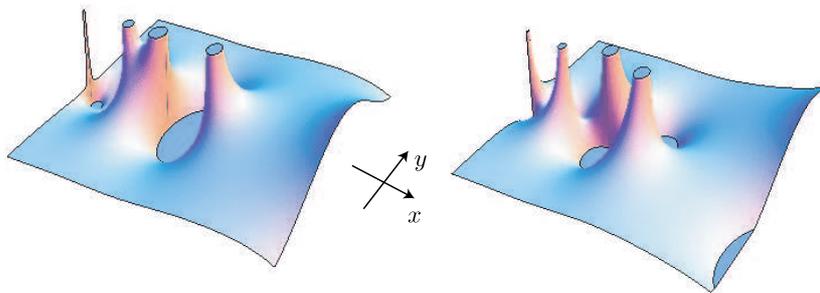


Figure 7.4: Partie réelle (à gauche) et imaginaire (à droite) de $\Gamma(z)$ pour $z = x + iy$.

□ ? Soit C le contour formé du demi-axe réel positif et du demi-axe imaginaire $\Im z > 0$, reliés par un grand quart de cercle à l'infini et par un petit quart de cercle contournant l'origine ; combien vaut l'intégrale $\int_C t^{z-1} e^{-t} dt$ ($0 < \Re z < 1$) ? En déduire que :

$$\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt = \int_0^{+\infty} (it)^{z-1} e^{-it} dt = e^{i\frac{\pi z}{2}} \int_0^{+\infty} t^{z-1} e^{-it} dt . \quad (7.34)$$

¹⁵En essence, rappelons que ce théorème se démontre comme suit. Par $\lim_{|z| \rightarrow +\infty} |f(z)| \rightarrow 0$, l'intégrale $\int_{C_\infty} \frac{f(\xi)}{\xi-z} d\xi$ est nulle, où C_∞ est une boucle à l'infini ; mais, par le théorème des résidus, cette intégrale vaut aussi $2i\pi \left[f(z) + \sum_k \frac{1}{z_k - z} \text{Res}[f, z_k] \right]$, où la somme se fait sur tous les résidus aux singularités de f . On en déduit le développement de $f(z)$ sur ses pôles : $f(z) = \sum_k \frac{1}{z-z_k} \text{Res}[f, z_k]$.

□ ? Appliquer ceci pour trouver le développement de $f(z) \stackrel{\text{déf}}{=} \frac{1}{z} \cot z$ (qui a un pôle d'ordre 2 en $z = 0$), et comparer au résultat obtenu au chapitre 6.

Posant $z = \frac{1}{n}$, $n > 1$, et $t^{\frac{1}{n}} = \xi$, en déduire que :

$$\int_0^{+\infty} t^{\frac{1}{n}-1} e^{-it} dt = n \int_0^{+\infty} e^{-i\xi^n} d\xi = e^{-i\frac{\pi}{2n}} \Gamma\left(\frac{1}{2n}\right) \quad (7.35)$$

et trouver les intégrales généralisant celles de Fresnel :

$$\int_0^{+\infty} \cos x^n dx = \frac{1}{n} \Gamma\left(\frac{1}{n}\right) \cos \frac{\pi}{2n}, \quad \int_0^{+\infty} \sin x^n dx = \frac{1}{n} \Gamma\left(\frac{1}{n}\right) \sin \frac{\pi}{2n} \quad (7.36)$$

que l'on a rencontrées au chapitre 6 dans le cas particulier $n = 2$.

7.1.2 La fonction Beta $B(p, q)$

À la fonction Γ d'Euler est associée une autre fonction importante, la fonction¹⁶ B . Dans les notations traditionnelles, sa définition est :

$$B(p, q) \stackrel{\text{déf}}{=} \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \quad (p, q \in \mathbb{C}) \quad (7.37)$$

Cette fonction admet diverses représentations intégrales, que l'on va maintenant établir. Quand p et q ont des parties réelles strictement positives, la définition (7.6) est utilisable ; en y faisant le changement de variable $t = x^2$, il vient¹⁷ :

$$\Gamma(p) \stackrel{\text{déf}}{=} \int_0^{+\infty} t^{p-1} e^{-t} dt = 2 \int_0^{+\infty} x^{2p-1} e^{-x^2} dx. \quad (7.39)$$

De même, $\Gamma(q) = 2 \int_0^{+\infty} y^{2q-1} e^{-y^2} dy$, de sorte que :

$$\Gamma(p)\Gamma(q) = 4 \int_{\mathbb{R}_+^2} x^{2p-1} y^{2q-1} e^{-(x^2+y^2)} dx dy \quad (7.40)$$

où \mathbb{R}_+^2 désigne le quadrant positif $x > 0$, $y > 0$ de \mathbb{R}^2 . En posant maintenant $x = r \cos \theta$, $y = r \sin \theta$, d'où $dx dy = r dr d\theta$, il vient :

$$\Gamma(p)\Gamma(q) = 4 \int_0^{+\infty} r^{2(p+q)-1} e^{-r^2} dr \int_0^{\pi/2} d\theta \cos^{2p-1} \theta \sin^{2q-1} \theta ; \quad (7.41)$$

utilisant (7.39), on voit que la première intégrale vaut $\frac{1}{2}\Gamma(p+q)$, d'où :

$$\Gamma(p)\Gamma(q) = 2\Gamma(p+q) \int_0^{\pi/2} \cos^{2p-1} \theta \sin^{2q-1} \theta d\theta. \quad (7.42)$$

¹⁶ B = bêta majuscule, appelée fonction d'Euler de première espèce.

¹⁷Au passage, on note une relation utile pour la suite :

$$\int_0^{+\infty} x^{a-1} e^{-x^2} dx = \frac{1}{2} \Gamma\left(\frac{a}{2}\right) \quad (\Re a > 0) \quad (7.38)$$

en particulier, $\Gamma(\frac{1}{2})$ est égal à l'intégrale gaussienne $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$.

Compte tenu de la définition (7.37), on trouve ainsi la représentation intégrale¹⁸ suivante de la fonction B :

$$B(p, q) = 2 \int_0^{\pi/2} \cos^{2p-1} \theta \sin^{2q-1} \theta \, d\theta \quad (\Re p > 0, \Re q > 0) \quad (7.43)$$

L'intégrale de Wallis $W(r)$ est un cas particulier :

$$W(r) \stackrel{\text{déf}}{=} \int_0^{\pi/2} \cos^r \theta \, d\theta = \int_0^{\pi/2} \sin^r \theta \, d\theta = \frac{1}{2} B\left(\frac{r+1}{2}, \frac{1}{2}\right) \quad (7.44)$$

En faisant le changement de variable $\cos^2 \theta = t$, on obtient une autre représentation intégrale :

$$B(p, q) = \int_0^1 t^{p-1} (1-t)^{q-1} \, dt = \int_0^1 (1-t)^{p-1} t^{q-1} \, dt \quad (\Re p > 0, \Re q > 0) \quad (7.45)$$

Ces relations supposent $\Re p > 0$ et $\Re q > 0$ (elles proviennent de (7.6)), et ceci se voit tout naturellement sur les intégrales (7.45) elles-mêmes, mais comme $\Gamma(z)$ est prolongeable à gauche, la définition première (7.37) montre que $B(p, q)$ est analytique partout où le second membre l'est. Les singularités de $B(p, q)$ se déduisent de celles de $\Gamma(z)$. Posant $p = -n + \xi$ ($n \in \mathbb{N}$) et compte tenu de (7.31) :

$$B(p = -n + \xi, q) = \frac{\Gamma(-n + \xi)\Gamma(q)}{\Gamma(-n + \xi + q)} = \left[\frac{(-1)^n}{n!} \frac{1}{p+n} + \mathcal{O}(\xi^0) \right] \frac{\Gamma(q)}{\Gamma(-n + \xi + q)}. \quad (7.46)$$

Tant que q n'est pas un entier négatif, le facteur de droite est analytique, d'où l'amorce du développement de Laurent de B autour de l'entier négatif $-n$:

$$B(p = -n + \xi, q) = \frac{(-1)^n}{n!} \frac{\Gamma(q)}{\Gamma(-n + q)} \frac{1}{p+n} + \mathcal{O}(\xi^0). \quad (7.47)$$

B a donc les mêmes pôles que Γ , le résidu au pôle simple $-n$ étant $\frac{(-1)^n}{n!} \frac{\Gamma(q)}{\Gamma(q-n)}$, soit $\frac{(-1)^n}{n!} (q-1)(q-2) \dots (q-n)$. Ce résidu s'annule si $q = 1, 2, \dots, n$, et c'est bien normal ; en effet, si $p = -n + \xi$, $\Gamma(p)$ se comporte comme $\frac{c-n}{\xi}$, $\Gamma(q) = (q-1)!$, $\Gamma(p+q) = \Gamma(-n+q+\xi)$ se comporte comme $\frac{c-(n-q)}{\xi}$; au total B a une limite finie, la singularité en $p = -n$ devenant éliminable dans ces conditions.

Formule des compléments

L'expression (7.45) permet de démontrer l'importante relation dite *formule dite des compléments* :

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z} \quad (7.48)$$

¹⁸ B admet bien d'autres représentations intégrales.

En effet, de (7.37), (7.45) et de $\Gamma(1) = 1$, on déduit $B(z, 1 - z) = \Gamma(z)\Gamma(1 - z)$ soit :

$$B(z, 1 - z) = \int_0^1 t^{z-1}(1-t)^{-z} dz = \int_0^{+\infty} \frac{u^{z-1}}{1+u} du \quad (0 < \Re z < 1) , \quad (7.49)$$

où l'intégrale de droite vient du changement de variable $t = \frac{u}{1+u}$. Cette dernière intégrale (d'ailleurs déjà rencontrée, voir chapitre 6, éq. (6.147)) se calcule par résidus en prenant le contour habituel (qui longe les deux bords de la coupure \mathbb{R}_+ et se ferme par une grande boucle, un grand cercle si on préfère). Le seul pôle est en $u = -1 = e^{i\pi}$, le résidu vaut $(e^{i\pi})^{z-1} = -e^{i\pi z}$, d'où¹⁹ :

$$(1 - e^{2i\pi z}) \int_0^{+\infty} \frac{u^{z-1}}{1+u} du = -2i\pi e^{i\pi z} \iff \int_0^{+\infty} \frac{u^{z-1}}{1+u} du = \frac{\pi}{\sin \pi z} . \quad (7.50)$$

De ceci et de (7.49), on tire la formule des compléments (7.48). Pour le moment, cette formule n'est démontrée que pour $0 < \Re z < 1$; si on écrit le premier membre de (7.48) non plus avec z mais avec $z + 1$, il devient :

$$\Gamma(z + 1)\Gamma(-z) = z\Gamma(z)\Gamma(-z) = -(-z)\Gamma(-z)\Gamma(z) = -\Gamma(-z + 1)\Gamma(z) , \quad (7.51)$$

et redonne le premier membre de (7.48) simplement changé de signe. Il en va de même pour le second puisque $\sin[\pi(z + 1)] = -\sin \pi z$. Si donc les deux membres de (7.48) sont égaux pour $0 < \Re z < 1$, ils le sont encore pour $1 < \Re z < 2$, et ainsi de suite. En écrivant cette même formule des compléments pour $z - 1$, la même gymnastique montre qu'elle est encore vraie pour $-1 < \Re z < 0$, etc. Au total, cette formule remarquable tient $\forall z$ où elle a un sens, soit pour $z \notin \mathbb{Z}$: les pôles du premier membre sont $z = -n$ ($n \in \mathbb{N}$) et $1 - z = -n'$, $n' \in \mathbb{N}$, soit tous les entiers positifs et négatifs, tout comme ceux du second membre. Notons que la formule (7.48) s'écrit aussi $B(z)B(1 - z) = \frac{\pi}{\sin \pi z}$.

En faisant $z = \frac{1}{2}$ dans (7.48), on trouve $[\Gamma(\frac{1}{2})]^2 = \pi$, d'où, $\Gamma(\frac{1}{2})$ étant visiblement donnée par une intégrale positive :

$$\boxed{\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}} \quad (7.52)$$

Ce résultat donne la valeur de l'intégrale gaussienne $I_{\text{Gauss}} = \int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$ puisque, d'après (7.39), $\Gamma(p = \frac{1}{2}) = I_{\text{Gauss}}$. À partir de ce résultat, la relation fonctionnelle (7.26) permet d'obtenir la valeur de $\Gamma(z)$ pour toutes les valeurs demi-entières positives :

$$\Gamma\left(n + \frac{1}{2}\right) = \left(n - \frac{1}{2}\right)\left(n - \frac{3}{2}\right) \dots \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{1.3.5. \dots (2n - 1)}{2^n} \sqrt{\pi} , \quad (7.53)$$

soit :

$$\boxed{\Gamma\left(n + \frac{1}{2}\right) = \frac{(2n)!}{2^{2n}n!} \sqrt{\pi} \quad (n \in \mathbb{N})} \quad (7.54)$$

¹⁹Noter aussi que, posant $u = e^t$, l'intégrale de droite dans (7.49) est $\int_{-\infty}^{+\infty} \frac{e^{zt}}{e^t + 1} dt$: c'est l'intégrale d'Euler rencontrée au chapitre 6 (voir (6.105)).

Utilisant à nouveau la formule des compléments (7.48), cette expression permet d'obtenir Γ pour les valeurs demi-entières négatives :

$$\boxed{\Gamma\left(-n + \frac{1}{2}\right) = (-1)^n \frac{4^n n!}{(2n)!} \sqrt{\pi} \quad (n \in \mathbb{N})} \quad (7.55)$$

Rassemblant ces deux résultats, on trouve :

$$\boxed{\Gamma\left(n + \frac{1}{2}\right)\Gamma\left(-n + \frac{1}{2}\right) = (-1)^n \pi} \quad (7.56)$$

La formule des compléments permet d'ailleurs d'obtenir une autre représentation intégrale de la fonction $\Gamma(z)$, ou plutôt de $\frac{1}{\Gamma(z)}$. En effet, à l'aide de l'expression (7.25), on a :

$$\frac{1}{\Gamma(z)} = \frac{\sin \pi z}{\pi} \Gamma(1-z) = \frac{\sin \pi z}{\pi [e^{2i\pi(1-z)} - 1]} \int_C t^{1-z-1} e^{-t} dt = \frac{ie^{i\pi z}}{2\pi} \int_C t^{-z} e^{-t} dt. \quad (7.57)$$

Faisant le changement de variable $t' = -t = e^{i\pi}t$, on a $t^{-z} = (e^{i\pi}t')^{-z} = e^{-i\pi z}t'^{-z}$, $dt = -dt'$ et le contour est changé en C' (voir fig. 7.5) :

$$\boxed{\frac{1}{\Gamma(z)} = \frac{1}{2i\pi} \int_{C'} t'^{-z} e^{-t'} dt'} \quad (7.58)$$

qui constitue la représentation de Hankel de la fonction $\frac{1}{\Gamma(z)}$.

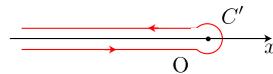


Figure 7.5: Contour C' utilisé pour la représentation de Hankel de la fonction $\frac{1}{\Gamma(z)}$ (voir (7.58)).

La formule des compléments permet aussi de montrer que $\Gamma(z)$ ne s'annule pour aucune valeur de z . En effet, partons de :

$$\Gamma(1-z) = \frac{\pi}{\sin \pi z} \frac{1}{\Gamma(z)} ; \quad (7.59)$$

Dire que $\Gamma(z)$ ne s'annule jamais revient à dire que $\Gamma(1-z)$ ne s'annule pas non plus, donc que le second membre de (7.59) n'est jamais nul, autrement dit que le produit $\pi^{-1} \sin \pi z \Gamma(z)$ n'est jamais infini. Comme $|\Gamma(z)|$ ne diverge que pour les valeurs $z \in -\mathbb{N}$, il suffit de trouver la limite de ce produit lorsque z tend vers un entier négatif ou nul $-n$. Compte tenu de la valeur du résidu $c_{-1, -n}$ (voir (7.32)), cette limite est égale à $\frac{1}{n!}$, et est donc finie.

Surface de la sphère dans \mathbb{R}^D

La fonction Γ (et/ou B) est également reliée à une quantité apparaissant ici et là en Physique²⁰ : la surface de la sphère dans l'espace à D dimensions \mathbb{R}^D . Pour montrer comment le calcul fonctionne et l'articulation des idées, commençons par traiter le cas $D = 3$.

On part de la relation (7.39), écrite pour $p = \frac{1}{2}$:

$$\Gamma\left(\frac{1}{2}\right) = 2 \int_0^{+\infty} e^{-x^2} dx = \int_{-\infty}^{+\infty} e^{-x^2} dx , \quad (7.60)$$

ce qui permet d'écrire :

$$\left[\Gamma\left(\frac{1}{2}\right)\right]^3 = \prod_{i=1}^3 \int_{-\infty}^{+\infty} e^{-x_i^2} dx_i \equiv \int_{\mathbb{R}^3} e^{-r^2} d^3r , \quad (7.61)$$

où on a bien pris soin de distinguer les variables muettes d'intégration pour écrire le cube de l'intégrale. Passant en coordonnées polaires pour effectuer l'intégration²¹, il vient :

$$\left[\Gamma\left(\frac{1}{2}\right)\right]^3 = \int_0^{+\infty} e^{-r^2} r^2 dr \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi \equiv \int_0^{+\infty} e^{-r^2} r^2 dr \int_{\text{sphère unité}} dS , \quad (7.62)$$

où dS note l'élément de surface de la sphère de rayon unité. En notant Ω_D la surface de cette sphère (c'est la valeur de l'angle solide dans \mathbb{R}^D), il vient donc (compte tenu de (7.38) et de (7.52)) :

$$\left[\Gamma\left(\frac{1}{2}\right)\right]^3 = \int_0^{+\infty} e^{-r^2} r^2 dr \Omega_3 = \frac{1}{2} \Gamma\left(\frac{3}{2}\right) \Omega_3 , \quad (7.63)$$

où on a utilisé (7.39) avec $2p - 1 = 2$, soit $p = \frac{3}{2}$, d'où :

$$\Omega_3 = \frac{\left[\Gamma\left(\frac{1}{2}\right)\right]^3}{\frac{1}{2} \Gamma\left(\frac{3}{2}\right)} = 2 \frac{\left[\Gamma\left(\frac{1}{2}\right)\right]^3}{\Gamma\left(\frac{1}{2} + 1\right)} = 2 \frac{\left[\Gamma\left(\frac{1}{2}\right)\right]^3}{\frac{1}{2} \Gamma\left(\frac{1}{2}\right)} = 4 \left[\Gamma\left(\frac{1}{2}\right)\right]^2 = 4\pi . \quad (7.64)$$

On retrouve ainsi la surface S_3 de la sphère de rayon R dans \mathbb{R}^3 , égale à $4\pi R^3^{-1} = 4\pi R^2$.

Le calcul se généralise sans peine à D quelconque. On écrit maintenant comme en (7.61) :

$$\left[\Gamma\left(\frac{1}{2}\right)\right]^D = \int_{\mathbb{R}^D} e^{-(x_1^2+x_2^2+\dots+x_D^2)} dx_1 dx_2 \dots dx_D , \quad (7.65)$$

de sorte que (7.62) devient :

$$\left[\Gamma\left(\frac{1}{2}\right)\right]^D = \int_0^{+\infty} e^{-r^2} r^{D-1} dr \Omega_D \equiv \frac{1}{2} \Gamma\left(\frac{D}{2}\right) \Omega_D \iff \Omega_D = 2 \frac{\left[\Gamma\left(\frac{1}{2}\right)\right]^D}{\Gamma\left(\frac{D}{2}\right)} , \quad (7.66)$$

²⁰Par exemple, s'agissant de calculer le nombre d'états d'énergie comprise entre E et $E + \delta E$ pour un gaz parfait de N particules classiques libres dans \mathbb{R}^3 , il faut connaître la surface de la sphère dans \mathbb{R}^{3N} , d'équation $\sum_{i=1}^N \frac{\vec{p}_i^2}{2m} = E$.

²¹ $x = r \cos \phi \sin \theta$, $y = r \sin \phi \sin \theta$, $z = r \cos \theta$, $d^3r = r^2 dr \sin \theta d\theta d\phi$.

d'où le résultat cherché²² :

$$\boxed{\Omega_D = \frac{2\pi^{D/2}}{\Gamma(\frac{D}{2})} \quad S_D = \frac{2\pi^{D/2}}{\Gamma(\frac{D}{2})} R^{D-1}} . \quad (7.67)$$

En particulier :

$$S_1 = 2R , \quad S_2 = 2\pi R , \quad S_3 = 4\pi R^2 , \quad S_4 = 2\pi^2 R^3 , \quad S_5 = \frac{8}{3}\pi^2 R^4 , \dots \quad (7.68)$$

En faisant $R = 1$ dans ces formules, on obtient l'angle solide $\Omega_{D=1,2,3,\dots}$ pour les dimensions d'espace "ordinaire". Notons que l'expression de ces éléments géométriques à l'aide de fonctions définies pour D quelconque (réel, pas forcément entier positif, et même complexe) permet de traiter des problèmes en dimension quelconque, pas forcément entière (ce que font parfois les physiciens, exemple : phénomènes critiques).

7.2 La fonction $\psi(z)$ d'Euler

Aussi appelée fonction digamma, la fonction $\psi(z)$ est la dérivée logarithmique de $\Gamma(z)$:

$$\boxed{\psi(z) \stackrel{\text{d'ef}}{=} \frac{\Gamma'(z)}{\Gamma(z)}} \quad (7.69)$$

Comme $\Gamma(z)$ ne s'annule pas, et que les pôles de $\Gamma'(z)$ sont les mêmes que ceux de $\Gamma(z)$, $\psi(z)$ est une fonction méromorphe ayant des pôles simples en $-n$ ($n \in \mathbb{N}$), avec un résidu égal à -1 .

Selon la définition de la fonction $B(p, q)$ et sa représentation intégrale (7.45), on peut écrire :

$$\frac{\Gamma(z-h)\Gamma(h)}{\Gamma(z)} = B(z-h, h) = \int_0^1 (1-t)^{z-h-1} t^{h-1} dt \quad (7.70)$$

avec $\Re h > 0$ et $\Re(z-h) > 0$. En écrivant $(1-t)^{z-h-1} = [(1-t)^{z-h-1} - 1] + 1$ et en intégrant le terme additif t^{h-1} , il vient :

$$\frac{\Gamma(z-h)\Gamma(h)}{\Gamma(z)} = \frac{1}{h} + \int_0^1 [(1-t)^{z-h-1} - 1] t^{h-1} dt . \quad (7.71)$$

En introduisant les développements :

$$(1-t)^{z-h-1} = (1-t)^{z-1} e^{-h \ln(1-t)} = (1-t)^{z-1} [1 - h \ln(1-t) + \dots] , \quad (7.72)$$

$$t^{h-1} = t^{-1} e^{h \ln t} = t^{-1} [1 + h \ln t + \dots] , \quad (7.73)$$

²²Dans \mathbb{R}^D , les volumes sont homogènes à L^D et les surfaces à L^{D-1} .

on voit que l'intégrale est $\int_0^1 [(1-t)^{z-1} - 1] t^{-1} dt + \mathcal{O}(h)$, d'où :

$$\frac{\Gamma(z-h)\Gamma(h)}{\Gamma(z)} = \frac{1}{h} + \int_0^1 [(1-t)^{z-1} - 1] t^{-1} dt + \mathcal{O}(h) . \quad (7.74)$$

Par ailleurs, le premier membre de (7.71) est $\frac{1}{\Gamma(z)}[\Gamma(z) - h\Gamma'(z) + \dots][\frac{1}{h} + c_0 + c_1h + \dots]$, compte tenu des propriétés analytiques de $\Gamma(z)$ établies plus haut. D'où :

$$\frac{1}{\Gamma(z)}[\Gamma(z) - h\Gamma'(z) + \dots][\frac{1}{h} + c_0 + c_1h + \dots] = \frac{1}{h} + \int_0^1 [(1-t)^{z-1} - 1] t^{-1} dt + \mathcal{O}(h) . \quad (7.75)$$

Les termes en h^{-1} disparaissent entre les deux membres ; le terme en h^0 donne l'égalité :

$$\frac{\Gamma'(z)}{\Gamma(z)} = \int_0^1 [1 - (1-t)^{z-1}] t^{-1} dt - c_0 . \quad (7.76)$$

En notant que $\frac{1}{t}$ est la somme de la série géométrique $\sum_{n \in \mathbb{N}} (1-t)^n$, l'intégrale s'écrit $\int_0^1 \sum_{n=0}^{+\infty} [(1-t)^n - (1-t)^{n+z-1}] dt$; profitant de la convergence uniforme de la série, on peut maintenant écrire :

$$\frac{\Gamma'(z)}{\Gamma(z)} = \sum_{n=0}^{+\infty} \left(\frac{1}{n+1} - \frac{1}{n+z} \right) - c_0 \quad (7.77)$$

ou encore, après remaniement du second membre :

$$\frac{\Gamma'(z)}{\Gamma(z)} + \frac{1}{z} = \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+z} \right) + 1 - c_0 . \quad (7.78)$$

En faisant $z = 1$, il vient $\psi(1) + 1 = 1 - c_0$, d'où le développement de $\psi(z)$:

$$\boxed{\psi(z) = \psi(1) - \frac{1}{z} + \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+z} \right)} . \quad (7.79)$$

On peut montrer que $\psi(1) = -\gamma$, où²³ γ est la constante d'Euler définie par :

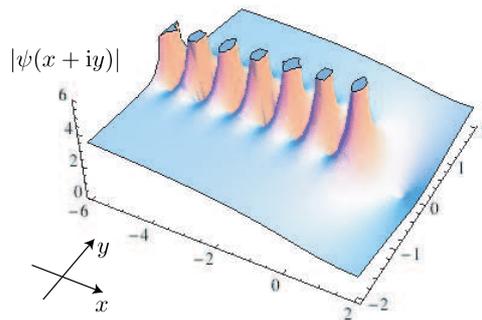
$$\boxed{\gamma \stackrel{\text{d\'ef}}{=} \lim_{N \rightarrow +\infty} \left(\sum_{n=1}^N \frac{1}{n} - \ln N \right) = 0.577\,215\,664\,901\,532\,860\dots} \quad (7.80)$$

La fonction digamma est visiblement reliée aux nombres harmoniques $H_{N,1}$ rencontrés au chapitre 1 à propos de la loi de Zipf (voir (1.164)) ; on a :

$$\psi(N) = H_{N-1,1} - \gamma . \quad (7.81)$$

²³ γ est aussi dite *constante d'Euler - Mascheroni*. Attention à la notation : dans certains ouvrages (par exemple [36]) ce qui est noté ici γ est noté là-bas \mathbf{C} .

Les mêmes auteurs posent ... $\gamma = e^{\mathbf{C}} = 1.781\,072\,417\,990\,197\,985\dots$

Figure 7.6: Graphe du module de la fonction digamma $\psi(x + iy)$.

En dérivant la relation (7.26), on établit sans peine que la fonction digamma satisfait l'équation fonctionnelle :

$$\boxed{\psi(z + 1) = \psi(z) + \frac{1}{z}} \quad (7.82)$$

De même, la dérivation de la formule des compléments (7.48) donne :

$$\boxed{\psi(1 - z) = \psi(z) + \pi \cot \pi z} \quad (7.83)$$

Par ailleurs, partant de (7.79), écrivant $\frac{1}{n} - \frac{1}{z+n} = \frac{1}{n} \left(1 - \frac{1}{1+\frac{z}{n}}\right)$, et en développant la dernière fraction en série entière ($|z| < 1$), on obtient :

$$\psi(z) = -\frac{1}{z} - \gamma + \sum_{n=1}^{+\infty} \frac{1}{n} \sum_{m=1}^{+\infty} (-1)^{m+1} \left(\frac{z}{n}\right)^m. \quad (7.84)$$

Il apparaît une série à double indice de terme général $u_{n,m} \stackrel{\text{déf}}{=} (-1)^{m+1} \frac{z^m}{n^{m+1}}$. Si la série totale est absolument convergente, on peut échanger l'ordre des sommations ; or $|u_{n,m}| = \frac{|z|^m}{n^{m+1}} < \frac{|z|^m}{n^2}$: ceci est le terme général de la double série égale au produit de $\sum_n \frac{1}{n^2}$, qui converge, et de $\sum_m |z|^m$, qui converge (uniformément) dans le disque $|z| = 1$. La double série étant absolument convergente, l'interversion des sommations donne :

$$\boxed{\psi(z) = -\frac{1}{z} - \gamma + \sum_{m=1}^{+\infty} (-1)^{m+1} \zeta(m+1) z^m \quad (|z| < 1)} \quad (7.85)$$

où $\zeta(m) \stackrel{\text{déf}}{=} \sum_{n \in \mathbb{N}^*} n^{-m}$ est la fonction de Riemann présentée dans la section suivante.

Occasionnellement, il est utile de définir la fonction *trigamma*, notée $\psi_1(z)$, définie comme la dérivée seconde de $\ln \Gamma(z)$:

$$\psi_1(z) \stackrel{\text{déf}}{=} \frac{d^2}{dz^2} \ln \Gamma(z) = \psi'(z) = \sum_{m=0}^{+\infty} \frac{1}{(z+n)^2} ; \quad (7.86)$$

$\psi_1(z)$ satisfait :

$$\psi_1(z+1) = \psi_1(z) - \frac{1}{z^2} . \quad (7.87)$$

Les fonctions $\psi_{n-1} \stackrel{\text{déf}}{=} \frac{d^n}{dz^n} \ln \Gamma(z)$ ainsi définies constituent la famille des *fonctions polygamma* ($\psi_0 \equiv \psi$). On a d'une façon générale :

$$\psi_m(z) = (-1)^{m+1} \int_0^{+\infty} \frac{t^m e^{-zt}}{1-e^{-t}} dt , \quad \psi_m(z) = (-1)^{m+1} m! \sum_{n=0}^{+\infty} \frac{1}{(z+n)^{m+1}} . \quad (7.88)$$

et :

$$\psi_m(z+1) = \psi_m(z) + (-1)^m \frac{m!}{z^{m+1}} . \quad (7.89)$$

Ces fonctions satisfont une relation de *multiplication* tout à fait remarquable ($m \geq 1$) :

$$k^{m+1} \psi_m(kz) = \sum_{n=0}^{k-1} \psi_m\left(z + \frac{n}{k}\right) ; \quad (7.90)$$

pour $m = 0$, on a :

$$k[\psi(kz) - \ln k] = \sum_{n=0}^{k-1} \psi\left(z + \frac{n}{k}\right) . \quad (7.91)$$

7.3 La fonction $\zeta(z)$ de Riemann

Une première définition [37] de $\zeta(z)$ est en effet²⁴ :

$$\boxed{\zeta(z) = \sum_{n=1}^{+\infty} \frac{1}{n^z} \quad (\Re z > 1)} \quad (7.94)$$

et c'est sous cette forme qu'elle apparaît souvent en Physique (rayonnement du corps noir, gaz de Fermi dégénéré, condensation de Bose, modèle d'Ising en dimension D traité par la méthode de Bethe - Peierls, etc). La série (7.94) est bien convergente pour $\Re z > 1$

²⁴On appelle *série de Dirichlet* toute série de la forme $\sum_{n \in \mathbb{N}} \frac{a_n}{n^z}$; on peut dire que la fonction $\zeta(z)$ ainsi définie est la plus simple des séries de Dirichlet. Une autre série de Dirichlet est fameuse :

$$\sum_{n=1}^{+\infty} \frac{\mu(n)}{n^z} \quad (\Re z > 1) \quad (7.92)$$

où $\mu(n)$ est la fonction de Möbius :

$$\mu(n) = \begin{cases} 1 & \text{si } n = 1 \\ (-1)^k & \text{si } n \text{ est le produit de } k \text{ nombres premiers distincts} \\ 0 & \text{si } n \text{ contient un facteur au carré} \end{cases} . \quad (7.93)$$

La somme de la série (7.92) est égale à ... $\frac{1}{\zeta(z)}$.

puisque $|n^{x+iy}| = |n^x e^{iy \ln n}| = |n^x|$. Par ailleurs, soit l'intégrale $I(z) \stackrel{\text{d\u00e9f}}{=} \int_0^{+\infty} \frac{t^{z-1}}{e^t-1} dt$; en \u00e9crivant $(e^t - 1)^{-1} = e^{-t} \sum_{n=0}^{+\infty} e^{-nt}$:

$$I(z) = \sum_{n=0}^{+\infty} \int_0^{+\infty} t^{z-1} e^{-(n+1)t} dt = \sum_{n=0}^{+\infty} (n+1)^{-z} \int_0^{+\infty} u^{z-1} e^{-u} du ; \quad (7.95)$$

on reconna\u00eet $\Gamma(z)$ \u00e0 droite, d'o\u00f9 une autre expression²⁵ de $\zeta(z)$:

$$\zeta(z) = \frac{1}{\Gamma(z)} \int_0^{+\infty} \frac{t^{z-1}}{e^t-1} dt \quad (\Re z > 1) . \quad (7.96)$$

Cette repr\u00e9sentation se pr\u00eate au prolongement analytique de $\zeta(z)$ \u00e0 gauche de la droite $\Re z = 1$, en utilisant la m\u00eame argumentation que celle utilis\u00e9e pour prolonger $\Gamma(z)$. On d\u00e9finit $F(z) \stackrel{\text{d\u00e9f}}{=} \int_C \frac{t^{z-1}}{e^t-1} dt$, o\u00f9 le contour C est le m\u00eame que celui de la figure 7.1, mais doit \u00e9viter les z\u00e9ros du d\u00e9nominateur, soit tous les points $\xi_k = 2ik\pi$, $k \in \mathbb{Z}$; comme auparavant on voit que $F(z)$ est une fonction enti\u00e8re. En d\u00e9formant le contour, on \u00e9tablit la relation $F(z) = (e^{2i\pi z} - 1)\Gamma(z)\zeta(z)$ qui, lue \u00e0 l'envers et parce que $F(z)$ est une fonction enti\u00e8re, fournit le prolongement analytique de $\zeta(z)$:

$$\zeta(z) = \frac{1}{(e^{2i\pi z} - 1)\Gamma(z)} \int_C \frac{\xi^{z-1}}{e^\xi - 1} d\xi \quad (7.97)$$

En \u00e9crivant $e^{-i\pi z} \xi^{z-1} = -(-\xi)^{z-1}$, et en utilisant la formule des compl\u00e9ments (7.48), le prolongement (7.97) prend la forme traditionnelle :

$$\boxed{\zeta(z) = \frac{i}{2\pi} \Gamma(1-z) \int_C \frac{(-\xi)^{z-1}}{e^\xi - 1} d\xi \quad (z \neq 1)} \quad (7.98)$$

L'int\u00e9grale \u00e9tant une fonction enti\u00e8re, les seuls p\u00f4les possibles de $\zeta(z)$ sont ceux de $\Gamma(1-z)$, soit les p\u00f4les simples $1 - z_n = -n$, $n \in \mathbb{N}$. Comme on sait que $\zeta(z)$ est holomorphe $\forall z$, $\Re z > 1$, le seul p\u00f4le possible²⁶ est $z_0 = 1$, les autres valeurs enti\u00e8res $z_{n>0}$ \u00e9tant en fait des singularit\u00e9s apparentes. Pour $z = 1$, l'int\u00e9grale de (7.98) est $\int_C \frac{1}{e^\xi - 1} d\xi$; par r\u00e9sidus, on voit qu'elle vaut $2i\pi$. Ainsi, le r\u00e9sidu de $\zeta(z)$ au p\u00f4le simple $z_0 = 1$ est $\lim_{z \rightarrow 1} \frac{i}{2\pi} (z-1)\Gamma(1-z) \times 2i\pi$; en utilisant (7.32), on trouve :

$$\boxed{\text{Res}[\zeta(z), z = 1] = 1} \quad (7.99)$$

On conna\u00eet les valeurs explicites de $\zeta(z)$ pour certaines valeurs remarquables de z . Certaines d'entre elles s'obtiennent en introduisant les nombres de Bernoulli B_n , eux-m\u00eames associ\u00e9s aux polyn\u00f4mes de Bernoulli $B_n(x)$, d\u00e9finis \u00e0 partir de leur fonction g\u00e9n\u00e9ratrice :

$$\frac{te^{tx}}{e^t-1} \stackrel{\text{d\u00e9f}}{=} \sum_{n=0}^{+\infty} B_n(x) \frac{t^n}{n!} \quad (|t| < 2\pi) . \quad (7.100)$$

²⁵On sait que $\Gamma(z)$ ne s'annule nulle part (voir p.297), de sorte que le pr\u00e9facteur $\frac{1}{\Gamma(z)}$ n'introduit aucune singularit\u00e9.

²⁶L'existence de ce p\u00f4le explique (suivant l'argument standard) l'impossibilit\u00e9 d'une condensation de Bose en dimension $D = 2$ (ou inf\u00e9rieure) pour un gaz parfait de bosons dont le nombre est conserv\u00e9.

Ce type de développement apparaît parfois en Physique, par exemple lorsqu'il s'agit de définir une série de perturbation pour l'opérateur d'évolution en Mécanique quantique assurant l'unitarité à tout ordre fini [38]. Les nombres de Bernoulli sont, par définition²⁷, $B_n \stackrel{\text{d\'ef}}{=} B_n(0)$, soit explicitement :

$$\frac{t}{e^t - 1} \stackrel{\text{d\'ef}}{=} \sum_{n=0}^{+\infty} B_n \frac{t^n}{n!} \quad (|t| < 2\pi) ; \quad (7.102)$$

les premiers B_n sont :

$$B_0 = 1, \quad B_1 = -\frac{1}{2}, \quad B_2 = \frac{1}{6}, \quad B_4 = -\frac{1}{30}, \quad B_6 = \frac{1}{42}, \dots \quad (7.103)$$

tous les B_{2p+1} , $p \geq 1$ étant nuls²⁸. Notons d'abord que pour toutes les valeurs entières de z , le contour C introduit dans les définitions ci-dessus (par exemple (7.98)), peut être remplacé par une boucle fermée autour de l'origine, puisque pour ces valeurs de z , ξ^{z-1} n'est pas multiforme. En prenant une boucle de rayon $R < 2\pi$, on peut intégrer terme à terme le développement (7.102) ; pour tout entier $-m \leq 0$, il vient :

$$\zeta(-m) = \frac{i}{2\pi} \Gamma(1+m) \sum_{n=0}^{+\infty} \frac{B_n}{n!} (-1)^{-m-1} \int_C \xi^{n-m-2} d\xi ; \quad (7.104)$$

par le théorème des résidus, l'intégrale vaut $2i\pi\delta_{n, m+1}$; on en déduit :

$$\boxed{\zeta(-m) = \frac{(-1)^m}{m+1} B_{m+1} \quad (m \in \mathbb{N})} \quad (7.105)$$

d'où :

$$\boxed{\zeta(0) = -\frac{1}{2} \quad \zeta(-2p) = 0 \quad \zeta(-2p+1) = -\frac{B_{2p}}{2p} \quad (p \in \mathbb{N}^*)} \quad (7.106)$$

²⁷ Attention : cette définition varie d'un auteur à l'autre ; (7.100) est celle de Gradshteyn et Ryzhik [36], p. xxix. Titchmarsh ([6], p. 152) écrit :

$$\frac{1}{e^t - 1} \stackrel{\text{d\'ef}}{=} \frac{1}{t} - \frac{1}{2} + \sum_{p=1}^{+\infty} (-1)^{p-1} B_p \frac{t^{2p-1}}{(2p)!} . \quad (7.101)$$

²⁸ On vérifie sans peine que la fonction $\phi(t) \stackrel{\text{d\'ef}}{=} \frac{t}{e^t - 1} + \frac{1}{2}t$ est paire puisque la différence $\phi(t) - \phi(-t)$ est nulle.

En particulier²⁹ :

$$\boxed{\zeta(-1) = -\frac{1}{12} \quad \zeta(-3) = \frac{1}{120} \quad \zeta(-5) = -\frac{1}{252}} \quad (7.108)$$

$\zeta(z)$ s'annule donc pour $z = -2, -4, \dots$, valeurs appelées “zéros triviaux” de $\zeta(z)$, pour les distinguer des autres zéros au centre de la fameuse conjecture de Riemann (voir plus loin, p. 308).

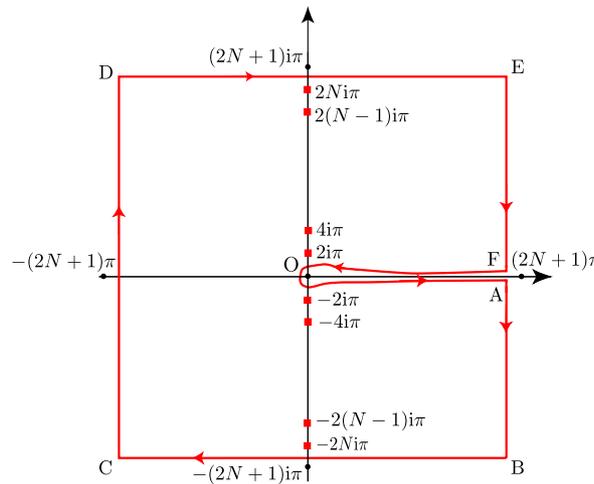


Figure 7.7: Fermeture du contour C de la figure 7.1 pour l'établissement de la relation fonctionnelle (7.112) de la fonction $\zeta(z)$.

La fonction $\zeta(z)$ satisfait une équation fonctionnelle qui peut d'ailleurs être utilisée pour effectuer le prolongement analytique. Soit le contour C_N dessiné sur la figure 7.7 ; par le théorème des résidus, on a :

$$\int_{C_N} \frac{(-\xi)^{z-1}}{e^\xi - 1} d\xi = -2i\pi \sum_{n=-N, n \neq 0}^{n=+N} \text{Res} \left[\frac{(-\xi)^{z-1}}{e^\xi - 1}, 2in\pi \right] . \quad (7.109)$$

²⁹Du Sauty ([34], p.211) évoque la perplexité de Hill à la lecture des carnets de Ramanujan, notamment lorsque celui-ci affirme avoir démontré que la somme des entiers naturels est égale à $-\frac{1}{12}$ (“*M. Ramanujan est tombé dans l'abîme du sujet, ô combien ardu, des séries divergentes*”) :

$$1 + 2 + 3 + 4 + \dots = -\frac{1}{12} \quad (7.107)$$

Sans aucune formation académique, Ramanujan avait réussi à prolonger $\zeta(z)$ en-deçà de $\Re z = 1$.

Comment diable la somme de nombres tous positifs peut-elle être négative ?

La fonction $\zeta(z)$ – et notamment le résultat $1^n + 2^n + \dots = \zeta(-n)$ joue un rôle important dans les problèmes de renormalisation en Théorie des champs ; pour une introduction limpide, voir [44] et [45].

Le résidu en $2in\pi$ ($n > 0$) est $\frac{(-2in\pi)^{z-1}}{e^{2in\pi}} = (2n\pi)^{z-1}e^{-i(z-1)\frac{\pi}{2}}$, celui en $-2in\pi$ ($n > 0$) est $\frac{(2in\pi)^{z-1}}{e^{-2in\pi}} = (2n\pi)^{z-1}e^{+i(z-1)\frac{\pi}{2}}$; d'où :

$$\int_{C_N} \frac{(-\xi)^{z-1}}{e^\xi - 1} d\xi = -2i\pi \sin \frac{\pi z}{2} (2\pi)^{z-1} \sum_{n=1}^N n^{z-1} . \quad (7.110)$$

Dans la limite $N \rightarrow +\infty$, le lacet reproduit $-2i \sin \pi z \zeta(z)\Gamma(z)$, cependant que l'intégrale le long de ABCDEF tend vers zéro. Après passage à la limite, il vient, supposant $\Re z < 0$ pour que la série soit convergente :

$$-2i \sin \pi z \zeta(z)\Gamma(z) = -2i\pi \sin \frac{\pi z}{2} (2\pi)^{z-1} \sum_{n=1}^{+\infty} n^{z-1} \quad (\Re z < 0) ; \quad (7.111)$$

$\zeta(1-z)$ apparaissant à droite, on obtient³⁰ :

$$\boxed{\zeta(1-z) = 2^{1-z} \pi^{-z} \cos \frac{\pi z}{2} \Gamma(z) \zeta(z)} \quad (7.112)$$

Une fois obtenue avec $\Re z < 0$, cette relation se prolonge immédiatement dans tout le plan. Elle est souvent exprimée à l'aide d'une autre fonction $\xi(z)$:

$$\xi(z) \stackrel{\text{déf}}{=} \frac{1}{2} z(z-1) \pi^{-z/2} \Gamma\left(\frac{z}{2}\right) \zeta(z) , \quad (7.113)$$

avec laquelle on a l'égalité symétrique :

$$\boxed{\xi(z) = \xi(1-z)} \quad (7.114)$$

La relation (7.112) permet notamment d'obtenir les valeurs de $\zeta(2m)$, $m \geq 1$ en posant $z = 1 - 2m + \varepsilon$ et en examinant la limite $\varepsilon \rightarrow 0$:

$$\zeta(2m) = 2^{2m} \pi^{2m-1} \zeta(1-2m) \lim_{\varepsilon \rightarrow 0} \left[\cos \frac{\pi(1-2m+\varepsilon)}{2} \Gamma(1-2m+\varepsilon) \right] . \quad (7.115)$$

En utilisant (7.31), et sachant que $(-1)^{m-1} B_{2m} = |B_{2m}|$, on trouve :

$$\boxed{\zeta(2m) = 2^{2m-1} \frac{\pi^{2m}}{(2m)!} |B_{2m}|} \quad (7.116)$$

En particulier :

$$\boxed{\zeta(2) = \frac{\pi^2}{6} \simeq 1.645 \quad \zeta(4) = \frac{\pi^4}{90} \simeq 1.082 \quad \zeta(6) = \frac{\pi^6}{945} \simeq 1.017} \quad (7.117)$$

On ne connaît pas d'expressions simples du même type pour $\zeta(2n+1)$.

³⁰Cette relation, démontrée par Riemann (1859), avait été anticipée par Euler. On la trouve – “pratiquement méconnaissable” selon Hardy ([35], p. 115) – dans les notes de Ramanujan.

La fonction $\zeta(z)$ peut aussi s'exprimer sous la forme dite *du produit d'Euler* (1737), utilisant le fait que tout entier n peut être décomposé de façon unique en ses facteurs premiers : $n = p_1^{r_1} p_2^{r_2} \dots p_{r_n}^{r_n}$. Par ailleurs, la quantité $\frac{1}{1-1/p^z}$ (p premier > 1) peut toujours être représentée par $\sum_{n \in \mathbb{N}} \frac{1}{(p^z)^n}$. Au total, pour $\Re z > 1$, on a l'égalité :

$$\sum_{n=1}^{+\infty} \frac{1}{n^z} = \sum_{n_2 \in \mathbb{N}} \frac{1}{(2^z)^{n_2}} \sum_{n_3 \in \mathbb{N}} \frac{1}{(3^z)^{n_3}} \sum_{n_5 \in \mathbb{N}} \frac{1}{(5^z)^{n_5}} \sum_{n_7 \in \mathbb{N}} \frac{1}{(7^z)^{n_7}} \sum_{n_{11} \in \mathbb{N}} \frac{1}{(11^z)^{n_{11}}} \dots \quad (7.118)$$

de sorte que la série (7.94) peut être écrite sous la forme :

$$\zeta(z) = \prod_{p \text{ premier}} \frac{1}{1-p^{-z}} \quad (\Re z > 1) \quad (7.119)$$

Une autre façon d'établir ce résultat est la suivante. On a (toujours avec $\Re z > 1$) :

$$\begin{aligned} \zeta(z)(1-2^{-z}) &= \left(1 + \frac{1}{2^z} + \frac{1}{3^z} + \dots\right) \left(1 - \frac{1}{2^z}\right) = \\ &= \left(1 + \frac{1}{2^z} + \frac{1}{3^z} + \dots\right) - \left(\frac{1}{2^z} + \frac{1}{4^z} + \frac{1}{6^z} + \dots\right) = 1 + \frac{1}{3^z} + \frac{1}{5^z} + \frac{1}{7^z} + \frac{1}{9^z} + \dots \end{aligned} \quad (7.120)$$

De même :

$$\begin{aligned} \zeta(z)(1-2^{-z})(1-3^{-z}) &= \left(1 + \frac{1}{3^z} + \frac{1}{5^z} + \frac{1}{7^z} + \frac{1}{9^z} + \dots\right) - \\ &\quad \left(\frac{1}{3^z} + \frac{1}{9^z} + \frac{1}{15^z} + \frac{1}{21^z} + \frac{1}{27^z} + \dots\right) \\ &= 1 + \frac{1}{5^z} + \frac{1}{7^z} + \frac{1}{11^z} + \dots \end{aligned} \quad (7.121)$$

et ainsi de suite :

$$\zeta(z)(1-2^{-z})(1-3^{-z})\dots(1-p_n^{-z})\dots \equiv \zeta(z) \prod_{n=1}^{\infty} (1-p_n^{-z}) = 1, \quad (7.122)$$

d'où l'expression (7.119), qui a permis à Euler de démontrer que la "série harmonique" formée uniquement avec les inverses des nombres premiers est, elle aussi, divergente (une façon de montrer qu'il y a une infinité de nombres premiers !).

Revenons brièvement sur la célèbre hypothèse de Riemann³¹ :

³¹Pour une brève revue récente, voir [40]. Le Clay Mathematics Institute offre un prix de 1 000 000 \$ à qui démontrera l'affirmation de Riemann. Hardy a démontré (1914) que la droite $\Re z = \frac{1}{2}$ contient une infinité de zéros R_k .

À l'heure actuelle, il est établi que les ~ 100 premiers milliards de zéros de $\zeta(z)$, R_k , ont effectivement une partie imaginaire égale à $\frac{1}{2}$... [41]. Sur ce problème, des physiciens ne sont pas inactifs, explorant l'une des pistes qui est de mettre en évidence un Hamiltonien (hermitique !) dont le spectre serait composé des $i(R_k - \frac{1}{2})$. D'autres approches physiques ont été suggérées, voir par exemple [42].

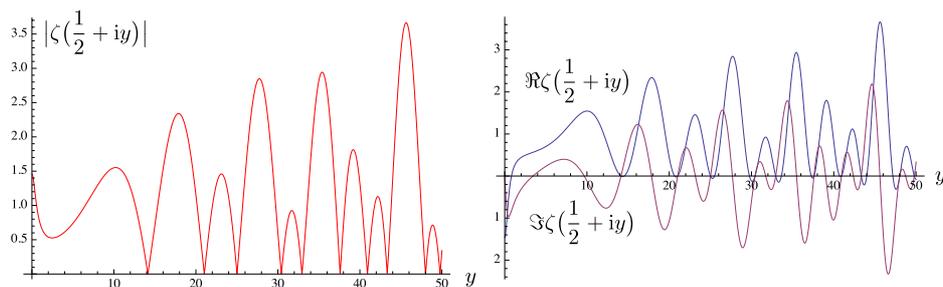


Figure 7.8: Module (à gauche) et parties réelle et imaginaire (à droite) de $\zeta(\frac{1}{2} + iy)$ quand y varie de 0 à 50.

“Tous les zéros de $\zeta(z)$ situés dans la bande $0 \leq \Re z \leq 1$ sont sur la droite $\Re z = \frac{1}{2}$ ”.

Tout d’abord, le produit d’Euler montre que $\zeta(z)$ n’a pas de zéros ayant une partie réelle supérieure à 1 ; en effet, $|p^z| = |p^{x+iy}| = |p^x p^{iy \ln p}| = p^x$, et quand $x > 1$, $p^x > 1$ ($p \geq 2$), tous les facteurs du produit d’Euler sont positifs. Par ailleurs, la relation fonctionnelle (7.112) montre qu’il n’y a pas de zéros à partie réelle négative, autres que les entiers pairs négatifs. Au total, tous les zéros (non-triviaux) sont dans la *bande critique* $0 \leq \Re z \leq 1$ (*critical strip*), où ils sont d’ailleurs répartis *symétriquement* par rapport à la droite $\Re z = \frac{1}{2}$ (*critical line*) comme le montre la même relation (7.112), ou (7.114).

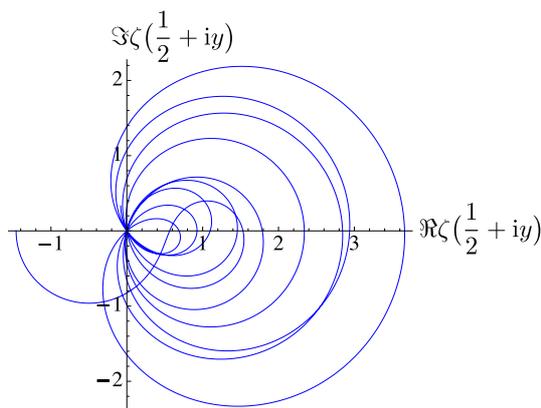


Figure 7.9: Lieu du point d’affixe $\zeta(\frac{1}{2} + iy)$ quand y varie de 0 à 50.

Concernant les nombres premiers, il existe une conjecture due à Legendre, et

rendue plus précise par Gauss³² selon laquelle :

$$\pi(x) \sim \int_2^x \frac{dt}{\ln t} \quad x \rightarrow +\infty \quad (7.123)$$

où $\pi(x)$ désigne traditionnellement le nombre de nombres premiers inférieurs à x . Le symbole \sim établit une simple relation de *proximité numérique* entre les deux membres, le premier étant visiblement une fonction étagée, alors que le second est manifestement (absolument) continue. Cela étant, Riemann a démontré la relation suivante :

$$\mathbf{P}(x) = x - \sum_{k \in \mathbb{N}} \frac{x^{R_k}}{R_k} - \ln 2\pi - \frac{1}{2} \ln(1 - x^{-2}) . \quad (7.124)$$

Dans cette écriture, R_k désigne un zéro non-trivial, et $\mathbf{P}(x) \stackrel{\text{d\u00e9f}}{=} \sum_{n \leq x} \Lambda(n)$ où $\Lambda(n)$ est la fonction de von Mangoldt³³ définie comme :

$$\Lambda(n) = \begin{cases} \ln p & \text{si } \exists k \in \mathbb{N}, n = p^k \\ 0 & \text{autrement} \end{cases} . \quad (7.126)$$

La relation (7.124) est l'un des ponts essentiels entre les zéros R_k de $\zeta(z)$ et la théorie des nombres premiers, dont la répartition garde encore une partie de ses mystères.

Enfin, outre sa relation intime avec la théorie des nombres premiers, la fonction $\zeta(z)$ intervient dans des sommes assez sidérantes ; un florilège³⁴ :

$$\sum_{n=2}^{\infty} \frac{1}{n} [\zeta(n) - 1] = 1 - \gamma , \quad \sum_{n=2}^{\infty} \frac{\zeta(2n)}{n 2^{2n} (2n+1)} = \ln \pi - 1 , \quad \text{etc...} \quad (7.128)$$

7.4 Méthode du col

La méthode du col est une technique permettant d'obtenir une approximation des intégrales du type :

$$I(\lambda) = \int_{-\infty}^{+\infty} e^{\lambda f(x)} dx , \quad (7.129)$$

³²Riemann fut tout près d'établir la preuve de la véracité de la conjecture de Gauss. La dernière étape fut franchie indépendamment par Hadamard et par de la Vallée Poussin (1896), après qu'ils eurent démontré que $\zeta(1 + iy)$ n'est jamais nul.

³³L'intervention de cette fonction provient de son étroite relation avec la fonction ζ . On peut en effet montrer que :

$$\ln \zeta(z) = \sum_{n=2}^{+\infty} \frac{\Lambda(n)}{\ln n} \frac{1}{n^z} \implies \frac{\zeta'(z)}{\zeta(z)} = - \sum_{n=1}^{+\infty} \frac{\Lambda(n)}{n^z} . \quad (7.125)$$

La fonction $\Lambda(n)$ joue un rôle important dans la théorie des séries de Dirichlet.

³⁴En fait, beaucoup de propriétés de la fonction $\zeta(z)$ ne sont que des cas particuliers de propriétés de la fonction de Riemann généralisée $\zeta(z, a)$:

$$\zeta(z, a) \stackrel{\text{d\u00e9f}}{=} \sum_{n=0}^{+\infty} \frac{1}{(a+n)^z} \quad (7.127)$$

De toute évidence : $\zeta(z) = \zeta(z, 1)$.

dans la limite où le paramètre³⁵ $\lambda > 0$ est grand. En Physique, une fois les échelles pertinentes identifiées, il conviendra de préciser ce que signifie “ λ est grand”, comme toujours. Les deux bornes $\pm\infty$ indiquées ne sont pas cruciales, loin de là, comme on le comprendra par la suite. En tout cas, on suppose que l'intégrale $I(\lambda)$ est bien définie par la relation (7.129) telle qu'elle est écrite. C'est le cas si $|e^{\lambda f(x)}|$ est sommable ; avec $\lambda > 0$ une condition suffisante est que $f(x)$ est bornée supérieurement et que $f(x) \rightarrow -\infty$ quand $x \rightarrow \pm\infty$.

La technique exposée ci-dessous permet aussi d'approximer des intégrales du genre :

$$L(\lambda) = \int_0^{+\infty} e^{-\lambda x} \phi(x) dx , \quad (7.130)$$

en remarquant que l'expression (7.130) s'écrit aussi :

$$L(\lambda) = \int_0^{+\infty} e^{\lambda[-x + \frac{1}{\lambda} \ln \phi(x)]} dx . \quad (7.131)$$

L'égalité (7.130) est en fait une transformation de Laplace (voir chapitre 9), d'où l'importance de la méthode du col pour les problèmes traités dans ce cadre.

La méthode du col est d'usage très fréquent en Physique. Par exemple, en Mécanique statistique, les questions centrales consistent à obtenir des grandeurs physiques (énergie, entropie, ...) dans la limite thermodynamique, lorsque le nombre de constituants (ou le nombre de degrés de liberté), N , tend vers l'infini³⁶. D'ailleurs, même si on ne prend pas formellement la limite thermodynamique, un système macroscopique (par nature) contient un nombre *énorme* de constituants (atomes) ($N \sim 10^{23}$). Si E désigne une grandeur extensive de ce système, on trouve systématiquement des expressions du type :

$$\Phi(N) = \int_0^{+\infty} e^{-\beta E} \phi(E) dE = N \int_0^{+\infty} e^{-N\beta\epsilon} \phi(N\epsilon) d\epsilon , \quad (7.132)$$

où ϵ est la grandeur intensive $\frac{E}{N}$.

Pour la même raison de fond, cette méthode joue un rôle de premier plan lorsqu'il s'agit d'étudier le comportement à grand temps, à grande distance, ... de quantités physiques (Physique à grande échelle). Par exemple, t désignant le temps, il est fréquent de devoir connaître la variation (au moins approximative) de fonctions définies comme suit :

$$f(t) = \int_0^{+\infty} e^{-\gamma t} F(\gamma\tau) d\gamma . \quad (7.133)$$

³⁵On peut toujours supposer λ positif. En effet, avec l'intégrale $\int_{-\infty}^{+\infty} e^{AF(x)} dx$, $A \in \mathbb{C}$, $A = |A|e^{i\alpha}$, il suffit de poser $|A| = \lambda$ et $e^{i\alpha}F = f$. Par ailleurs, quand A est imaginaire pur et F réelle, on parle plutôt de *méthode de la phase stationnaire* (c'est ce qui permet de définir la vitesse de groupe, en Optique ou en Mécanique quantique).

³⁶Il faut simultanément faire tendre vers l'infini (et pas n'importe comment) les différentes variables extensives. Par exemple, pour un système enfermé dans un conteneur de volume V , la limite thermodynamique s'effectue en faisant $N \rightarrow +\infty$, $V \rightarrow +\infty$ avec $\frac{N}{V} = C^{\text{ste}} =$ densité (moyenne).

où la fonction F est connue, où γ est l'inverse d'un temps et où τ est une échelle de temps caractéristique apparaissant spontanément dans le problème posé. En posant $\gamma\tau = x$, $f(t)$ se réécrit comme :

$$f(t) = \frac{1}{\tau} \int_0^{+\infty} e^{-\frac{t}{\tau}x} F(x) dx . \quad (7.134)$$

Dans cette expression, c'est le rapport $\frac{t}{\tau}$ qui (au signe près) joue le rôle du paramètre λ de (7.129) : la méthode du col permettra de trouver (approximativement) $f(t)$ pour des valeurs telles que $t \gg \tau$, c'est-à-dire aux grands temps (comparés à l'échelle de temps "propre" τ). La même analyse peut être refaite en remplaçant le temps par une coordonnée : on pourra de façon analogue obtenir le comportement aux grandes distances (toujours relativement à une échelle naturelle, de longueur cette fois). Ces remarques permettent d'être convaincu que la méthode du col fait partie de l'arsenal de base du physicien. Ajoutons que, si cette méthode est relativement universelle, on connaît bien d'autres techniques pour trouver des approximations asymptotiques d'intégrales ; en particulier, le lemme de Watson est d'un emploi courant ([5], § 6.4, [47], p. 446). Il s'énonce comme suit : soit la fonction $F(t)$ de la forme $f(t^a)t^b$, avec $a > 0$, $b > -1$, $f(x)$ une fonction continue, ayant un développement de Maclaurin pour $|x| < \delta$. Si de plus il existe des constantes M et c telles que $|F(t)| \leq Me^{ct}$ quand $t \rightarrow +\infty$, alors :

$$\boxed{\int_0^{+\infty} F(t) e^{-zt} dt \sim \sum_{n=0}^{+\infty} a_n \Gamma(na + b + 1) z^{-(na+b+1)}} \quad (7.135)$$

quand $|z| \rightarrow \infty$ avec $|\text{Arg } z| \leq \frac{\pi}{2} - \alpha$, $0 < \alpha < \frac{\pi}{2}$ et $a_n = \frac{1}{n!} f^{(n)}(0)$. Pour en savoir plus sur les approximations asymptotiques d'intégrales, la lecture de Bender et Orszag ([5], chapitre 6, un *must* en cette matière et beaucoup d'autres) est vivement recommandée.

Plus généralement, comme on va le voir, la méthode du col s'applique aux intégrales de la forme :

$$I(\lambda) = \int_C e^{\lambda f(z)} dz , \quad (7.136)$$

où λ est toujours grand et pris (conventionnellement) positif, et où C est un certain chemin dans le plan complexe, d'extrémités A et B (C va de A à B). $f(z)$ est une fonction analytique, sauf en des points dûment identifiés comme formant l'ensemble de ses singularités.

Pour saisir l'idée première de la méthode du col, revenons au cas où tout est réel – l'intégrale étant plus précisément de la forme (7.129) –, et supposons que la fonction $f(x)$ présente un seul maximum en un certain point x_0 ; alors $e^{\lambda f(x)}$ possède³⁷ un maximum et un seul, en x_0 . Quand λ est très grand, la fonction exponentielle d'une part exalte le maximum, d'autre part "écrase" l'intégrand en dehors de x_0 . Au total, la fonction sous l'intégrale (7.129) présente une sorte de résonance très marquée près de x_0 , imposant l'idée intuitive selon laquelle la valeur de l'intégrale est essentiellement due à ce qui se passe presque exclusivement au voisinage de ce maximum très pointu (voir figure 7.10).

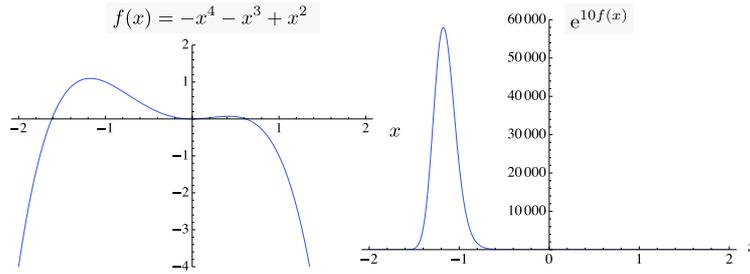


Figure 7.10: Illustration des variations comparées d'une fonction $f(x)$ et de $e^{\lambda f(x)}$ quand λ est grand et positif.

Ces considérations s'illustrent aussi avec une vraie gaussienne, en prenant simplement $f(x) = -(x - x_0)^2$, dont le graphe est une parabole maximum en x_0 . Maintenant, l'exponentielle $e^{\lambda f(x)} \equiv e^{-\lambda(x-x_0)^2}$ est une gaussienne, très fine si $\lambda \gg 1$ (son extension autour de x_0 est d'ordre $\frac{1}{\sqrt{\lambda}}$). Il n'y a pas de doute que, pour $\int_{-\infty}^{+\infty} e^{\lambda f(x)} dx$, c'est le voisinage de x_0 qui donne l'essentiel de la valeur de l'intégrale – que l'on sait ici calculer exactement (c'est une intégrale gaussienne !).

De façon plus quantitative et générale, et revenant à une fonction $f(x)$ admettant un développement de Taylor autour de son unique³⁸ maximum, on écrit :

$$f(x) = f(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) + \dots \quad (f''(x_0) < 0) . \quad (7.137)$$

Il n'y a pas de terme linéaire en $x - x_0$ puisque $f'(x_0) = 0$ (condition de stationnarité), et la valeur $f''(x_0)$ est bien *négative* puisqu'il s'agit d'un *maximum*. En laissant tomber les termes au-delà du second, on écrit :

$$I(\lambda) \simeq \int_{-\infty}^{+\infty} e^{\lambda[f(x_0) - \frac{1}{2}(x-x_0)^2 |f''(x_0)|]} dx ; \quad (7.138)$$

utilisant alors³⁹ $\int_{-\infty}^{+\infty} e^{-ax^2} dx = \left(\frac{\pi}{a}\right)^{\frac{1}{2}}$, (7.138) donne :

$$I(\lambda) \simeq e^{\lambda f(x_0)} \sqrt{\frac{2\pi}{\lambda |f''(x_0)|}} = e^{\lambda f(x_0)} \sqrt{\frac{2\pi}{-\lambda f''(x_0)}} . \quad (7.139)$$

En définitive, la méthode intuitive ainsi construite consiste à aller chercher le(s) maximum(a) de la fonction dans l'argument de l'exponentielle, et à faire un développement quadratique afin de retomber sur des intégrales gaussiennes. C'est ce résultat majeur que la méthode du col généralise au champ complexe.

³⁷Affirmation intuitive évidente, qui se confirme en notant que $(e^{\lambda f(x)})' = \lambda f'(x) e^{\lambda f(x)}$.

³⁸Cette hypothèse n'est pas vraiment restrictive : il suffit de faire le même travail près de chaque maximum quand il y en a plusieurs.

³⁹Cette égalité est vraie $\forall a \in \mathbb{C}, \Re a > 0$, mais se prolonge analytiquement en tournant continûment la demi-droite d'intégration quand a acquiert un argument non-nul. La coupure de la racine carrée $a^{1/2}$ est \mathbb{R}_- .

L'idée-clé étant donnée, revenons à la forme générale (7.136), et posons comme d'habitude $f(z) = u(x, y) + iv(x, y)$. L'intégrand est alors $e^{\lambda u(x, y)} e^{i\lambda v(x, y)}$. Pour les raisons exposées ci-dessus, on s'attend à ce que la valeur de l'intégrale soit dominée par le voisinage des points du contour d'intégration où la partie *réelle* $u(x, y)$ est grande, à condition toutefois que, *en plus*, les termes impliqués aient tous à peu près la même phase – d'où l'idée intuitive qu'il faut cheminer de sorte que $v(x, y)$ varie *lentement*⁴⁰. Dans le cas contraire, si la partie imaginaire varie vite, $e^{\lambda f(z)}$ possède un facteur de phase $e^{i\lambda v(x, y)}$ oscillant rapidement de sorte que la somme locale de $e^{\lambda f(z)}$ contient certes des termes de grand module (puisque u est grand) mais ceux-ci vont se compenser partiellement à cause de la grande variabilité du facteur de phase $e^{i\lambda v(x, y)}$ et ses alternances de signe rapides : comme $e^{i\lambda v} = \cos \lambda v + i \sin \lambda v$, si v varie vite, la sommation intègre des lignes trigonométriques variant rapidement, et le résultat net est petit, quel que soit le module, petit ou grand.

La méthode consiste ainsi d'abord à déformer le contour initial pour le faire passer par le(s) point(s) remarquable(s) où $u(x, y)$ est extremum ($u'_x = u'_y = 0$) – et $v(x, y)$ l'est tout autant en vertu des conditions de Cauchy - Riemann –, et en gardant l'œil comme toujours sur les singularités de f . Un premier pas consiste ainsi à écrire :

$$I(\lambda) = \int_{C'} e^{\lambda f(z)} dz, \quad (7.140)$$

où C' est un avatar du contour initial C , passant par le(s) point(s) où $\Re f(z)$ est extremum, obtenu en évitant soigneusement les singularités de f (contournement des pôles, passage en lisière des coupures, ...) – exploitant toujours la possibilité de déformer le chemin (à extrémités fixées) sans changer la valeur de l'intégrale (à ce stade, aucune approximation n'a encore été faite). Il est clair que la détermination des singularités de $e^{\lambda f(z)}$ est un préalable absolu.

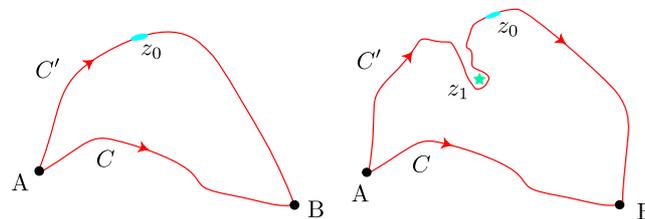


Figure 7.11: Déformations du contour initial C en un contour C' passant par un extremum de $f(z)$, dans l'hypothèse où la région balayée ne contient aucune singularité (à gauche), ou en évitant une singularité z_1 (à droite).

Par un point stationnaire, noté z_0 , il passe autant de chemins que l'on veut, de sorte que la question est maintenant de trouver un critère de sélection. En ce point, $f(z)$ est stationnaire ($f'(z_0) = 0$), mais il ne peut s'agir que d'un col pour chacune des surfaces Σ_u et Σ_v représentant $u(x, y)$ et $v(x, y)$

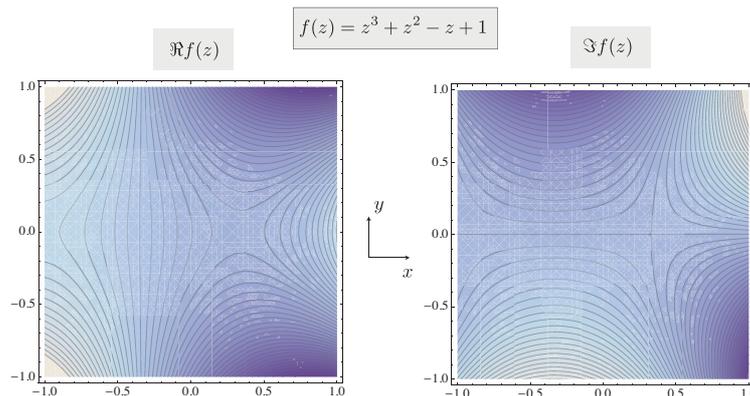


Figure 7.12: Courbes de niveau des parties réelle (à gauche) et imaginaire (à droite) d'une certaine fonction holomorphe. Ces deux familles de courbes sont orthogonales, en conséquence des conditions de Cauchy - Riemann. Les lignes iso- $\Im f(z)$ sont donc les chemins de plus grande pente sur la surface Σ_u .

On va montrer que le meilleur chemin dans le plan \mathbb{C} est la projection du chemin sur la surface Σ_u le long duquel la partie réelle de f (localement gaussienne par rapport à l'écart au col) croît et décroît le plus vite. Ce chemin sur Σ_u , dit de plus grande pente (et passant donc par z_0), est orthogonal aux courbes de niveau de Σ_u , définies par $u(x, y) = C^{\text{ste}}$ (voir un exemple sur la figure 7.12, à gauche). Par ailleurs, on sait que, en conséquence des conditions de Cauchy - Riemann, les familles de courbes $u(x, y) = C^{\text{ste}}$ et $v(x, y) = C^{\text{ste}}$ sont orthogonales les unes aux autres. Au total, le chemin de plus grande pente sur Σ_u est aussi celui où la partie imaginaire de $f(z)$ est constante. Il en résulte qu'un tel chemin n'est envisageable en tant que substitut du chemin initial que si, sur ce dernier, la partie imaginaire de $f(z)$ prend la même valeur en ses deux extrémités, ce que l'on suppose pour la suite. On indiquera en temps utile comment la méthode peut néanmoins être utilisée lorsque tel n'est pas le cas (voir remarque 2 p. 318).

Au voisinage de z_0 , où $f'(z_0) = 0$, le développement de Taylor commence comme $f(z) = f(z_0) + \frac{1}{2}(z - z_0)^2 f''(z_0) + \dots$. En admettant qu'il n'y a pas de singularité dans le voisinage de z_0 , on peut même rectifier localement le chemin, pour en faire un segment de droite incliné de θ par rapport à l'axe Ox – une modification qui fait perdre l'invariance *globale* de la partie imaginaire, mais qui est sans dommage réel sur la qualité de l'approximation pourvu que λ soit assez grand. Posons dès lors $z - z_0 = se^{i\theta}$, où $s \in \mathbb{R}$ est l'abscisse (curviligne) réelle⁴¹ le long du chemin du plan rectifié localement (voir figure 7.13) ; si on ajoute que s est négatif au départ de z_0 vers A et positif de z_0 vers B, l'angle θ définit une droite *orientée*, et est maintenant déterminé à 2π près (et non plus à π près).

⁴⁰Il peut même arriver que la phase $e^{i\lambda v(x, y)}$ prend la même valeur aux extrémités du chemin ; on prend alors celui où elle est constante d'un bout à l'autre (voir aussi la Remarque 2 p. 318).

⁴¹Attention ! s est l'abscisse curviligne réelle (positive ou négative), ce n'est pas le *module* de l'écart

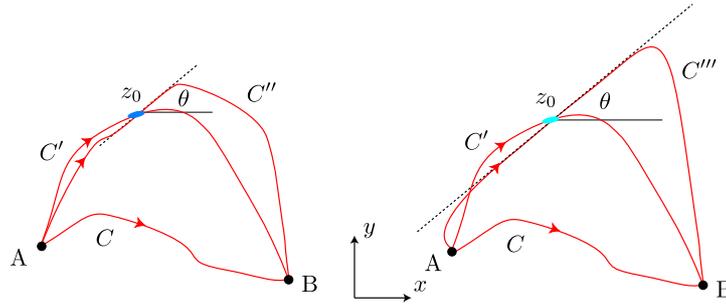


Figure 7.13: Contour C'' passant par l'extremum de $f(z)$ en z_0 et dont la tangente en z_0 fait l'angle θ avec Ox , correspondant au chemin de plus grande pente sur Σ_u . En l'absence de singularité à proximité, on peut encore déformer C'' en C''' pour le superposer le plus possible avec la droite tangente en pointillés.

Posant enfin $f''(z_0) = \rho e^{i\phi}$ ($\rho > 0$), on a :

$$f(z) = f(z_0) + \frac{1}{2}s^2\rho e^{i(2\theta+\phi)} + \dots, \quad (7.141)$$

d'où :

$$u(x, y) \simeq u(x_0, y_0) + \frac{1}{2}s^2\rho \cos(2\theta+\phi), \quad v(x, y) \simeq v(x_0, y_0) + \frac{1}{2}s^2\rho \sin(2\theta+\phi). \quad (7.142)$$

Puisque l'on veut suivre un chemin où la partie réelle passe par un *maximum*, il faut $\cos(2\theta+\phi) < 0$; par ailleurs, il faut que sur ce chemin la partie imaginaire varie le moins possible : elle donne un facteur de phase qui, amplifié par le paramètre λ , provoque des oscillations très rapides qui "tuent" la contribution à l'intégrale venant du maximum de la partie réelle. Pour éviter ce phénomène parasite qui dégraderait l'approximation recherchée, il faut prendre naturellement $\sin(2\theta+\phi) = 0$, auquel cas la partie imaginaire est (localement) constante, d'où le choix $2\theta+\phi = 0$ ou π . Compte tenu de la condition $\cos(2\theta+\phi) < 0$, on doit prendre $2\theta+\phi = \pi$, donnant $\cos(2\theta+\phi) = -1$, valeur maximale pour un cosinus : comme le montre l'expression développée de u , (7.142), c'est bien le chemin de *plus grande pente* (passant par z_0) pour la partie réelle ; par construction, c'est aussi celui qui se recolle au chemin de phase constante.

Ce chemin correspond à un contour C'' dans \mathbb{C} dont la tangente en z_0 fait l'angle θ avec l'axe réel, avec $\theta = -\frac{\phi}{2} + \frac{\pi}{2} + k\pi$, ou encore⁴² :

$$\theta = -\frac{\phi}{2} \pm \frac{\pi}{2}; \quad (7.143)$$

⁴² $z - z_0$.

Le signe \pm correspond en fait aux deux sens de parcours possibles du chemin de plus grande pente. Les deux directions opposées définies en (7.143) donnent la ligne à suivre afin de rencontrer un maximum pour $\Re f$. Le long des (deux) directions $\cos(2\theta+\phi) = +1$, $\Re f$ a un minimum, soit $2\theta+\phi = 0$ (2π), ou encore $\theta = -\frac{\phi}{2}$, $\pi - \frac{\phi}{2}$: ces deux directions sont bien orthogonales à celles définies en (7.143), et croisent ces dernières au col z_0 .

selon ces deux directions (opposées), l'écart quadratique en s pour la partie imaginaire v est nul, ce qui montre que, à cet ordre, v est constant le long de cette direction. L'intégrand $e^{\lambda f(x,y)}$ a ainsi une phase constante (à cet ordre) : en éliminant les interférences destructives, c'est bien en suivant un tel chemin que l'on ramasse l'essentiel de la valeur de $I(\lambda)$ pour λ grand⁴³. Au total, le long de ce chemin, et dans le voisinage de z_0 , la partie imaginaire de f ne varie pas à s^3 près, l'écart entre $f(z)$ et $f(z_0)$ étant quadratique et réel :

$$f(z) \simeq u(x_0, y_0) - \frac{1}{2}\rho s^2 + iv(x_0, y_0) + \mathcal{O}((z-z_0)^3) \equiv f(z_0) - \frac{1}{2}\rho s^2 + \mathcal{O}((z-z_0)^3) , \quad (7.144)$$

le signe $-$ devant $\frac{1}{2}\rho s^2$ venant de $\cos(2\theta + \phi) = -1$ (voir (7.142)).

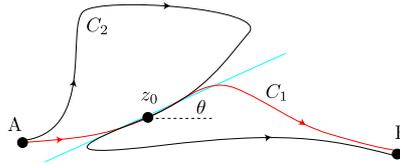


Figure 7.14: Les intégrales le long des deux contours sont égales (on peut manifestement superposer C_1 et C_2 par des déformations continues), tant que l'on ne fait aucune approximation pour les évaluer.

L'intégrale (7.136) s'écrit alors :

$$I(\lambda) \simeq e^{\lambda f(z_0)} \int_{C'''} e^{-\lambda \frac{1}{2} s^2 \rho} dz , \quad (7.145)$$

où C''' est le contour déformé suivi par la variable muette d'intégration z pour la faire passer par z_0 conformément au chemin de plus grande pente sur la surface Σ_u représentant la partie réelle. Ceci étant précisé, et toute singularité évitée, on peut encore déformer C''' en l'étirant de façon à le superposer le plus possible à la droite tangente en z_0 définie par l'angle θ , d'où le contour C'''' sur la figure 7.13 ; alors, l'intégrale s'écrit sous forme réelle⁴⁴ ($dz = e^{i\theta} ds = e^{-i\frac{\phi}{2}} e^{i\frac{\pi}{2}} ds$). Les bornes pour s sont indéfinies, mais l'approximation la plus naturelle à ce stade est de les rejeter à l'infini, ce qui revient à approximer f par une vraie gaussienne d'un bout à l'autre de la tangente. Il vient alors :

$$I(\lambda) \simeq \int_{-\infty}^{+\infty} e^{\lambda f(z_0) - \frac{1}{2}\lambda |f''(z_0)| s^2} e^{-i\frac{\phi}{2} + i\frac{\pi}{2}} ds = i e^{\lambda f(z_0)} e^{-i\frac{\phi}{2}} \sqrt{\frac{2\pi}{\lambda |f''(z_0)|}} ; \quad (7.146)$$

⁴³On retrouve d'ailleurs à ce moment l'idée de *phase stationnaire*, expliquant pourquoi le centre d'un paquet d'ondes se déplace à la vitesse (de groupe) $v_g = \left(\frac{d\omega}{dk}\right)_{k_0}$, k_0 étant le vecteur d'onde majoritaire de la décomposition du paquet en ondes planes.

⁴⁴Si on prenait pour θ l'autre possibilité $\theta = -\frac{\phi}{2} - \frac{\pi}{2}$ contenue dans (7.143), le résultat serait *au total* le même (à condition de ne faire aucune approximation) : tant que l'on ne franchit pas de singularité, le contour C_1 et le contour C_2 en *spaghetti* partant de A et arrivant en B après avoir suivi la tangente en sens inverse sont homotopes (voir figure 7.14). En revanche, et c'est évidemment le cas en pratique, seule une portion de chaque contour intervient ; il n'y a toutefois pas d'ambiguïté de signe résultant d'une incertitude sur le sens de parcours : il suffit se s'ancrer sur le chemin de phase constante *partant* de A et *arrivant* en B. Quoi qu'il en soit, tout signe incertain peut être tranché en raisonnant par continuité dans un cas extrême.

En rapprochant la phase et la racine carrée, on voit que cette expression peut s'écrire sous la forme de la détermination de la fonction $(\#)^{\frac{1}{2}}$ qui prend une valeur positive sur l'axe réel positif :

$$I(\lambda) \simeq i e^{\lambda f(z_0)} \left(\frac{2\pi}{\lambda f''(z_0)} \right)^{\frac{1}{2}} = e^{\lambda f(z_0)} \left(\frac{2\pi}{-\lambda f''(z_0)} \right)^{\frac{1}{2}}, \quad (7.147)$$

qui est identique à (7.139). Cette coïncidence n'est pas fortuite, mais provient d'un résultat déjà mentionné, à savoir que, pour a complexe, on a encore :

$$G(a) \equiv \int_{-\infty}^{+\infty} e^{-ax^2} dx = \left(\frac{\pi}{a} \right)^{\frac{1}{2}}, \quad (7.148)$$

où $-\pi < \text{Arg } a < +\pi$ et où la branche de la racine carrée est celle qui vaut 1 en $z = 1$: $(1)^{\frac{1}{2}} = 1$. (7.148) résulte notamment du fait que, de par sa définition par l'intégrale ci-dessus, G possède la propriété de réflexion de Schwarz $G(a^*) = [G(a)]^*$. De plus, en redéfinissant convenablement les grandeurs du problème, on voit facilement que le résultat (7.147) tient aussi pour $\lambda \in \mathbb{C}$. Notons que, à ce stade, rien ne permet d'évaluer l'erreur commise en approximant l'intégrale par l'expression (7.147) ; d'une façon générale, la méthode du col donne essentiellement le premier terme du développement asymptotique.

Clairement, et c'est une vérification après coup, l'expression approchée (7.147) doit reproduire le résultat *exact* $\forall \lambda \in \mathbb{C}$ pour l'intégrale gaussienne $\int_{-\infty}^{+\infty} e^{-\lambda x^2} dx$. En prenant $C = C' = C'' =]-\infty, +\infty[$, et $f(z) = -z^2$ (alors $z_0 = 0$ et $f''(z_0) = -2$), (7.147) donne $i \left(\frac{2\pi}{-2\lambda} \right)^{1/2} = \sqrt{\pi} \lambda^{-1/2}$, comme il se doit (la racine carrée étant proprement définie comme on l'a vu).

◆ *Remarques*

1. La méthode du col peut faire imaginer des situations apparemment paradoxales, comme dans l'exemple suivant (voir figure 7.15). Soit à intégrer une fonction $f(x)$ le long du segment de \mathbb{R} allant des points A à B, et supposons que $f(x)$ est positive le long de AB. L'intégrale de A à B, I_{AB} , est positive, et certainement plus grande strictement que l'intégrale allant de tout point intermédiaire C à B, I_{CB} . Supposons aussi que $f(x)$ prend les mêmes valeurs en ordre de grandeur partout sur AB, de sorte que, par exemple, I_{CB} est égale à la moitié de I_{AB} .

Supposons maintenant que l'on utilise la méthode du col pour calculer chacune de ces deux intégrales, ce qui revient à envisager les deux contours C_1 et C_2 (on suppose évidemment qu'il n'y a pas de singularité en jeu) ; afin de fixer les idées, admettons que ce qui vient du col en z_0 est positif. Par construction, le col est le même pour les deux intégrales (c'est la même fonction $f(z)$ partout !), de sorte que, si l'on s'en tient à la formule ci-dessus, on trouve que les deux intégrales sont égales, alors que l'une est en gros la moitié de l'autre...

Il n'y a évidemment pas d'incohérence : en réalité, l'intégrand de I_{CB} prend certes de grandes valeurs en module au voisinage de z_0 , mais l'arc en "S" constituant le

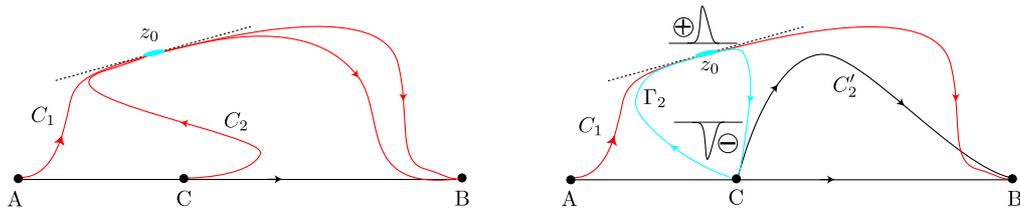


Figure 7.15: Illustration de la résolution du paradoxe auquel peut conduire la manipulation trop rapide de la formule du col.

premier tronçon de C_2 doit donner, au total, une grande contribution *négative* qui d'une part compense celle ramassée en z_0 , d'autre part donne l'opposé de ce qui vient du segment AC pour I_{AB} .

La partie de droite de la même figure complète l'explication : l'intégrale I_{CB} est aussi égale à l'intégrale le long du chemin C'_2 , ou du chemin $\Gamma_2 \cup C'_2$, la boucle Γ_2 donnant zéro. Comme elle est nulle, précisément, toute grande contribution venant du voisinage de z_0 est compensée par ce qui se passe partout ailleurs sur Γ_2 . En quelque sorte, ce qui vient de la portion de Γ_2 loin du col est l'image en négatif de ce qui se passe près de z_0 (il y a comme un effet d'*ombrage* qui efface l'importance du col pour l'intégrale sur Γ_2).

2. La méthode du col est parfois délicate à mettre en œuvre, en raison des subtilités de l'analyse complexe et de l'interdépendance entre les propriétés locales et les propriétés intégrales (ce qui se passe en un point conditionne ce qui se passe ailleurs). En cas de difficulté, le principe d'action suivant permet de surmonter les difficultés et, finalement, résume bien l'esprit de fond de la méthode (pour plus de détails, et des exemples, voir [5], chapitre 6).

On a vu que s'il fallait privilégier les endroits où la fonction est "grande", il faut aussi les balayer de façon à éviter des interférences destructives liées à une trop grande variabilité de la phase – d'où le choix du chemin où la partie imaginaire est *localement* constante. Le moyen extrême et radical d'éviter cet effet parasite est de suivre exclusivement dans le plan des lignes $\Im f(z) = \alpha = C^{\text{ste}}$ – ceci n'étant *a priori* possible que si la partie imaginaire prend les mêmes valeurs aux extrémités du chemin d'intégration. Ce choix étant fait, la phase $e^{i\lambda\alpha}$ est en facteur d'une intégrale portant sur $e^{\lambda\Re f(z)}$, dont l'analyse locale au maximum rencontré *sur cette ligne* en z_0 permettra de trouver une expression approchée (répétons que ce maximum est un col : dans les directions perpendiculaires à la ligne $\Im f(z) = C^{\text{ste}}$, $\Re f(z)$ est *minimum* en z_0).

Cette technique est d'ailleurs parfois applicable même si la phase ne prend pas la même valeur aux extrémités. Il est en effet parfois possible de déformer le contour pour le faire passer par le col, puis de le découper en tronçons où, sur chacun d'entre eux, la phase est constante, ces tronçons étant reliés par des arcs où la phase varie mais dont la contribution est finalement sous-dominante, voire carrément nulle. Schématiquement : on découpe C en $C_1 \cup C_2 \cup C_3$; la phase prend une valeur constante α_1 au départ puis tout au long de C_1 , et la valeur constante α_3 sur C_3 .

C_2 est par exemple un arc à l'infini où la phase varie de α_1 à α_3 mais où l'intégrale est nulle. Ensuite, pour chacune des intégrales restantes, on effectue l'analyse comme ci-dessus, chacune portant alors sur une quantité du genre $e^{i\lambda\alpha_i} e^{\lambda\Re f(z)}$.

- Il est le plus souvent possible de “pousser” le développement au-delà de la contribution obtenue par les moyens ci-dessus, afin de trouver au moins le premier terme sous-dominant. Les formules correspondantes sont pour le moins laborieuses à établir (voir par exemple [5], p. 272). \blacklozenge

Bien sûr, si la fonction $f(x)$ possède plusieurs cols en différents points d'affixes z_j , il faut *a priori* additionner les contributions correspondantes :

$$I(\lambda) \simeq \sum_j e^{\lambda f(z_j)} \left(\frac{2\pi}{-\lambda f''(z_j)} \right)^{\frac{1}{2}}. \quad (7.149)$$

En fait, et à condition que les $f''(z_j)$ soient toutes du même ordre de grandeur, un seul terme suffit souvent (c'est le terme *dominant*), celui correspondant à la plus grande des parties réelles⁴⁵ $\Re f(z_j)$. En effet, si $z_{j_{\max}}$ est l'affixe correspondante, on a :

$$I(\lambda) \simeq \left(\frac{2\pi}{-\lambda f''(z_{j_{\max}})} \right)^{\frac{1}{2}} e^{\lambda f(z_{j_{\max}})} \left[1 + \sum_{j \neq j_{\max}} e^{-\lambda[f(z_{j_{\max}}) - f(z_j)]} \left(\frac{f''(z_{j_{\max}})}{f''(z_j)} \right)^{\frac{1}{2}} \right]; \quad (7.150)$$

dans le crochet, on voit que chaque terme de la somme est en module exponentiellement petit devant 1. Bien sûr, si les valeurs des dérivées secondes sont franchement différentes les unes des autres, il faut effectuer au cas par cas une analyse plus fine. Il faut aussi être moins brutal quand l'effet recherché est une modification *qualitative* : par exemple l'apparition d'une partie imaginaire dans une quantité physique purement réelle dans une approximation que l'on cherche précisément à raffiner⁴⁶.

Donnons maintenant deux exemples classiques et importants d'application de la méthode du col.

Formule de Stirling

Le premier concerne la fonction $\Gamma(z)$ et est très utile en pratique, en Mécanique statistique, en Théorie des probabilités, etc. Le résultat est général, mais exprimé pour $z \in \mathbb{N}$, il est connu sous le nom de *formule de Stirling*.

⁴⁵ $|e^{x+iy}| = e^x$. Le cas des fonctions de Bessel traité p. 321 est un exemple où, au contraire, deux cols doivent simultanément pris en compte.

⁴⁶Un exemple : en théorie quantique élémentaire, les énergies d'un système lié sont parfaitement définies (et sont bien sûr réelles). Un traitement plus élaboré du système couplé (indissociablement) matière+champ électromagnétique fait apparaître des corrections qui, si elles sont numériquement relativement petites, se traduisent par l'apparition d'une partie imaginaire finie pour les “énergies”, donnant la durée de vie des états excités (c'est du moins ce que l'on trouve dans l'approche la plus simple, souvent appelée *approximation du pôle*). Typiquement, en pareil cas, cette correction (imaginaire) provient de l'un des termes de la somme présente dans (7.150) ; ne pas en tenir compte au motif qu'elle est petite en module ferait passer à côté du but recherché (voir aussi le résultat final (7.250) de la section 7.7).

Reprenons la définition (7.6) et écrivons :

$$\Gamma(\lambda) = \int_0^{+\infty} z^{\lambda-1} e^{-z} dz = \int_0^{+\infty} e^{-z+(\lambda-1)\ln z} dz \equiv \int_0^{+\infty} e^{\lambda f(z)} dz, \quad (7.151)$$

avec $f(z) = \frac{1}{\lambda}[-z + (\lambda - 1)\ln z]$; $f'(z_0) = 0$ si $z_0 = \lambda - 1$, d'où $f''(z_0) = -\frac{1}{\lambda(\lambda-1)}$. Dans ce cas, (7.147) donne^{47, 48} :

$$\Gamma(\lambda) \simeq \sqrt{2\pi} (\lambda - 1)^{\lambda-\frac{1}{2}} e^{-(\lambda-1)} \simeq \sqrt{2\pi} \lambda^{\lambda-\frac{1}{2}} e^{-\lambda} \quad (\lambda \in \mathbb{C}, |\lambda| \gg 1), \quad (7.152)$$

où la dernière expression résulte d'un développement limité au plus bas ordre en λ^{-1} :

$$(\lambda - 1)^{\lambda-\frac{1}{2}} = \lambda^{\lambda-\frac{1}{2}} e^{(\lambda-\frac{1}{2})\ln(1-\frac{1}{\lambda})} = \lambda^{\lambda-\frac{1}{2}} e^{(\lambda-\frac{1}{2})(-\frac{1}{\lambda})} \simeq \lambda^{\lambda-\frac{1}{2}} e^{-1+\mathcal{O}(\frac{1}{\lambda})}, \quad (7.153)$$

d'où (7.152), étant toujours entendu que la racine carrée $\lambda^{\frac{1}{2}}$ correspond à la branche admettant comme coupure le demi-arc réel négatif (ce qui, pour $\lambda \in \mathbb{R}_+$ donne bien un réel positif – la moindre des choses pour une approximation de $\Gamma(\lambda)$ qui est dans ce cas un réel positif !).

On sait que $\Gamma(N + 1) = N!$ si $N \in \mathbb{N}$; la première expression approchée de $\Gamma(\lambda)$, (7.152), donne $N! \simeq \sqrt{2\pi} N^{N+\frac{1}{2}} e^{-N}$, d'où l'approximation de la factorielle, usuellement écrite comme :

$$N! \simeq \left(\frac{N}{e}\right)^N \sqrt{2\pi N} \quad (\text{formule de Stirling}) \quad (7.154)$$

en particulier :

$$\ln N! \simeq N \ln \frac{N}{e} + \mathcal{O}(\ln N) \simeq N \ln N - N \quad (N \gg 1). \quad (7.155)$$

Pour $N = 10$, on a $N! = 3\,628\,800$; la formule de Stirling donne⁴⁹ $10! \simeq 3\,598\,696$, soit une erreur relative voisine de 0.8%. L'approximation plus rustique $\ln N! \simeq N \ln N$ est d'usage courant en Mécanique statistique (elle fait notamment ressortir naturellement l'extensivité de l'entropie d'un système macroscopique), où le nombre N est alors gigantesque (c'est essentiellement le nombre d'Avogadro). Elle permet aussi de montrer dans quelles conditions la distribution de probabilité de Poisson et la distribution gaussienne tendent l'une vers l'autre (voir chapitre 1).

L'expression au second membre de (7.152) est le terme dominant du développement de $\Gamma(\lambda)$ quand $|\lambda| \gg 1$. Plus généralement, pour $|z| \gg 1$, on peut s'attendre (et c'est le cas) à l'existence d'un développement du genre :

$$\Gamma(z) \sim \sqrt{2\pi} z^{z-\frac{1}{2}} e^{-z} \left(1 + \frac{A}{z} + \frac{B}{z^2} + \dots\right). \quad (7.157)$$

⁴⁷On se souvient que (7.147) est aussi vrai quand $\lambda \in \mathbb{C}$ – voir le commentaire final après (7.148).

⁴⁸Comme indiqué antérieurement, la méthode du col donne en fait le premier terme du développement asymptotique. Dans l'équation (7.152), on doit en toute rigueur remplacer le signe \simeq par \sim .

⁴⁹Pour $N = 1000$, cette approximation conduit à une erreur plus petite que 10^{-3} :

$$(1000 \ln 1000 - 1000) / \ln(1000!) \simeq 0.999\,26. \quad (7.156)$$

Que dire si $N \sim 10^{23}$...

En utilisant la relation fonctionnelle (7.15) et en y injectant le développement (7.157) pour $\Gamma(z)$ et $\Gamma(z+1)$, l'identification des termes de même puissance en z permet de trouver (un peu laborieusement) les coefficients A, B, \dots :

$$\Gamma(z) \sim \sqrt{2\pi} z^{z-\frac{1}{2}} e^{-z} \left(1 + \frac{1}{12z} + \frac{1}{288z^2} + \dots \right) \quad (7.158)$$

Comportement des fonctions de Bessel ordinaires à l'infini

Le second exemple concerne la fonction de Bessel ordinaire J_n , $n \in \mathbb{Z}$, dont l'une des représentations intégrales est :

$$J_n(t) \stackrel{\text{déf}}{=} \frac{1}{2\pi} \int_0^{2\pi} e^{-in\theta} e^{it \sin \theta} d\theta ; \quad (7.159)$$

c'est un cas illustrant ce qu'il convient de faire quand il y a plusieurs points-cols.

Il est certain que si t est très grand, cette fonction prend des petites valeurs, puisque le deuxième facteur de phase oscille alors très vite dès que $|\theta|$ ou $|\pi - \theta|$ est à peine supérieur à zéro. La question est de trouver le comportement précis de J_n pour les grandes valeurs de t – plus précisément pour $|t| \gg |n|$, comme le révèle le calcul suivant quand on le fait en détail. La première astuce consiste à remarquer que J_n peut s'écrire sous la forme d'une intégrale le long du cercle de rayon unité dans \mathbb{C} , parcouru dans le sens positif ; en effet, en posant $z = e^{i\theta}$, on voit que :

$$J_n(t) = \frac{1}{2i\pi} \int_{C_{r=1}} \frac{1}{z^{n+1}} e^{\frac{t}{2}(z-\frac{1}{z})} dz . \quad (7.160)$$

L'intégrand est de la forme $e^{tf(z)}$, avec $f(z) = \frac{1}{2}(z - \frac{1}{z}) - \frac{1}{t}(n+1) \ln z$; la dérivée de f s'annule pour z solution de l'équation :

$$z^2 - \frac{2}{t}(n+1)z + 1 = 0 , \quad (7.161)$$

dont les racines sont voisines de $\pm i$ quand $|n+1| \ll |t|$. Il y a donc deux points-cols $z_{0\pm} \simeq \pm i$. Le calcul montre que les deux nombres $f(z_{0\pm})$ ont le même module et qu'il en va de même pour les dérivées secondes $f''(z_{0\pm})$, de sorte que les contributions des deux points doivent être retenues sur un pied d'égalité. Par ailleurs, les développements de Taylor au voisinage de $z_{0\pm}$ révèlent que les chemins de plus grande pente sont des droites inclinées de $\pm \frac{\pi}{4}$ par rapport à Ox . Au total, tous calculs faits, la formule (7.149) donne :

$$J_n(t) \simeq \sqrt{\frac{2}{\pi t}} \cos \left(t - n \frac{\pi}{2} - \frac{\pi}{4} \right) \quad (|t| \gg |n|) \quad (7.162)$$

À nouveau, il faut en toute rigueur remplacer \simeq par \sim , car il s'agit là du premier terme du développement asymptotique de $J_n(t)$ de $J_n(t)$. On retiendra que, dans les conditions

précisées, cette fonction de Bessel oscille avec une amplitude décroissant comme $t^{-\frac{1}{2}}$, donc très lentement. On note également que les cols en $\pm i$ pour z correspondent au voisinage de $\pm \frac{\pi}{2}$ pour la variable d'intégration θ : ce sont bien les points où la phase ultra-rapide dans (7.159) est *stationnaire* ; cette remarque permet aussi de comprendre physiquement l'origine de la restriction $|t| \gg |n|$: dans le cas contraire, la phase rapide serait dans le premier facteur, $e^{in\theta}$ (d'ailleurs, on a utilisé le fait que t est grand devant n pour écrire $z_{0\pm} \simeq \pm i$). Cette restriction entre n et t se voit aussi en sachant que $J_n(t)$ admet un développement du type⁵⁰ $(t/2)^n S_n(t)$ où S_n est une série entière (convergente pour $|\arg t| < \pi$), telle que $S_n(0) = 1/n!$; pour $n \gg 1$, la fonction $J_n(t)$ est donc très écrasée à l'origine, et ne décolle que pour $(t/2)^n \sim n!$, soit (utiliser la formule de Stirling, (7.154)) $t \sim \frac{2n}{e} (2\pi n)^{\frac{1}{2n}} \simeq n$. Évidemment, le comportement oscillatoire amorti obtenu en (7.162) ne saurait être valide dans la région $t \lesssim n$.

? Reconstituer ce calcul dans tous ces détails.

7.5 Développements asymptotiques

7.5.1 Idées générales et définitions

Le plus souvent en Physique, on ne sait trouver que des solutions approchées au problème que l'on se pose. Pour cela, il est essentiel d'analyser ce dernier et d'identifier un (parfois plusieurs) petit paramètre, ε , par rapport auquel on essaie de trouver une quantité d'intérêt sous la forme d'un développement en ε , dont les termes successifs sont de plus en plus petits grâce au fait que ε est réputé petit.

Comme déjà souligné à plusieurs reprises, ce petit paramètre *doit être sans dimension physique*⁵¹, c'est-à-dire une combinaison de grandeurs physiques constituant au total un nombre pur. Tout naturellement, les grandeurs apparaissant dans ε constituent les *échelles* physiques caractéristiques du problème posé. Par exemple, en Électrodynamique quantique (la version quantique de l'interaction entre le champ électromagnétique et la matière), le petit paramètre fondamental est une constante remarquable, la constante de structure fine α définie comme :

$$\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c} \quad , \quad (7.163)$$

où e est la charge de l'électron, ε_0 la permittivité du vide, \hbar la constante historique⁵² de Planck h divisée par 2π et c la vitesse de la lumière. Il est facile de vérifier par l'analyse dimensionnelle que α est bien un nombre pur, dont la valeur numérique (connue en fait

⁵⁰Les égalités qui suivent restent vraies pour $n \in \mathbb{C}$.

⁵¹tout comme l'argument de n'importe quelle fonction apparaissant dans un problème de Physique, à l'exception de la fonction monôme x^α .

⁵²Celle que Planck introduisit en 1900 pour résoudre les problèmes gravissimes à propos du rayonnement du corps noir ($h \simeq 6,6 \times 10^{-34} \text{J.s}^{-1}$).

avec une extraordinaire précision) est :

$$\alpha \simeq \frac{1}{137} \simeq 7 \times 10^{-4} \ll 1 . \quad (7.164)$$

Autre exemple : lorsque l'on recherche les corrections relativistes pour la description d'un phénomène impliquant des vitesses qui ne sont pas si petites que cela devant c – mais que l'on ne veut pas pour autant se compliquer la vie en recourant à la théorie relativiste complète –, on effectue des développements selon le rapport $\frac{v}{c}$, où v désigne une vitesse typique du problème. C'est ce rapport qui, en la circonstance, joue le rôle du petit paramètre ε . Ainsi, on trouve que le déplacement Doppler est donné, à l'ordre le plus bas, par $\Delta\nu = \nu_0 \frac{v}{c} \cos \theta$, dans des notations évidentes⁵³. Enfin, l'analyse de la limite thermodynamique d'un système à N constituants (atomes, molécules, ...) procède le plus souvent en effectuant des développements asymptotiques par rapport à N ; ainsi, on trouve que les grandeurs extensives, l'énergie E par exemple, ont le bon comportement : $E(N) = NE_1(1 + \dots)$ où E_1 est indépendant de N et où les \dots notent des corrections tendant vers zéro dans la limite $N \rightarrow +\infty$ (ce sont par exemple des corrections de surface qui, pour des forces à courte portée, tendent vers zéro comme $N^{-\frac{D-1}{D}}$ dans \mathbb{R}^D , donc très lentement ; c'est l'énormité du nombre d'Avogadro qui assure le succès de la Mécanique statistique – $N^{-2/3}$ ne décroît pas bien vite !).

S'agissant de trouver un développement dont les termes sont de plus en plus petits au fur et à mesure que l'on pousse le développement, le choix le plus naturel – en tout cas le plus élémentaire – consiste à mettre sur pied un développement en *puissances* de ε . Ainsi, désignant par X la quantité physique d'intérêt, l'idée est alors de trouver les quantités X_n dans l'expression :

$$X = X_0 + X_1\varepsilon + X_2\varepsilon^2 + \dots . \quad (7.165)$$

Si $0 < \varepsilon \ll 1$, et si les X_n ne sont pas trop grands, on peut espérer avoir une bonne représentation de X avec un petit nombre de termes. En Physique, une telle façon de procéder est généralement appelée méthode des perturbations. X_0 désigne la valeur (ou l'expression) de X quand $\varepsilon = 0$: c'est ce que l'on appelle l'approximation d'ordre zéro, que l'on connaît en général et qui, en principe et intuitivement, définit un bon point de départ pour une description à ε fini mais petit.

Écrire un développement tel que (7.165) soulève un certain nombre de questions, dont la toute première est : un tel développement existe-t-il ? En général, le physicien ne sait pas répondre à cette question – sauf dans les cas triviaux où le problème est soluble directement, et alors le recours aux méthodes de perturbations est artificiel, ou

⁵³Il y a toutefois une différence essentielle entre les limites non-relativiste et non-quantique : l'approche quantique \rightarrow classique est toujours *singulière*, les développements étant toujours de nature asymptotique. Schématiquement, cela tient à ce que des quantités du genre $e^{-A/\hbar}$ jouent un rôle central, et que la fonction $e^{1/z}$ a une singularité essentielle à l'origine, où il n'existe pas de série de Taylor.

À l'inverse, $(1 - v^2/c^2)^{1/2}$ possède un développement de Taylor en $c = \infty$, convergent $\forall v < c$, une condition pas vraiment gênante pour la Physique.

en tout cas un peu académique. Bien évidemment, l'existence d'un tel développement en puissances n'est jamais garantie en soi ; la seule chose que l'on sait est que :

$$\lim_{\varepsilon \rightarrow 0} X = X_0 \quad , \quad (7.166)$$

et il existe autant de fonctions que l'on veut, représentant l'écart $X - X_0$, respectant cette prescription. Par exemple, on pourrait avoir⁵⁴ :

$$X = X_0 + X_1 \varepsilon \ln \varepsilon \quad . \quad (7.167)$$

La correction tend bien vers zéro avec ε , mais n'est certainement pas développable en série entière autour de $\varepsilon = 0$. Il en irait de même si :

$$X = X_0 + X_e e^{-1/\varepsilon} \quad , \quad (7.168)$$

et on pourrait multiplier les exemples à l'infini. La caractéristique de corrections de ces derniers types est d'être *singulières* ; ni $\ln x$, ni $e^{-1/x}$ n'admettent de développement en série entière autour de $x = 0$ car ce point est ultra-singulier : c'est un point de branchement pour $\ln z$, une singularité essentielle pour $e^{-1/x}$. Les développements de Taylor hypothétiques ont l'un des coefficients tous infinis, l'autre des coefficients tous nuls. En pareil cas, on parle de perturbations *singulières* et d'autres traitements sont parfois possibles⁵⁵ – et souhaitables.

L'impossibilité usuelle de répondre à de telles questions n'interdit pas pour autant de persister dans l'idée d'écrire des développements en puissances, qui auront donc forcément, dans le cas où un vrai développement en série entière n'existe pas, un caractère un peu particulier. Il peut s'agir tout simplement de sommes partielles de séries divergentes, ou de séries dont tous les coefficients sont soit nuls, soit infinis⁵⁶ ; quelques exemples illustrant les acrobaties que l'on sait faire avec ces objets un peu exotiques seront mentionnés ultérieurement. Quoi qu'il en soit, de tels développements sont généralement dits *asymptotiques* et, convenablement manipulés, constituent un outil d'analyse extrêmement utile et performant.

Avant de définir précisément ce qu'est un développement asymptotique, une remarque s'impose. Jusqu'à présent, on a parlé de *petit* paramètre, noté ε , qui peut être l'écart $z - z_0$ entre la variable et une valeur de référence z_0 . Les mêmes questions se posent quand un *grand* paramètre a été identifié, noté A . Alors, ce que l'on tente (avec toutes les réserves déjà mentionnées) est un développement en puissances *inverses* de A :

$$X = X_0 + X_1 \frac{1}{A} + X_2 \frac{1}{A^2} + \dots \quad , \quad (7.169)$$

⁵⁴C'est par exemple le cas des métaux, où l'énergie du gaz (quantique) d'électrons admet, dans la limite des hautes densités, un développement de ce type, le petit paramètre ε étant le volume (moyen) occupé par un électron, compté en unité a_0^3 ($a_0 \simeq 0,53 \text{ \AA}$).

⁵⁵Par exemple le traitement WKB pour un système quantique qui n'est pas trop loin de la limite classique.

⁵⁶Tous les coefficients de Taylor à droite de $\ln x$ en $x = 0$ sont infinis, tous ceux de $e^{-1/x}$ sont nuls. Ceci se comprend bien intuitivement : $\ln x$ va si lentement vers $-\infty$ quand $x \rightarrow 0_+$ que si on tentait un développement (infini) en puissances de x , il faudrait y mettre des coefficients infiniment grands. De même, $e^{-1/x}$ va si vite vers zéro quand $x \rightarrow 0_+$ qu'un développement en puissances devrait avoir des coefficients infiniment petits.

X_0 étant alors la valeur de X pour $A = \infty$.

La référence à un développement en puissances n'est d'ailleurs pas une obligation – un tel développement, même asymptotique, peut tout simplement ne pas exister (un symptôme usuel : les coefficients c_n sont nuls, ou infinis...). De fait, l'idée essentielle est d'introduire des fonctions simples de référence donnant une bonne approximation de la fonction dans un voisinage donné, et permettant d'établir une *relation d'équivalence* entre deux fonctions, notée $f(z) \sim g(z)$. On écrira ainsi :

$$\boxed{f(z) \sim g(z) \iff \lim_{z \rightarrow z_0} \frac{f(z)}{g(z)} = 1} \quad (7.170)$$

On voit ainsi que la notion de développement asymptotique généralise celle de développement de Taylor : d'une part, il ne s'agit pas forcément d'un développement polynomial par rapport à la variable, d'autre part tous les développements de Taylor satisfont les conditions précisées ci-après, lesquelles servent de définition à un développement asymptotique. Par exemple, la fonction $f(x) \stackrel{\text{déf}}{=} \frac{x \ln |x|}{1+e^x}$ n'a pas de développement de Taylor au voisinage de $x = 0$, mais on peut écrire (par exemple) :

$$f(x) = \frac{x}{2} \ln |x| - \frac{x^2}{4} \ln |x| + R_2(x) \quad (7.171)$$

où le reste $R_2(x)$ est tel que $\lim_{x \rightarrow 0} \frac{R_2(x)}{x^2 \ln |x|} = 0$. On peut ainsi écrire :

$$f(x) \sim g(x) \quad \text{avec} \quad g(x) \stackrel{\text{déf}}{=} \frac{x}{2} \ln |x| - \frac{x^2}{4} \ln |x| . \quad (7.172)$$

Il convient de mentionner que l'obtention d'un développement asymptotique est parfois une affaire très difficile. Pour prendre un problème déjà rencontré – l'estimation de Gauss du nombre de nombres premiers plus petits que x , (7.123) –, rappelons qu'il s'est écoulé un siècle avant que Hadamard et de la Vallée - Poussin ne démontrent la conjecture de Gauss (voir p. 309).

Pour simplifier, on ne considère dans la suite que des développements asymptotiques en puissances (c'est d'ailleurs le cas le plus fréquent en pratique). Fixons les idées en prenant le cas où la variable prend de grandes valeurs ; quelle que soit la fonction $f(z)$, il est toujours possible d'écrire un développement du genre :

$$f(z) = \sum_{n=0}^N \frac{c_n}{z^n} + R_N(z) ; \quad (7.173)$$

tant que l'on n'a rien dit sur le reste $R_N(z)$, une telle écriture est évidemment vide de sens. Toute la question est donc de savoir ce que fait $R_N(z)$, à N fixé, quand $|z| \rightarrow +\infty$ (une fois que l'on s'est donné les moyens de calculer les c_n). On dit que $f(z)$ admet un développement asymptotique (à grand z) s'il existe de fait des coefficients c_n tels que :

$$\boxed{\lim_{z \rightarrow \infty} \left[z^N \left(f(z) - \sum_{n=0}^N \frac{c_n}{z^n} \right) \right] = 0} \quad (7.174)$$

soit :

$$\boxed{\lim_{z \rightarrow +\infty} [z^N R_N(z)] = 0} \quad (7.175)$$

le polynôme en $\frac{1}{z}$ écrit explicitement constitue le développement asymptotique de $f(z)$.

Il est essentiel de noter que la sommation contient un nombre *fini* de termes. Le sens de la condition (7.174) est le suivant : quand z est grand, la somme finie donne une valeur correcte de la fonction $f(z)$ à une erreur d'ordre z^{-N} près. Il s'agit d'une approximation *numérique*, en aucune façon une approximation en tant que fonction, au sens où on l'entend habituellement quand on considère une série convergente tronquée en un polynôme de degré donné⁵⁷. De fait, il peut arriver que, à z fixé, la somme $\sum_{n=0}^N \frac{c_n}{z^n}$ ait un comportement épouvantable quand $N \rightarrow +\infty$, en particulier divergent ou que la série entière correspondante ait un rayon de convergence *nul* ; en pareil cas, l'impérieuse nécessité de garder partout N fini saute aux yeux. Le point essentiel est que, pour un z donné assez grand en module, il existe une valeur "raisonnable" (pas trop grande) de N , soit N_ε , pour laquelle le polynôme de degré N donne de f une valeur numérique très précise (à ε près) – N_ε dépendant bien évidemment de ε . Il faut retenir le fait important que la condition (7.174) ne dit strictement rien sur le reste de la série tronquée au rang N (qui peut être ... infini), reste à ne pas confondre avec la quantité notée $R_N(z)$ dans les écritures générales (et qui n'est que la différence entre la fonction et le polynôme qui va approximer celle-ci). La somme apparaissant dans un développement asymptotique étant finie, la question de la convergence ne se pose pas ; noter toutefois que l'usage fréquent est de parler de *série asymptotique*, somme infinie le plus souvent *formelle* puisque la série en question est, en règle générale, divergente. Signalons que certains auteurs (Whittaker et Watson, par exemple, [7], § 8.2) réservent justement le qualificatif *asymptotique* aux développements qui sont l'amorce d'une série divergente. En pareil cas, on a :

$$\lim_{|z| \rightarrow +\infty} z^N R_N(z) = 0 \quad (N \text{ fixé}) , \quad \lim_{N \rightarrow +\infty} R_N(z) = \infty \quad (|z| \text{ fixé}) \quad (7.177)$$

Le symbole spécial d'équivalence \sim est utilisé pour bien marquer le caractère particulier de tels développements⁵⁸ :

$$\boxed{f(z) \sim \sum_{n=0}^N \frac{c_n}{z^n} + \mathcal{O}(z^{-N})} \quad (7.178)$$

L'écriture $\mathcal{O}(z^{-N})$ signifie très précisément que le reste $R_N(z)$ satisfait (7.175).

⁵⁷ Ainsi, quand on écrit, pour z grand devant 1 :

$$e^{1/z} \simeq \sum_{n=0}^N \frac{z^{-n}}{n!} . \quad (7.176)$$

la condition (7.174) est *aussi* satisfaite, mais c'est parce qu'une autre condition, bien plus contraignante l'est aussi : la série dont la somme finie de (7.176) est une troncature est une série convergente, pour tout z de module fini. Répétons que, au contraire, la condition (7.174) ne dit strictement rien de la série dont $\frac{c_n}{z^n}$ est le terme général.

⁵⁸ Le symbole \sim a été introduit par Poincaré. Quand de (mauvaises) habitudes se sont installées, on omet souvent l'écriture du reste $\mathcal{O}(z^{-N})$; le \mathcal{O} évoque le fait que z est *grand*.

On définit de la même façon le développement asymptotique pour les petites valeurs de la variable ; $f(z)$ admet un développement asymptotique à petit z si la condition suivante est satisfaite :

$$\lim_{z \rightarrow 0} \left[\frac{1}{z^N} \left(f(z) - \sum_{n=0}^N d_n z^n \right) \right] = 0 \quad (7.179)$$

Définissant le reste $R_N(z)$ par $f(z) = \sum_{n=0}^N d_n z^n + R_N(z)$, ceci se lit aussi :

$$\lim_{z \rightarrow 0} \left[\frac{1}{z^N} R_N(z) \right] = 0 \quad (7.180)$$

Les mêmes remarques sur N fini, etc, restent de vigueur, et on écrit alors :

$$f(z) \sim \sum_{n=0}^N d_n z^n + \mathcal{O}(z^N) \quad (7.181)$$

où le symbole \mathcal{O} rappelle le fait – si vraiment il le faut (!) – que le développement tient à *petit* z , et que $R_N(z)$ satisfait $\lim_{z \rightarrow 0} \left[\frac{1}{z^N} R_N(z) \right] = 0$.

En résumé, il est utile de bien marquer la différence entre un développement de Taylor (convergent dans un disque D de rayon R centré sur z_0) et un développement asymptotique :

- développement de Taylor :

$$f(z) = \sum_{n=0}^N c_n (z - z_0)^n + \varepsilon_N(z) \quad , \quad z \text{ fixé } \in D, \quad \lim_{N \rightarrow +\infty} \varepsilon_N(z) = 0 \quad , \quad (7.182)$$

avec $c_n = \frac{1}{n!} \left(\frac{d^n}{dz^n} f(z) \right)_{z_0}$.

- Développement asymptotique :

$$f(z) = \sum_{n=0}^N c_n (z - z_0)^n + R_N(z) \quad , \quad N \text{ fixé}, \quad \lim_{z \rightarrow z_0} (z - z_0)^{-N} R_N(z) = 0 \quad . \quad (7.183)$$

Noter qu'un développement de Taylor est un cas particulier de développement asymptotique ; en effet, la fonction $\varepsilon_N(z)$ dans (7.182) est :

$$\varepsilon_N(z) = \sum_{n=N+1}^{+\infty} c_n (z - z_0)^n \equiv (z - z_0)^{N+1} \sum_{p=0}^{+\infty} c_{p+N+1} (z - z_0)^p \quad ; \quad (7.184)$$

et satisfait visiblement $\lim_{z \rightarrow z_0} (z - z_0)^{-N} \varepsilon_N(z) = 0$.

Le chapitre VIII de la *Mécanique céleste* de Poincaré [46] commence par le commentaire suivant^a :

“Il y a entre les géomètres et les astronomes une sorte de malentendu au sujet de la signification du mot convergence. Les géomètres préoccupés de la parfaite rigueur et souvent trop indifférents à la longueur de calculs inextricables dont ils conçoivent la possibilité, sans songer à les entreprendre effectivement, disent qu’une série est convergente quand la somme des termes tend vers une limite déterminée, quand même les premiers termes diminueraient très lentement. Les astronomes, au contraire, ont coutume de dire qu’une série converge quand les 20 premiers termes, par exemple, diminuent très rapidement, quand même les termes suivants devraient croître indéfiniment. Ainsi, pour prendre un exemple simple, considérons les deux séries qui ont pour terme général :

$$\frac{1\,000^n}{n!} \qquad \frac{n!}{1\,000^n}$$

Les géomètres diront que la première converge, et même qu’elle converge rapidement, parce que le millionième terme est beaucoup plus petit que le 999 999ème ; mais ils regarderont la seconde comme divergente, parce que le terme général peut croître au-delà de toute limite. Les astronomes, au contraire, regarderont la première série comme divergente, parce que les 1 000 premiers termes vont en croissant ; et la seconde comme convergente, parce que les 1 000 premiers termes vont en décroissant et que cette décroissance est d’abord très rapide. Les deux règles sont légitimes : la première, dans les recherches théoriques ; la seconde, dans les applications numériques. Toutes deux doivent régner, mais dans deux domaines séparés et dont il importe de bien connaître les frontières”.

^aCité en http://fr.wikipedia.org/wiki/Théorie_des_perturbations.

“The series is divergent ; therefore we may be able to do something with it”

(Oliver Heaviside)

Avant de donner quelques exemples, il convient de signaler un fait important : le développement asymptotique d’une fonction donnée est unique, mais des fonctions différentes peuvent avoir des développements asymptotiques identiques – d’où le symbole \sim évoquant une classe d’équivalence. Cette possibilité est évidemment à rapprocher du fait que ces développements expriment essentiellement des coïncidences *numériques* entre une fonction et son développement asymptotique, et surtout pas *a priori* une coïncidence fonctionnelle même approximative. Plus formellement, remarquons qu’il existe des fonctions $L(z)$ dont le développement asymptotique en puissances est identiquement nul : c’est le cas quand $\lim_{|z| \rightarrow \infty} z^n L(z) = 0$ quel que soit n (par exemple e^{-z} quand $z \in \mathbb{R}_+$). Il en résulte que le développement asymptotique d’une fonction $f(z)$ est aussi celui de $f(z) + \lambda L(z)$, quelle que soit la constante λ ... À l’inverse, l’unicité du développement asymptotique d’une fonction se voit comme suit ; disposant des deux développements $\sum_n \frac{c_n}{z^n}$ et $\sum_n \frac{c'_n}{z^n}$, la condition (7.175) donne :

$$\lim_{|z| \rightarrow \infty} \left(\sum_{n=1}^N \frac{c_n}{z^n} - \sum_{n=1}^N \frac{c'_n}{z^n} \right) = 0 \tag{7.185}$$

qui ne peut être satisfaite que si $c'_n = c_n \forall n$.

On écrit parfois des bêtises⁵⁹ quand, ayant trouvé un développement du genre $X(\varepsilon) \sim 1 - \varepsilon$, on en déduit (!) $X(\varepsilon) = e^{-\varepsilon} \dots$. En pareil cas – et à défaut d’informations sur les propriétés d’analyticité de la fonction cherchée $X(\varepsilon)$ –, on aurait pu tout autant affirmer que $X(\varepsilon) = \frac{1}{1+\varepsilon}$, ou $X(\varepsilon) = 1 - \ln(1 + \varepsilon)$, etc. Il est bien clair que ces fonctions ont des propriétés analytiques franchement différentes : e^{-z} est une fonction entière, $\frac{1}{1+z}$ est une fonction méromorphe, $1 - \ln(1 + z)$ est une fonction multiforme ! Tous ces *caveat*, franchement triviaux, sont explicitement mentionnés pour attirer la vigilance de l’utilisateur. Il est bien évident que s’agissant de raisonner avec une fonction $f(z)$ dont on connaît le développement asymptotique, toutes les informations *a priori* sur les propriétés analytiques de f sont les bienvenues et constituent de précieux garde-fous.

Trois résultats sont utiles à connaître, énoncés ici sans démonstration :

1. les développements asymptotiques de la somme et du produit de deux fonctions s’obtiennent en faisant la somme et le produit des développements asymptotiques de ces deux fonctions ;
2. l’intégration terme à terme du développement asymptotique d’une fonction produit celui de sa primitive qui s’annule à l’infini⁶⁰ ; l’existence du développement :

$$f(z) \sim \sum_{n=2}^N \frac{c_n}{z^n} + \mathcal{O}(z^{-N}) , \quad (7.186)$$

entraîne que l’on a :

$$F(z) \equiv \int_z^{+\infty} f(z) dz \sim \sum_{n=2}^N \frac{c_n}{(n-1)z^{n-1}} + \mathcal{O}(z^{-(N-1)}) . \quad (7.187)$$

S’il existe un tout premier terme ($n = 1$) en $\frac{1}{z}$ dans $f(z)$, alors l’intégration produit un terme en $\ln z$ (à une constante additive près) ; l’opération est légitime (en général), mais la primitive n’a pas alors un développement asymptotique de la classe de ceux introduits antérieurement (qui ne contiennent que des puissances entières de la variable).

3. en revanche, la dérivation terme à terme du développement asymptotique ne produit pas en général le développement asymptotique de la dérivée $f'(z)$ – voir l’exemple de $e^{-x} \sin e^x$, discuté par Whittaker et Watson ([7], § 8.31, p. 153).

Comme toujours, l’intégration est une opération inoffensive, alors que la dérivation est toujours dangereuse.

Clairement, la notion de développement asymptotique est reliée aux différents traitements que l’on peut élaborer à propos des séries divergentes. Le plus connu est celui dû à Borel, qui repose sur le théorème suivant, également dû à Borel. Soit une

⁵⁹y compris dans des articles publiés dans des revues de Physique prestigieuses...

⁶⁰Bien noter que $f(z)$ est un IP d’ordre au moins égal à deux quand $z \rightarrow \infty$.

fonction $f(z) = \sum_{n \in \mathbb{N}} a_n z^n$ analytique pour $|z| \leq r < R$ et la fonction de Borel associée $\phi(z)$:

$$\boxed{\phi(z) \stackrel{\text{d\'ef}}{=} \sum_{n \in \mathbb{N}} a_n \frac{z^n}{n!}} \quad (7.188)$$

par les inégalités de Cauchy pour la dérivée d'ordre n , $|f^{(n)}(0)| \leq \frac{n!M}{R^n}$ (voir (4.86)), on a $|a_n R^n| < M$, avec M indépendant de n . Il en résulte que $\phi(z)$ satisfait :

$$|\phi(z)| < \sum_{n \in \mathbb{N}} M \frac{|z|^n}{n! R^n} = M e^{|z|/R} , \quad (7.189)$$

et plus généralement :

$$|\phi^{(n)}(z)| < \frac{M}{R^n} e^{|z|/R} \quad (n \in \mathbb{N}) . \quad (7.190)$$

Soit maintenant la fonction :

$$g(z) \stackrel{\text{d\'ef}}{=} \int_0^{+\infty} e^{-t} \phi(z t) dt ; \quad (7.191)$$

l'intégrand est borné par $M e^{-t} e^{|z|t/R} = M e^{-(1-\frac{|z|}{R})t}$, de sorte que l'intégrale existe $\forall |z| < R$, assurant que, dans les mêmes conditions, la fonction $g(z)$ est analytique. Une intégration par parties donne :

$$g(z) = [-e^{-t} \phi(z t)]_0^{+\infty} + z \int_0^{+\infty} e^{-t} \phi'(z t) dt ; \quad (7.192)$$

répétant cette opération, on obtient :

$$g(z) = \sum_{m=0}^n z^m [-e^{-t} \phi^{(m)}(z t)]_0^{+\infty} + z^{n+1} \int_0^{+\infty} e^{-t} \phi^{(n+1)}(z t) dt . \quad (7.193)$$

Comme $e^{-t} \phi^{(m)}(z t)|_{t=0} = a_m$ et que $\lim_{t \rightarrow +\infty} e^{-t} \phi^{(m)}(z t) = 0$ si $|z| < R$, il vient :

$$g(z) = \sum_{m=0}^n a_m z^m + R_n , \quad R_n \stackrel{\text{d\'ef}}{=} z^{n+1} \int_0^{+\infty} e^{-t} \phi^{(n+1)}(z t) dt . \quad (7.194)$$

Le reste est tel que :

$$|R_n| < |z|^{n+1} \int_0^{+\infty} e^{-t} \frac{M}{R^{n+1}} e^{|z|t/R} dt = M \frac{|z|^{n+1}}{R^{n+1}} \frac{1}{1 - \frac{|z|}{R}} \quad (7.195)$$

L'expression ci-dessus à droite tend vers zéro quand $n \rightarrow +\infty$ puisque $\frac{|z|}{R} < 1$, d'où $g(z) = \sum_{m \in \mathbb{N}} a_m z^m \equiv f(z) \forall |z| < R$, soit finalement $f(z) = \int_0^{+\infty} e^{-t} \phi(z t) dt$. Noter que comme $\int_0^{+\infty} t^m e^{-t} dt = m!$, ce résultat de Borel peut être rapidement retrouvé en invoquant la convergence uniforme de la série $\sum_{n \in \mathbb{N}} a_m (z t)^m / m!$, et en montrant que la légitimité de l'intégration terme à terme d'une série uniformément convergente est sauvegardée pour une intégrale impropre de cette sorte (voir [6], § 1.77). Au total, le théorème de Borel peut s'énoncer :

Soit une fonction analytique dans le disque $|z| < R$ admettant le développement en série entière $\sum_{n \in \mathbb{N}} a_n z^n$, et sa fonction de Borel associée $\phi(z)$ définie en (7.188). Alors⁶¹ :

$$f(z) = \int_0^{+\infty} e^{-t} \phi(z t) dt \quad (|z| < R) \quad (7.196)$$

Si l'intégrale existe pour des z en dehors du disque de rayon R , la série $\sum_{n \in \mathbb{N}} a_n z^n$ est dite, par définition, *Borel-sommable*. Par exemple, la série $\sum_{n \in \mathbb{N}} z^n$, qui est divergente si $|z| \geq 1$, est Borel-sommable : sa fonction de Borel associée est $\phi(z) = e^z$ et l'intégrale $\int_0^{+\infty} e^{-t} e^{zt} dt = \frac{1}{1-z}$ quel que soit $z \neq 1$.

Il existe d'autres procédés⁶² pour manipuler les séries divergentes ; l'une, proposée par Euler, permet de donner un sens à la somme de la série :

$$S = 1 - 1 + 1 - 1 + 1 - \dots \quad (7.197)$$

en la définissant comme une limite à gauche⁶³ :

$$S \stackrel{\text{déf}}{=} \lim_{x \rightarrow 1^-} \frac{1}{1+x} = \frac{1}{2} \quad (7.198)$$

Cette définition est fortement suggérée par le théorème d'Abel⁶⁴. Clairement, le problème de la sommation des séries divergentes est étroitement lié à la notion de prolongement analytique (voir section 5.5).

7.5.2 Un exemple classique

Présentons maintenant un exemple à propos d'une fonction apparaissant souvent en Physique, exemple abordé de diverses façons. Soit la fonction :

$$E(x) \stackrel{\text{déf}}{=} \int_x^{+\infty} \frac{e^{-(t-x)}}{t} dt = \int_0^{+\infty} \frac{e^{-t}}{t+x} dt \quad (7.199)$$

⁶¹En définitive, la transformation de Borel est la série obtenue en appliquant une transformée de Laplace terme à terme à la série initiale. Lorsque le calcul des transformées de Laplace peut se faire terme à terme, la sommation au sens de Borel donne le même résultat que la sommation usuelle des séries, mais la somme de Borel est définie dans de nombreux cas où cette dernière ne l'est pas. La similitude avec la transformation de Laplace saute aux yeux.

⁶²Pour une contribution assez récente, voir [48].

⁶³On peut utiliser cette régularisation pour sommer les résidus dans le calcul de l'intégrale présentée p. 253.

⁶⁴Un théorème d'Abel parmi d'autres : soit la série $S(z) \stackrel{\text{déf}}{=} \sum_{m \in \mathbb{N}} a_m z^m$ de rayon de convergence R . Si la série $\sum_{m \in \mathbb{N}} a_m$ est convergente, alors $\forall x \in [0, 1]$, $\lim_{x \rightarrow 1-0} \sum_{m \in \mathbb{N}} a_m x^m = \sum_{m \in \mathbb{N}} a_m$. L'idée est que, la somme $S(z) \stackrel{\text{déf}}{=} \lim_{N \rightarrow +\infty} S_N(z)$ étant une fonction continue pour $|z| < R$, si de surcroît la somme $S(z=1)$ existe, alors elle doit coïncider avec la limite $\lim_{x \rightarrow 1-0} S(x)$, obtenue en venant de l'intérieur du disque de convergence.

où, pour ne pas compliquer les choses, x est supposé réel et positif. Cette fonction a clairement une singularité en $x = 0$ puisqu'alors l'intégrale diverge logarithmiquement. Par ailleurs, elle tend vers zéro si $x \rightarrow +\infty$, mais n'est pas bien définie pour $x < 0$, où elle exige une régularisation laissant pressentir une partie principale de Cauchy et l'apparition d'une partie imaginaire ; on peut donc anticiper l'existence d'une singularité à l'infini pour $E(x)$, un point qui se précisera par la suite sur la forme plus maniable $I(g)$ définie en (7.207).

◆ *Remarque*

La fonction $E(x)$ définie en (7.199) est reliée simplement à l'une des fonctions spéciales répertoriées et dont les propriétés sont connues, la fonction exponentielle intégrale, notée $Ei(x)$ et définie comme :

$$Ei(x) \stackrel{\text{déf}}{=} \int_{-\infty}^x \frac{e^t}{t} dt . \tag{7.200}$$

En effet, on vérifie sans peine que $E(x) = -e^x Ei(-x)$. Sur l'expression intégrale (7.200), on voit que $Ei(-x)$ est bien définie quel que soit $x < 0$. En revanche, si on prend $x \in \mathbb{R}_+$, l'intégrale devient impropre, puisque l'intervalle d'intégration inclut le point $t = 0$, où l'intégrand diverge.

D'un autre côté, si x est un nombre complexe z , l'intégrale est toujours bien définie, à condition de déformer le contour de façon à éviter le point $t = 0$ (rappelons que déformer le contour est un procédé standard pour effectuer un prolongement analytique) : quand x se rapproche de l'origine (qui constitue un obstacle), on peut passer au-dessus de celle-ci et allonger le contour en le faisant passer juste au-dessus de l'axe réel, ou passer au-dessous en complétant le contour juste au-dessous de \mathbb{R}_+ ; ceci donne finalement deux contours complexes conjugués l'un de l'autre. Il en résulte que, pour $\Re z > 0$, $Ei(z)$ a une partie imaginaire finie, même si z est infiniment proche du demi-axe réel positif. Sur les deux contours conjugués ci-dessus, l'intégrand est réel donc l'intégrale a des valeurs complexes conjuguées. Ainsi, pour $x > 0$, $Ei(x + i0) = [Ei(x - i0)]^*$: $Ei(z)$ a donc une coupure le long de \mathbb{R}_+ ; compte tenu de l'égalité $E(x) = -e^x Ei(-x)$, $E(x)$ en a une le long de \mathbb{R}_- . $Ei(z)$ et $E(x)$ sont d'autres exemples de fonctions multiformes. ◆

On va maintenant établir le développement asymptotique de $E(x)$ pour x grand. Une intégration par parties ($u = \frac{1}{t}$, $dv = e^{-(t-x)} dt$) permet d'écrire :

$$\int_x^{+\infty} \frac{e^{-(t-x)}}{t} dt = e^x \int_x^{+\infty} \frac{e^{-t}}{t} dt = e^x \left[\left. \frac{-e^{-t}}{t} \right|_x^{+\infty} - \int_x^{+\infty} \frac{e^{-t}}{t^2} dt \right] = \frac{1}{x} - e^x \int_x^{+\infty} \frac{e^{-t}}{t^2} dt ; \tag{7.201}$$

en continuant de la sorte les intégrations par parties successives, on trouve ainsi⁶⁵ :

$$E(x) = \frac{1}{x} - \frac{1!}{x^2} + \dots + (-1)^{N-1} \frac{(N-1)!}{x^N} + R_N , \tag{7.202}$$

⁶⁵Le premier terme de (7.202) se lit $\frac{0!}{x}$.

avec :

$$R_N = (-1)^N N! \int_x^{+\infty} \frac{e^{-(t-x)}}{t^{N+1}} dt, \quad (7.203)$$

Pour le reste R_N de ce développement, une nouvelle intégration par parties donne la majoration :

$$|R_N| \equiv N! \int_x^{+\infty} \frac{e^{-(t-x)}}{t^{N+1}} dt = \frac{N!}{x^{N+1}} - (N+1)! \int_x^{+\infty} \frac{e^{-(t-x)}}{t^{N+2}} dt < \frac{N!}{x^{N+1}}; \quad (7.204)$$

on peut aussi majorer directement en notant que $e^{-(t-x)} \leq 1$ et $\frac{1}{t^{N+1}} \leq \frac{1}{x^{N+1}}$ quand $x \leq t$. Il apparaît ainsi que le reste R_N est ici plus petit que le premier terme omis⁶⁶. Par ailleurs, cette majoration montre que :

$$x^N \left| E(x) - \sum_{n=1}^N \frac{(-1)^n (n-1)!}{x^n} \right| = x^N |R_N| < \frac{N!}{x}, \quad (7.205)$$

de sorte que, à N fixé, la condition (7.174) est satisfaite. En définitive, on peut écrire :

$$\boxed{E(x) \sim \sum_{n=1}^N (-1)^{n-1} \frac{(n-1)!}{x^n} + \mathcal{O}(x^{-N}) \quad (x \gg 1)} \quad (7.206)$$

et ceci est le développement asymptotique cherché. Pour $x = 10$, la somme avec $N = 6$ est égale à 0.091 592, alors que $E(10) = 0.091\ 563 \dots$, d'où une erreur relative voisine de 3×10^{-4} . L'approximation par un polynôme en $\frac{1}{x}$ de degré fixe et égal à N est donc clairement un moyen puissant pour calculer $E(x)$ quand $x \gg 1$. Inversement, il ne faudrait pas s'amuser à écrire le reste $R_N(x)$ sous la forme de la série $\sum_{n=N+1}^{+\infty} (-1)^{n-1} \frac{(n-1)!}{x^n} \dots$ qui est monstrueusement divergente !

Présentons le même exemple, en suivant une démarche plus intuitive, comme le font souvent les physiciens ; il s'agit en fait d'un exemple célèbre connu sous le nom de série d'Euler. Soit la fonction $I(g)$ définie comme :

$$I(g) \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{e^{-t/g}}{t+1} dt \quad (g > 0) \quad (7.207)$$

Cette fonction tend visiblement vers zéro quand $g \in \mathbb{R}$ tend vers zéro par valeurs positives. À l'inverse, telle qu'elle est définie, $I(g)$ est à première vue infinie⁶⁷ si on y fait brutalement $g = 0_-$: de toute évidence, le point $g = 0$ est un point pour le moins singulier. Notons aussi que la définition (7.207) définit en fait $I(g)$ dans le demi-plan de droite $\Re g > 0$, puisque l'intégrale converge $\forall g \in \mathbb{C}, \Re g > 0$. On peut donc prolonger immédiatement et considérer d'emblée :

$$\boxed{I(g) \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{e^{-t/g}}{t+1} dt \quad (\Re g > 0)} \quad (7.208)$$

⁶⁶En utilisant la formule de Stirling (7.154), on voit que $\frac{N!}{x^{N+1}} \ll 1$ impose essentiellement $N \ll x$.

⁶⁷Cette affirmation, qui semble frappée au coin du bon sens, est pourtant fautive (voir p. 336).

$I(g)$ est visiblement reliée à la fonction $E(x)$ introduite en (7.199) ; posant $t = gt'$, il vient $I(g) = \int_0^{+\infty} \frac{e^{-t'}}{t'+\frac{1}{g}} dt'$, soit :

$$\boxed{I(g) = E\left(x = \frac{1}{g}\right)} \tag{7.209}$$

Sur la définition (7.208), on voit aussi que I a la symétrie de réflexion :

$$I(g^*) = [I(g)]^* . \tag{7.210}$$

Enfin, il est facile de trouver un majorant pour le module de $I(g)$:

$$|I(g)| \leq \int_0^{+\infty} \left| \frac{e^{-t/g}}{t+1} \right| dt < \int_0^{+\infty} |e^{-t/g}| dt = \int_0^{+\infty} e^{-t \cos \phi/|g|} dt = \frac{|g|}{\cos \phi} \tag{7.211}$$

où on a écrit $g = |g| e^{i\phi}$, $-\frac{\pi}{2} < \phi < \frac{\pi}{2}$.

Supposons que g représente une constante de couplage (la constante de structure fine par exemple), positive mais très petite devant 1. En regardant l'intégrand de (7.207), on peut dire qu'il est exponentiellement petit pour $t \gtrsim g$, c'est-à-dire qu'il s'annule très vite sur un intervalle à droite de l'origine d'amplitude $\sim g$, donc par hypothèse très petit devant 1. On fait alors une opération illicite : bien que $\frac{1}{1+t}$ ne soit développable en série entière que si $|t| < 1$, et en dépit du fait que l'intervalle d'intégration en t est $[0, +\infty[$, on écrit sauvagement (7.207) sous la forme :

$$I(g) = \int_0^{+\infty} e^{-t/g} \sum_{n=0}^{+\infty} (-1)^n t^n dt ; \tag{7.212}$$

ne reculant plus maintenant devant rien, on commute sommation et intégration pour intégrer terme à terme. En utilisant le fait que⁶⁸ $\int_0^{+\infty} t^n e^{-at} dt = \frac{n!}{a^{n+1}}$ ($\Re a > 0$), cette suite d'opérations audacieuses donne :

$$\boxed{I(g) = \sum_{n=0}^{+\infty} (-1)^n n! g^{n+1}} \tag{7.213}$$

La série est épouvantablement divergente (le terme général augmente comme la factorielle) ; néanmoins, tronquée à un ordre fini, N , elle constitue le développement asymptotique de $I(g)$ puisque la condition (7.179) est visiblement satisfaite⁶⁹. Clairement, la procédure utilisée ci-dessus est pour le moins discutable, mais on peut faire les choses beaucoup plus sérieusement, en écrivant simplement le développement limité :

$$\frac{1}{1+t} = \sum_{n=0}^{N-1} (-1)^n t^n + \frac{(-1)^N t^N}{1+t} \tag{7.214}$$

⁶⁸En posant $at = t'$, on voit que cette intégrale est $a^{-(n+1)} \int_0^{+\infty} t'^n e^{-t'} dt' = a^{-(n+1)} \Gamma(n+1)$.

⁶⁹Tout ce qui a été explicitement montré pour $E(x)$ – voir(7.206) – vaut pour $I(g)$ en vertu de (7.209).

et en reportant dans (7.208). Il vient alors :

$$I(g) = \sum_{n=0}^{N-1} (-1)^n \int_0^{+\infty} t^n e^{-t/g} dt + (-1)^N \int_0^{+\infty} \frac{t^N}{1+t} e^{-t/g} dt . \quad (7.215)$$

L'intégrale à droite est bornée par $\int_0^{+\infty} t^N e^{-t/g} dt = g^{N+1} N!$, produisant un reste $R_N(g)$ tel que $g^{-N} R_N(g)$ tend vers zéro quand $g \rightarrow 0$. Au total⁷⁰ :

$$I(g) = \sum_{n=0}^{N-1} (-1)^n g^{n+1} n! + \mathcal{O}(g^N) \quad (7.216)$$

En revenant à la série divergente (7.213) obtenue par la méthode *sauvage*, on voit que la règle est simple : on ne garde que les N premiers termes de cette série, et on oublie le reste...

En conclusion, une expression telle que :

$$I_N(g) \stackrel{\text{déf}}{=} \sum_{n=0}^N (-1)^n n! g^{n+1} \quad (7.217)$$

est tout à fait capable de reproduire *numériquement* $I(g)$ avec une très grande précision (voir fig. 7.16), sur un intervalle en g dépendant bien sûr de la valeur de N retenue. Ici, les procédés illicites utilisés ne conduisent pas à une réponse qualitativement incorrecte – pourvu que l'on réagisse convenablement en notant que la série (7.213) est divergente –, au contraire de ce qui se passe pour (7.221). La raison fondamentale est qu'ici, la singularité de l'intégrand (7.207) est en $t = -1$, *i.e.* en-dehors de l'intervalle d'intégration.

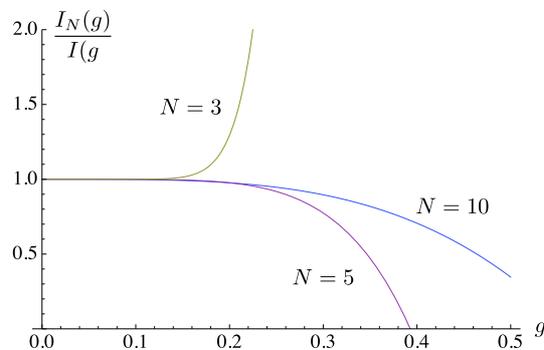


Figure 7.16: Variation du rapport $\frac{I_N(g)}{I(g)}$ avec $N = 3, 5, 10$.

On a anticipé le fait que $g = 0$ est un point singulier, compte tenu du fait que pour $g > 0$, la série en puissances entières de $I(g)$ a un rayon de convergence nul (donc

⁷⁰Clairement, ceci est en accord avec (7.206), compte tenu de (7.209).

I n'est pas analytique en $g = 0$). Avec la relation $I(g) = -E(\frac{1}{g}) = -e^{\frac{1}{g}} \text{Ei}(-\frac{1}{g})$, et selon les précisions données dans la Remarque p.332, $I(g)$ est une fonction multiforme ayant une coupure le long de \mathbb{R}_- . Ainsi, quand g est "presque" réel négatif, on a :

$$I(-|g| \pm i0) = I_1 \pm iI_2 \quad , \quad (7.218)$$

où I_2 est une quantité positive. Le diagnostic intuitif initial $I(0_+) = 0$ et $I(0_-) = +\infty$ était donc erroné (en fait, I n'est pas définie pour g strictement réel négatif). Il est fréquent que l'inattention prévienne l'infini réel quand c'est une partie imaginaire qui est cachée quelque part.

Ainsi présenté cet exemple est une autre illustration de la *sommation des séries divergentes*⁷¹, dont l'intérêt remarquable en Physique a été relevé semble-t-il pour la première fois par H. Poincaré (voir encadré p.328). Il convient de retenir l'idée-maîtresse : en prenant un nombre de termes fini (plutôt petit) d'une série divergente, on peut obtenir une approximation bien meilleure que par l'usage de séries certes convergentes, mais convergeant si lentement qu'il faudrait calculer un très grand nombre de termes pour obtenir une précision comparable à celle donnée par le développement asymptotique. La série de terme général $\frac{g^{-n}}{n!}$ converge pour tout g , $|g| > 0$ (c'est $e^{\frac{1}{g}}$!), la série $(-1)^n n! g^n$ diverge violemment, mais si $g = 10^{-3}$, cette dernière série *tronquée* fournit avec assez peu de termes une très bonne valeur numérique pour $I(g)$, alors qu'il faut un nombre astronomique de termes pour que la première donne une bonne approximation de $e^{\frac{1}{g}}$. Sur ces questions, les remarques de Titchmarsh ([6], § 7.1 à 7.3) sont particulièrement pénétrantes.

⁷¹Même pour les séries convergentes, il est bien connu que, d'un point de vue strictement numérique, il faut faire preuve d'une certaine habileté. Par exemple, on a (série harmonique alternée) :

$$\ln 2 = \sum_{n=1}^{+\infty} \frac{(-1)^n}{n} \quad ; \quad (7.219)$$

cette série converge avec une lenteur désespérante (essayer avec une calculette programmable). D'un autre côté, en développant $\ln(1-x)$ autour de $x = \frac{1}{2}$, on voit facilement que :

$$\ln 2 = \sum_{n=1}^{+\infty} \frac{1}{n 2^n} \quad , \quad (7.220)$$

Il saute aux yeux que si l'objectif est de calculer numériquement $\ln 2$, mieux vaut utiliser la représentation (7.220), qui converge à toute vitesse en comparaison de (7.219).

La raison de ces rapidités de convergence si différentes est intuitivement évidente : dans un cas on utilise un développement en série centré sur le bord du disque de convergence, dans l'autre le centre est largement à l'intérieur de ce disque.

7.6 De l'importance des singularités apparemment innocentes

Il arrive aussi que l'inattention fasse passer à côté d'un résultat important. Soit par exemple l'intégrale :

$$L(\beta E_0) \stackrel{\text{déf}}{=} \int_0^{+\infty} \frac{e^{-\beta E}}{E_0 - E} dE = \int_0^{+\infty} \frac{e^{-\beta E_0 x}}{\beta E_0 - x} dx, \quad (7.221)$$

où $0 < \beta^{-1} \ll E_0$. L'intégrand chute exponentiellement sur un intervalle $\sim \frac{1}{\beta E_0} \ll 1$ et on peut à nouveau tenter un développement asymptotique comme pour la fonction $E(x)$ ci-dessus. Ceci tire un trait sur la singularité de l'intégrand en E_0 , qui se produit très loin à droite, là où ce dernier est exponentiellement petit, mais qui néanmoins donne une partie imaginaire finie à la fonction L . C'est ainsi que Vlasov a raté l'instabilité de Landau observée dans un plasma... Plus précisément, la loi de dispersion à laquelle on arrive alors est, sous forme adimensionnée :

$$\frac{\omega^2}{\omega_p^2} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{e^{-X^2}}{1 - \lambda X^2} dX. \quad (7.222)$$

$\omega_p \stackrel{\text{déf}}{=} \sqrt{\frac{n_0 q^2}{\varepsilon_0 m}}$ est une fréquence caractéristique du plasma, fonction de la charge q , de la masse m des particules, et de la densité (moyenne) n_0 du plasma ; le facteur gaussien de l'intégrand vient de la distribution des vitesses de Maxwell. Enfin, le paramètre λ est égal à $\sqrt{\frac{2}{m}} k_B T \frac{k}{\omega}$, k étant le nombre d'onde. Dans la limite $\lambda \ll 1$, correspondant au cas où la vitesse thermique est très petite devant la vitesse de phase $\frac{\omega}{k}$, on peut être tenté de se dire qu'un développement de la fraction rationnelle est licite. Ce faisant, il reste à calculer quelques intégrales gaussiennes, et on trouve :

$$\frac{\omega^2}{\omega_p^2} \sim 1 + 3 \frac{k_B T}{m \omega_p^2} k^2 + \dots; \quad (7.223)$$

le second membre donne le comportement de la loi de dispersion des plasmons aux petits k (parabolique avec un *gap*, comme dans une théorie de champ massif). Un tel calcul, apparemment justifiable sur des bases naïves (tout est exponentiellement petit à droite dès que $X \gtrsim \lambda^{-1}$), est cependant fondamentalement incorrect, puisqu'il efface complètement une singularité.

En fait, il doit être clair que, telle qu'elle est écrite, une expression du genre (7.221) a un défaut majeur, puisqu'il y a un point, situé sur le segment d'intégration, où la divergence de l'intégrand exige une régularisation – laquelle ne peut être arbitraire, ou en tout cas conduire à différentes réponses physiques. Une telle expression, apparaissant dans un problème de Physique, est toujours le signal d'alarme que quelque chose a été mal formulé en amont ; ce qui a été oublié ici est l'écrantage de la loi de Coulomb dans le vide, caractérisé par une longueur (*longueur de Debye*, $l_D \propto T^{1/2}$) transformant le potentiel coulombien nu en un potentiel de Yukawa, dont la transformée de Fourier est analytique

à l'origine. Cela étant fait, une partie imaginaire apparaît dans la relation de dispersion, qui est précisément à l'origine de l'instabilité de Landau. L'effet est hautement singulier, et varie comme $e^{-C/T}$: il est strictement nul à température nulle.

7.7 Un exercice pour en finir

À titre d'illustration (pour plus de détails, voir [50]), traitons un exemple qui apparaît en Théorie des champs et en Mécanique statistique (notamment à propos des phénomènes critiques), et qui rassemble diverses méthodes exposées ci-dessus : on va d'abord trouver un développement asymptotique, puis montrer comment fonctionne un procédé de prolongement analytique, et enfin utiliser la méthode du col.

Soit la fonction⁷² $Z(g)$:

$$Z(g) = \int_{-\infty}^{+\infty} e^{-\beta(\frac{1}{2}x^2 + \frac{1}{4}gx^4)} dx \quad (7.225)$$

où β est un paramètre réel positif, et où g est pour l'instant considéré comme un réel positif⁷³ (c'est par exemple une constante de couplage physique). Des quantités du genre de Z apparaissent dans les problèmes de *tunnelling* et permettent aussi de construire des formalismes effectifs où une énergie libre acquiert une partie imaginaire finie, que l'on interprète comme une durée de vie d'état métastable. Plus schématiquement, la fonction $Z(g)$ peut être vue comme la "fonction de partition" d'un oscillateur anharmonique, dont l'énergie potentielle est $V(x)/V_0 = \frac{1}{2}x^2 + \frac{g}{4}x^4$. Si g est positif, il n'y a qu'un état (position) d'équilibre, en $x = 0$. Si $g < 0$, $V(x \rightarrow \pm\infty) = -\infty$, il s'agit d'un double puits renversé et la position $x = 0$ devient métastable (état discret (!) face à un *continuum* accessible à énergie constante par effet tunnel). Plus généralement, le type d'intégrale (7.225) intervient dans les modèles dits ϕ^4 : la variable x devient un champ $\phi(x)$ et l'intégrale devient une *intégrale fonctionnelle*, usuellement notée⁷⁴ $\int dx \rightarrow \int \mathcal{D}\phi$; c'est pourquoi dans la version élémentaire avec la simple variable x , on parle de "champ zéro-dimensionnel".

Notons d'abord que pour $g = 0$, Z est donné par une intégrale gaussienne et de toute évidence :

$$Z(0) = \sqrt{\frac{2\pi}{\beta}} . \quad (7.226)$$

⁷²À nouveau, il s'agit d'une fonction spéciale connue par ailleurs ; en effet, l'intégrale (7.225) peut être prise, pour $g > 0$, comme définition d'une certaine fonction de Bessel, usuellement notée K_ν , avec $\nu = -\frac{1}{4}$; plus précisément :

$$Z(g) = \frac{1}{\sqrt{2g}} e^{\frac{\beta}{8g}} K_{\frac{1}{4}}\left(\frac{\beta}{8g}\right) . \quad (7.224)$$

Cela étant, et comme K_ν est prolongeable dans le demi-plan de gauche, $Z(g)$ s'en trouve de même *ipso facto* prolongée. K_ν a une coupure sur le demi-axe réel négatif : c'est bien ce que va donner l'approximation de $Z(g)$ obtenue plus loin, voir (7.250).

⁷³En posant $x^2 = \frac{X}{\beta}$, on voit que Z est de la forme $\frac{1}{\sqrt{\beta}} \Xi\left(\frac{g}{\beta}\right)$ où $\Xi(\xi) = \int_0^{+\infty} e^{-\frac{\xi}{2} - \xi \frac{x^4}{4}} dx$.

⁷⁴Pour une splendide introduction dans le cadre quantique, voir [49].

Le but du calcul suivant est de trouver une approximation de $Z(g)$ quand g est petit devant 1. Quand $0 < g \ll 1$, la première chose qui vient à l'esprit est de tenter un développement de $Z(g)$ en puissances de g . Ainsi, en développant l'exponentielle et en échangeant sans état d'âme intégration et sommation⁷⁵, on obtient :

$$Z(g) = \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\beta x^2} e^{-\frac{1}{4}\beta g x^4} dx = \sum_{n=0}^{+\infty} \frac{(-\beta g/4)^n}{n!} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\beta x^2} x^{4n} dx . \quad (7.227)$$

En posant $\frac{\beta x^2}{2} = u$ dans l'intégrale, on retrouve la définition (7.6) de la fonction Γ , d'où :

$$Z(g) = \sqrt{\frac{2}{\beta}} \sum_{n=0}^{+\infty} (-1)^n \frac{\Gamma(2n + \frac{1}{2})}{n! \beta^n} g^n , \quad (7.228)$$

qui redonne bien (7.226) en $g=0$ puisque $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Cette série a un rayon de convergence nul, le rapport de deux termes consécutifs étant $\simeq \frac{4ng}{\beta}$ ($n \gg 1$). En revanche, la somme limitée à N termes constitue le développement asymptotique de $Z(g)$:

$$\boxed{Z(g) \sim \sqrt{\frac{2}{\beta}} \sum_{n=0}^N (-1)^n \frac{\Gamma(2n + \frac{1}{2})}{n! \beta^n} g^n} \quad (7.229)$$

puisque la condition (7.179) est satisfaite. En effet, on a :

$$\int_{-\infty}^{+\infty} dX e^{-X^2} e^{-aX^4} = \int_{-\infty}^{+\infty} dX e^{-X^2} \sum_{n=0}^N (-1)^n \frac{a^n X^{4n}}{n!} + R_N , \quad (7.230)$$

avec $R_N \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} e^{-X^2} [e^{-aX^4} \sum_{n=0}^N (-1)^n \frac{a^n X^{4n}}{n!}] dX$; comme il y a un nombre fini de termes dans la somme, on peut échanger immédiatement sommation et intégration, d'où :

$$\int_{-\infty}^{+\infty} e^{-X^2} e^{-aX^4} dX = \sum_{n=0}^N (-1)^n \frac{a^n}{n!} \int_{-\infty}^{+\infty} X^{4n} e^{-X^2} dX + R_N . \quad (7.231)$$

⁷⁵Comme la série de l'exponentielle e^z converge uniformément pour tout $|z|$ fini, une telle opération serait légitime si l'intégrale était prise entre deux bornes *finies*. Le fait que l'intégrale est *impropre* peut conduire à des séries de rayon de convergence nul : c'est le cas ici pour $Z(g)$, mais ce n'est pas systématique.

En pratique, il est souvent possible de voir d'emblée où se trouvent certaines singularités d'une fonction définie par une intégrale impropre, et c'est ce qui permet de prévoir si un développement en série sera du genre série entière ou développement asymptotique.

Par exemple, soit l'intégrale $I(a) \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} e^{-(1+a)x^2} dx$, qui est visiblement holomorphe si $\Re a > -1$. On peut calculer $I(a)$ en l'écrivant $\int_{-\infty}^{+\infty} e^{-x^2} e^{-ax^2} dx$, en développant en série e^{-ax^2} et en intégrant terme à terme. On trouve alors le développement en puissances entières de $\sqrt{\pi}(1+a)^{-\frac{1}{2}}$ autour de $a=0$, comme il se doit, dont le rayon de convergence est égal à 1 (la première singularité est en $a=-1$).

À l'inverse, l'intégrale $Z(g)$ définie en (7.224) a visiblement une singularité en $g=0$, puisqu'elle diverge si $g < 0$, et en particulier si $g=0_-$. Cette simple constatation permet de prévoir qu'il ne peut exister un développement en série entière dans un voisinage de l'origine du plan de g .

Après intégration, la somme finie donne $\sum_{n=0}^N (-1)^n \frac{a^n}{n!} \Gamma(2n + \frac{1}{2})$. Quant au reste R_N , il a pour expression :

$$R_N = \int_{-\infty}^{+\infty} dX e^{-X^2} \sum_{n=N+1}^{+\infty} (-1)^n \frac{a^n}{n!} X^{4n} ; \quad (7.232)$$

la somme de la série alternée est inférieure en module au module du premier terme, soit $\frac{a^{N+1}}{(N+1)!} X^{4(N+1)}$, d'où :

$$|R_N| < \frac{a^{N+1}}{(N+1)!} \int_{-\infty}^{+\infty} X^{4(N+1)} e^{-X^2} dX = \frac{a^{N+1}}{(N+1)!} \Gamma(2N + \frac{5}{2}) , \quad (7.233)$$

donc $|R_N| < \text{nombre} \times a^{N+1}$, assurant que la condition (7.179) est satisfaite ; elle se traduit explicitement ici par :

$$\lim_{g \rightarrow 0} \left[\frac{1}{g^N} \left(Z(g) - \sum_{n=0}^N (-1)^n \frac{\Gamma(2n + \frac{1}{2})}{n! \beta^n} g^n \right) \right] = 0 . \quad (7.234)$$

Le polynôme (7.229), pour g petit, donne une excellente approximation numérique de Z – à condition toutefois bien sûr de considérer un nombre suffisant de termes (N est une fonction de g , définie par le degré d'approximation souhaité).

La divergence de la série (7.228) est un symptôme caractéristique du fait que $Z(g)$ a une singularité en $g = 0$ (autre façon d'exprimer le fait qu'il n'existe pas de série entière convergente dans un disque fini), ce que l'on peut deviner sur la forme intégrale (7.225) en remarquant que pour $g < 0$ l'intégrale n'a pas de sens⁷⁶. Pour savoir ce qu'il en est plus précisément, il faut prolonger Z par continuité, et c'est ce que l'on va faire en modifiant *continûment* le contour d'intégration en relation avec le déplacement continu de g à partir du demi-axe \mathbb{R}_+ . À ce stade, on peut de toute façon remarquer que, sans rien changer dans la définition (7.225), la fonction $Z(g)$ est en fait bien définie pour $g \in \mathbb{C}$, pourvu que $\Re g > 0$. $Z(g)$ est donc certainement une fonction analytique quel que soit g dans le demi-plan de droite. De surcroît, on sait bien que la ligne d'intégration n'est pas rigidement fixée une fois pour toutes ; en effet, réécrivant (7.225) comme l'intégrale complexe :

$$Z(g) = \int_C e^{-\beta(\frac{1}{2}z^2 + \frac{1}{4}gz^4)} dz , \quad (7.235)$$

il est loisible de déformer le contour C sans changer le résultat, à condition de respecter les règles connues (ne pas franchir des singularités de l'intégrand, qui est ici une fonction entière de z). Par exemple, on peut incliner la droite d'un certain angle θ et refermer par deux arcs de cercle à l'infini pour revenir sur l'axe réel en $\pm\infty$, en respectant $-\frac{\pi}{2} < \text{Arg}(gz^4) < +\frac{\pi}{2}$, de sorte que les contributions des deux arcs tendent exponentiellement vers zéro quand $|z| \rightarrow +\infty$. Autrement dit, l'argument de g étant donné, α , la droite d'intégration doit se trouver dans le "cône" $-\frac{\alpha}{4} \pm \frac{\pi}{8}$.

⁷⁶On a envie de dire que Z est infinie pour $g < 0$, mais une fois encore cette conclusion naïve est fautive : Z a une coupure le long de \mathbb{R}_- et, tout simplement, n'est pas définie pour $g \in \mathbb{R}_-$.

Montrons maintenant que $Z(g)$ a une coupure sur le demi-axe réel négatif, ce qui identifie la singularité à l'origine comme un point de branchement. Supposons que g , partant d'un point de \mathbb{R}_+ , se déplace dans le plan vers $\Im g > 0$ et acquiert un argument $\alpha > 0$ ($g = |g|e^{i\alpha}$). Comme on vient de le dire, plus α croît, plus il faut incliner la droite d'intégration ; en la mettant dans l'axe du "cône", l'intégrale (7.225) devient :

$$Z(g) = \int_{C_{-\alpha/4}} dz e^{-\beta(\frac{1}{2}z^2 + \frac{1}{4}gz^4)} , \tag{7.236}$$

où $C_{-\alpha/4}$ est donc la droite inclinée de $-\frac{\alpha}{4}$ par rapport à l'axe réel (voir fig. 7.17). Quand g arrive sur l'axe réel négatif (en venant toujours de $\Im g > 0$), la droite est inclinée de $-\pi/4$, c'est la seconde bissectrice.

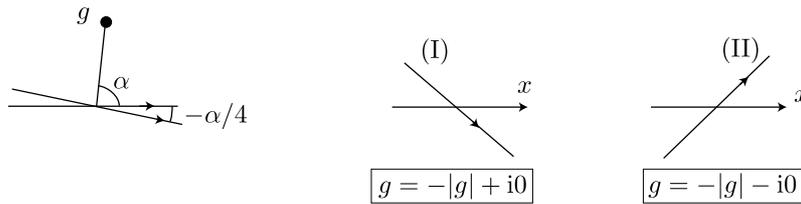


Figure 7.17: À gauche : rotation du contour d'intégration pour (7.236) quand l'argument de g est positif et égal à α . À droite : les deux bissectrices (I) et (II) correspondent respectivement à $-|g| + i0$ ($\alpha = +\pi$) et $-|g| - i0$ ($\alpha = -\pi$).

Inversement, si g part de \mathbb{R}_+ et se déplace du côté $\Im g < 0$ (l'argument α est donc négatif), la droite s'incline de $-\frac{\alpha}{4}$, et tourne cette fois dans le sens positif. Quand g arrive finalement sur \mathbb{R}_- en venant par en-dessous, la droite d'intégration est la première bissectrice. Au total, pour $g = -|g| \pm i0$, les deux chemins requis sont conjugués l'un de l'autre et en deux points complexes conjugués, l'intégrand prend des valeurs complexes conjuguées ($z = e^{\mp i\frac{\pi}{4}}s$, $s \in \mathbb{R}$). Il en résulte que les deux intégrales le long de $C_{\pm\pi/4}$ sont complexes conjuguées l'une de l'autre, d'où :

$$Z(-|g| + i0) = [Z(-|g| - i0)]^* . \tag{7.237}$$

Ceci est vrai quel que soit le module de g , autrement dit – et si, bien sûr, la partie imaginaire n'est pas nulle –, la fonction $Z(g)$ est multiforme et a une coupure le long de \mathbb{R}_- . À la réflexion, cette relation de symétrie n'est pas surprenante : vraie de toute évidence pour $\Re g > 0$ sur la définition initiale de $Z(g)$, elle doit l'être aussi pour le prolongement analytique de Z ; la différence essentielle réside dans le fait que la partie imaginaire de Z est non nulle si $g < 0$ comme on va le voir, alors qu'elle est nulle si g est positif comme le montre de toute évidence la définition première (7.225).

De fait, le point important (physiquement !) est que, quand $\Re g < 0$, $Z(g)$ a une partie imaginaire finie même si $\Im g$ est arbitrairement petit. En effet, par exemple pour $g = -|g| + i0$, on a $z = se^{-i\pi/4}$, où s varie de $-\infty$ à $+\infty$, l'intégrale donnant

$Z(g = -|g| + i0)$ est :

$$Z(g) = \int_{C_{-\frac{\pi}{4}}} e^{-\beta(\frac{1}{2}z^2 + \frac{1}{4}gz^4)} dz = \int_{-\infty}^{+\infty} e^{-\beta[\frac{1}{2}(se^{-i\frac{\pi}{4}})^2 + \frac{1}{4}g(se^{-i\frac{\pi}{4}})^4]} e^{-i\frac{\pi}{4}} ds, \quad (7.238)$$

soit :

$$Z(g) = e^{-i\frac{\pi}{4}} \int_{-\infty}^{+\infty} e^{-\beta(-\frac{1}{2}s^2 - \frac{1}{4}gs^4)} ds = e^{-i\frac{\pi}{4}} \int_{-\infty}^{+\infty} e^{-\beta(-\frac{1}{2}s^2 + \frac{1}{4}|g|s^4)} ds. \quad (7.239)$$

Il n'y a pas de doute que cette intégrale existe et est bien définie, et que sa partie imaginaire n'est pas nulle en général. En outre, si $g = 0_+$, on a $Z(0_+) = e^{-i\frac{\pi}{4}} \sqrt{\frac{\pi}{-\frac{1}{2}\beta}}$ soit $Z(0) = \sqrt{\frac{2\pi}{\beta}}$, comme il se doit : le prolongement de $Z(g)$ doit évidemment prendre la même valeur en $g = 0_+$ que la fonction de départ (voir (7.226)), où cette dernière est déjà bien définie !

L'objectif est maintenant de trouver une approximation de $Z(g)$ quand g est "négatif", mais très petit en module. Dans ces conditions, le terme en z^4 ne devient important que pour les très grandes valeurs de $|z|$, plus précisément $|gz^4| \gtrsim |z^2|$, soit $|z| \gtrsim \frac{1}{\sqrt{|g|}}$; ceci suggère de faire le changement de variable $\xi = \sqrt{|g|}z$, ce qui ne fera plus apparaître que des variables d'ordre unité. Ceci fait, on obtient (g est encore quelconque) :

$$Z(g) = \int_{C_{-\frac{\pi}{4}}} dz e^{-\beta(\frac{1}{2}z^2 + \frac{1}{4}gz^4)} = |g|^{-\frac{1}{2}} \int_{-\infty}^{+\infty} d\xi e^{-\frac{\beta}{|g|}(\frac{1}{2}\xi^2 + e^{-i\alpha}\frac{1}{4}\xi^4)} \equiv |g|^{-\frac{1}{2}} I(\lambda); \quad (7.240)$$

l'intégrale $I(\lambda)$ est de la forme $\int_C e^{\lambda f(\xi)} d\xi$ avec :

$$\lambda = \frac{\beta}{|g|}, \quad f(\xi) = -\frac{\xi^2}{2} - e^{-i\alpha}\frac{\xi^4}{4}; \quad (7.241)$$

si dès lors on suppose $\frac{\beta}{|g|} \gg 1$, on peut appliquer la méthode du col pour trouver une approximation de $Z(g)$.

Les extrema de f surviennent pour $f'(\xi) = 0$ soit :

$$-\xi - e^{-i\alpha}\xi^3 = 0 \iff \xi = 0 \text{ ou } \xi = \pm e^{i\frac{\alpha+\pi}{2}} \equiv \pm\xi_0. \quad (7.242)$$

Il y a donc ici trois points-cols ; la dérivée seconde étant $f''(\xi) = -1 - 3e^{-i\alpha}\xi^2$, aux différents points-cols on a :

$$f(0) = 0, \quad f(\xi_0) = \frac{e^{i\alpha}}{4}; \quad f''(0) = -1, \quad f''(\xi_0) = -2, \quad (7.243)$$

d'où :

$$\xi \simeq 0 : f(\xi) \simeq -\frac{1}{2}\xi^2, \quad \xi \simeq \pm\xi_0 : f(\xi) \simeq \frac{e^{i\alpha}}{4} + \frac{1}{2}(2)(\xi \pm \xi_0)^2. \quad (7.244)$$

Au voisinage des différents extrema, l'intégrand est donc approximativement :

$$\xi \simeq 0 : e^{\frac{\beta}{|\beta|}(-\frac{1}{2}\xi^2)} , \quad \xi \simeq \pm\xi_0 : e^{\frac{\beta}{|\beta|}[\frac{e^{i\alpha}}{4}+(\xi\pm\xi_0)^2]} . \quad (7.245)$$

Il en résulte que pour $\xi \simeq 0$, le chemin de plus grande pente (où $\Im f$ est localement constante) est tangent à l'axe réel, cependant que pour $\xi \simeq \pm\xi_0$, la plus grande pente est perpendiculaire à l'axe réel. L'estimation par la méthode du col consiste maintenant à déformer $C_{-\alpha/4}$ (voir fig. 7.18, à gauche) : partant de $(-\infty e^{-i\alpha/4})$ avec $\alpha > 0$ pour fixer les idées, on tire peu à peu sur le contour vers le bas⁷⁷, pour le faire "plonger" et arriver en descendant à la verticale en $-\xi_0$. Après avoir franchi le col en $-\xi_0$, on remonte vers l'axe réel de façon à arriver horizontalement en $z = 0$ par la gauche, on dépasse ce point avant de remonter pour redescendre à la verticale de $+\xi_0$ et finalement rejoindre graduellement l'asymptote $(+\infty e^{-i\alpha/4})$.

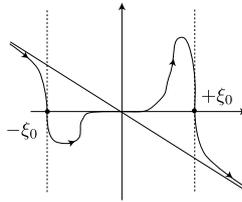


Figure 7.18: Contour utilisé pour le calcul approché de $I(\lambda)$ par la méthode du col ; la figure est tracée pour $\alpha = \pi$ ($g = -|g| + i0$), auquel cas ξ_0 est réel (et vaut $+1$).

L'approximation du col permet alors d'écrire :

$$I(\lambda) \simeq \int_{\downarrow, -\xi_0} e^{\lambda[\frac{e^{i\alpha}}{4}+(\xi+\xi_0)^2]} d\xi + \int_{\rightarrow, 0} e^{-\lambda\frac{\xi^2}{2}} d\xi + \int_{\downarrow, +\xi_0} e^{\lambda[\frac{e^{i\alpha}}{4}+(\xi-\xi_0)^2]} d\xi . \quad (7.246)$$

En introduisant $\xi = \pm\xi_0 + iy$ pour $\xi \simeq \pm\xi_0$ et $\xi = x$ pour $\xi \simeq 0$ et rejetant les bornes à l'infini autour de chaque col, il vient :

$$I(\lambda) \simeq \int_{+\infty}^{-\infty} e^{\lambda(\frac{e^{i\alpha}}{4}-y^2)} idy + \int_{-\infty}^{+\infty} e^{-\lambda\frac{x^2}{2}} dx + \int_{+\infty}^{-\infty} e^{\lambda(\frac{e^{i\alpha}}{4}-y^2)} idy . \quad (7.247)$$

En effectuant finalement les intégrales gaussiennes réelles, il vient ainsi, pour $\alpha > 0$ (mais $\alpha < \pi$) :

$$I(\lambda) \simeq \sqrt{\frac{2\pi}{\lambda}} - 2ie^{\lambda\frac{e^{i\alpha}}{4}} \sqrt{\frac{\pi}{\lambda}} . \quad (7.248)$$

Si α est négatif (mais plus grand que $-\pi$), le sens des intégrales en y est inversé, et la partie imaginaire change de signe. Au total :

$$I(\lambda) \simeq \sqrt{\frac{2\pi}{\lambda}} - (\operatorname{sgn} \alpha) 2ie^{\lambda\frac{e^{i\alpha}}{4}} \sqrt{\frac{\pi}{\lambda}} , \quad (7.249)$$

⁷⁷Se remémorant l'ambiguïté de signe relevée lors de la présentation générale, on peut dire que le choix du contour est (aussi) dicté par le goût de la simplicité, et pour prendre à contre-pied l'adage militaro-bureaucratique "Pourquoi faire simple quand on peut faire compliqué ?".

d'où, revenant à $g < 0$:

$$Z(g = -|g| \pm i0) \simeq \sqrt{\frac{2\pi}{\beta}} \left(1 \mp i\sqrt{2} e^{-\frac{\beta}{4|g|}} \right) \quad (7.250)$$

Bien sûr, la partie imaginaire de Z est exponentiellement petite, mais il ne faut pas la négliger⁷⁸, car c'est elle qui apporte l'effet désiré : la partie imaginaire qui apparaît dans l'énergie libre $F = -k_B T \ln Z$ pourra être interprétée comme une durée de vie d'état métastable. Par ailleurs, l'approximation du col ne donne pas de correction à la partie réelle, par rapport à sa valeur en $g = 0$. Enfin, pour $g > 0$, la même approximation donne $Z \in \mathbb{R}$: alors les cols sont en $\xi = 0, \pm i$, tous les chemins de plus grande pente sont parallèles à l'axe réel, les intégrales gaussiennes venant des cols en $\pm i$ sont complexes conjuguées (celle du col en $\xi = 0$ est réelle), et le résultat net est un nombre réel, comme il se doit.

⁷⁸En revanche, toute correction exponentiellement petite dans la partie *réelle* peut être omise – au moins dans un premier temps. Dans le contexte de la Physique atomique, une telle correction serait le *Lamb shift*, tandis que la partie imaginaire donnerait (essentiellement) l'inverse de la durée de vie de l'atome à l'état excité.

Chapitre 8

Analyse de Fourier

“Perhaps the most surprising thing about mathematics
is that it is so surprising.”

(Edward Charles TITCHMARSH, 1899–1963)

<i>Le but de ce chapitre est de faire des rappels sur les séries de Fourier et d'introduire la transformation de Fourier, avant d'en présenter les premières applications.</i>
--

L'analyse de Fourier (ou analyse harmonique) joue un rôle de tout premier plan en Physique, notamment pour l'étude des systèmes linéaires. Par définition, ces systèmes sont régis par des équations linéaires, caractérisées par le fait que toute combinaison linéaire quelconque de solutions est encore solution. Cette propriété est de première importance, et assure que l'ensemble des solutions peut être muni d'une structure d'espace vectoriel, dont une base est formée par un ensemble complet de solutions linéairement indépendantes. Bien évidemment, les équations régissant le comportement des grandeurs physiques d'intérêt ne sont pas toujours linéaires, loin de là (en fait, la linéarité est plutôt l'exception). Les infinies richesses de comportement des systèmes réels sont dues presque toujours aux *non-linéarités* – pour ne citer qu'un exemple, célèbre : l'application logistique et ses extravagances, notamment la cascade de Feigenbaum qu'elle introduit, l'un des scénarios reconnus de la transition vers le chaos.

Pour fixer les idées, considérons la dynamique d'une particule classique de masse m se déplaçant sur une droite (un seul degré de liberté, l'abscisse x comptée à partir d'une certaine origine). Toute la dynamique est décrite par une équation différentielle exprimant la relation entre la coordonnée x , la vitesse $\frac{dx}{dt} \equiv \dot{x}$, l'accélération $\frac{d^2x}{dt^2} \equiv \ddot{x}$ et la force totale F agissant sur la particule :

$$m\ddot{x} = F \quad , \quad (8.1)$$

Dans le cas le plus courant¹, F ne dépend que de la coordonnée et est une certaine fonction $F(x)$. À une dimension d'espace, on peut toujours poser que $F(x)$ est la dérivée d'une fonction, $F(x) = -V'(x)$; autrement dit, à une dimension et pour des forces "ordinaires", on peut toujours supposer que la force dérive d'un potentiel au sens usuel², et $V(x)$ désigne alors l'énergie potentielle de la particule dans le champ de forces considéré. Alors, (8.1) devient :

$$\boxed{m\ddot{x} = -V'(x)} \quad (8.2)$$

Dans le cas général, $V'(x)$ ne dépend pas linéairement de x : c'est une fonction quelconque, dont les zéros (quand ils existent) définissent les extrema de $V(x)$, donc les positions d'équilibre ($V'(x_0) = 0 \iff$ force nulle). Si $V''(x_0) > 0$, l'équilibre est stable (force de *rappel*), il est instable si $V''(x_0) < 0$. Par exemple, si l'énergie potentielle est de la forme :

$$V(x) = V_0 \left[\left(\frac{x}{a}\right)^4 - g\left(\frac{x}{a}\right)^2 + \lambda\frac{x}{a} \right] \quad (V_0 > 0, g > 0, \lambda > 0, a > 0) \quad (8.3)$$

$V'(x)$ est un polynôme du troisième degré. (8.3) décrit un champ de force où la particule a trois positions d'équilibre si g est assez grand (trouver la condition précise), dont l'une est instable (laquelle ?).

L'équation (8.2) n'est linéaire que si $V(x)$ est au plus quadratique en x , ce qui contient les deux cas physiquement importants :

1. particule dans un champ de force constant (champ de pesanteur terrestre dans l'approximation locale, particule chargée dans un champ électrique constant, ...) ; alors $V(x) \propto x$ et $V'(x) = C^{\text{ste}}$
2. $V(x) \propto x^2$ (oscillateur harmonique), et alors $V'(x) \propto x$.

Lorsque $V(x)$ est une fonction quelconque, le système est intrinsèquement non-linéaire et possède en général une grande richesse de comportements, dont l'étude détaillée est parfois très difficile. D'un autre côté, si on se contente d'examiner le mouvement près d'un point d'équilibre, on peut adopter une hypothèse de *petites oscillations*, qui restitue *de facto* un cadre linéaire à l'intérieur duquel le système devient (localement) harmonique.

¹Ceci exclut le cas d'une particule chargée dans un champ électromagnétique, soumise à la force de Lorentz $\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$, et celui d'une force de frottement dépendant de la vitesse.

²À une dimension, il n'y a qu'un seul chemin allant directement d'un point à un autre. La question de savoir si le travail $W_{A \rightarrow B}$ de la force dépend du chemin suivi entre A et B ne se pose donc même pas. $W_{A \rightarrow B} \stackrel{\text{déf}}{=} \int_{x_A}^{x_B} F(x) dx$ est ainsi la différence des valeurs d'une primitive, c'est $-[V(x_B) - V(x_A)]$, d'où $F(x) = -V'(x)$ pour tout champ de forces dans \mathbb{R} .

Dans \mathbb{R}^3 , le travail $dW = F_x dx + F_y dy + F_z dz$ ne dépend pas du chemin suivi ssi il existe une fonction $-V(x, y, z)$ dont les trois composantes de la force sont les dérivées partielles par rapport à x, y et z respectivement ; dès lors, $dW = -\vec{\nabla} V \cdot d\vec{r}$ et $W_{A \rightarrow B} = V(A) - V(B)$ ne dépend que des valeurs de la fonction V aux extrémités du chemin.

En particulier, en dimension $D = 2$, le travail est indépendant du chemin suivi si les composantes F_x et $-F_y$ de la force sont respectivement les parties réelle u et imaginaire v d'une fonction holomorphe $f = u + iv$ (ou si $F_x = v$ et $F_y = -u$).

En effet, si x_m désigne l'une des positions d'équilibre stable, et si on se contente de chercher les petits écarts³ $|x - x_m| \ll a$, alors il suffit de faire un développement de Taylor autour de x_m :

$$V(x) \simeq V(x_m) + \frac{1}{2}(x - x_m)^2 V''(x_m) . \quad (8.4)$$

En posant $V''(x_m) = m\omega_m^2$, on fait apparaître la pulsation relative au minimum de V considéré (la fréquence correspondante s'appelle *fréquence de fond de puits*). Si le fond de puits est assez plat (oscillateur mou), la fréquence de fond de puits est petite, et inversement ; en reportant l'expression simplifiée (8.4) au second membre de (8.2), on obtient un oscillateur harmonique. Bien sûr, tout ceci ne vaut que pour les petites oscillations⁴, et la condition de validité du développement tronqué (8.4) doit être précisée au cas par cas ; elle s'exprimera souvent comme⁵ :

$$|(x - x_m)V'''(x_m)| \ll |V''(x_m)| . \quad (8.5)$$

L'exemple traité ci-dessus repose sur une équation *différentielle*, mais l'analyse harmonique conserve ses droits et son importance quand les équations dynamiques sont des équations aux dérivées partielles, à condition toujours qu'elles soient linéaires⁶ ; c'est le cas par exemple pour un système quantique, régi par l'équation de Schrödinger, ou pour les phénomènes de diffusion classique libre (diffusion d'une densité de particules, diffusion de la chaleur⁷), ou pour les champs régis par les équations de Maxwell. Enfin, l'analyse harmonique est d'usage presque systématique pour tous les systèmes traités dans un cadre de réponse linéaire, comme les systèmes électroniques au voisinage de leur point de fonctionnement statique, ou les systèmes macroscopiques caractérisables par une susceptibilité (Théorie de la réponse linéaire), que ce soit pour qualifier leur comportement sous l'effet d'un champ statique (magnétique par exemple) ou d'une sonde les "éclairant" pour voir ce qu'ils sont (diffusion de la lumière, des rayons X, des neutrons, ...). Ces quelques exemples devraient suffire à convaincre de l'extrême importance et de l'omniprésence de l'analyse harmonique en Physique, esquissée dans la suite de ce chapitre.

³ a désigne généralement une longueur caractéristique de $V(x)$, c'est la grandeur qui figure dans l'exemple donné en (8.3).

⁴Penser au pendule simple, à qui il n'est pas interdit de faire des tours complets, si on le lance avec une vitesse suffisante (laquelle ?) !

⁵Cette condition est d'ordre purement mathématique et porte exclusivement sur les caractéristiques de la fonction $V(x)$. Dans d'autres contextes, elle devra être complétée par des considérations physiques. Par exemple, si la particule est soumise à l'agitation thermique et que le minimum considéré est séparé de son(s) "premier(s) voisin(s)" par une barrière d'énergie ΔE , l'approximation ci-dessus ne vaut que si $k_B T \ll \Delta E$.

⁶C'est pourquoi l'analyse de Fourier peut encore jouer un rôle important pour étudier la stabilité des points fixes d'une application non linéaire.

⁷Pour la diffusion, l'hypothèse de Fick (proportionnalité entre courant et gradient) est nécessaire. De même, pour l'équation de la chaleur, il faut d'une part pouvoir définir la notion d'équilibre local (la température est une grandeur essentiellement statistique) et faire à nouveau une hypothèse de linéarité (courant d'énergie proportionnel au gradient de température). Dans les deux cas, le courant est bien sûr dirigé en sens contraire du gradient.

Historiquement, c'est Fourier (à la suite de Bernoulli) qui a introduit les séries trigonométriques (1807) pour trouver les solutions de l'équation de la chaleur, en tous points analogues à l'équation de la diffusion (résultant de la loi de Fick) :

$$\boxed{\frac{\partial P}{\partial t} = D \frac{\partial^2 P}{\partial x^2}} \quad (8.6)$$

Il s'agit d'une équation linéaire dont la solution générale peut s'obtenir en combinaison linéaire de solutions particulières. Il existe visiblement des solutions particulières de la forme $e^{\Phi(x,t)}$, où Φ est une forme linéaire de x et t . Posant plus précisément (pour la simple commodité) $\Phi(x,t) = \omega t + ikx$, on voit que ω et k doivent satisfaire l'égalité :

$$\boxed{\omega = -Dk^2} \quad (8.7)$$

qui constitue la relation de dispersion $\omega(k)$ des modes propres de cette équation. Une solution générale⁸ est ainsi de la forme $P(x,t) = \sum_k p_k e^{-Dk^2 t + ikx}$, les constantes p_k dépendant des conditions aux limites et de la condition initiale. Par exemple, si la diffusion est confinée sur le segment $[0, a] \subset \mathbb{R}$ entre deux murs parfaitement réfléchissants, les valeurs de k sont restreintes aux multiples entiers de $q = \frac{\pi}{a}$. En effet, dans ces conditions, le courant de probabilité est nul en $x=0, a$ (tout ce qui arrive repart), ce qui se traduit par $\partial_x P|_{x=0,a} = 0$ à tout instant ; il faut donc satisfaire les deux séries d'égalités $\sum_k ik p_k e^{-Dk^2 t} = 0$ et $\sum_k ik p_k e^{-Dk^2 t} e^{ika} = 0$. Les exponentielles $e^{-Dk^2 t}$ étant linéairement indépendantes, la première série donne $k(p_k - p_{-k}) = 0$ d'où $p_k = p_{-k}$, la deuxième donnant alors $p_k \sin ka = 0$, d'où $k = \text{entier} \times \frac{\pi}{a}$, la solution triviale $p_k = 0$ étant évidemment exclue car sans intérêt. C'est bien la condition aux limites de courant nul qui *quantifie* les valeurs possibles de k . À ce stade, la solution est de la forme :

$$\boxed{P(x,t) = \sum_{n \in \mathbb{N}} p_n e^{-Dn^2 q^2 t} \cos nqx} \quad (0 \leq x \leq a) \quad (8.8)$$

les p_n s'obtiennent par calage sur l'état initial $P(x, t=0)$ supposé connu, en inversant l'équation $P(x, 0) = \sum_{n \in \mathbb{N}} p_n \cos nqx$ grâce à la relation d'orthogonalité des cosinus $\int_0^a \cos n\pi \frac{x}{a} \cos n'\pi \frac{x}{a} dx = \frac{a}{2} \delta_{nn'}$; les coefficients p_n ressortent alors sous la forme :

$$p_n = \frac{2}{a} \int_0^a \cos nqx P(x, 0) dx, \quad (8.9)$$

ce qui achève la détermination complète de $P(x, t)$. L'expression (8.8) est visiblement une série de Fourier en x , la fonction étant restreinte à la période fondamentale $[0, a]$.

Si l'analyse de Fourier constitue un outil de grande utilité pour la résolution des équations linéaires, il ne faut pas toutefois en conclure qu'elle est sans importance pour les systèmes non-linéaires. D'une part elle permet, après linéarisation, d'élucider la dynamique au voisinage des points fixes, d'autre part, une fois *résolues* les équations dynamiques d'un système non-linéaire par d'autres méthodes, rien n'empêche d'effectuer

⁸Une telle équation aux dérivées partielles a bien d'autres solutions qui ne sont pas de cette forme.

l'analyse harmonique des solutions ainsi trouvées. Par exemple, une fois obtenues les trajectoires du problème de Kepler par intégration directe de l'équation fondamentale de la Dynamique, il est fort possible (et très instructif) de déterminer les séries de Fourier des mouvements liés, périodiques par nature, – on trouve d'ailleurs des développements remarquables, les séries de Kapteyn.

8.1 Rappels sur les séries de Fourier

Soit $f(\phi)$ une fonction à valeurs réelles définie sur \mathbb{R} , sauf éventuellement en un certain nombre de points. $f(\phi)$ est dite périodique s'il existe un nombre θ , appelé *période* tel que l'on ait :

$$\boxed{f(\phi + p\theta) = f(\phi) \quad \forall p \in \mathbb{Z}} \quad (8.10)$$

Il est bien clair que si θ est une période, il en va de même de $10\theta, -3\theta, \dots$. Usuellement, on appelle *période* la plus petite période positive, ici notée θ , et la fonction est dite θ -périodique.

Si ϕ a la dimension d'une longueur, θ est aussi une longueur (p est un *nombre* !). f est par exemple l'énergie potentielle $V(x)$ d'une particule ayant la coordonnée x , se déplaçant dans un paysage répétitif dans l'espace (réseau) admettant la période spatiale a : quand on translate le paysage de $a, 2a, -5a, \dots$ il se superpose à lui-même, d'où $V(x + pa) = V(x)$. Techniquement parlant, c'est le cas quand $V(x)$ est de la forme⁹ :

$$V(x) = \sum_{n \in \mathbb{Z}} v(x - na) , \quad (8.11)$$

où $v(x)$ est une fonction donnée (en forme de puits par exemple) ; en effet :

$$V(x + pa) = \sum_{n \in \mathbb{Z}} v(x + pa - na) \stackrel{(n' = n - p)}{=} \sum_{n' \in \mathbb{Z}} v(x - n'a) \equiv V(x) . \quad (8.12)$$

f peut aussi désigner la variation spatiale à temps fixé de la composante d'une onde plane $\cos(\omega t - \vec{k} \cdot \vec{r})$, ou $\cos(\omega t - kx)$. La période est alors une longueur, appelée *longueur d'onde* λ , grandeur à laquelle on associe traditionnellement le nombre d'onde k :

$$k \stackrel{\text{d'éf}}{=} \frac{2\pi}{\lambda} ; \quad (8.13)$$

à une dimension d'espace, on a alors :

$$f(x + p\lambda) = f(x) \quad \forall p \in \mathbb{Z} . \quad (8.14)$$

Lorsque la variable désigne le temps, il est naturel de noter $f(t)$ la valeur de la fonction f en t , et la (plus petite) période est souvent notée T . $f(t)$ décrit alors un

⁹Dans un cristal, $v(x)$ est l'énergie potentielle due au champ de force d'un seul atome, celui situé à l'origine par exemple.

phénomène qui se répète identiquement à lui-même (cycle) au bout de l'intervalle de temps T :

$$f(t+T) = f(t) \implies f(t+pT) = f(t) \quad \forall p \in \mathbb{Z} . \quad (8.15)$$

À la période T , on associe la fréquence ν et la pulsation ω définies comme suit :

$$\nu = \frac{1}{T} , \quad \omega = \frac{2\pi}{T} = 2\pi\nu ; \quad (8.16)$$

La fréquence ν est donc le nombre de cycles effectués par unité de temps, et s'exprime en Hertz (Hz) ou ses multiples (kHz, MHz, ...). La pulsation ω s'exprime en radians par seconde (rad/s). À l'aide des fréquence ν , période T et longueur d'onde λ , une onde plane $\cos(\omega t - kx)$ s'écrit $\cos 2\pi(\nu t - \frac{x}{\lambda})$ ou encore $\cos 2\pi(\frac{t}{T} - \frac{x}{\lambda})$.

Il est bien clair que, *stricto sensu*, la notion même de fonction périodique est une vue de l'esprit pour le physicien et repose sur un saut conceptuel. Très précisément, on dira en Physique qu'une fonction est périodique si elle est définie sur un intervalle qui peut être rendu arbitrairement grand par rapport à la plus petite période physique définie au préalable. Aucun phénomène ne dure infiniment longtemps¹⁰ : le physicien considèrera que la fonction $f(t)$ est T -périodique¹¹ si elle décrit un phénomène durant très longtemps par rapport à T , autrement dit on doit avoir le temps de compter un très grand nombre de périodes¹² (voir fig. 8.1). De la même façon, pour un cristal cubique *macroscopique* de côté L , le potentiel $V(x)$ sera considéré comme périodique puisque $L \gg a$, côté de la maille cristalline ($\frac{L}{a} \sim \mathcal{N} \sim 10^{23}$). Le saut conceptuel à l'infini s'en trouve légitimé et avec la définition (8.10), le domaine de définition de f s'étend de $-\infty$ à $+\infty$, en-dehors bien sûr des éventuels points où f présente des singularités.

Cela étant précisé, la notion de série de Fourier s'introduit tout naturellement quand on considère une fonction périodique ; pour fixer les idées, soit t la variable temps, $f(t)$ étant alors dite T -périodique. En effet, on connaît par ailleurs deux types de fonctions périodiques élémentaires, les sinus et les cosinus ; plus précisément ici, on peut construire deux familles, $s_n(t) = \sin n\omega t$ et $c_n(t) = \cos n\omega t$ ($n \in \mathbb{N}$), où $\omega = \frac{2\pi}{T}$, fonctions qui ont aussi pour période T (pour elles, ce n'est pas la plus petite période), satisfaisant également (8.15). La question qui vient alors à l'esprit est : dans quelle mesure peut-on *représenter* (i.e. exprimer) une fonction périodique telle que f sous la forme d'une somme (série ?) – c'est-à-dire d'une combinaison linéaire – des s_n et des c_n ? Explicitement, la

¹⁰De même, aucun cristal réel n'est infiniment grand.

¹¹En pareil cas, on évitera la terminologie *quasi-périodique*, un terme le plus souvent utilisé pour qualifier une fonction contenant des termes dont les périodes individuelles sont dans un rapport *irrationnel*, par exemple $\cos \phi + \cos \sqrt{2}\phi$.

¹²Pour une telle fonction, il n'existe pas de série de Fourier – la fonction n'est pas *strictement* périodique –, mais on peut généraliser tout ceci en considérant la transformée de Fourier de f , $F(\omega)$, qui est en quelque sorte la version continue du spectre de Fourier $\{f_n\}_n$. En pareilles circonstances, $F(\omega)$ est une fonction très pointue (en forme de résonance), le maximum survenant en $\omega_0 = \frac{2\pi}{T}$ et ayant une largeur $\simeq \frac{1}{\text{durée}} \ll \omega_0$.

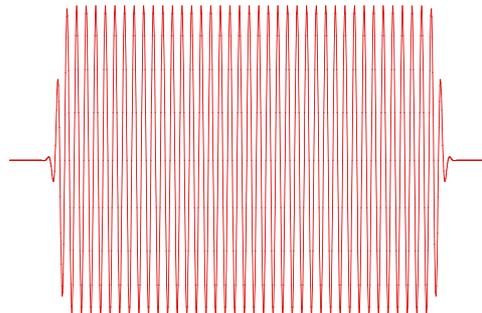


Figure 8.1: Allure d'une fonction considérée comme presque exactement périodique parce que sa durée est très grande devant l'intervalle de temps séparant deux oscillations consécutives

question est de préciser les conditions permettant d'écrire une expression telle¹³ :

$$f(t) = c + \sum_{n \in \mathbb{N}^*} (a_n \cos n\omega t + b_n \sin n\omega t) , \quad (8.17)$$

où c , les a_n et les b_n constituent par définition les coefficients de Fourier de la fonction f . Une série telle que (8.17) est un exemple¹⁴ de *série trigonométrique*. Si la fonction $f(t)$ est paire, seuls les cosinus apparaissent, si elle est impaire, il n'y a que des sinus ; cette propriété de symétrie se traduit ainsi par :

$$f(-t) = f(t) : b_n = 0 \quad \forall n , \quad f(-t) = -f(t) : c = 0, a_n = 0 \quad \forall n \in \mathbb{N} \quad (8.19)$$

En outre, comme la moyenne sur une période des sinus et des cosinus est nulle, la constante c est la moyenne (ici temporelle) de f , soit $c = \frac{1}{T} \int_0^T f(t) dt \equiv \bar{f}$.

¹³La sommation se fait sur \mathbb{N}^* puisque les sinus et cosinus ont une parité déterminée et que, de toute évidence, mieux vaut n'utiliser que des fonctions linéairement indépendantes.

¹⁴Toutes les séries trigonométriques ne sont pas de cette forme. Par exemple, la célèbre fonction de Weierstrass :

$$\mathcal{W}(x) = \sum_{n \in \mathbb{N}} b^n \cos(a^n \pi x) \quad (8.18)$$

est un autre exemple de série trigonométrique. Sa célébrité tient au fait qu'il s'agit du premier exemple historiquement publié de fonction continue partout et nulle part dérivable (on a retrouvé des traces de tels "monstres" dans des notes non publiées de Riemann, plusieurs décennies auparavant) : en 1872, Weierstrass a démontré que si $|b| < 1$, si a est un entier impair et si $ab > 1 + \frac{3\pi}{2}$, la dérivée n'existe pour aucune valeur de x , alors que pour $|b| < 1$, \mathcal{W} est visiblement continue (somme d'une série uniformément convergente). Par la suite, Hardy (1916) a montré que cette propriété est vraie pour $0 < b < 1$, $ab \geq 1$ et $a > 1$, pas forcément entier impair.

Un autre exemple de séries trigonométriques : les séries de Gauss, de la forme $\sum_{n \in \mathbb{Z}} f_n e^{in^2 x}$ qui sont, elles aussi, des fonctions 2π -périodiques. Une série de Gauss est comme une série de Fourier dont on aurait supprimé les harmoniques dont le rang n'est pas un carré parfait ; si les f_n décroissent lentement, la variation en x de la série est extrêmement irrégulière (voir fig. 8.2, mais faire soi-même l'expérience numérique avec $f_n = q^n$ et q légèrement inférieur à 1). Dans le cas le plus "simple" où $f_n = 1 \forall n$, on obtient la fonction $\vartheta_3(x)$ de Jacobi.

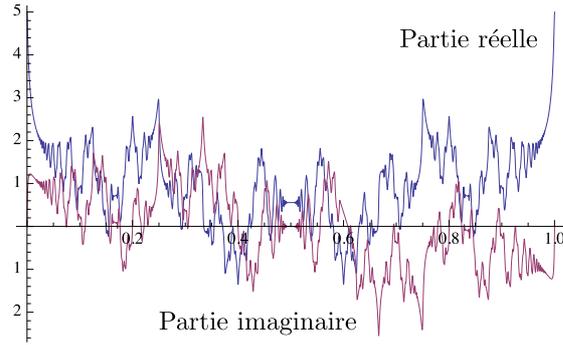


Figure 8.2: Série de Gauss $\sum_n q^n e^{in^2 x}$ avec $q = 0.8$; la variable en abscisse est $\frac{x}{2\pi}$.

Un développement tel que (8.17) est remarquable au sens où il exprime $f(t)$ en combinaison linéaire des lignes trigonométriques élémentaires contenant exclusivement la pulsation fondamentale et toutes ses harmoniques (entier $\times \omega$) – d'où la terminologie *analyse harmonique*. Il ne fait pas de doute que, au vu de (8.17), la condition caractéristique de *périodicité* de f , (8.10), est satisfaite. En outre, si un tel développement existe, il faudra en temps utile établir des formules permettant de calculer les coefficients c , a_n et b_n .

À l'aide des formules d'Euler, il est possible de mettre la série (8.17) sous une forme plus compacte, et souvent beaucoup plus commode en pratique :

$$f(t) = \sum_{n=-\infty}^{+\infty} f_n e^{in\omega t} \quad (8.20)$$

où la série au second membre doit être comprise comme :

$$\lim_{N \rightarrow \infty} \sum_{n=-N}^{+N} f_n e^{in\omega t} . \quad (8.21)$$

En prenant les complexes conjugués et en changeant la variable muette de n en $-n$, on voit que :

$$[f(t)]^* = \sum_{n=-\infty}^{+\infty} f_{-n}^* e^{in\omega t} ; \quad (8.22)$$

il en résulte que pour une fonction f à valeurs réelles, les coefficients de Fourier f_n ont la propriété de symétrie :

$$f(t) \in \mathbb{R} \forall t \quad : \quad f_n = f_{-n}^* \quad (8.23)$$

L'ensemble des coefficients $\{f_n\}$ constitue le *spectre de Fourier* de $f(t)$. Les relations entre les divers coefficients sont :

$$c = f_0 , \quad a_n = f_n + f_{-n} , \quad b_n = i(f_n - f_{-n}) . \quad (8.24)$$

Abordons maintenant la question de trouver les coefficients du développement de Fourier, la fonction $f(t)$ étant donnée. Pour cela, supposons d'abord que la série de Fourier (8.20) est uniformément convergente, auquel cas sa somme, la fonction $f(t)$, est continue. Dans ces conditions, le calcul des coefficients f_n est immédiat. En effet, multiplions (8.20) par $e^{-im\omega t}$ et intégrons membre à membre en profitant de la convergence uniforme pour permuter intégration et sommation infinie :

$$\int_0^T dt e^{-im\omega t} f(t) = \int_0^T dt e^{-im\omega t} \sum_{n=-\infty}^{+\infty} f_n e^{in\omega t} = \sum_{n=-\infty}^{+\infty} \int_0^T dt f_n e^{i(n-m)\omega t} . \quad (8.25)$$

On a, si $n \neq m$:

$$\int_0^T e^{i(n-m)\omega t} dt = \frac{e^{i(n-m)\omega T} - 1}{i(n-m)\omega} = 0 , \quad (8.26)$$

puisque $\omega T = 2\pi$ et que $e^{in2\pi} = 1 \forall n \in \mathbb{Z}$, cependant que si $n = m$, $\int_0^T dt = T$, d'où¹⁵ :

$$\int_0^T e^{i(n-m)\omega t} dt = T\delta_{nm} , \quad (8.27)$$

où δ_{nm} est le symbole de Kronecker qui vaut 1 si $n = m$, 0 autrement. Revenant à (8.25), il vient :

$$\int_0^T dt e^{-im\omega t} f(t) = \sum_{n=-\infty}^{+\infty} f_n T\delta_{nm} = T f_m , \quad (8.28)$$

d'où la formule d'inversion donnant les coefficients du développement (8.20) :

$$\boxed{f_m = \frac{1}{T} \int_0^T e^{-im\omega t} f(t) dt} \quad (8.29)$$

Ceci permet, pour une série uniformément convergente dont la somme est de ce fait une fonction continue, de calculer les coefficients de Fourier f_m et d'écrire explicitement la série de Fourier. Les coefficients de la série en sinus et cosinus se déduisent facilement de (8.29) et on trouve :

$$c_0 = \frac{1}{T} \int_0^T dt f(t) , \quad a_n = \frac{2}{T} \int_0^T dt f(t) \cos n\omega t , \quad b_n = \frac{2}{T} \int_0^T dt f(t) \sin n\omega t . \quad (8.30)$$

Enfin, toujours quand la convergence est uniforme, on peut écrire :

$$\begin{aligned} \int_0^T f^*(t) f(t) dt &= \int_0^T dt \sum_{(n, n') \in \mathbb{Z}^2} f_n^* f_{n'} e^{i(n' - n)\omega t} = \\ &= \sum_{(n, n') \in \mathbb{Z}^2} f_n^* f_{n'} \int_0^T e^{i(n' - n)\omega t} dt = \sum_{(n, n') \in \mathbb{Z}^2} f_n^* f_{n'} T\delta_{nn'} = T \sum_{n \in \mathbb{Z}} |f_n|^2 . \end{aligned} \quad (8.31)$$

¹⁵La même relation est encore vraie quand on remplace les bornes par $t_0, t_0 + T$.

La relation :

$$\boxed{\int_0^T |f(t)|^2 dt = T \sum_{n=-\infty}^{+\infty} |f_n|^2} \quad (8.32)$$

porte le nom de *relation de Bessel - Parseval - Plancherel*. Le second membre (au facteur T près) évoque le carré de la norme d'un vecteur de composantes f_n d'un espace vectoriel sur le corps \mathbb{C} . En conséquence, l'intégrale du premier membre peut recevoir la même interprétation : c'est ce qui est fait couramment quand on considère, en Mécanique quantique, l'intégrale du module carré d'une fonction d'onde, $\int_{\mathbb{R}^3} |\Psi(\vec{r}, t)|^2 d^3r$. On voit même que, en raisonnant avec deux fonctions distinctes $f(t)$ et $g(t)$, on a :

$$\int_0^T f^*(t)g(t) dt = \int_0^T dt \sum_{(n, n') \in \mathbb{Z}^2} f_n^* g_{n'} e^{i(n'-n)\omega t} = \sum_{(n, n') \in \mathbb{Z}^2} f_n^* g_{n'} T \delta_{nn'} = T \sum_{n \in \mathbb{N}} f_n^* g_n, \quad (8.33)$$

une relation qui cette fois évoque un *produit scalaire* dans un espace vectoriel de dimension infinie dénombrable, défini sur le corps des complexes, exprimé en terme des composantes des vecteurs sur une base orthonormée. Tout ceci provient du fait que l'ensemble des fonctions $e_n(x) \stackrel{\text{d'ef}}{=} \frac{1}{\sqrt{2\pi}} e^{inx}$, $n \in \mathbb{Z}$, constitue de fait une base (orthonormée) pour l'espace des fonctions admettant un développement de Fourier ; le facteur de normalisation $\frac{1}{\sqrt{2\pi}}$ élimine le facteur 2π disgracieux dans (8.36). Ces points seront repris dans le chapitre ??.

Il est utile de récrire les relations importantes ci-dessus en termes de variables adimensionnées¹⁶ ; on a :

$$\boxed{f(x) = \sum_{n=-\infty}^{+\infty} f_n e^{inx} \quad f_n = \frac{1}{2\pi} \int_0^{2\pi} e^{-inx} f(x) dx} \quad (8.34)$$

$$c_0 = \frac{1}{2\pi} \int_0^{2\pi} dx f(x), \quad a_n = \frac{1}{\pi} \int_0^{2\pi} dx f(x) \cos nx, \quad b_n = \frac{1}{\pi} \int_0^{2\pi} dx f(x) \sin nx. \quad (8.35)$$

et :

$$\boxed{\int_0^{2\pi} |f(x)|^2 dx = 2\pi \sum_{n=-\infty}^{+\infty} |f_n|^2} \quad (8.36)$$

L'existence de l'intégrale assure que la série est convergente, donc nécessairement :

$$\lim_{n \rightarrow +\infty} |f_n| = 0, \quad (8.37)$$

d'où il résulte que les limites de $\Re f_n$ et $\Im f_n$ sont séparément nulles :

$$\Re \left(\lim_{n \rightarrow +\infty} \int_0^{2\pi} f(x) e^{-inx} dx \right) = 0, \quad \Im \left(\lim_{n \rightarrow +\infty} \int_0^{2\pi} f(x) e^{-inx} dx \right) = 0; \quad (8.38)$$

¹⁶De fait, il suffit de remplacer formellement partout la période T par 2π et ω par 1.

dans le cas où f est à valeurs réelles, ceci se traduit comme :

$$\lim_{n \rightarrow +\infty} \int_0^{2\pi} f(x) \cos nx \, dx = 0 \quad , \quad \lim_{n \rightarrow +\infty} \int_0^{2\pi} f(x) \sin nx \, dx = 0 \quad . \quad (8.39)$$

Prenons maintenant le problème sous un autre angle. Soit $f(t)$ une fonction T -périodique à variation bornée, localement sommable mais pas forcément continue¹⁷. Pour cette fonction, on peut calculer des coefficients de Fourier selon la recette (8.29) et écrire la série $\sum_{n \in \mathbb{Z}} f_n e^{in\omega t}$. La question est : quelle est la relation entre la série de Fourier ainsi construite (évidemment supposée convergente), et la fonction $f(t)$? On ne peut espérer que la série converge partout vers $f(t)$, puisque modifier f sur un ensemble de mesure nulle ne change pas l'intégrale (8.29) et donc laisse les coefficients f_n inaltérés : au mieux, une fonction et sa série de Fourier sont *presque partout* égales.

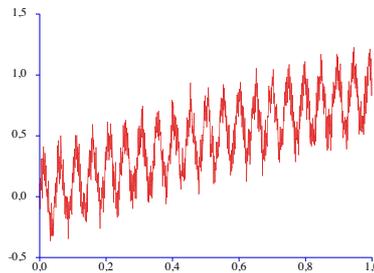


Figure 8.3: La fonction discutée par Picard (voir note 18).

De fait, le théorème de Dirichlet (1829, pas facile à démontrer) établit que pour toute fonction bornée monotone par morceaux¹⁸, la suite des sommes partielles de Fourier

¹⁷Une fonction en créneau, en dents de scie égoïne, etc.

¹⁸Un exemple de fonction bornée mais non monotone par morceaux : $\cos \frac{1}{x}$ dans un voisinage de l'origine.

L'expression de ce théorème varie d'un auteur à l'autre. La formulation ci-dessus est celle de Piskounov, tome II, p. 338. Schwartz (MMP, p. 172) exige la propriété pour f d'être à *variation bornée* – voir la définition qu'il en donne –, et ajoute plus loin que “les fonctions bornées rencontrées dans la pratique sont à variation bornée”, tout en précisant qu'il existe des fonctions *continues* à variations non bornées.

Picard donne l'exemple suivant d'une fonction continue sur l'intervalle $[-\pi, +\pi]$, mais dont la série de Fourier ne converge pas. Soit la fonction $F(\phi)$ définie comme :

$$f(\phi) = \sum_{p=1}^{\infty} \frac{1}{p^2} \sin(2p^3 + 1) \frac{\phi}{2} \quad , \quad F(\phi) = \begin{cases} f(\phi) & \text{si } 0 < \phi < \pi \\ f(-\phi) & \text{si } -\pi < \phi < 0 \end{cases} \quad (8.40)$$

La série donnant $f(\phi)$ converge uniformément, donc la somme $f(\phi)$ est bien continue, donc bornée, et $f(0) = 0$. La fonction F est aussi continue et bornée – en revanche, en raison de ses rapides oscillations, elle n'est pas à variation bornée.

La série de Fourier de F ne contient que les cosinus (la fonction est paire) :

$$F(\phi) = a_0 + a_1 \cos x + a_2 \cos 2x + \dots \quad (8.41)$$

converge simplement partout vers la demi-somme :

$$\boxed{\sum_{n=-\infty}^{+\infty} f_n e^{in\omega t} = \frac{1}{2}[f(t+0) + f(t-0)] \quad \forall t} \quad (8.42)$$

La démonstration réside essentiellement dans l'important résultat :

Soit $f(x)$ une fonction à variation bornée dans l'intervalle $[a, b]$. Alors l'intégrale suivante, dite de Dirichlet :

$$I(\lambda) = \frac{1}{\pi} \int_a^b f(x) \frac{\sin \lambda x}{x} dx \quad (8.43)$$

a les limites suivantes quand $\lambda \rightarrow \infty$:

$$I(+\infty) = \begin{cases} 0 & \text{si } a \text{ et } b \text{ sont de même signe} \\ \frac{1}{2} f(0+) & \text{si } a = 0 \text{ et } b > 0 \\ \frac{1}{2} f(0-) & \text{si } a < 0 \text{ et } b = 0 \end{cases} \quad (8.44)$$

Il en résulte en particulier que :

$$I(+\infty) = \frac{1}{2} [f(0+) + f(0-)] \quad \text{si } a < 0 < b \quad (8.45)$$

La démonstration utilise également une relation importante (mais facile, et dont il est bon de connaître l'existence) entre la somme partielle :

$$S_N(t) = \sum_{n=-N}^{+N} f_n e^{in\omega t} \quad (8.46)$$

et la fonction $f(t)$. Cette relation s'établit comme suit. Compte tenu de (8.29), on a :

$$S_N(t) = \sum_{n=-N}^{+N} \frac{1}{T} \int_0^T e^{-in\omega t'} f(t') dt' e^{in\omega t} = \frac{1}{T} \int_0^T \sum_{n=-N}^{+N} e^{in\omega(t-t')} f(t') dt' ; \quad (8.47)$$

la somme géométrique se calcule et on trouve :

$$S_N(t) = \frac{1}{T} \int_0^T D_N(t-t') f(t') dt' , \quad D_N(t) = \frac{\sin(N + \frac{1}{2})\omega t}{\sin \frac{1}{2}\omega t} ; \quad (8.48)$$

la somme S_N se rencontre dans les problèmes de diffraction par un réseau (fini) de N atomes, un certain déphasage spatial élémentaire entre deux ondes diffusées jouant le rôle de ωt . La figure 8.4 donne D_N en fonction de $\frac{\omega t}{2}$ pour deux (petites !) valeurs de N (imaginer ce qui se passe avec $N \sim 10^{23} \dots$).

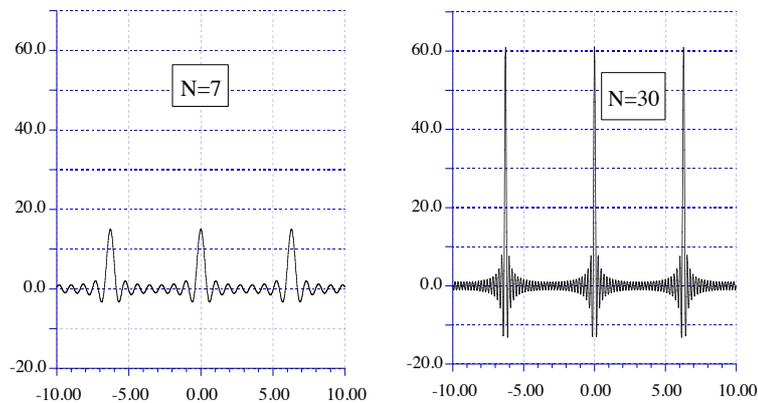


Figure 8.4: Variations de $D_N(x) = \frac{\sin(2N+1)x}{\sin x}$ pour deux valeurs de N .

Ainsi, le théorème de Dirichlet affirme que la série de Fourier converge partout simplement vers la moyenne arithmétique des valeurs à gauche et à droite. Ceci veut dire que en tout point où f est continue et vaut $f(t)$, la série de Fourier est égale à $f(t)$. En un point de discontinuité, la série donne la demi-somme de la limite à droite et de la limite à gauche. De façon sans doute un peu abusive, mais passée dans les mœurs, on dit encore que la série *est* la série de Fourier de la fonction $f(t)$, bien que la coïncidence n'ait pas lieu partout : la série et $f(t)$ sont *presque partout* égales¹⁹. Il faut bien avouer que, en Physique, ces subtilités de convergence ne jouent pas souvent un rôle : une fonction discontinue représente en fait une fonction variant très vite sur une échelle hors de portée des capacités d'observation, ou non pertinente pour la question posée ; la restitution de la bonne échelle doit supprimer toute ambiguïté.

Intuitivement, on sent bien que la série tronquée S_N définit une *approximation* de f . Une mesure de l'écart (l'erreur) est constituée par la moyenne sur une période du carré Δ_N^2 de la distance entre f et S_N , soit :

$$\Delta_N^2 = \frac{1}{T} \int_0^T |f(t) - S_N(t)|^2 dt = \frac{1}{T} \int_0^T \sum_{|n| \geq N+1} |f(t) e^{-in\omega t}|^2 dt = \sum_{|n| \geq N+1} |f_n|^2. \quad (8.49)$$

Soit $S_N = a_0 + a_1 + \dots + a_N$ la somme partielle de cette série de Fourier en $\phi = 0$. En choisissant $n = 2^{p'} - 1$, Picard montre que $S_{2^{p'} - 1} > \frac{1}{2\pi} \frac{1}{p'^2} \log(2^{p'} + 1)$. Pour p' grand, $\log(2^{p'} + 1) \sim p' \log 2$, donc $S_{2^{p'} - 1} > \frac{p' \log 2}{2\pi}$ qui croît sans limite. La suite des sommes partielles S_n ne peut donc converger : la série de Fourier de F en $\phi = 0$ diverge ! (É. Picard, *Traité d'Analyse*).

¹⁹Ceci sous-entend que sur une période, $f(t)$ a un nombre fini de discontinuités.

Mesurée à l'aune de Δ_N , il est clair que l'approximation est d'autant meilleure²⁰ que N est grand ($\Delta_{N+1} \leq \Delta_N$). À N donné, l'erreur est d'autant plus faible que les coefficients f_n décroissent vite en module en fonction de n , c'est-à-dire que la fonction a un grand nombre de dérivées continues (voir ci-dessous, notamment (8.63)). Par exemple, soit la fonction :

$$f(t) \stackrel{\text{d\u00e9f}}{=} t^2, \quad -\pi \leq t \leq +\pi \quad (8.50)$$

et p\u00e9riodis\u00e9e sur \mathbb{R} selon $f(t + p2\pi) = f(t)$ ($p \in \mathbb{Z}$). On trouve facilement la s\u00e9rie de Fourier correspondante²¹ :

$$f(t) = \frac{\pi^2}{3} + 2 \sum_{n \in \mathbb{Z}^*} \frac{(-1)^n}{n^2} e^{int} = \frac{\pi^2}{3} + 4 \sum_{n=1}^{+\infty} \frac{(-1)^n}{n^2} \cos nt \quad (8.53)$$

cette s\u00e9rie converge uniform\u00e9ment, en accord avec le fait que la fonction $f(t)$ d\u00e9finie en (8.50) est continue. L'approximation correspondant \u00e0 une troncature \u00e0 N termes est simplement :

$$f_{\text{ap}}^{(N)}(t) = \frac{\pi^2}{3} + 2 \sum_{n=-N}^{+N} \frac{(-1)^n}{n^2} e^{int} = \frac{\pi^2}{3} + 4 \sum_{n=1}^N \frac{(-1)^n}{n^2} \cos nt ; \quad (8.54)$$

La diff\u00e9rence $f(t) - f_{\text{ap}}^{(N)}(t)$ est ici une s\u00e9rie altern\u00e9e, dont la somme est en module plus petite que $\frac{4}{(N+1)^2}$.

La figure 8.5 compare f et $f_{\text{ap}}^{(N)}$ et les d\u00e9riv\u00e9es correspondantes. On note au passage que les coefficients de Fourier de $f(x)$ d\u00e9croissent comme n^{-2} , et que ceux de la d\u00e9riv\u00e9e d\u00e9croissent corr\u00e9lativement comme n^{-1} : il est visible sur la figure que, \u00e0 N donn\u00e9, l'erreur est plus importante pour la d\u00e9riv\u00e9e que pour la fonction elle-m\u00eame (et que dire des d\u00e9riv\u00e9es d'ordre sup\u00e9rieur ?!). En outre, la s\u00e9rie de Fourier de la d\u00e9riv\u00e9e ne converge pas uniform\u00e9ment (et c'est normal : la d\u00e9riv\u00e9e n'est pas continue).

De fait, \u00e9tablissons maintenant une relation importante entre la s\u00e9rie de Fourier d'une fonction et celles de ses d\u00e9riv\u00e9es. Soit $f(t)$ une fonction une fois contin\u00fament d\u00e9rivable. En partant de (8.29) et en effectuant une int\u00e9gration par parties, il vient :

$$f_m = \frac{1}{T} \int_0^T e^{-im\omega t} f(t) dt = \frac{1}{T} \left[\frac{e^{-im\omega t}}{-im\omega} f(t) \right]_0^T + \frac{1}{im\omega T} \int_0^T e^{-im\omega t} f'(t) dt . \quad (8.55)$$

²⁰Toutefois, l'optimisme qui peut en r\u00e9sultier doit \u00eatre quelque peu temp\u00e9r\u00e9 par un ph\u00e9nom\u00e8ne surprenant, le *ph\u00e9nom\u00e8ne de Gibbs* – voir p. 362.

²¹De ce d\u00e9veloppement, on peut d\u00e9duire des \u00e9galit\u00e9s importantes. Par exemple, en faisant $t = \pi$ puis $t = 0$, on trouve :

$$\pi^2 = \frac{\pi^2}{3} + 4 \sum_{n=1}^{+\infty} \frac{(-1)^n}{n^2} (-1)^n \iff \sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6} , \quad (8.51)$$

$$0 = \frac{\pi^2}{3} + 4 \sum_{n=1}^{+\infty} \frac{(-1)^n}{n^2} \iff \sum_{n=1}^{+\infty} \frac{(-1)^n}{n^2} = -\frac{\pi^2}{12} , \quad \text{etc}, \quad (8.52)$$

o\u00f9 l'on voit appara\u00eetre des valeurs particuli\u00e8res de la fonction $\zeta(z)$ de Riemann.

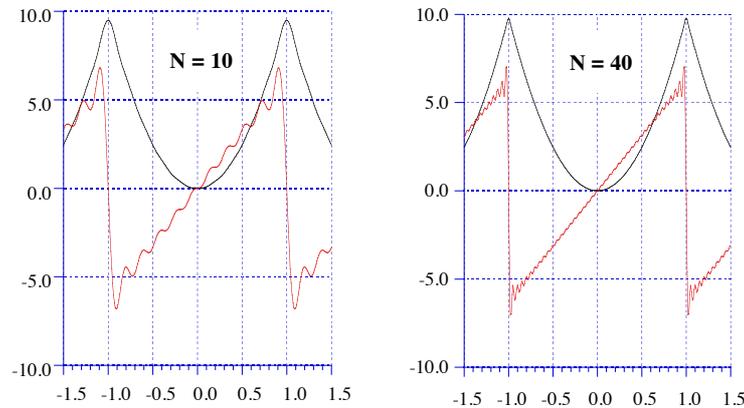


Figure 8.5: Comparaison entre d'une part la fonction (8.50) et sa dérivée, d'autre part leurs "séries" de Fourier tronquées à 10 et 40 termes.

Le terme tout intégré est nul (puisque $\omega T = 2\pi$, que m est entier et que f est T -périodique), et il reste :

$$f_m = \frac{1}{im\omega T} \int_0^T e^{-im\omega t} f'(t) dt . \quad (8.56)$$

(8.56) se lit aussi :

$$\frac{1}{T} \int_0^T e^{-im\omega t} f'(t) dt = im\omega f_m , \quad (8.57)$$

où l'on voit apparaître, suivant (8.29), le coefficient de Fourier $f_m^{(1)}$ de la dérivée f' ; d'où la règle pratique :

$$\boxed{f(t) \rightarrow f_m \implies f'(t) \rightarrow f_m^{(1)} = im\omega f_m} \quad (8.58)$$

ainsi, pour avoir le coefficient de Fourier f_m de la dérivée, il suffit de multiplier celui de la fonction par $im\omega$. Ce résultat, important pour les applications, se retrouve de tête en dérivant (8.20) terme à terme, une opération dont la légitimité repose sur une hypothèse contraignante de convergence uniforme de la série des dérivées, ce que n'exige pas la démonstration ci-dessus utilisant une intégration par parties (mais qui a toutefois supposé la fonction une fois continûment dérivable). Comme toujours, l'intégration est une opération sûre, alors que celle de dérivation peut être hasardeuse.

Par exemple, soit la fonction sinus redressé $f(t) = |\sin \omega t|$; il n'est pas difficile de

trouver les coefficients de Fourier et d'écrire la série²² :

$$f(t) \stackrel{\text{d'éf}}{=} |\sin \omega t| = \frac{2}{\pi} \sum_{n \in \mathbb{N}^*} \frac{\sin^2 n\omega t}{n^2 - \frac{1}{4}} = -\frac{2}{\pi} \sum_{n \in \mathbb{Z}} \frac{1}{4n^2 - 1} e^{2in\omega t} ; \quad (8.60)$$

cette série converge uniformément. En revanche, la dérivée $f'(t)$ a un saut à chaque fois que $\omega T = n\pi$; sa série de Fourier, obtenue suivant la règle (8.58), est :

$$f'(t) = \frac{\omega}{i\pi} \sum_{n \in \mathbb{Z}} \frac{n}{n^2 - \frac{1}{4}} e^{2in\omega t} ; \quad (8.61)$$

cette série ne converge pas uniformément, elle vaut zéro pour $\omega t = p\pi$ ($p \in \mathbb{Z}$), qui est bien la demi-somme à gauche et à droite de $\frac{d}{dt}|\sin \omega t|$ en $\omega t = p\pi$.

Pour une fonction k -fois continûment dérivable, la règle se généralise en :

$$\boxed{f(t) \rightarrow f_m \implies f^{(k)}(t) \rightarrow f_m^{(k)} = (im\omega)^k f_m} \quad (8.62)$$

En lisant cette relation à l'envers, on voit que plus une fonction est un grand nombre de fois dérivable, plus ses coefficients de Fourier décroissent vite. En effet, désignant par M_k un majorant du module de la k^e dérivée, $|f^{(k)}(t)| \leq M_k$, on a :

$$|f_m| \leq \left(\frac{1}{m|\omega|}\right)^k \int_0^T |f^{(k)}(t)| \frac{dt}{T} = \frac{M_k}{|m\omega|^k} . \quad (8.63)$$

De ceci résulte immédiatement un résultat important : si $f(t)$ est deux fois continûment dérivable, alors sa série de Fourier converge uniformément²³. En effet, faisant $k = 2$ dans (8.63), on voit que $|f_m| \leq \frac{M_2}{k^2}$. Il en résulte que la série des modules est bornée par une série convergente :

$$\left| \sum_{n=-\infty}^{+\infty} f_n e^{in\omega t} \right| \leq \sum_{n=-\infty}^{+\infty} |f_n| \leq \sum_{n=-\infty}^{+\infty} \frac{M_2}{|n\omega|^2} . \quad (8.64)$$

La majoration est indépendante de t , ce qui établit la convergence uniforme.

Remarquons enfin que (8.20) n'est rien d'autre qu'un développement de Laurent dans la variable $z = e^{i\omega t}$:

$$f(t) = \phi(e^{i\omega t}) \quad \text{avec} \quad \phi(z) = \sum_{n=-\infty}^{+\infty} f_n z^n , \quad (8.65)$$

²²Avec la méthode exposée au chapitre 6 ($\pi \cot \pi z$ & Co), on pourra vérifier que :

$$\sum_{n \in \mathbb{Z}} \frac{1}{4n^2 - 1} = 0 . \quad (8.59)$$

Par ailleurs, $f(t)$ a pour plus petite période $\frac{\pi}{\omega}$ (et non $\frac{2\pi}{\omega}$), d'où l'apparition des seules valeurs *paires* de n dans la série (8.60).

²³Noter que la réciproque est fautive : la fonction sinus redressé (8.60) a une série de Fourier uniformément convergente, mais elle n'est pas deux fois continûment différentiable. Le fait est que sa dérivée n'est pas continue, et que la dérivée seconde n'est pas bornée : dans ces conditions, la majoration exprimée en (8.64) ne tient plus ou, si on préfère, est illusoire ($M_2 = +\infty$!).

et que la relation d'inversion dans (8.29) s'écrit alors :

$$f_m = \frac{1}{2i\pi} \int_{C_{r=1}} z^{-(m+1)} \phi(z) dz , \quad (8.66)$$

où l'intégration se fait en tournant dans le sens positif le long du cercle de rayon unité entourant l'origine de \mathbb{C} (l'intégrale au second membre de (8.66) résulte directement du changement de variable $z = e^{i\omega t}$) – sans préjuger d'éventuelles déformations du contour compatibles avec les singularités de $\phi(z)$.

Avant de continuer, examinons quelques exemples illustrant quelques points développés ci-dessus.

• Exemples

1. Soit la fonction T -périodique :

$$f(t) = \begin{cases} +1 & \text{si } 0 < t < \frac{T}{2} \\ -1 & \text{si } \frac{T}{2} < t < T \end{cases} , \quad (8.67)$$

et prolongée par périodes sur tout l'axe réel. Cette fonction (impair) est discontinue en $t = p \frac{T}{2}$ ($p \in \mathbb{Z}$). Les formules (8.30) permettent d'écrire l'égalité suivante (à comprendre comme une égalité *presque partout*, comme déjà mentionné) :

$$f(t) = \frac{4}{\pi} \left(\sin \omega t + \frac{\sin 3\omega t}{3} + \frac{\sin 5\omega t}{5} + \dots \right) = \frac{4}{\pi} \sum_{n=0}^{+\infty} \frac{\sin(2n+1)\omega t}{(2n+1)} . \quad (8.68)$$

Par définition de la somme d'une série, on a :

$$f(t) = \lim_{N \rightarrow +\infty} f_N(t) , \quad f_N(t) = \frac{4}{\pi} \sum_{n=0}^N \frac{\sin(2n+1)\omega t}{(2n+1)} . \quad (8.69)$$

Les possibles bizarreries de convergence d'une série de Fourier sont patentes sur cet exemple : si on fait brutalement²⁴ $t = \frac{T}{2}$ dans la série (8.68) au second membre, chacun des termes est nul ($\sin(2n+1)\pi = 0$!), alors que, par sa définition même, la fonction $f(t)$ est discontinue en ces points (où donc on ne sait pas ce qu'elle vaut) ; quoi qu'il en soit, on voit que la série est bien égale à la demi-somme des valeurs à gauche et à droite d'une discontinuité de f .

En outre, cet exemple illustre de façon exemplaire un phénomène assez surprenant²⁵, appelé *phénomène de Gibbs* : juste à gauche ou à droite d'une discontinuité de f , toute somme partielle $f_N(t)$ présente des oscillations autour de la valeur de $f(t)$. L'amplitude maximale survient en un point t_N et

²⁴Procédé expéditif qui résume l'opération précise : on considère la suite *numérique* des sommes partielles $f_N(t = n \frac{T}{2})$, suite nulle puisque chaque élément est nul et on prend la limite ; la limite d'une suite dont tous les éléments sont nuls est égale à zéro.

²⁵à propos duquel Michelson a dû s'arracher quelques cheveux.

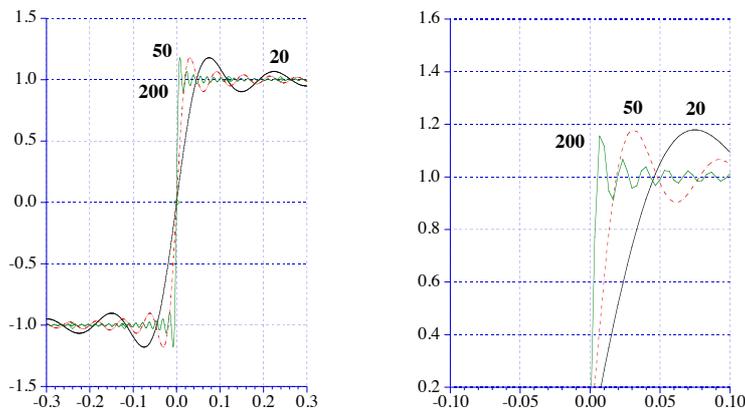


Figure 8.6: Illustration du phénomène de Gibbs.

vaut un certain δ_N . Quand N tend vers l'infini, t_N tend vers 0 (modulo $\frac{T}{2}$), mais δ_N ne tend pas vers zéro ! Par exemple, juste à droite de $t = 0$, la série tronquée à N termes (qui est clairement infiniment dérivable) présente des oscillations amorties autour de la valeur 1 ; son *maximum maximorum* survient en $t_N = \frac{\pi}{2N+2}$ et excède 1 d'environ 20 % dans la limite $N \rightarrow +\infty$. L'écart est donc fini, même pour la série ; il faut toutefois noter que l'abscisse où cet écart persiste vaut $0+$ à la limite²⁶. Ceci mis à part, il reste que plus la série tronquée contient un grand nombre de termes, meilleure est la représentation de f (un ampli large bande reproduit mieux une fonction en créneaux qu'un ampli à bande étroite²⁷), comme on l'a déjà noté. Par ailleurs, clairement la série des dérivées ne converge pas (au moins au sens usuel !), puisque si on dérive terme à terme on trouve :

$$\frac{4\omega}{\pi} \sum_{n=0}^{+\infty} \cos(2n+1)\omega t, \quad (8.70)$$

une série dont le sens mérite pour le moins d'être clarifié. D'un autre côté, sur la définition (8.67), on voit que $f'(t) = 0$ partout sauf si $t = p\frac{T}{2}$ ($p \in \mathbb{Z}$), auquel cas f' n'est pas définie. On verra plus loin comment comprendre tout ceci.

²⁶Pour plus de détails, voir [8], ou [2], section 11.3.

²⁷d'où, pour un ampli audio, l'importance d'une bande passante large (comparée à [20, 20 000] Hz !) pour une bonne restitution des transitoires (un coup d'archet, une percussion,...)

2. Considérons maintenant la fonction $f(t)$ définie comme :

$$f(t) = \frac{\pi - \omega_0 t}{2}, \quad 0 < t < T = \frac{2\pi}{\omega_0} \quad (8.71)$$

et prolongée périodiquement sur \mathbb{R} :

$$f(t + nT) = f(t) \quad \forall n \in \mathbb{Z}, \quad 0 < t < T. \quad (8.72)$$

Le graphe de cette fonction – impaire – est constitué d’une infinité de segments parallèles de pente négative, allant de $+\frac{\pi}{2}$ à $-\frac{\pi}{2}$ pour $t = 0, T$, modulo T . La dérivée $f'(t)$ vaut $-\frac{\omega_0}{2}$ si $t \neq nT$, et n’est pas définie en $t = nT$. Toutefois, si on imagine que les discontinuités sont arrondies sur un très petit intervalle autour de nT (en pensant par exemple à la série tronquée à un nombre N fini de termes), on admet l’idée que près d’un tel point f'_N prend de très grandes valeurs positives, puisque la fonction passe de $-\frac{\pi}{2}$ à $+\frac{\pi}{2}$ alors que t augmente à peine de part et d’autre de nT : le rapport des accroissements $\frac{\delta f}{\delta t}$ est donc très grand (sans que l’on sache d’ailleurs précisément ce qu’il vaut puisque δt n’est pas vraiment défini). Ce point mérite d’être approfondi (voir plus loin), et produit d’ailleurs une relation rencontrée souvent en Physique, notamment à propos des problèmes de diffraction par un réseau *infini*²⁸.

Le calcul des coefficients de Fourier suivant (8.30) permet d’écrire la série de Fourier représentant la fonction $f(t)$ définie en (8.71) et (8.72) ; s’agissant à nouveau d’une fonction impaire, $f(-t) = -f(t)$, seuls les sinus apparaissent dans la série trigonométrique :

$$f(t) = \sum_{n=1}^{+\infty} b_n \sin n\omega_0 t, \quad b_n = \frac{2}{T} \int_0^T \frac{\pi - \omega_0 t}{2} \sin n\omega_0 t \, dt = \frac{1}{n} \quad (8.73)$$

Considérons maintenant la série tronquée à N termes :

$$f_N(t) = \sum_{n=1}^N \frac{1}{n} \sin n\omega_0 t. \quad (8.74)$$

La dérivée de f_N est :

$$f'_N(t) = \omega_0 \sum_{n=1}^N \cos n\omega_0 t = \frac{\omega_0}{2} \sum_{n=-N}^N e^{in\omega_0 t} - \frac{\omega_0}{2} \quad (8.75)$$

soit :

$$f'_N(t) + \frac{\omega_0}{2} = \frac{\omega_0}{2} \text{III}_N(\omega_0 t), \quad (8.76)$$

où on a posé :

$$\text{III}_N(x) \stackrel{\text{déf}}{=} \sum_{n=-N}^N e^{inx}; \quad (8.77)$$

²⁸Elle sert notamment à démontrer la condition de von Laue, équivalente à celle de Bragg, $n\lambda = 2d \sin \theta$ dans les notations traditionnelles.

$\text{III}_N(x)$ est une somme géométrique qui se calcule facilement :

$$\text{III}_N(x) = \frac{\sin(2N+1)\frac{x}{2}}{\sin\frac{x}{2}}. \quad (8.78)$$

d'où²⁹ :

$$f'_N(t) + \frac{\omega_0}{2} = \frac{\sin\left[(2N+1)\frac{\omega_0 t}{2}\right]}{\sin\frac{\omega_0 t}{2}} \equiv \frac{\pi\omega_0}{2}\tilde{\delta}_N\left(\frac{\omega_0 t}{2}\right), \quad \tilde{\delta}_N(x) = \frac{\sin(2N+1)x}{\pi\sin x}. \quad (8.79)$$

Si $N \gg 1$, $\tilde{\delta}_N(x)$ est une fonction 2π -périodique présentant un maximum très pointu à chaque fois que $x = p2\pi$, $p \in \mathbb{Z}$; par exemple, $\tilde{\delta}_N(0) = \frac{2N+1}{\pi} \gg 1$, et $\tilde{\delta}_N(\frac{\pi}{2N+1}) = 0$. L'aire sous la courbe représentant $\tilde{\delta}_N$ s'estime en assimilant chaque partie pointue à un triangle et on trouve $\frac{1}{\pi} \frac{1}{2}(2N+1) \frac{2\pi}{2N+1} = 1$ pour chaque résonance. D'ailleurs, en revenant à la définition (8.75), on a :

$$\int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} \left(f'_N(t) + \frac{\omega_0}{2}\right) dt = \frac{\omega_0}{2} \sum_{n=-N}^N \int_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} e^{in\omega_0 t} dt = \frac{\omega_0}{2} \sum_{n=-N}^N T\delta_{n0} = \pi, \quad (8.80)$$

de sorte que l'aire sous $\tilde{\delta}_N(x)$, calculée sur un intervalle d'une période, est *exactement* égale à 1.

Il en résulte que pour toute fonction $\phi(x)$ définie en $x = 0$ et variant sur une échelle d'ordre 1 (donc très lentement à l'échelle $\frac{1}{N}$), on a :

$$\int_{-\pi}^{+\pi} \phi(x)\tilde{\delta}_N(x) dx \simeq \phi(x=0) \int_{-\pi}^{+\pi} \tilde{\delta}_N(x) dx = \phi(0). \quad (8.81)$$

C'est ainsi que, passant à la limite $N \rightarrow +\infty$ sous l'intégrale, s'introduit intuitivement la "fonction" de Dirac $\delta(x)$, *restreinte pour l'instant à l'intervalle* $[-\pi, +\pi]$, dont la règle d'usage est :

$$\boxed{\int_a^b \phi(x)\delta(x) dx = \phi(0) \quad (-\pi \leq a < 0 < b \leq \pi)} \quad (8.82)$$

De par sa définition (8.79), la fonction $\tilde{\delta}_N$ est 2π -périodique et il en va de même de sa limite δ : le raisonnement ci-dessus peut être reproduit au voisinage de chaque valeur de x de la forme $\text{entier} \times 2\pi$. Il en résulte que la limite de $\tilde{\delta}_N$, définie sur tout l'axe réel, est en fait une série de pics équidistants, ce que l'on peut écrire précisément, après calage des bonnes variables à partir de (8.75) :

$$\lim_{N \rightarrow \infty} \sum_{n=-N}^N e^{in\omega_0 t} = 2\pi \sum_{p \in \mathbb{Z}} \delta(\omega_0 t - p2\pi) = \sum_{p \in \mathbb{Z}} \delta\left(\frac{t}{T} - p\right). \quad (8.83)$$

²⁹La fonction $\tilde{\delta}_N(x)$ introduite en (8.79) est essentiellement le noyau de Dirichlet D_N défini en (8.48). De même, la fonction $\text{III}_N(x)$ est égale à $D_N(\frac{x}{\omega})$.

En termes de variables sans dimension³⁰ :

$$\boxed{\sum_{n \in \mathbb{Z}} e^{inx} = 2\pi \sum_{p \in \mathbb{Z}} \delta(x - p2\pi) = \sum_{p \in \mathbb{Z}} \delta\left(\frac{x}{2\pi} - p\right) \equiv \text{III}\left(\frac{x}{2\pi}\right)} \quad (8.84)$$

où $\text{III}(x) \stackrel{\text{déf}}{=} \sum_{n \in \mathbb{Z}} \delta(x - n)$ est le “peigne de Dirac”, représenté par un symbole évocateur³¹. •

Suivant une idée déjà discutée, la série tronquée après le N^{e} terme, f_N , constitue une approximation de f , d’autant meilleure que N est grand. De la même façon, f'_N est une approximation de $f'(t)$. Visiblement, f'_N est nulle le plus souvent (“presque partout”, dans le langage commun) mais présente des pics très marqués à chaque fois que $t = \text{entier} \times T$. À la limite $N = +\infty$, f' est nulle presque partout (cette fois au sens du Mathématicien) ; f' est non-nulle aux seuls points de discontinuité de f . On retiendra que si une fonction présente un saut Δf en un point t_0 , alors sa dérivée est de la forme $f'_{\text{reg}}(t) + \Delta f \delta(t - t_0)$, où $f'_{\text{reg}}(t)$ est la dérivée calculée par les moyens ordinaires ; plus précisément, on a :

$$\boxed{f'(t) = f'_{\text{reg}}(t) + [f(t_0 + 0) - f(t_0 - 0)] \delta(t - t_0)} \quad (8.86)$$

Ce résultat ressortira plusieurs fois dans la suite, sous une forme ou sous une autre. Intuitivement, la fonction de Dirac exprime le fait que, la fonction f étant discontinue en t_0 , sa “dérivée” y est infinie (variation finie sur un intervalle de mesure nulle).

Revenons maintenant au développement (8.20) ; il peut être considéré comme une transformation (linéaire) qui, à une suite infinie dénombrable de coefficients $\{f_n\}_n$ fait correspondre³² la fonction $f(t)$ égale³³ à la somme de la série au second membre de (8.20) :

$$\{f_n\}_n \xrightarrow{\text{série de Fourier}} f(t) . \quad (8.87)$$

Dans la même optique, la recette (8.29) constitue la relation inverse :

$$f(t) \xrightarrow{\text{formule intégrale}} \{f_n\}_n \quad (8.88)$$

³⁰De (8.81), on voit que $\tilde{\delta}_N(ax) = \frac{1}{|a|} \tilde{\delta}_N(x)$ ($a \in \mathbb{R}$), une relation qui est préservée à la limite : $\delta(ax) = \frac{1}{|a|} \delta(x)$.

³¹De la même façon, le physicien – fort de ses idiosyncrasies et sachant plus ou moins vaguement que la justification existe en termes de Théorie des distributions – dira sans hésiter que la série obtenue en dérivant (8.68) terme à terme représente le peigne de Dirac alterné, trouvant sans états d’âme (après un petit calcul) :

$$\frac{4\omega}{\pi} \sum_{n=0}^{+\infty} \cos(2n+1)\omega t = 2 \sum_{n=-\infty}^{+\infty} (-1)^n \delta\left(t - n\frac{T}{2}\right) . \quad (8.85)$$

³²En Physique, on appelle souvent *champs* certaines fonctions d’une grandeur physique. Un champ peut être vu comme l’ensemble (infini non-dénombrable) des composantes d’un vecteur – c’est pourquoi on dit d’un champ qu’il possède une infinité de degrés de liberté. C’est le cas en Mécanique quantique où les valeurs (complexes !) de la fonction d’onde $\Psi(x, t)$ (champ des amplitudes de probabilité) sont les différentes composantes du vecteur $|\Psi(t)\rangle$ dans la représentation dite *représentation-q*.

³³l’égalité étant toujours à prendre avec les subtilités énoncées plus haut.

qui permet d'associer à f la suite de nombres f_n . Pour une fonction périodique, il y a une dissymétrie entre les relations de passage dans un sens et dans l'autre : sommation discrète dans un sens, intégration (sommation continue) dans l'autre sens. Quoi qu'il en soit, on peut dire que (8.29) est la formule d'inversion associée à (8.20), ou que ces deux relations sont inverses l'une de l'autre.

◆ *Remarques*

1. La relation de définition du peigne de Dirac, (8.84) :

$$\text{III}\left(\frac{x}{2\pi}\right) \stackrel{\text{d\'ef}}{=} 2\pi \sum_{p \in \mathbb{Z}} \delta(x - p2\pi) = \sum_{n \in \mathbb{Z}} e^{inx} \equiv \left(\sum_{n \in \mathbb{Z}} f_n e^{inx} \right)_{f_n=1}, \quad (8.89)$$

peut bien être interprétée comme un développement en série de Fourier, celui de la "fonction" périodique $f(x)$ définie par (8.89), mais les coefficients f_n ne tendent pas vers zéro (ils sont strictement constants !), contrairement à ce qui se passe pour une vraie fonction (voir (8.37)). La raison en est que f n'est pas une fonction, mais une *distribution*.

2. La fonction $\tilde{\delta}_N(x)$ introduite en (8.79) engendre le peigne de Dirac dans la limite $N \rightarrow +\infty$ et doit être distinguée de la "fonction" de Dirac $\delta(x)$, dont il est utile de rappeler rapidement ici comment elle s'introduit intuitivement (voir chapitre 2 pour plus de détails). $\delta(x)$ est la limite symbolique d'une fonction non-périodique, et dont la règle opérationnelle de base est :

$$\boxed{\int_{-\infty}^{+\infty} \phi(x) \delta(x - x_0) dx = \phi(x_0)} \quad (8.90)$$

(comparer avec (8.81), où les bornes sont $\pm\pi$). $\delta(x)$ peut être définie en considérant la limite d'une suite de fonctions (le choix est vaste), par exemple la suite :

$$\delta_n(x - x_0) = \sqrt{\frac{n}{\pi}} e^{-n(x-x_0)^2}. \quad (8.91)$$

Chacune de ces fonctions est une gaussienne présentant un et un seul pic en x_0 de largeur d'ordre $\frac{1}{\sqrt{n}}$ et telle que $\int_{-\infty}^{+\infty} \delta_n(x-x_0) dx = 1$. Il en résulte que l'association avec une bonne fonction $\phi(x)$ donne :

$$\int_{-\infty}^{+\infty} \phi(x) \delta_n(x - x_0) dx \simeq \phi(x_0); \quad (8.92)$$

à la limite, c'est bien la règle (8.90) qui est retrouvée ; rappelons que l'on peut tout autant prendre la suite de lorentziennes :

$$\delta_n(x - x_0) = \frac{1}{n\pi} \frac{1}{(x - x_0)^2 + \frac{1}{n^2}}. \quad (8.93)$$

Toutes ces suites constituent des précurseurs de la fonction de Dirac, au sens où l'on peut énoncer (intuitivement) le résultat (symbolique en toute rigueur) :

$$\text{“ } \lim_{n \rightarrow +\infty} \delta_n(x - x_0) = \delta(x - x_0) \text{ ”} . \quad (8.94)$$

Rappelons enfin qu'il est d'ailleurs fréquent de considérer des précurseurs dépendant *continûment* d'un paramètre. Ainsi on écrira par exemple :

$$\text{“ } \lim_{\varepsilon \rightarrow 0} \frac{1}{\sqrt{2\pi\varepsilon}} e^{-\frac{(x-x_0)^2}{2\varepsilon}} = \lim_{\varepsilon \rightarrow 0} \frac{1}{\pi} \frac{\varepsilon}{(x-x_0)^2 + \varepsilon^2} = \dots = \delta(x - x_0) \text{ ”} . \quad (8.95)$$

Noter que de (8.90) on déduit $\int_{-\infty}^{+\infty} x^\nu \phi(x) \delta(x) dx = 0 \forall \nu > 0, \forall \phi$, d'où l'égalité au niveau des “fonctions” :

$$\boxed{x^\nu \delta(x) = 0 \quad (\nu > 0)} \quad (8.96)$$

◆

8.2 Transformation de Fourier

La théorie mathématique de la transformation de Fourier est difficile. On se contentera ici d'une version élémentaire, suffisante pour la plupart des applications en Physique, reportant le lecteur aux ouvrages spécialisés pour en savoir (beaucoup) plus sur les subtilités inhérentes à ce genre de transformation ; on peut dire que les subtilités rencontrées pour les séries de Fourier deviennent encore plus profondes puisqu'il s'agit maintenant d'une application *intégrale*.

8.2.1 Définition et formule d'inversion

La transformation de Fourier est en quelque sorte la généralisation au cas des fonctions non-périodiques de la notion de série de Fourier.

Pour bien montrer intuitivement la filiation entre la série de Fourier et l'intégrale de Fourier, récrivons (8.20) en appelant maintenant ω_0 la pulsation propre, et en écrivant quelque peu différemment la série de Fourier³⁴ :

$$f(t) = \sum_{n=-\infty}^{+\infty} f_n e^{-in\omega_0 t} , \quad f(t + nT) = f(t) \quad (\omega_0 T = 2\pi) . \quad (8.97)$$

³⁴Ces modifications inessentiels sont juste destinées à retomber sur ses pieds vis-à-vis des détails purement conventionnels choisis pour écrire précisément la transformation intégrale de Fourier.

Par ailleurs, rien n'interdit de voir les f_n comme les valeurs d'une certaine fonction \tilde{f} (inconnue) pour les valeurs $n\omega_0$ de sa variable ω . Cela étant réalisé et admis, (8.97) s'écrit tout autant :

$$f(t) = \sum_{n=-\infty}^{+\infty} \tilde{f}(\omega = n\omega_0) e^{-in\omega_0 t} . \quad (8.98)$$

Pour l'instant, l'espace des pulsations (*a priori* dense (continu)) est en un sens discrétisé (quantifié) : seules les pulsations ω multiples entiers de ω_0 apparaissent. Imaginons maintenant que la période $T = \frac{2\pi}{\omega_0}$ devienne de plus en plus grande, ce qui équivaut à densifier les points $n\omega_0$ sur l'axe réel ; la différence $(n+1)\omega_0 - n\omega_0$ joue le rôle d'un petit incrément $\delta\omega$. À la limite $T \rightarrow +\infty$, on conçoit que la somme (8.98) devienne une intégrale^{35,36}, moyennant l'introduction d'une certaine fonction F associée à \tilde{f} , $F(n\omega_0) \stackrel{\text{déf}}{=} \frac{2\pi}{\omega_0} \tilde{f}(n\omega_0)$:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{-i\omega t} d\omega . \quad (8.99)$$

Dans cette limite, la période T étant infinie, la fonction f n'est pas périodique ; d'ailleurs, selon (8.99) :

$$f(t + \tau) = \int_{-\infty}^{+\infty} F(\omega) e^{i\omega(t+\tau)} d\omega \quad (8.100)$$

et il n'existe pas de temps τ tel que :

$$\int_{-\infty}^{+\infty} F(\omega) e^{i\omega(t+\tau)} d\omega = \int_{-\infty}^{+\infty} F(\omega) e^{i\omega t} d\omega . \quad (8.101)$$

Plus précisément, le passage naturel de la série de Fourier à l'intégrale de Fourier peut s'explicitier comme suit. Soit $f(t)$ une fonction non-périodique, continue et à variation bornée ; définissons la fonction $f_T(t)$ T -périodique telle que :

$$f_T(t) \stackrel{\text{déf}}{=} f(t) , \quad \forall t \in \left[-\frac{T}{2}, +\frac{T}{2}\right] , \quad f_T(t+T) = f_T(t) . \quad (8.102)$$

f_T étant périodique, on peut écrire sa série de Fourier :

$$f_T(t) = \sum_{n \in \mathbb{Z}} c_n e^{-in\omega_0 t} , \quad c_n = \frac{1}{T} \int_{-T/2}^{+T/2} f_T(t) e^{in\omega_0 t} dt , \quad (8.103)$$

³⁵Ce scénario – très intuitif – est d'usage fréquent en Physique théorique, où on passe souvent d'une représentation discrète à une représentation continue, pour l'espace par exemple. Quand on parle d'une théorie *sur réseau*, c'est juste pour dire que, par commodité, l'espace continu \mathbb{R}^D a été remplacé par un réseau discret de points \mathbb{Z}^D . Bien évidemment, l'équivalence entre les deux descriptions ne saute pas toujours aux yeux : entre \mathbb{Z}^D et \mathbb{R}^D , il y a toute (!?) la différence existant entre le dénombrable et le non-dénombrable.

³⁶Attention : au vu de (8.98), \tilde{f} a la même dimension que f . En revanche, ω étant l'inverse d'un temps, la transformée F dans (8.99) a pour dimension $[f] \times \text{temps}$. En termes géométriques : on remplace le bâton $f_n \stackrel{\text{déf}}{=} \tilde{f}(n\omega_0)$ par un *rectangle* vertical de base ω_0 et de hauteur $F(n\omega_0)$.

soit, avec $\frac{1}{T} = \frac{\omega_0}{2\pi}$:

$$f_T(t) = \sum_{n \in \mathbb{Z}} \frac{\omega_0}{2\pi} F(n\omega_0) e^{-in\omega_0 t}, \quad F(\omega) \stackrel{\text{d\u00e9f}}{=} \int_{-\frac{T}{2}}^{+\frac{T}{2}} f_T(t') e^{i\omega t'} dt'. \quad (8.104)$$

La somme \u00e0 gauche dans (8.104) est une somme de Darboux qui, dans la limite $T \rightarrow +\infty$, soit $\omega_0 \rightarrow 0$, engendre l'int\u00e9grale de Riemann ; par ailleurs, dans la m\u00eame limite, les deux fonctions $f(t)$ et $f_T(t)$ co\u00efncident ; il vient ainsi :

$$T = +\infty : \quad f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{-i\omega t} d\omega. \quad (8.105)$$

Cette introduction r\u00e9solument intuitive conduit \u00e0 se poser la question de la d\u00e9finition pr\u00e9cise d'une transformation int\u00e9grale du genre (8.99). Pour sacrifier \u00e0 la tradition, et prenant des notations g\u00e9n\u00e9rales, on part d'une certaine fonction donn\u00e9e, $f(x)$, et on d\u00e9finit sa transform\u00e9e de Fourier $F(k)$ par la relation suivante³⁷ :

$$F(k) \stackrel{\text{d\u00e9f}}{=} \int_{-\infty}^{+\infty} f(x) e^{ikx} dx \quad (8.108)$$

Dans cette \u00e9criture, x et k sont des quantit\u00e9s r\u00e9elles (appel\u00e9es *variables conjugu\u00e9es*), mais f et F sont *a priori* des fonctions \u00e0 valeurs complexes. Si $f(x)$ est une fonction paire (*resp.* impaire), $F(k)$ est r\u00e9elle (*resp.* imaginaire pure). Le cas \u00e9ch\u00e9ant, on utilisera la notation compacte :

$$F = \mathcal{F}[f] \iff f \xrightarrow{\mathcal{F}} F \quad (8.109)$$

pour d\u00e9signer la transform\u00e9e de Fourier. S'il est utile de pr\u00e9ciser les variables, on \u00e9crira :

$$F(k) = \mathcal{F}[f(x)](k) \iff f(x) \xrightarrow{\mathcal{F}} F(k) \quad (8.110)$$

Avant de continuer, signalons que la d\u00e9finition de l'int\u00e9grale de Fourier varie d'un auteur \u00e0 l'autre, un peu en fonction de sa sp\u00e9cialit\u00e9 favorite (les \u00e9lectriciens n'ont pas les m\u00eames

³⁷Si x est une longueur, k est l'inverse d'une longueur, un nombre d'onde par exemple. Si la variable de f est le temps, t , on \u00e9crit tout naturellement :

$$F(\omega) \stackrel{\text{d\u00e9f}}{=} \int_{-\infty}^{+\infty} f(t) e^{i\omega t} dt. \quad (8.106)$$

La transformation int\u00e9grale de Fourier permet de r\u00e9soudre commod\u00e9ment l'\u00e9quation (8.6) de la diffusion, cette fois sur \mathbb{R} . La solution est :

$$P(x, t) = \int_{-\infty}^{+\infty} p(k) e^{-Dk^2 t} e^{ikx} dk, \quad (8.107)$$

(comparer avec (8.8)), o\u00f9 la fonction $p(k)$ se d\u00e9duit de l'\u00e9tat initial $P(x, 0)$ (comment ?). Noter que la question de l'existence de $\mathcal{F}[P(x, t)]$ ne se pose pas ici puisque $P(x, t)$ est positive et int\u00e9grable.

habitudes que les probabilistes³⁸). D'abord, le signe de l'argument de l'exponentielle est purement conventionnel et on aurait pu aussi définir $F(k)$ avec e^{-ikx} . Par ailleurs, pour des raisons qui seront plus claires dans la suite, on trouve aussi des définitions ayant des facteurs additionnels³⁹ ; par exemple :

$$F_1(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{ikx} dx, \quad F_2(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(x) e^{ikx} dx. \quad (8.111)$$

Enfin, outre ces variations de facteur global, on trouve aussi :

$$F_3(\kappa) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{2i\pi\kappa x} dx. \quad (8.112)$$

Ces variations sont purement conventionnelles, les différentes transformées étant trivialement reliées les unes aux autres (par exemple, $F_1(k) = F_3(\kappa = \frac{k}{2\pi})$). Toutefois, une fois choisie une convention, il faut s'y tenir jusqu'au bout, notamment dans l'écriture des relations inverses qui permettent d'exprimer f en fonction de F . Dans la suite, on adopte la définition (8.108) ; il faut bien sûr savoir jongler avec les définitions en cas de besoin, et pouvoir passer rapidement de l'une à l'autre si nécessaire.

Cela étant précisé, il convient de donner les conditions dans lesquelles une fonction donnée $f(x)$ admet une transformée de Fourier, c'est-à-dire d'énoncer les conditions dans lesquelles l'intégrale de définition (8.108) existe. Il s'agit d'une intégrale impropre qui s'étend entre $\pm\infty$, de sorte que la fonction f doit déjà posséder les bonnes propriétés à l'infini⁴⁰. En outre, f ne doit pas posséder de singularités trop sévères sur l'axe réel. Il s'avère que les conditions générales sont difficiles à écrire ; on se bornera ici à énoncer une condition *suffisante* d'existence : si la fonction f a une limite en $|x| = \infty$ et est absolument intégrable⁴¹, alors sa transformée de Fourier existe.

³⁸La définition (8.108) est celle utilisée en Théorie des probabilités, qui associe à toute distribution de probabilité d'une aléatoire X sa *fonction caractéristique* par l'égalité $\phi(t) \stackrel{\text{déf}}{=} \langle e^{itX} \rangle$ (voir chapitre 1). Quand une densité $p(x)$ (au sens large) existe, $\phi(t)$ s'exprime comme l'intégrale de Fourier $\int_{-\infty}^{+\infty} p(x) e^{itx} dx$; la normalisation conventionnelle des probabilités à l'unité entraîne que $\phi(0) = 1$, d'où la préférence des probabilistes pour la définition (8.108). De surcroît, celle-ci conduit à une relation très simple entre la transformée de Fourier $F(k)$ et la transformée de Laplace $F_L(z)$ de $Y(t)f(t)$: $F(k) = F_L(z = -ik)$, $Y(t)$ étant la fonction de Heaviside.

³⁹En Mécanique quantique, c'est par une convention du genre F_1 (et avec $p = \hbar k$) que l'on passe de la représentation- q à la représentation- p et inversement.

⁴⁰Quand f est de signe constant à partir d'un certain x_0 (positif par exemple), il suffit que f décroisse assez vite pour que l'intégrale impropre existe ; ceci revient à dire essentiellement que f est en fait sommable en module (quand il n'y a pas par ailleurs d'autres singularités sur l'axe \mathbb{R} , exigeant une régularisation à la *Cauchy* par exemple). Sans cette propriété de signe constant, la décroissance à zéro n'est pas nécessaire pour que la fonction soit sommable – voir les intégrales de Fresnel. Se souvenir aussi du fait que, même pour une fonction de signe constant, l'intégrabilité n'exige pas la décroissance à zéro en général – d'où la différence de fond entre convergence d'une série numérique et convergence d'une intégrale.

⁴¹Attention ! Dans le langage conventionnel des mathématiciens, une telle fonction est dite *intégrable*, tout simplement. En cas d'ambiguïté, mieux vaut préciser les choses : *intégrable* si $\int_{-\infty}^{+\infty} f(x) dx$ existe, *absolument intégrable* si $\int_{-\infty}^{+\infty} |f(x)| dx$ existe. L'espace des fonctions absolument intégrables sur \mathbb{R} est traditionnellement noté $\mathbb{L}_1(\mathbb{R})$. Cette définition vaut aussi pour les fonctions à valeurs complexes, et c'est alors le module du nombre complexe f qu'il convient de considérer.

En effet, puisque f a une limite et est absolument intégrable, cette limite est nulle. Soit alors la suite $F_X \stackrel{\text{déf}}{=} \int_{-\infty}^X f(x) e^{ikx} dx$; on a $|F_{X+\Delta X} - F_X| = |\int_X^{X+\Delta X} f(x) e^{ikx} dx|$, d'où (prenant $\Delta X > 0$) $|F_{X+\Delta X} - F_X| \leq \int_X^{X+\Delta X} |f(x)| dx$: avec ΔX quelconque mais fixé, cette dernière intégrale tend vers zéro quand $X \rightarrow +\infty$, montrant que la suite F_X converge :

$$F_X(k) \stackrel{\text{déf}}{=} \int_{-\infty}^X f(x) e^{ikx} dx : \quad \exists \lim_{X \rightarrow +\infty} F_X(k) . \quad (8.113)$$

De surcroît, cette limite est majorée puisque $\forall k \in \mathbb{R}$:

$$|F_X(k)| = \left| \int_{-\infty}^X f(x) e^{ikx} dx \right| \leq \int_{-\infty}^X |f(x) e^{ikx}| dx = \int_{-\infty}^X |f(x)| dx \leq \int_{-\infty}^{+\infty} |f(x)| dx . \quad (8.114)$$

Par hypothèse, la dernière intégrale est finie, assurant que $F(k)$, limite établie de la suite $F_X(k)$, est elle aussi finie.

Il convient de se souvenir que cette condition est suffisante mais pas nécessaire : il existe des fonctions non absolument intégrables qui ont pourtant une transformée de Fourier, par exemple⁴² $f(x) = \frac{\sin x}{x}$:

$$\lim_{R \rightarrow +\infty} \int_{-R}^{+R} \left| \frac{\sin x}{x} \right| dx = +\infty , \quad (8.115)$$

et pourtant :

$$\int_{-\infty}^{+\infty} \frac{\sin x}{x} e^{ikx} dx = \pi [Y(k+1) - Y(k-1)] , \quad (8.116)$$

où Y est la fonction de Heaviside⁴³, ou encore⁴⁴ :

$$\int_{-\infty}^{+\infty} e^{i\alpha x^2} e^{ikx} dx = e^{i\frac{\pi}{4}} \sqrt{\frac{\pi}{\alpha}} e^{i\frac{k^2}{4\alpha}} \quad (\alpha \in \mathbb{R}_+) . \quad (8.118)$$

La question naturelle venant maintenant à l'esprit est la suivante : étant donné une fonction $f(x)$ admettant une transformée de Fourier conformément à (8.108), existe-t-il un moyen de retrouver $f(x)$ à partir de $F(k)$? Autrement dit, dans le droit fil de

⁴²(8.116) se démontre en utilisant $\mathcal{P} \int_{-\infty}^{+\infty} \frac{1}{x} e^{ikx} dx = i\pi \operatorname{sgn} k$, comme le montre le calcul par résidus (voir (6.86)), et en remarquant que $Y(k) = \frac{1}{2}(1 + \operatorname{sgn} k)$. En sous-produit, on retombe sur l'intégrale connue $\int_{-\infty}^{+\infty} \frac{\sin x}{x} dx = \pi$.

⁴³La combinaison $Y(k+1) - Y(k-1)$ est la "fonction-porte", qui vaut 1 entre -1 et $+1$ et zéro ailleurs. La fonction $\frac{\sin x}{x}$ est appelée *sinus cardinal*.

⁴⁴Le résultat exprimé en (8.118) a été obtenu précédemment comme exemple de prolongement analytique de l'intégrale gaussienne (voir le rappel dans la note 50, p.373). Quand $k = 0$, le premier membre est aussi égal à $2(C + iS)$ où C et S sont les intégrales de Fresnel.

Autre exemple de fonction non absolument intégrable et ayant pourtant une transformée de Fourier : la fonction d'Airy $\operatorname{Ai}(x)$ (que l'on retrouvera au chapitre 9) ; c'est l'une des solutions de l'équation $y''(x) - xy(x) = 0$, et a la représentation intégrale :

$$\operatorname{Ai}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i(xt + \frac{t^3}{3})} dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ixt} e^{-i\frac{t^3}{3}} dt . \quad (8.117)$$

La deuxième écriture montre que $\mathcal{F}[\operatorname{Ai}](k) = e^{-i\frac{k^3}{3}}$.

(8.109), on se pose la question de la transformation de Fourier *inverse* :

$$f = \mathcal{F}^{-1}[F] \iff F \xrightarrow{\mathcal{F}^{-1}} f, \quad (8.119)$$

L'établissement de la formule d'inversion repose sur le *théorème de la convergence dominée* dû à Lebesgue, dont l'énoncé est rappelé par commodité (voir aussi p. 21) :

Soit une suite de fonctions $f_n(x)$ intégrables convergeant simplement⁴⁵ vers la fonction $f(x)$. S'il existe une fonction intégrable positive $g(x)$ telle que $|f_n(x)| \leq g(x) \forall x$, alors :

$$\lim_{n \rightarrow +\infty} \int_{-\infty}^{+\infty} f_n(x) dx = \int_{-\infty}^{+\infty} \lim_{n \rightarrow +\infty} f_n(x) dx \equiv \int_{-\infty}^{+\infty} f(x) dx, \quad (8.120)$$

Ainsi, ce théorème stipule les conditions permettant d'échanger l'ordre des opérations intégration et limite⁴⁶. En notant $[\sharp, \flat] \equiv \sharp \flat - \flat \sharp$ le commutateur des deux opérations (comme on le fait en Mécanique quantique pour les opérateurs), le théorème précise les conditions où :

$$\left[\lim_{n \rightarrow +\infty}, \int_{-\infty}^{+\infty} \right] = 0 \quad (8.121)$$

sans rien préjuger de la convergence uniforme de la suite de fonctions $f_n(x)$.

Montrons maintenant, avec des hypothèses contraignantes pour simplifier la démonstration⁴⁷, que la transformation inverse de Fourier s'écrit comme suit :

$$\boxed{f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(k) e^{-ikx} dk} \quad (8.122)$$

Pour cela, considérons la fonction $I_n(x)$ définie par l'intégrale suivante :

$$I_n(x) \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} F(k) e^{-(a\frac{k}{n})^2} e^{-ikx} dk \equiv \int_{-\infty}^{+\infty} u_n(k, x) dk \quad (a > 0, n \in \mathbb{N}^*) . \quad (8.123)$$

La suite de fonctions $u_n(k, x)$ satisfait :

$$|u_n(k, x)| = |F(k) e^{-(a\frac{k}{n})^2} e^{-ikx}| = |F(k) e^{-(a\frac{k}{n})^2}| \leq |F(k)| . \quad (8.124)$$

⁴⁵ce qui signifie que pour tout x fixé, f_n calculée au point x tend vers f calculée au point x quand n tend vers l'infini.

⁴⁶On sait qu'une telle inversion est légitime pour une intégrale sur un intervalle fini et s'il y a convergence uniforme. Ici, d'une part l'intégrale est impropre, d'autre part la convergence uniforme n'est pas requise.

⁴⁷Notamment en admettant d'emblée que $f(x)$ est *absolument* intégrable, condition suffisante comme on l'a vu de l'existence de $F(k)$. La suite est donc typiquement une démonstration *faible* d'un résultat qui reste vrai quand certaines d'entre elles sont abandonnées.

On sait déjà que, comme $f(x)$ est supposée absolument intégrable, $F(k)$ est bornée en module, $|F(k)| \leq M$ (voir (8.114)) ; supposons de plus⁴⁸ que $F(k)$ est elle aussi *absolument intégrable*. La suite u_n convergeant vers $F(k)e^{-ikx}$, le théorème de la convergence dominée, faisant jouer par $|F(k)|$ le rôle de $g(x)$, permet alors d'écrire :

$$\lim_{n \rightarrow +\infty} I_n(x) = \int_{-\infty}^{+\infty} F(k) e^{-ikx} dk . \quad (8.125)$$

Ceci est vrai quelque soit x : la suite de fonctions $I_n(x)$ converge donc (au moins) simplement vers une certaine fonction $I(x)$.

Examinons maintenant d'une autre façon le second membre de (8.123). Par la définition (8.108) de F , et supposant réunies toutes les conditions permettant d'utiliser le théorème de Fubini⁴⁹, ce second membre s'écrit s'écrit :

$$\int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} f(x') e^{ikx'} dx' \right) e^{-\frac{a^2 k^2}{n^2}} e^{-ikx} dk = \int_{-\infty}^{+\infty} f(x') dx' \int_{-\infty}^{+\infty} dk e^{ik(x'-x)} e^{-\frac{a^2 k^2}{n^2}} . \quad (8.126)$$

L'intégration interne en k se fait aisément⁵⁰, donnant :

$$\int_{-\infty}^{+\infty} f(x') \frac{n\sqrt{\pi}}{a} e^{-\frac{n^2}{4a^2}(x-x')^2} dx' ; \quad (8.130)$$

posant $x' - x = \frac{2a}{n}\xi$, il vient :

$$\begin{aligned} \frac{n\sqrt{\pi}}{a} \int_{-\infty}^{+\infty} f\left(x + \frac{2a\xi}{n}\right) e^{-\xi^2} \frac{2a}{n} d\xi &= 2\sqrt{\pi} \int_{-\infty}^{+\infty} f\left(x + \frac{2a\xi}{n}\right) e^{-\xi^2} d\xi \equiv \\ &= 2\sqrt{\pi} \int_{-\infty}^{+\infty} v_n(x, \xi) d\xi . \end{aligned} \quad (8.131)$$

⁴⁸L'espace $\mathbb{L}_1(\mathbb{R})$ n'est pas stable par \mathcal{F} .

⁴⁹Voir par exemple [2], section 3.4. Essentiellement, ce théorème dit que si les fonctions $x \rightarrow \phi(x, y)$ (*resp.* $y \rightarrow \phi(x, y)$) est intégrable en x pour presque tout y (*resp.* en y pour presque tout x) et si $\phi(x, y)$ est absolument intégrable sur \mathbb{R}^2 , alors on peut échanger l'ordre des intégrations en x et y pour calculer l'intégrale double $\iint_{\mathbb{R}^2} \phi(x, y) dx dy$. Ici, $f(x')$ est supposée absolument intégrable (condition suffisante d'existence de $\mathcal{F}[f]$) ; le facteur gaussien est aussi absolument intégrable. Quant à l'intégrale double, c'est juste la définition de I_n , qui est supposée exister depuis le début.

⁵⁰On rappelle que :

$$\int_{-\infty}^{+\infty} e^{-ax^2+bx} dx = \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a}} \quad (-\pi < \arg a < +\pi) . \quad (8.127)$$

En prenant b imaginaire pur, on obtient la transformée de Fourier de la Gaussienne :

$$\int_{-\infty}^{+\infty} e^{-ax^2} e^{ikx} dx = \sqrt{\frac{\pi}{a}} e^{-\frac{k^2}{4a}} \quad (-\pi < \arg a < +\pi) . \quad (8.128)$$

Dans la limite $a \rightarrow 0$, on trouve enfin (voir (8.91), p. 366, avec $x_0 = 0$ et $n = \frac{1}{4a}$) :

$$\int_{-\infty}^{+\infty} e^{ikx} dx = 2\pi\delta(k) , \quad (8.129)$$

une autre façon de trouver la représentation intégrale de la fonction de Dirac, (8.138). Noter enfin que cette même limite s'obtient en prenant $a \in \mathbb{C} \setminus \mathbb{R}_-$, en particulier a imaginaire pur (faire le lien avec le propagateur quantique $U(x, t)$ d'un Hamiltonien quadratique, qui doit se déduire à $\delta(x)$ si $t \rightarrow 0$).

Supposons maintenant que la fonction f est non seulement absolument intégrable mais également bornée, $|f(x)| \leq M'$. Alors, $v_n(x, \xi)$ est l'élément général d'une suite de fonctions toutes bornées par $g(\xi) \equiv M' e^{-\xi^2}$, suite qui pour $n \rightarrow +\infty$ converge simplement vers $f(x) e^{-\xi^2}$. Le théorème de la convergence dominée permet alors d'affirmer que :

$$\lim_{n \rightarrow +\infty} 2\sqrt{\pi} \int_{-\infty}^{+\infty} f\left(x + \frac{2a\xi}{n}\right) e^{-\xi^2} d\xi = 2\sqrt{\pi} \int_{-\infty}^{+\infty} f(x) e^{-\xi^2} d\xi = 2\sqrt{\pi} f(x) \int_{-\infty}^{+\infty} e^{-\xi^2} d\xi = 2\pi f(x) , \quad (8.132)$$

d'où :

$$\lim_{n \rightarrow +\infty} I_n(x) = 2\pi f(x) . \quad (8.133)$$

Comparant ce résultat avec (8.125), il vient :

$$2\pi f(x) = \int_{-\infty}^{+\infty} e^{-ikx} F(k) dk , \quad (8.134)$$

ce qui établit la formule d'inversion (8.122). La démonstration a été faite en supposant $f(x)$ bornée (et intégrable), deux conditions suffisantes mais non nécessaires (voir exemples ci-dessous, p. 378).

Cela étant, il vaut aussi la peine de raisonner, comme suit, ne serait-ce qu'à titre mnémotechnique. En combinant les deux formules (8.108) et (8.134), on peut écrire pour "bonne" toute fonction $f(x)$:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk e^{-ikx} \int_{-\infty}^{+\infty} dx' f(x') e^{ikx'} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dx' f(x') \int_{-\infty}^{+\infty} dk e^{ik(x'-x)} . \quad (8.135)$$

Se souvenant alors de la relation d'usage de la "fonction" de Dirac :

$$\int_{-\infty}^{+\infty} dx' f(x') \delta(x' - x) = f(x) , \quad (8.136)$$

on voit que l'on a⁵¹ :

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ik(x'-x)} dk = \delta(x' - x) , \quad (8.138)$$

soit :

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} 1 \times e^{-ikx} dk = \delta(x) ; \quad (8.139)$$

⁵¹On a déjà utilisé de fait cette relation. Par ailleurs, se souvenir que δ est une "fonction" paire ; posant $k = -k'$ dans l'intégrale de (8.138), on a aussi :

$$\delta(x' - x) = \frac{1}{2\pi} \int_{+\infty}^{-\infty} e^{i(-k')(x'-x)} d(-k') = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ik'(x-x')} dk' = \delta(x - x') . \quad (8.137)$$

La relation (8.139) peut d'ailleurs s'établir en remplaçant 1 dans l'intégrale par $e^{-\frac{1}{2}\varepsilon k^2}$; l'intégrale gaussienne se calcule aisément et vaut $\frac{1}{\sqrt{2\pi\varepsilon}} e^{-\frac{x^2}{2\varepsilon}}$, qui est un précurseur gaussien de $\delta(x)$.

cette relation, établie intuitivement et sans grande rigueur, montre que $\delta(x)$ est la \mathcal{F}^{-1} de la fonction égale à 1 partout. Ceci se confirme en contemplant l'égalité⁵² $\int_{-\infty}^{+\infty} e^{ikx} \delta(x) dx = 1$, qui exprime le fait que $\mathcal{F}[\delta(x)] = 1$; au total :

$$\mathcal{F}[\delta(x)] = 1, \quad \mathcal{F}^{-1}[1] = \delta(x), \quad (8.140)$$

des résultats qui ont été établis dans le cadre de la Théorie des distributions (voir chapitre 2, notamment (2.123)). Plus généralement, on a :

$$\mathcal{F}[\delta(x - x_0)] = e^{ikx_0}, \quad \mathcal{F}^{-1}[e^{ikx_0}] = \delta(x - x_0). \quad (8.141)$$

Faisant toujours fi de tout état d'âme, on en déduit les relations réciproques :

$$\mathcal{F}[e^{-ik_0x}] = 2\pi\delta(k - k_0), \quad \mathcal{F}^{-1}[\delta(k - k_0)] = \frac{1}{2\pi}e^{-ik_0x}. \quad (8.142)$$

C'est en jouant avec ces résultats que l'on peut écrire la transformée de Fourier de la fonction de Weierstrass introduite dans la note 14 p. 351 (qui est continue et absolument intégrable si $b < 1$) :

$$\mathcal{F}[\mathcal{W}(x)](k) = \sum_{n \in \mathbb{N}} b^n \mathcal{F}[\cos(a^n \pi x)] = \pi \sum_{n \in \mathbb{N}} b^n [\delta(k + a^n \pi) + \delta(k - a^n \pi)]. \quad (8.143)$$

À nouveau, il s'agit d'un peigne de Dirac irrégulier, dont les amplitudes s'écrasent exponentiellement, et dont les points de concentration forment une suite convergente si $a < 1$ (là où on sait que \mathcal{W} est dérivable), ou une suite exponentiellement divergente ($a > 1$), et alors on sait que \mathcal{W} n'est nulle part dérivable.

Ces calculs symboliques permettent aussi d'établir rapidement une formule, dite *formule de sommation de Poisson*. En effet soit une fonction $f(x)$ et sa transformée de Fourier $F(k)$. On peut alors écrire, supposant satisfaites toutes les hypothèses autorisant les diverses opérations successives :

$$\sum_{n \in \mathbb{Z}} f(x - na) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} \int_{-\infty}^{+\infty} e^{-ik(x-na)} F(k) dk = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk e^{-ikx} F(k) \sum_{n \in \mathbb{Z}} e^{in ka}; \quad (8.144)$$

on retrouve à droite le peigne de Dirac $\text{III}(ka)$ défini en (8.84), de sorte que :

$$\sum_{n \in \mathbb{Z}} f(x - na) = \int_{-\infty}^{+\infty} dk e^{-ikx} F(k) \sum_{p \in \mathbb{Z}} \delta(ka - p 2\pi), \quad (8.145)$$

d'où la formule de sommation de Poisson :

$$\boxed{\sum_{n \in \mathbb{Z}} f(x - na) = \frac{1}{a} \sum_{p \in \mathbb{Z}} F\left(p \frac{2\pi}{a}\right) e^{-ip \frac{2\pi}{a} x}} \quad (8.146)$$

⁵²qui résulte de la règle d'usage de $\delta(x)$.

Ceci n'est autre que la décomposition de Fourier d'une fonction périodique (voir la Remarque 5 p. 379) ; noter que si $a \rightarrow 0$, l'égalité $\sum_{n \in \mathbb{Z}} a f(x - na) = \sum_{p \in \mathbb{Z}} F(p \frac{2\pi}{a}) e^{-ip \frac{2\pi}{a} x}$ donne $\int_{\mathbb{R}} f(x) dx = F(0)$ puisque, dans cette limite, la série au second membre se réduit au terme $p = 0$. En particulier, pour $x = 0$, (8.146) devient :

$$\boxed{\sum_{n \in \mathbb{Z}} f(na) = \frac{1}{a} \sum_{p \in \mathbb{Z}} F(p \frac{2\pi}{a})} \quad (8.147)$$

Par exemple, avec $f(x) = e^{-k_0|x|}$, donc $F(k) = \frac{2k_0}{k^2 + k_0^2}$, cette formule donne :

$$\sum_{n \in \mathbb{Z}} e^{-|n|k_0 a} = \frac{ak_0}{2\pi^2} \sum_{p \in \mathbb{Z}} \frac{1}{p^2 + (\frac{k_0 a}{2\pi})^2}, \quad (8.148)$$

que l'on peut vérifier en sommant le premier membre (série géométrique, égale à $\cot \frac{k_0 a}{2}$), et le second par la méthode $\pi \cot \pi z$ & Co vue au chapitre 6. Avec des fonctions gaussiennes⁵³, la formule sommatoire de Poisson donne⁵⁴ :

$$\sum_{n \in \mathbb{Z}} e^{-n^2 \alpha^2} = \sqrt{\frac{\pi}{\alpha^2}} \sum_{p \in \mathbb{Z}} e^{-p^2 \frac{\pi^2}{\alpha^2}}, \quad (8.151)$$

une égalité pas si évidente que cela. En particulier, elle montre que si $\alpha \ll 1$, alors $\sum_{n \in \mathbb{Z}} e^{-n^2 \alpha^2} \simeq \sqrt{\frac{\pi}{\alpha^2}}$, un résultat qui ne saute pas aux yeux en regardant le premier membre : s'il est clair qu'il prend de grandes valeurs pour $\alpha \ll 1$, on ne sait pas comment il se comporte en α ; la formule de Poisson permet d'élucider aisément ce point. Par ailleurs, on voit que l'on passe d'une gaussienne large dans l'espace direct à une gaussienne étroite dans l'espace réciproque, premier exemple de la propriété d'échelle discutée ci-dessous p. 383.

On devine que la formule sommatoire de Poisson joue un rôle important pour définir les *accélérateurs de convergence*, procédés permettant de réduire le temps de calcul numérique d'une série : il saute aux yeux que si $\alpha \gg 1$, la série au premier membre de (8.151) converge très vite, celle qui est au second membre converge très lentement (et inversement si $\alpha \ll 1$). Ces deux séries sont égales : s'agissant de les calculer numériquement, une expression s'impose à l'autre selon la valeur de α . La méthode d'Ewald, utilisée

⁵³La transformée de Fourier d'une gaussienne est une gaussienne : $\mathcal{F}[e^{-\lambda^2 x^2}] = \sqrt{\frac{\pi}{\lambda^2}} e^{-\frac{k^2}{4\lambda^2}}$. Noter que si $\alpha = \sqrt{\pi}$, (8.151) est une identité triviale !

⁵⁴Le premier membre de (8.151) est $\vartheta_3(0, e^{-\alpha^2})$ où ϑ_3 est l'une des fonctions de Jacobi, définie comme :

$$\vartheta_3(x, q) \stackrel{\text{d'ef}}{=} \sum_{n \in \mathbb{Z}} q^{n^2} e^{2in x} \quad (|q| < 1); \quad (8.149)$$

l'égalité (8.151) est ainsi équivalente à :

$$\vartheta_3(0, e^{-\lambda^2}) = \frac{\sqrt{\pi}}{\lambda} \vartheta_3(0, e^{-\frac{\pi^2}{\lambda^2}}) \iff \vartheta_3(0, q) = \sqrt{\frac{\pi}{-\ln q}} \vartheta_3(0, e^{\frac{\pi^2}{\ln q}}). \quad (8.150)$$

notamment pour calculer aisément des énergies de réseau en présence de forces à longue portée, utilise intensément ce type de procédure.

◆ *Remarques*

1. Il est évidemment possible de définir la transformée de Fourier d'une fonction discontinue ; par exemple, si $f(x)$ a un saut fini en $x = x_0$, sa transformée de Fourier $F(k)$ se définit naturellement comme :

$$\mathcal{F}[f](k) \equiv F(k) = \int_{-\infty}^{x_0} e^{ikx} f(x) dx + \int_{x_0}^{+\infty} e^{ikx} f(x) dx , \quad (8.152)$$

et est parfaitement définie dans l'hypothèse où chaque intégrale existe. Réciproquement, la fonction $\tilde{f}(x)$ obtenue par transformation inverse :

$$\tilde{f}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ikx} F(k) dk , \quad (8.153)$$

est bien définie, et coïncide avec $f(x) \forall x \neq x_0$. En x_0 , où \tilde{f} existe, la comparaison n'a pas de sens puisque $f(x_0)$ n'existe pas (l'usage est alors de dire que f et \tilde{f} sont *presque partout* égales). Tout comme pour les séries de Fourier, des exemples permettent de réaliser que $\tilde{f}(x_0)$ est égal à la demi-somme des valeurs de $f(x)$ à gauche et à droite (voir ci-dessous). Ceci reste évidemment vrai si $f(x)$ possède un nombre fini de sauts finis.

Ce type de résultat doit être manié avec prudence en Physique, puisque la notion de discontinuité résulte toujours d'une idéalisation. Le cas échéant, il convient de revenir en arrière pour préciser la nature physique de la "discontinuité" et éviter ainsi de dire des âneries⁵⁵.

Illustrons ce qui précède par deux exemples. Soit d'abord la fonction $f(x) \stackrel{\text{déf}}{=} e^{-k_0|x|}$ ($k_0 > 0$) ; sa transformée se calcule par les moyens élémentaires (on décompose en deux l'intégrale pour expliciter le terme contenant $|x|$ et on utilise une primitive connue) et vaut :

$$F(k) = \int_{-\infty}^{+\infty} e^{ikx} e^{-k_0|x|} dx = \frac{2k_0}{k^2 + k_0^2} . \quad (8.154)$$

Dans ce cas, f et F sont toutes deux dans $\mathbb{L}_1(\mathbb{R})$. Soit ensuite l'exponentielle unilatérale $g(x) \stackrel{\text{déf}}{=} Y(x) e^{-k_0x}$ où $Y(x)$ est la fonction de Heaviside : $Y(x < 0) = 0$, $Y(x > 0) = 1$; la transformée de Fourier de $g(x)$ est :

$$G(k) = \int_{-\infty}^{+\infty} e^{ikx} Y(x) e^{-k_0|x|} dx = \int_0^{+\infty} e^{ikx} Y(x) e^{-k_0|x|} dx = \frac{1}{-ik + k_0} . \quad (8.155)$$

⁵⁵Par exemple, si la variable est le temps, le Principe de causalité n'est pas bien loin. Il serait stupide d'écrire qu'un certain effet $e(t)$ de la cause $f(t)$ est de la forme $C \frac{1}{2} [f(t+0) - f(t-0)]$! En pareil cas, ce qu'il faut considérer, c'est bien évidemment la limite à *droite*, et écrire $e(t) = Cf(t-0)$.

Cette fonction n'est pas dans $\mathbb{L}_1(\mathbb{R})$: le module de $G(k)$ est $(k^2 + k_0^2)^{-\frac{1}{2}}$ et n'est pas intégrable. Cet exemple suffit à montrer que $\mathbb{L}_1(\mathbb{R})$ n'est pas stable par \mathcal{F} .

Ces deux exemples lèvent le voile sur une tendance : plus une fonction est régulière à l'origine $x = 0$, plus sa transformée décroît vite à l'infini : $f(x)$ est continue (mais non partout dérivable), et $F(k)$ décroît comme k^{-2} . Au contraire, $g(x)$ est discontinue, et $G(k)$ décroît seulement comme k^{-1} .

Que donne la formule d'inversion de Fourier ? Pour $F(k)$, il faut calculer l'intégrale :

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ikx} \frac{2k_0}{k^2 + k_0^2} dk . \quad (8.156)$$

Par résidus, on trouve qu'elle vaut $e^{-k_0|x|}$, $\forall x \in \mathbb{R}$, qui est très exactement $f(x)$, partout y compris en $x = 0$, conformément au résultat général exprimé par la formule d'inversion pour un couple dans $\mathbb{L}_1^2(\mathbb{R})$. Pour $G(k)$, il faut trouver l'intégrale :

$$\tilde{f}(x) \stackrel{\text{d\'ef}}{=} \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ikx} \frac{1}{-ik + k_0} dk . \quad (8.157)$$

Le calcul par résidus donne 0 si $x < 0$, $e^{-k_0|x|}$ si $x > 0$. Pour $x = 0$, et si on décide de compléter la définition de $\tilde{f}(x)$ en décrétant que $\tilde{f}(0)$ est le nombre égal à l'intégrale où l'on fait $x = 0$, le calcul direct est immédiat puisque l'on connaît une primitive ; on a :

$$\lim_{R \rightarrow +\infty} \frac{1}{2\pi} \int_{-R}^{+R} \frac{1}{-ik + k_0} dk = -\frac{1}{2i\pi} \lim_{R \rightarrow +\infty} \ln(k + ik_0) \Big|_{-R}^{+R} = -\frac{1}{2i\pi} (-i\pi) = \frac{1}{2} . \quad (8.158)$$

La formule d'inversion redonne donc une fonction presque partout égale à l'original, sauf au point de discontinuité où elle donne la demi-somme des valeurs à gauche et à droite.

2. L'existence de $\mathcal{F}[f]$ n'exige pas que f soit bornée, une hypothèse qui a été utilisée pour donner une démonstration faible de la formule d'inversion (une fonction non bornée peut être absolument intégrable : $x^{-\frac{1}{2}}e^{-x}$ n'est pas bornée sur \mathbb{R}_+). Autre exemple : la fonction $f(x) = \ln(1 + \frac{a^2}{x^2})$ n'est pas bornée sur \mathbb{R} (mais est intégrable), et a une transformée :

$$\mathcal{F}\left[\ln\left(1 + \frac{a^2}{x^2}\right)\right](k) = \frac{2\pi}{|k|} (1 - e^{-|k|a}) \quad (a > 0) . \quad (8.159)$$

Ce résultat peut s'obtenir en effectuant une intégration par parties suivie d'une intégration par résidus⁵⁶.

Enfin, la fonction $\frac{Y(x)}{\sqrt{x}}$ n'est ni bornée ni intégrable (elle décroît trop lentement à l'infini). Elle a cependant une transformée de Fourier, égale⁵⁷ à $(i\pi/k)^{1/2}$, où la

⁵⁶La transformée est finie en $k = 0$, et vaut $\int_{-\infty}^{+\infty} \ln(1 + \frac{a^2}{x^2}) dx = 2\pi a$, comme il se doit.

⁵⁷Plus généralement $\mathcal{F}[Y(x)x^{-\nu}](k) = (i/k)^{1-\nu}\Gamma(1-\nu)$, $0 < \Re\nu < 1$, comme on le voit en déformant le contour pour le confondre avec l'axe imaginaire dans le plan \mathbb{C} de x .

branche de la racine carrée est coupée⁵⁸ sur \mathbb{R}_- . La divergence de la transformée en $k = 0$ est le symptôme de la décroissance lente de l'original quand $x \rightarrow +\infty$. Le calcul direct de $\frac{1}{2\pi} \int_{-\infty}^{+\infty} (i\pi/k)^{1/2} e^{-ikx} dk$ donne 0 si $x < 0$ et $x^{-1/2}$ si $x > 0$, soit $Y(x)x^{-\frac{1}{2}}$, montrant que la formule d'inversion de Fourier fonctionne bien encore dans ce cas.

3. Les couples de variables x et k , ou t et ω , apparaissant dans une transformation de Fourier sont souvent appelées *variables conjuguées*. Cette terminologie est notamment fréquente en Mécanique Quantique, où x est une coordonnée et où k est un nombre d'onde associé à l'impulsion⁵⁹ p ($p = \hbar k$). Les fonctions d'onde des deux représentations- q et $-p$ sont transformées de Fourier l'une de l'autre.
4. Il est d'usage, en Physique, de parler d'espace *direct* pour désigner l'espace où se promène le point associé à la variable x , et d'espace *réciroque* pour désigner celui de la variable k – tout comme les cristallographes parlent d'espace direct et d'espace réciroque. Il est clair que ces dénominations sont relatives, et n'ont de signification que par rapport à un présupposé mental (souvent implicite)
5. La formule de sommation de Poisson, (8.146), a une interprétation simple. Le premier membre de cette égalité est en fait une certaine fonction a -périodique⁶⁰ $\varphi(x) \stackrel{\text{déf}}{=} \sum_{n \in \mathbb{Z}} f(x - na)$, car de toute évidence (on pose $n - 1 = n'$ dans la sommation) :

$$\sum_{n \in \mathbb{Z}} f(x - na + a) = \sum_{n' \in \mathbb{Z}} f(x - n'a) \iff \varphi(x + a) = \varphi(x) . \quad (8.160)$$

En tant que telle, $\varphi(x)$ admet une décomposition en série de Fourier :

$$\varphi(x) = \sum_{p \in \mathbb{Z}} \varphi_p e^{ip \frac{2\pi}{a} x} . \quad (8.161)$$

La formule de Poisson (8.147) ne fait qu'exprimer l'existence de cette série de Fourier, et donne les coefficients $\varphi_p = \frac{1}{a} F(p \frac{2\pi}{a})$. \blacklozenge

8.2.2 Propriétés de la transformation de Fourier

De la définition (8.108) résultent immédiatement quelques propriétés importantes en pratique. On note toujours :

$$f(x) \xrightarrow{\mathcal{F}} F(k) , \quad F = \mathcal{F}[f] ; \quad F(k) \xrightarrow{\mathcal{F}^{-1}} f(x) , \quad f = \mathcal{F}^{-1}[F] . \quad (8.162)$$

⁵⁸C'est ainsi que l'on a $(i/k)^{1/2} = e^{i\frac{\pi}{4}}/\sqrt{k}$ si $k > 0$ et $(i/k)^{1/2} = e^{-i\frac{\pi}{4}}/\sqrt{-k}$ si $k < 0$. Visiblement, si $k = i$, la transformée de Fourier vaut $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

⁵⁹À ce propos, toujours en Mécanique quantique, on parle plus précisément de *grandeurs incompatibles*.

⁶⁰De la même façon, toute fonction de la forme $\sum_{n \in \mathbb{Z}} f(t - nT)$ représente un signal T -périodique.

■ **Linéarité** L'intégrale étant une opération linéaire, on a de toute évidence :

$$\boxed{\mathcal{F}[\lambda f + \mu g] = \lambda \mathcal{F}[f] + \mu \mathcal{F}[g]} \quad (8.163)$$

où λ et μ sont des *constantes* (indépendantes de la variable x), éventuellement paramétrées par d'autres variables du problème examiné.

La transformation de Fourier est donc bien adaptée à la résolution d'équations *linéaires*, mais il faut garder une chose en tête : quand on l'utilise à cette fin, on ne peut trouver par construction que des solutions ayant une transformée de Fourier. Il en résulte que si l'équation à résoudre possède des solutions n'admettant pas de transformée, on ne peut pas toujours les obtenir par ce moyen, même en s'en remettant à une forme de calcul symbolique qui permet parfois de s'en tirer – voir ci-après quelques manipulations iconoclastes.

■ **Translation** Étant donné une fonction $f(x)$, sa translatée de a est par définition la fonction⁶¹ :

$$\boxed{(\mathcal{T}_a f)(x) \stackrel{\text{déf}}{=} f(x - a)} \quad (8.164)$$

La transformée de $\mathcal{T}_a f$ est ainsi :

$$\mathcal{F}[\mathcal{T}_a f](k) = \int_{-\infty}^{+\infty} e^{ikx} f(x - a) dx = e^{ika} \int_{-\infty}^{+\infty} e^{ikx'} f(x') dx' = e^{ika} \mathcal{F}[f](k) . \quad (8.165)$$

d'où :

$$\boxed{\mathcal{F}[\mathcal{T}_a f](k) = e^{ika} \mathcal{F}[f](k)} \quad (8.166)$$

Cette propriété très simple est utilisée pour énoncer des résultats forts et importants, fondés sur l'invariance galiléenne de l'espace (physique) par translation. Par exemple, si $f(x)$ est l'attribut d'un système ayant une symétrie de translation de a , alors on a $f(x - a) = f(x)$: il en résulte que la transformée de Fourier satisfait $(e^{ika} - 1)\mathcal{F}[f] = 0$, ce qui signifie que $\mathcal{F}[f] = 0$ sauf aux points tels que $ka = \text{entier} \times 2\pi$, soit $k = \text{entier} \times \frac{2\pi}{a}$. Ceci était à attendre : $f(x - a) = f(x)$ signifie que la fonction f est a -périodique et qu'elle est décomposable en série de Fourier. En pareil cas, l'intégrale de Fourier dégénère en *série* de Fourier.

Cette propriété est utile pour résoudre certaines équations. Avant de donner un exemple, notons que l'égalité (8.166) est aussi vraie quand il s'agit de transformées de Fourier de distributions (tempérées). En effet, revenant à (2.44) avec $a \rightarrow 1$ et $b \rightarrow -a$, on définit d'abord la translatée de a d'une distribution f comme :

$$\langle \mathcal{T}_a f, \phi \rangle \equiv \langle f(x - a), \phi \rangle = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n(X) \phi(X + a) dX ; \quad (8.167)$$

⁶¹Géométriquement : on "prend" le graphe de $f(x)$ et on le déplace en bloc de a vers la droite (si $a > 0$).

puis, repartant de la définition (2.119), on peut écrire :

$$\langle \mathcal{F}[\mathcal{T}_a f], \Phi \rangle = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} \mathcal{F}[\mathcal{T}_a f_n](k) \Phi(k) dk = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} e^{ika} F_n(k) \Phi(k) dk , \quad (8.168)$$

d'où par identification :

$$\boxed{\mathcal{F}[\mathcal{T}_a f](k) = e^{ika} \mathcal{F}[f](k)} \quad (8.169)$$

mais cette fois pour une *distribution* f .

• Exemple

Soit l'équation aux différences finies :

$$f(x) = \lambda f(x - a) \quad (8.170)$$

où $\lambda, a \in \mathbb{R}_+$. Cette équation dit que la valeur de la fonction en x est λ fois sa valeur au point décalé de a à gauche. Si $\lambda > 1$, ceci veut dire que f croît quand x augmente, le contraire se produit si $\lambda < 1$. Ceci permet de deviner peu à peu la solution de (8.170) : il y a sûrement un facteur $\lambda^{\frac{x}{a}}$ dans f , puisque la fonction est multipliée par λ quand x augmente de a ; on peut aussi ajouter un facteur périodique en a , qui ne change pas si x varie de a . En définitive, cette équation possède des solutions du genre :

$$f_n(x) = \lambda^{\frac{x}{a}} e^{in\frac{2\pi}{a}x} = e^{(\ln \lambda)\frac{x}{a}} e^{in\frac{2\pi}{a}x} \quad (n \in \mathbb{Z}) . \quad (8.171)$$

L'équation étant linéaire, toute combinaison linéaire est encore solution ; il en résulte que :

$$f(x) = \lambda^{\frac{x}{a}} \sum_{n \in \mathbb{Z}} C_n e^{in\frac{2\pi}{a}x} , \quad (8.172)$$

où les C_n sont des constantes quelconques, est la solution (la plus générale ?) de cette équation ; $f(x)$ ressort ainsi comme le produit d'une série de Fourier par le facteur exponentiel $e^{(\ln \lambda)\frac{x}{a}}$, qui diverge en $+\infty$ si $\lambda > 1$ ($\ln \lambda > 0$), en $-\infty$ si $\lambda < 1$. Les fonctions du type $f(x)$ ne sauraient avoir une transformée de Fourier !

Que donne ici pourtant la transformation de Fourier ? En notant $F = \mathcal{F}[f]$ et appliquant la règle (8.166) (ou plutôt (8.169)), l'équation se transforme en :

$$F(k) = \lambda e^{ika} F(k) \iff (e^{ika} - 1) F(k) = 0 . \quad (8.173)$$

Cette équation dit que F est nulle partout, sauf aux points où $\lambda e^{ika} = 1$, où F est indéterminée, et qui sont de mesure nulle. La seule solution semble donc être $F = 0$. Faut-il pour autant jeter Fourier dans ce cas ? La réponse est non, à condition de relâcher un peu de rigueur formelle, de s'en remettre au savoir-faire et de garder en tête l'aspect *symbolique* des opérations effectuées – quitte à reporter à plus tard une vraie justification.

En effet, acceptant les relations (8.140) et (8.141) telles qu'elles sont, on peut raisonner comme suit. (8.173) dit très précisément que $F = 0$ tant que $\lambda e^{ika} \neq 1$,

mais quand $\lambda e^{ika} = 1$, $F(k)$ est indéterminée. On est donc très tenté d'écrire $F(k)$ en combinaison linéaire (avec des coefficients à trouver mais pour l'instant quelconques) des fonctions de Dirac concentrées aux zéros de $\lambda e^{ika} - 1$, soit pour les valeurs de k égales à $k_m = 2m\frac{\pi}{a} + \frac{i}{a} \ln \lambda$ ($m \in \mathbb{Z}$) :

$$F(k) = \sum_m A_m \delta(k - k_m) , \quad (8.174)$$

où les A_m sont des constantes quelconques ; en effet, comme $(k - k_m)\delta(k - k_m) = 0$ (voir (8.96)), n'importe quelle combinaison linéaire satisfait (8.173). En appliquant la règle d'usage de la fonction δ – et passant outre le fait que les k_m sont ici complexes⁶² (!!!) –, la formule d'inversion (8.142) incite à écrire :

$$\mathcal{F}^{-1}[F](x) = \sum_m A_m \frac{1}{2\pi} e^{-ik_m x} = \sum_m \frac{A_m}{2\pi} e^{-im\frac{2\pi}{a}x + \frac{x}{a} \ln \lambda} = \lambda^{\frac{x}{a}} \sum_m \frac{A_m}{2\pi} e^{-im\frac{2\pi}{a}x} . \quad (8.175)$$

À des notations près, ceci est bien la fonction $f(x)$ écrite en (8.172) et construite intuitivement ; cette coïncidence permet de suspecter que, malgré son caractère indéniablement hérétique, cette procédure de résolution peut recevoir sa légitimité.

Remarquons qu'il y a *a priori* une infinité de constantes encore indéterminées. Par analogie avec une équation différentielle, on peut dire qu'une équation aux différences finies telle que (8.170) est une équation différentielle d'ordre infini : intuitivement, si on invoque pour $f(x)$ la possibilité d'un développement de Taylor, le second membre est $\lambda \sum_{n \in \mathbb{N}} \frac{(-a)^n}{n!} f^{(n)}(x)$, qui contient *toutes* les dérivées.

Clairement, la justification de la procédure symbolique ci-dessus consiste d'abord à prendre l'équation $f(x) = \lambda f(x - a)$ au sens des distributions, puis à utiliser (8.169). Il convient aussi de donner un sens à $\delta(x - \alpha)$ avec $\alpha \in \mathbb{C}$; ceci peut se faire en introduisant la suite de fonctions $f_n(x) \stackrel{\text{déf}}{=} \sqrt{\frac{n}{\pi}} e^{-n(x-\alpha)^2}$, et à utiliser la méthode du col – qui donne le résultat *exact* pour l'intégrale gaussienne (voir p. 317) – pour obtenir (en l'absence de toute singularité du prolongement $\phi(z)$) :

$$\langle \delta(x - \alpha), \phi \rangle = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} \sqrt{\frac{n}{\pi}} e^{-n(x-\alpha)^2} \phi(x) dx = \phi(\alpha) \quad (8.176)$$

On constate une fois encore que le calcul symbolique présente l'avantage d'une économie d'écriture auquel il serait dommage de renoncer. •

■ **Modulation** Étant donnée une fonction $f(x)$, sa modulée à $k_0 \in \mathbb{R}$ est par définition $e^{-ik_0 x} f(x)$. On a :

$$\mathcal{F}[e^{-ik_0 x} f(x)](k) = \int_{-\infty}^{+\infty} e^{i(k-k_0)x} f(x) dx , \quad (8.177)$$

⁶²Intuitivement : on définit $\delta(z - z_0) \stackrel{\text{déf}}{=} \delta(x - x_0)\delta(y - y_0)$ et on fait du Fourier à deux dimensions.

soit :

$$\boxed{\mathcal{F}[e^{-ik_0x} f](k) = \int_{-\infty}^{+\infty} e^{i(k-k_0)x} f(x) dx = \mathcal{F}[f](k - k_0)} \quad (8.178)$$

La modulation dans l'espace direct revient donc à une translation dans l'espace réciproque – un résultat dual de celui obtenu à propos de la transformée de Fourier d'une fonction translatée. Si x est le temps t , on voit qu'un facteur oscillatoire $e^{-i\omega_0 t}$ pour $f(t)$ supplémentaire décale⁶³ toutes les pulsations de ω_0 pour $F(\omega)$. Clairement, (8.178) vaut aussi quand f est une distribution.

■ **Dilatation** Un changement d'échelle⁶⁴ sur la variable d'une fonction $f(x)$ définit une nouvelle fonction f_λ telle que $f_\lambda(x) = f(\lambda x)$, $\lambda \in \mathbb{R}_+$. On a :

$$\mathcal{F}[f_\lambda(x)](k) = \int_{-\infty}^{+\infty} e^{ikx} f(\lambda x) dx = \frac{1}{\lambda} \int_{-\infty}^{+\infty} e^{ik \frac{x'}{\lambda}} f(x') dx' = \frac{1}{\lambda} \mathcal{F}[f]\left(\frac{k}{\lambda}\right), \quad (8.179)$$

soit :

$$\mathcal{F}[f_\lambda](k) = \frac{1}{\lambda} \mathcal{F}[f]\left(\frac{k}{\lambda}\right) \quad (\lambda \in \mathbb{R}_+) \quad (8.180)$$

Dans le cas où λ est négatif, le changement de variable inverse le sens de parcours de l'intégrale d'où, revenant au sens "naturel", l'apparition d'un signe – transformant le λ en facteur en $|\lambda|$. En définitive :

$$\boxed{\mathcal{F}[f_\lambda](k) = \frac{1}{|\lambda|} \mathcal{F}[f]\left(\frac{k}{\lambda}\right) \quad (\lambda \in \mathbb{R})} \quad (8.181)$$

Cette relation permet de découvrir une propriété semi-quantitative importante d'un couple de Fourier (f, F) . Supposons pour fixer les idées que la fonction $f(x)$ a l'allure d'une courbe en cloche ; par construction, le graphe de la fonction f_λ a le même aspect mais, si $\lambda > 1$, est plus ramassé (f_λ est plus fine). Le phénomène inverse se produit de $\mathcal{F}[f]$ à $\mathcal{F}[f_\lambda]$: en effet, et toujours avec l'hypothèse $\lambda > 1$, (8.180) dit que le graphe de $\mathcal{F}[f_\lambda]$ est plus large que celui de $\mathcal{F}[f]$. En définitive, plus f est étroite, plus $\mathcal{F}[f]$ est large. Si on introduit la largeur typique de f , Δx , et celle, Δk , de F , on a une relation du genre :

$$\boxed{\Delta x \Delta k \sim 2\pi} \quad (8.182)$$

le facteur 2π est certainement un peu conventionnel, mais il est traditionnel, étant entendu que de toute façon le symbole \sim introduit un certain flou dans cette relation ; en acceptant d'anticiper $\lambda = \frac{h}{p}$, (8.182) est l'un des avatars de ce qui est usuellement désigné par *Principe d'incertitude de Heisenberg*, $\Delta x \Delta p_x \sim h$. Ces considérations sont

⁶³En Mécanique quantique, où seules comptent les *différences* des pulsations, l'adjonction d'un tel facteur global ne change rien aux prévisions physiques. En un sens, c'est aussi dire qu'en Physique seules comptent les *différences* d'énergie.

⁶⁴En anglais *scaling*. Il peut s'agir d'une dilatation ou d'une contraction selon que $\lambda > 1$ ou $\lambda < 1$.

fréquemment utilisées dans les raisonnements d'ordre de grandeur, et gardent un caractère semi-quantitatif ; par exemple, la structure précise des fonctions peut présenter des particularités qui échappent à une telle analyse : la fonction-porte a des flancs raides et bien définis, mais sa transformée de Fourier, le sinus cardinal, présente des oscillations. Il n'en demeure pas moins vrai que la fonction-porte "décroit vite" (à support borné, elle est strictement nulle au-delà de valeurs finies), alors que le sinus cardinal décroît très lentement, essentiellement comme $\frac{1}{x}$.

L'égalité (8.180) est aussi vraie pour les distributions ; la définition d'une distribution dilatée se pose d'elle-même en revenant à (2.44) avec $a = \lambda$ et $b = 1$:

$$\langle f_\lambda, \phi \rangle \equiv \langle f(\lambda x), \phi \rangle = \lim_{n \rightarrow +\infty} \frac{1}{|\lambda|} \int_{\mathbb{R}} f_n(X) \phi\left(\frac{X}{\lambda}\right) dX . \quad (8.183)$$

Cela étant, la définition (2.119) donne :

$$\langle \mathcal{F}[f_\lambda], \Phi \rangle = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} \mathcal{F}[f_{\lambda n}](k) \Phi(k) dk = \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} \frac{1}{|\lambda|} F_n\left(\frac{k}{\lambda}\right) \Phi(k) dk . \quad (8.184)$$

d'où, pour une distribution :

$$\boxed{\mathcal{F}[f_\lambda](k) = \frac{1}{|\lambda|} \mathcal{F}[f]\left(\frac{k}{\lambda}\right) \quad (\lambda \in \mathbb{R})} \quad (8.185)$$

■ Conjugaison

$$\mathcal{F}[f^*] = \int_{-\infty}^{+\infty} e^{ikx} f^*(x) dx = \left[\int_{-\infty}^{+\infty} e^{-ikx} f(x) dx \right]^* = F^*(-k) . \quad (8.186)$$

Cette propriété joue un rôle important en Mécanique quantique, dans la discussion de l'invariance par renversement du temps. Elle est également invoquée pour déterminer les propriétés de symétrie de la susceptibilité relative à une grandeur réelle. En effet, pour une fonction f à valeurs réelles ($f^* = f$), (8.186) donne :

$$f(x) \in \mathbb{R} \quad \forall x \quad \iff \quad F(k) = F^*(-k) . \quad (8.187)$$

■ Dérivation Un résultat très utile en pratique est le lien entre les transformées de Fourier d'une fonction et de ses dérivées ; sans surprise, on va retrouver (à des modifications de détail près) une relation du même genre que dans la présentation des séries de Fourier.

Soit f une fonction dérivable dont la dérivée est notée f' ; la transformée de cette dernière, dans l'hypothèse où elle existe, est par définition :

$$\mathcal{F}[f'] = \int_{-\infty}^{+\infty} e^{ikx} f'(x) dx . \quad (8.188)$$

Une intégration par parties donne :

$$\mathcal{F}[f'] = [f(x)e^{ikx}]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} ik e^{ikx} f(x) dx . \quad (8.189)$$

Supposons que $\lim_{x \rightarrow \pm\infty} f(x) = 0$, une propriété assurée⁶⁵ si f et sa dérivée f' sont absolument intégrables ; le terme tout intégré est alors nul, et il vient, avec les conventions adoptées⁶⁶ pour la définition de \mathcal{F} :

$$\boxed{\mathcal{F}[f'] = -ik F(k)} \quad (8.190)$$

Quand on se souvient de la formule d'inversion $f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(k) e^{-ikx} dk$, le résultat (8.190) est équivalent à dire que l'on peut dériver sous le signe \int (une bonne façon de le mémoriser) :

$$f'(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} (-ik) e^{-ikx} F(k) dk \iff \mathcal{F}[f'] = -ikF(k) . \quad (8.191)$$

Le même travail donne ($f^{(m)}(x) = \frac{d^m f}{dx^m}$) :

$$\boxed{\mathcal{F}[f^{(m)}] = (-ik)^m F(k)} \quad (8.192)$$

Ces propriétés jouent un rôle essentiel pour la résolution des équations différentielles (linéaires). Comme à l'opération de dérivation dans l'espace direct se trouve associée la multiplication par $(-ik)$ dans l'espace réciproque, une équation différentielle se transforme par Fourier en une équation *algébrique*, ce qui simplifie la résolution⁶⁷ des variables. Indépendamment de cette commodité technique, le passage en Fourier se prête bien à une discussion générale (causalité, stabilité des solutions, etc.). On sait que les égalités (8.192) sont aussi vraies pour les distributions (voir (2.109)).

Au passage, on note que⁶⁸ :

$$(-ik)^m F(k) = \int_{-\infty}^{+\infty} e^{ikx} f^{(m)}(x) dx \iff |F(k)| \leq \frac{1}{|k^m|} \int_{-\infty}^{+\infty} |f^{(m)}(x)| dx \quad (8.193)$$

Ainsi, plus une fonction est dérivable et à dérivées intégrables, plus sa transformée de Fourier décroît vite à l'infini⁶⁹ – et inversement (avec les bonnes hypothèses sur $F(k)$) :

$$|f(x)| \leq \frac{1}{2\pi|x^m|} \int_{-\infty}^{+\infty} |F^{(m)}(k)| dk \quad (8.194)$$

⁶⁵On a $\int_0^x f'(x') dx' = f(x) - f(0)$; si f' est absolument intégrable, on en déduit que $f(x)$ a une limite en $x = \pm\infty$. Comme $f(x)$ est aussi absolument intégrable, ces deux limites doivent être nulles.

⁶⁶Attention au signe !

⁶⁷De même, une équation aux dérivées partielles avec N variables indépendantes que l'on Fourier-transforme sur $N - 1$ variables devient une simple équation différentielle sur la variable laissée en l'état.

⁶⁸En particulier, si la dérivée seconde $f''(x)$ est absolument intégrable, $F(k)$ l'est aussi puisqu'alors $|F(k)| \leq \frac{C}{k^2}$.

⁶⁹Le même phénomène a été rencontré à propos des séries de Fourier.

De telles propriétés seront brièvement rediscutées dans la section 8.3, où on verra aussi que si $f(x)$ décroît à l'infini plus vite que toute puissance x^{-n} , n entier positif (fonction dite à *décroissance rapide*), alors $F(k)$ est infiniment dérivable et :

$$F^{(m)}(k) = \int_{-\infty}^{+\infty} (ix)^m e^{ikx} f(x) dx . \quad (8.195)$$

• Exemples

1. À titre de premier exemple, soit à résoudre l'équation :

$$f'(x) - k_0 f(x) = 0 \quad (8.196)$$

avec $f(0)$ donné, et k_0 réel, histoire de fixer les idées. Il s'agit de l'équation différentielle linéaire homogène du premier ordre la plus simple que l'on puisse imaginer. La solution est (trivialement !) $f(x) = f(0)e^{k_0 x}$, une fonction qui n'a visiblement pas de transformée de Fourier au sens usuel, quel que soit k_0 réel (et même complexe).

Prenons le problème autrement. Comme l'équation est linéaire, la transformation de Fourier est *a priori* utile : on va voir que cette approche, indéniablement iconoclaste au premier abord, fournit néanmoins la solution. Avec $F(k) \stackrel{\text{déf}}{=} \mathcal{F}[f]$, la transformation de Fourier appliquée à (8.196) donne :

$$\mathcal{F}[f'(x) - k_0 f(x)] = 0 \iff -ikF(k) - k_0 F(k) = 0 \iff (k - ik_0)F(k) = 0 , \quad (8.197)$$

compte tenu de la règle (8.190). On voit bien que l'équation différentielle a été transformée par \mathcal{F} en une équation algébrique.

L'équation (8.197) signifie que $F(k)$ est nulle tant que $k \neq ik_0$, ce qui incite à poser $F(k) \propto \delta(k - ik_0)$, à nouveau sans s'arrêter au fait que la "fonction" de Dirac n'a été définie que pour les valeurs réelles de sa "variable". Cela étant fait, et selon (8.142), il vient $f(x) \propto e^{-i(ik_0)x} = e^{k_0 x} \dots$ qui est bien la solution de (8.196) ! L'aspect symbolique de la "dérivation" doit sauter aux yeux, sa justification rapide reprenant les mêmes arguments que ceux développés pour l'exemple p. 381.

2. Un autre exemple, tout aussi trivial. Soit à résoudre :

$$f''(x) - k_0^2 f(x) = 0 \quad (8.198)$$

avec $f(0)$ et $f'(0)$ donnés. La solution est $f(x) = f(0) \cosh k_0 x + \frac{f'(0)}{k_0} \sinh k_0 x$. D'un autre côté, la transformée de Fourier de (8.198) est :

$$(-ik)^2 F(k) - k_0^2 F(k) = 0 \iff (k^2 + k_0^2) F(k) = 0 . \quad (8.199)$$

Le même argument suggère de dire que $F(k)$ est une combinaison linéaire de $\delta(k \pm ik_0)$, ce qui par \mathcal{F}^{-1} donne $f(x)$ sous la forme d'une combinaison linéaire de $e^{\pm k_0 x}$; à nouveau, la démarche, *apparemment* très désinvolte, donne de fait le bon résultat (quelque soit k_0 d'ailleurs, réel ou complexe).

3. Considérons maintenant une équation *inhomogène*, qui montrera à nouveau l'efficacité du calcul symbolique qu'autorise un peu d'audace. Soit l'équation différentielle linéaire d'ordre N :

$$\sum_{m=0}^N a_m f^{(m)}(x) = \phi(x) \quad (8.200)$$

où les a_m sont des constantes, où $f^{(m)}$ est la dérivée m^e de f et où $\phi(x)$ fait office de source. Prenant la transformée de Fourier de cette équation, on trouve ;

$$\sum_{m=0}^N a_m (-ik)^m F(k) = \Phi(k) \iff P(k)F(k) = \Phi(k) \quad (8.201)$$

où F et Φ sont les transformées de f et ϕ , et où $P(k)$ est un polynôme de degré N en k . La grande vertu de \mathcal{F} est de transformer une équation différentielle en équation algébrique, qui se résout (en principe) à vue. Naïvement, (8.201) donne immédiatement la seule et unique solution $f_{\text{part}}(x)$:

$$F(k) = \frac{\Phi(k)}{P(k)}, \quad f_{\text{part}}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk \frac{\Phi(k)}{P(k)} e^{-ikx}. \quad (8.202)$$

Visiblement, il y a quelque chose qui ne va pas : la résolution complète de l'équation (8.200) doit introduire N constantes d'intégration : où sont-elles ? En fait, la méthode naïve a exclu tous les cas où le facteur de F dans (8.201) est égal à zéro. L'audace consiste à nouveau à considérer tous les $\delta(k - k_p)$ où les k_p sont les zéros du polynôme $P(k)$, au nombre de N selon le théorème fondamental de l'algèbre (mais pas forcément tous réels...). La solution la plus générale de (8.201) n'est pas (8.202) mais la "fonction" :

$$\frac{\Phi(k)}{P(k)} + \sum_{p=1}^N C_p \delta(k - k_p), \quad (8.203)$$

où les C_p sont des constantes arbitraires à ce stade ; la combinaison linéaire apparaissant dans (8.203) exprime la structure d'espace vectoriel de l'ensemble des solutions de l'équation homogène. Par inversion, on en déduit maintenant :

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk \frac{\Phi(k)}{P(k)} e^{-ikx} + \sum_{p=1}^N \frac{C_p}{2\pi} e^{-ik_p x}. \quad (8.204)$$

Les constantes C_p jouent le rôle des constantes d'intégration et seront trouvées par inversion du système linéaire $N \times N$ obtenu en écrivant explicitement les N conditions "initiales", par exemple en résolvant le système linéaire :

$$\sum_{p=1}^N (-ik_p)^m C_p = 2\pi f^{(m)}(0) - \int_{-\infty}^{+\infty} dk (-ik)^m \frac{\Phi(k)}{P(k)}, \quad (8.205)$$

avec $m = 0, 1, 2, \dots, N - 1$. Tous les seconds membres sont en principe connus quand le problème est bien posé ; ce système permet donc bien de trouver toutes les constantes C_p .

On voit ainsi que la solution naïve ne donne qu'une solution particulière de l'équation, notée $f_{\text{part}}(x)$, cependant que la combinaison linéaire des $\delta(k - k_p)$ donne la solution la plus générale de l'équation homogène. Au total, l'expression (8.204) correspond bien à la règle connue donnant la construction de la solution la plus générale d'une équation différentielle telle que (8.200) :

$$\text{solution générale} = \text{solution particulière} + \text{solution générale de l'équation homogène} . \quad (8.206)$$

Le jeu ci-dessus est clairement formel, notamment parce que les N racines k_p du polynôme $P(k)$ sont *a priori* des nombres complexes. •

Compte tenu de la symétrie des formules donnant \mathcal{F} et \mathcal{F}^{-1} , on peut s'attendre à ce que les différentes propriétés ci-dessus aient leur équivalent dans l'autre sens⁷⁰. Par exemple, si on translate la transformée $F(k)$ d'une fonction $f(x)$, $\mathcal{T}_{k_0}F(k) = F(k - k_0)$, la transformée inverse est (comparer avec (8.166)) :

$$\mathcal{F}^{-1}[\mathcal{T}_{k_0}F] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ikx} F(k - k_0) dk = e^{-ik_0x} \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ik'x} F(k') dk' \equiv e^{-ik_0x} \mathcal{F}^{-1}[F] . \quad (8.207)$$

De même :

$$\mathcal{F}^{-1}[F^{(m)}] = (ix)^m \mathcal{F}^{-1}[F] , \quad (8.208)$$

et ainsi de suite.

■ **Convolution** On définit la convolution de deux fonctions, notée $f * g$, par l'opération suivante :

$$(f * g)(x) \stackrel{\text{déf}}{=} \int_{-\infty}^{+\infty} f(x')g(x - x') dx' . \quad (8.209)$$

La transformée de Fourier de la convolution $f * g$ est :

$$\mathcal{F}[f * g] = \int_{-\infty}^{+\infty} dx e^{ikx} \int_{-\infty}^{+\infty} dx' f(x')g(x - x') . \quad (8.210)$$

Supposant que f et g sont absolument intégrables sur \mathbb{R} , ainsi que le produit $f(x')g(x - x')$ sur \mathbb{R}^2 , le théorème de Fubini s'applique ; insérant dans l'intégrand $e^{-ikx'} e^{ikx'}$:

$$\mathcal{F}[f * g] = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dx' e^{ikx'} f(x') e^{ik(x-x')} g(x - x') . \quad (8.211)$$

⁷⁰Cette affirmation doit toutefois être nuancée, ou en tout cas être utilisée à bon escient : on sait que la transformation de Fourier ne laisse pas toujours invariante certaines propriétés, par exemple celle d'être absolument intégrable. Affirmer que ce qui est vrai dans un sens l'est aussi dans l'autre repose pourtant sur une telle invariance.

Posant $x - x' = x''$ dans la deuxième intégrale, il vient :

$$\mathcal{F}[f * g] = \int_{-\infty}^{+\infty} dx' e^{ikx'} f(x') \int_{-\infty}^{+\infty} dx'' e^{ikx''} g(x'') \equiv F(k)G(k) . \quad (8.212)$$

D'où le résultat important, exprimant que la transformée de la convolution est simplement le *produit* des transformées de Fourier :

$$\boxed{\mathcal{F}[f * g] = \mathcal{F}[f] \mathcal{F}[g]} \quad (8.213)$$

On peut aussi montrer que la transformée de Fourier du produit de deux fonctions est la convolution de leurs transformées de Fourier :

$$\mathcal{F}[fg] = \frac{1}{2\pi} \mathcal{F}[f] * \mathcal{F}[g] \iff \mathcal{F}^{-1}[\mathcal{F}[f] * \mathcal{F}[g]](t) = 2\pi f(t)g(t) . \quad (8.214)$$

La démonstration est beaucoup plus délicate, et on se contentera d'une "preuve" un peu sauvage. La transformée du produit est :

$$\mathcal{F}[fg] = \int_{-\infty}^{+\infty} dx e^{ikx} f(x)g(x) . \quad (8.215)$$

Maintenant représentons chaque fonction par la formule de Fourier inverse ; il vient :

$$\mathcal{F}[fg](k) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} dx e^{ikx} \int_{-\infty}^{+\infty} dk' e^{-ik'x} F(k') \int_{-\infty}^{+\infty} dk'' e^{-ik''x} G(k'') . \quad (8.216)$$

En rassemblant les exponentielles, il apparaît $e^{i(k-k'-k'')x}$, dont l'intégration sur x donne $2\pi\delta(k-k'-k'')$:

$$\mathcal{F}[fg](k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk' F(k') \int_{-\infty}^{+\infty} dk'' G(k'') \delta(k-k'-k'') \quad (8.217)$$

et la règle d'usage de la fonction de Dirac donne alors :

$$\mathcal{F}[fg](k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(k') G(k-k') \equiv \frac{1}{2\pi} (F * G)(k) dk' \equiv \frac{1}{2\pi} (\mathcal{F}[f] * \mathcal{F}[g])(k) . \quad (8.218)$$

• Exemples

1. Le théorème de convolution est très utile pour résoudre des équations (linéaires) où figure un noyau intégral ω , du genre :

$$f(x) = \phi(x) + \int_{-\infty}^{+\infty} dx' \omega(x-x')f(x') . \quad (8.219)$$

Une telle équation dit (par son terme intégral) que la valeur de f en un point x est conditionnée par les valeurs en des points différents, l'importance des

points situés dans le voisinage de x dépendant des valeurs (grandes ou petites) de $\omega(x - x')$. Si x est le temps, on parlera de système ayant de la mémoire ; si x est l'espace, il peut s'agir d'un système où les interactions ne sont pas *locales*⁷¹.

L'utilité du théorème de convolution saute aux yeux dans cet exemple. En effet, avec $\mathcal{F}[f] = F$, $\mathcal{F}[\phi] = \Phi$ et $\mathcal{F}[\omega] = \Omega$, la transformation de Fourier appliquée à (8.219) donne⁷² :

$$F(k) = \Phi(k) + \Omega(k)F(k) \iff F(k) = \frac{\Phi(k)}{1 - \Omega(k)} . \quad (8.220)$$

de sorte qu'une solution de (8.219) apparaît immédiatement sous la forme intégrale :

$$f_{\text{part}}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk \frac{\Phi(k)}{1 - \Omega(k)} e^{-ikx} . \quad (8.221)$$

Cette intégrale, en pratique, se calcule par résidus, en prenant en compte les singularités de $\Phi(k)$ et les zéros de $1 - \Omega(k)$. L'expression (8.221) n'est pas la solution générale de (8.219) : si on ajoute à $f_{\text{part}}(x)$ n'importe quelle solution de l'équation homogène $f(x) = \int_{-\infty}^{+\infty} dx' \omega(x - x')f(x')$, la somme ainsi formée est encore solution de (8.219) ; cette dernière équation est une équation aux valeurs propres généralisée, l'action du noyau $\omega(x - x')$ ("matrice continue") se bornant à multiplier la fonction par le scalaire 1 :

$$\int_{-\infty}^{+\infty} dx' \omega(x - x')f(x') = 1 \times f(x) . \quad (8.222)$$

D'ailleurs, la méthode semble inepte dans le cas où $\phi(x) \equiv 0$, auquel cas l'équation (8.219) est homogène en f . En fait, le savoir-faire permet encore de s'en tirer ; en effet, quand $\phi \equiv 0$, (8.220) devient :

$$F(k) = \Omega(k)F(k) \iff [1 - \Omega(k)]F(k) = 0 , \quad (8.223)$$

et l'on peut écrire $F(k) = \sum_p C_p \delta(k - k_p)$ où les C_p sont des constantes quelconques et où les k_p sont les zéros de $1 - \Omega(k)$; d'où, selon (8.142) :

$$f(x) = \sum_p \frac{C_p}{2\pi} e^{-ik_p x} . \quad (8.224)$$

⁷¹Le plus souvent, la non-localité (au sens large : temporelle, spatiale,...) d'une interaction résulte d'une sommation (trace) partielle sur des degrés de liberté dont la détermination détaillée n'est pas requise, par exemple "ignorer" la vitesse quand on ne s'intéresse qu'à la position.

C'est aussi le cas quand on s'intéresse exclusivement à la dynamique d'un (petit) système en interaction avec un autre : un système quantique couplé à un bain, ou à un objet macroscopique ("l'appareil de mesure" ?), ou un système canonique quand le réservoir est pris en compte uniquement par la température T qu'il impose au système d'intérêt.

⁷²Le passage d'une expression à l'autre dans (8.220) impose visiblement $1 - \Omega(k) \neq 0$.

À titre de vérification, calculons avec cette solution le second membre de (8.219) avec $\phi \equiv 0$:

$$\int_{-\infty}^{+\infty} dx' \omega(x-x') \sum_p \frac{C_p}{2\pi} e^{-ik_p x'} = \sum_p \frac{C_p}{2\pi} \int_{-\infty}^{+\infty} dx' \omega(x-x') \times e^{-ik_p x} e^{ik_p(x-x')} = \sum_p \frac{C_p}{2\pi} e^{-ik_p x} \Omega(k_p) . \quad (8.225)$$

Comme $\Omega(k_p) = 1$ par définition des k_p , ce second membre est bien la fonction $f(x)$ donnée en (8.224). Au total, la solution générale de (8.219) est la somme de $f_{\text{part}}(x)$, (8.221), et de $f(x)$, (8.224).

2. Le théorème de convolution joue également un rôle important en Traitement du signal. D'une façon générale, soit un signal $e(t)$ injecté à l'entrée d'une "boîte noire", dont le fonctionnement est supposé linéaire sur toute la gamme de fréquence. Cette boîte est caractérisée par une *fonction d'appareil*, qui exprime sa façon de réagir à un signal d'entrée de fréquence ω donnée ; à la sortie, on récupère un certain signal $s(t)$.



Figure 8.7: Entrée et sortie d'une boîte noire.

En notant par des capitales les transformées de Fourier ($E = \mathcal{F}[e]$, etc.), on a :

$$S(\omega) = \int_{-\infty}^{+\infty} A(\omega - \omega') E(\omega') d\omega' . \quad (8.226)$$

Ici, $A(\omega)$ est une fonction caractéristique de l'appareil⁷³ ; la relation (8.226) exprime que l'on accumule en ω , tout ce que la boîte a la capacité de ramasser à $\omega - \omega'$, capacité représentée par la fonction $A(\omega - \omega')$. Pour une boîte parfaite, on a $A(\omega) = \delta(\omega)$, donc $S(\omega) = E(\omega)$, et $s(t) = e(t)$: le signal de sortie est alors l'exacte réplique du signal d'entrée. Plus généralement, par le théorème de convolution, (8.226) dit que :

$$s(t) = 2\pi a(t)e(t) , \quad (8.227)$$

où $a(t)$ est la transformée de Fourier inverse de $A(\omega)$ (pour la boîte parfaite, $a(t)$ est la constante $\frac{1}{2\pi}$).

⁷³La fonction de réponse d'un dispositif est la fonction $\chi(\omega)$ décrivant la réponse en fréquence de l'appareil : le signal de sortie est relié à l'entrée selon $S(\omega) = \chi(\omega)E(\omega)$. L'appareil est parfait s'il n'introduit aucune distorsion en fréquence, soit quand $\chi(\omega)$ est une constante (fonction parfaitement plate). Ceci est évidemment une vue de l'esprit : de façon plus réaliste $\hat{\chi}$ ressemble à une fonction-porte, par exemple $\hat{\chi} = \Delta\omega^{-1}[\Theta(\omega + \frac{\Delta\omega}{2}) - \Theta(\omega - \frac{\Delta\omega}{2})]$; le signal transmis sera donc d'autant meilleur que la largeur de bande $\Delta\omega$ est cette fois *grande* par rapport à τ^{-1} , où τ désigne toujours l'échelle de temps caractéristique de $e(t)$. La relation $S(\omega) = \hat{\chi}(\omega)E(\omega)$ signifie que $s(t) = \int_{-\infty}^{+\infty} \chi(t-t')e(t')dt'$, où $\chi(t) = \mathcal{F}^{-1}\hat{\chi}(\omega)$, avec $\chi(t < 0) = 0$ pour la causalité.

En règle générale, $A(\omega)$ est une fonction de largeur typique $\Delta\omega$, que l'on peut toujours supposer normalisée à l'unité et donc de hauteur $\sim \Delta\omega^{-1}$. Par exemple, si $A(\omega)$ est la fonction-porte⁷⁴ $\Delta\omega^{-1}[\Theta(\omega + \frac{\Delta\omega}{2}) - \Theta(\omega - \frac{\Delta\omega}{2})]$, alors $a(t) = \frac{1}{\pi t \Delta\omega} \sin \frac{t \Delta\omega}{2}$, et la relation explicite entre entrée et sortie est (voir (8.214)) :

$$s(t) = \frac{2}{t \Delta\omega} \sin \frac{t \Delta\omega}{2} e(t) . \quad (8.228)$$

Pour que le signal après la boîte soit le moins déformé possible, il faut évidemment que l'on ait $\tau \Delta\omega \ll 1$, où τ est une échelle caractéristique de variation de l'entrée $e(t)$. À la limite $\Delta\omega = 0+$ (boîte parfaite, $A(\omega) = \delta(\omega)$, $a(t) = \frac{1}{2\pi} \forall t$), on a strictement $s(t) = e(t)$.

3. Traitons enfin un autre exemple, qui sera l'occasion de démontrer un résultat jouant un rôle important en Physique (et aussi, comme on l'a vu au chapitre 1, en Théorie des probabilités) : la convolution de deux gaussiennes est une gaussienne. Soit :

$$G_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} . \quad (8.229)$$

$G_\sigma(x)$ est la loi normale pour une variable aléatoire centrée⁷⁵ (i.e. de moyenne nulle), et d'écart-type égal à σ .

Soit maintenant la convolution de deux telles gaussiennes ; par définition :

$$(G_{\sigma_1} * G_{\sigma_2})(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \frac{1}{\sqrt{2\pi}\sigma_2} \int_{-\infty}^{+\infty} e^{-\frac{(x-x')^2}{2\sigma_1^2}} e^{-\frac{x'^2}{2\sigma_2^2}} dx' . \quad (8.230)$$

Un calcul un peu laborieux permet de calculer explicitement l'intégrale du second membre et de montrer que⁷⁶ le résultat est :

$$(G_{\sigma_1} * G_{\sigma_2})(x) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} e^{-\frac{x^2}{2(\sigma_1^2 + \sigma_2^2)}} . \quad (8.232)$$

Ce résultat s'obtient aussi, et beaucoup plus rapidement, avec le théorème de convolution. En effet, soit $\mathcal{G}_\sigma = \mathcal{F}[G_\sigma]$, qui se calcule aisément :

$$\mathcal{G}_\sigma(k) = \int_{-\infty}^{+\infty} e^{ikx} G_\sigma(x) dx = e^{-\frac{1}{2}\sigma^2 k^2} , \quad (8.233)$$

soit :

$$\mathcal{F}[G_\sigma](k) = e^{-\frac{1}{2}\sigma^2 k^2} \iff \mathcal{F}^{-1}[e^{-\frac{1}{2}\sigma^2 k^2}](x) = G_\sigma(x) . \quad (8.234)$$

⁷⁴Voir note 43.

⁷⁵Tout ce qui va être établi se généralise immédiatement au cas d'une variable de moyenne non nulle, x_0 , en utilisant notamment le théorème de translation, et en raisonnant avec $X = x - x_0$, qui est une variable centrée.

⁷⁶Pour des gaussiennes centrées respectivement en x_1 et x_2 , la convolution est :

$$\frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} e^{-\frac{[x - (x_1 + x_2)/2]^2}{2(\sigma_1^2 + \sigma_2^2)}} . \quad (8.231)$$

On note que $\mathcal{G}_\sigma(k=0) = 1$, en conséquence du fait que $\int_{-\infty}^{+\infty} G_\sigma(x) dx = 1$. Le théorème de convolution permet maintenant d'écrire :

$$\mathcal{F}[G_{\sigma_1} * G_{\sigma_2}](k) = e^{-\frac{1}{2}\sigma_1^2 k^2} \times e^{-\frac{1}{2}\sigma_2^2 k^2} = e^{-\frac{1}{2}(\sigma_1^2 + \sigma_2^2)k^2} . \quad (8.235)$$

Prenant alors la \mathcal{F}^{-1} :

$$(G_{\sigma_1} * G_{\sigma_2})(x) = \mathcal{F}^{-1} \left[e^{-\frac{1}{2}(\sigma_1^2 + \sigma_2^2)k^2} \right] = G_{(\sigma_1^2 + \sigma_2^2)^{\frac{1}{2}}}(x) , \quad (8.236)$$

où la dernière égalité résulte de (8.234). Au total, on retrouve bien (8.232), venant du calcul direct de la convolution. Ainsi se trouve établi le résultat majeur : la convolution de deux gaussiennes est une gaussienne.

Il s'agit là d'un résultat très important, déjà obtenu dans le chapitre 1, exprimé en affirmant que la somme de deux variables aléatoires gaussiennes (indépendantes) est une variable gaussienne (la gaussienne est *stable* par l'addition⁷⁷). Par ailleurs, l'équation (8.230) peut maintenant se lire comme suit :

$$G_{(\sigma_1^2 + \sigma_2^2)^{\frac{1}{2}}}(x) = \int_{-\infty}^{+\infty} G_{\sigma_1}(x') G_{\sigma_2}(x - x') dx' . \quad (8.237)$$

Cette relation est un cas très particulier de l'équation de chaîne de Bachelier - Chapman - Kolmogorov, qui est l'équation centrale de la théorie des processus stochastiques markoviens, conduisant notamment, pour les processus à diffusion lente, à l'équation de la diffusion $\partial_t P = \partial_{x^2} P$ obtenue phénoménologiquement par application de la loi de Fick. •

8.3 Propriétés asymptotiques

La question est ici de toucher du doigt la relation entre les propriétés⁷⁸ à l'infini d'une fonction $f(x)$ et le comportement près de $k = 0$ de sa transformée $F(k) = \mathcal{F}[f](k)$. Il s'agit en fait de l'autre facette des résultats énoncés ci-dessus (voir par exemple (8.193), et (8.194)), concernant l'existence des dérivées $f^{(m)}$ et la décroissance de $F(k)$ à l'infini, et renvoie également aux propriétés duales de la fonction caractéristique $\phi(t)$ d'une "densité" de probabilité $p(x)$ (voir section 1.8).

Quand toutes les intégrales existent, de (8.195) on déduit :

$$F^{(m)}(0) = i^m \int_{-\infty}^{+\infty} x^m f(x) dx \quad (8.238)$$

⁷⁷Rappelons que cette propriété remarquable est le fondement du théorème central limite, d'importance cruciale rencontré au chapitre 1, et que la loi de Gauss n'est pas la seule loi stable par l'addition (voir sous-section 1.8.3).

⁷⁸Compte tenu de la symétrie intrinsèque de la transformation de Fourier, tout ce qui suit vaut aussi quand on échange simultanément f et F , et x et k – avec les réserves que l'on sait, liées au fait que $\mathbb{L}_1(\mathbb{R})$ n'est pas stable par \mathcal{F} .

Chacune de ces intégrales existe de fait si $f(x)$ décroît à l'infini plus vite que $x^{-(m+\alpha)}$, avec⁷⁹ avec $\alpha > 1$. En d'autres termes, la décroissance rapide de $f(x)$ à l'infini assure l'existence des dérivées en $k = 0$ de sa transformée de Fourier : en pareil cas, $F(k)$ est analytique en $k = 0$.

Comme déjà mentionné, il y a une dualité entre les propriétés des partenaires d'un couple de Fourier, dans l'hypothèse où les conditions d'existence sont satisfaites pour f et pour F (répétons que ceci n'est pas assuré dans le passage de f à F en général⁸⁰). Deux exemples regardés en détail permettent de bien se convaincre qu'une telle relation existe, et donnent d'ailleurs l'occasion de revenir sur des résultats importants déjà rencontrés dans le chapitre 1.

Soit la gaussienne (normalisée) $G(x)$ définie en (8.229), et dont la transformée est $\mathcal{G}_\sigma(k)$ donnée en (8.233). Considérant l'expression intégrale de $\mathcal{G}_\sigma(k)$, on peut développer l'exponentielle e^{ikx} en série entière et intégrer terme à terme :

$$\mathcal{G}_\sigma(k) = \sum_{n=0}^{+\infty} \frac{(ik)^n}{n!} M_n, \quad M_n = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} x^n e^{-\frac{x^2}{2\sigma^2}} dx. \quad (8.239)$$

On sait que les M_n sont appelés les *moments* de la fonction⁸¹, ici ce sont les moments de la gaussienne, qui sont tous définis compte tenu de la décroissance ultra-rapide de $G_\sigma(x)$ à l'infini. Par symétrie (parité de l'intégrand) seuls les moments pairs sont non-nuls⁸² ; il faut ainsi calculer :

$$M_{2p} = \frac{2}{\sqrt{2\pi}\sigma} \int_0^{+\infty} x^{2p} e^{-\frac{x^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}\sigma} \int_0^{+\infty} u^{p-\frac{1}{2}} e^{-\frac{u}{2\sigma^2}} du = \frac{1}{\sqrt{\pi}} (2\sigma^2)^p \Gamma(p + \frac{1}{2}), \quad (8.240)$$

d'où :

$$\mathcal{G}_\sigma(k) = \frac{1}{\sqrt{\pi}} \sum_{p=0}^{+\infty} \frac{(ik)^{2p}}{(2p)!} (2\sigma^2)^p \Gamma(p + \frac{1}{2}). \quad (8.241)$$

Sachant que⁸³ $\Gamma(p + \frac{1}{2}) = \frac{\sqrt{\pi}}{2^p} (2p-1)!!$, on trouve finalement que :

$$\mathcal{G}_\sigma(k) = \sum_{p=0}^{+\infty} \frac{(-1)^p}{p!} \left(\frac{k^2 \sigma^2}{2} \right)^p, \quad (8.242)$$

en accord avec (8.233), puisque ceci n'est autre que le développement en série de $e^{-\frac{1}{2}k^2\sigma^2}$. Il ne s'agissait pas de retrouver autrement ce résultat, mais de montrer que si tous les

⁷⁹Il ne suffit pas de dire : " $f(x)$ décroît à l'infini plus vite que $x^{-(m+1)}$ " ; penser à la fonction $\frac{1}{x \ln x}$ (dont la primitive est $\ln(\ln x)$) qui n'est visiblement pas intégrable. Cette façon de parler est toutefois d'usage courant.

⁸⁰Notamment, la propriété de sommabilité en module n'est pas forcément préservée par la transformation de Fourier : la fonction-porte est absolument sommable, mais pas sa transformée de Fourier, le sinus cardinal.

⁸¹Si x est considérée comme une variable aléatoire, le premier moment M_1 est la moyenne de x , l'écart quadratique de x étant la combinaison $M_2 - M_1^2$.

⁸²Bien sûr, ceci n'est vrai que pour une variable *centrée*, dont la valeur moyenne est nulle.

⁸³Cette relation se démontre à partir de $\Gamma(z+1) = z\Gamma(z)$ et sachant que $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

moments M_n de f existent, alors $F(k)$ est une fonction analytique admettant en tant que telle un développement en série entière de k , comme exhibé en (8.242). L'existence de tous les moments de f exige une décroissance rapide à l'infini de $f(x)$, et le développement en série entière donne aisément le comportement de $F(k)$ près de $k = 0$ ($k\sigma \ll 1$).

Comme cas extrême opposé, considérons la lorentzienne⁸⁴ (aussi appelée fonction de Cauchy) :

$$C_a(x) = \frac{1}{\pi} \frac{a}{x^2 + a^2} . \quad (8.243)$$

Tous les moments impairs de cette fonction paire sont nuls par symétrie, du moins si on les définit comme $\lim_{A \rightarrow +\infty} \frac{1}{\pi} \int_{-A}^{+A} dx \frac{ax^{2p+1}}{x^2+a^2}$. En revanche, tous les moments pairs M_{2p} , $p \geq 1$ n'existent pas⁸⁵, car $C_a(x)$ décroît trop lentement à l'infini. La transformée de Fourier se calcule facilement par résidus et vaut⁸⁶ :

$$\mathcal{C}_a(k) = e^{-|k|a} . \quad (8.244)$$

En raison de la présence du module $|k|$, $\mathcal{C}_a(k)$ n'est pas une fonction analytique (holomorphe) ; en particulier, il n'existe pas de développement de Taylor près de l'origine (la dérivée première est discontinue en $k = 0$!). C'est la décroissance lente de $C_a(x)$ à l'infini (cause première de l'absence des moments) qui induit une singularité à l'origine pour la transformée de Fourier. Les mêmes observations valent pour le couple de fonctions (8.159). Il apparaît ainsi que l'origine et le point à l'infini sont en quelque sorte duaux l'un de l'autre par la transformation de Fourier⁸⁷.

Bien sûr, on peut échanger les rôles. Par exemple, soit la fonction :

$$\gamma_\xi(x) = \frac{1}{2\xi} e^{-\frac{|x|}{\xi}} , \quad \mathcal{F}[\gamma_\xi] \equiv \Gamma(k) = \frac{1}{1 + \xi^2 k^2} . \quad (8.245)$$

Maintenant, c'est le tour de $f(x)$ de ne pas avoir un développement de Taylor à l'origine, lié au fait que $F(k)$ décroît trop lentement à l'infini.

Ces remarques ont en fait une portée générale, et sont à rapprocher de la majoration (8.194), qui exprime bien qu'une fonction qui décroît vite a une transformée de Fourier ayant beaucoup de dérivées – une façon commune (et un peu vague) de traduire les phénomènes observés sur les deux cas particuliers ci-dessus.

8.4 Généralisation en dimension quelconque

Jusqu'à présent, on a considéré la transformation de Fourier unidimensionnelle. La généralisation à \mathbb{R}^3 est immédiate. Soit une fonction $\mathbb{R}^3 \rightarrow \mathcal{D}$, domaine de \mathbb{C} ; avec

⁸⁴Dans les deux exemples considérés, les fonctions sont définies avec les bons facteurs conduisant à une intégrale $\int_{-\infty}^{+\infty} f(x) dx$ égale à 1, avec $f = G_\sigma$ et $f = C_a$. On observera que les transformées de Fourier ont corrélativement pour valeur 1 en $k = 0$.

⁸⁵Si on les définit plus précisément comme $\lim_{A \rightarrow +\infty} \frac{1}{\pi} \int_{-A}^{+A} dx \frac{ax^{2p}}{x^2+a^2}$ ($p \geq 1$), on peut dire qu'ils sont tous infinis.

⁸⁶À nouveau, $\mathcal{C}_a(k = 0) = 1$ en conséquence du fait que l'intégrale de $C_a(x)$ est égale à 1.

⁸⁷La même dualité sera retrouvée pour la transformation de Laplace (chapitre 9).

les conventions adoptées depuis le début, sa transformée de Fourier est par définition :

$$F(\vec{k}) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^3} f(\vec{r}) e^{i\vec{k}\cdot\vec{r}} d^3r . \quad (8.246)$$

L'établissement de la formule d'inversion procède essentiellement de la même façon que pour \mathbb{R} , et on trouve :

$$f(\vec{r}) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} F(\vec{k}) e^{-i\vec{k}\cdot\vec{r}} d^3k . \quad (8.247)$$

De fait, les formules et/ou propriétés établies précédemment se transposent sans difficulté. En particulier :

$$\delta(\vec{r}) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} e^{-i\vec{k}\cdot\vec{r}} d^3k = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} e^{i\vec{k}\cdot\vec{r}} d^3k = \prod_{u=x,y,z} \frac{1}{2\pi} \int_{\mathbb{R}} e^{ik_u u} dk_u , \quad (8.248)$$

où $\delta(\vec{r})$ est définie comme⁸⁸ :

$$\delta(\vec{r}) = \delta(x)\delta(y)\delta(z) . \quad (8.250)$$

On démontre aussi sans peine que :

$$\mathcal{F}[\vec{\nabla} f(\vec{r})] = -i\vec{k} \mathcal{F}[f(\vec{r})] , \quad (8.251)$$

relation que l'on retrouve immédiatement en dérivant mentalement (8.247) membre à membre.

Tout ceci se généralise à une fonction de \mathbb{R}^D dans \mathbb{C} :

$$F(\vec{k}) \stackrel{\text{déf}}{=} \int_{\mathbb{R}^D} f(\vec{r}) e^{i\vec{k}\cdot\vec{r}} d^D r , \quad f(\vec{r}) = \frac{1}{(2\pi)^D} \int_{\mathbb{R}^D} F(\vec{k}) e^{-i\vec{k}\cdot\vec{r}} d^D k . \quad (8.252)$$

En particulier pour la translation, avec $\mathcal{T}_{\vec{a}} f(\vec{r}) \stackrel{\text{déf}}{=} f(\vec{r} - \vec{a})$, on trouve :

$$\boxed{\mathcal{F}[\mathcal{T}_{\vec{a}} f](\vec{k}) = e^{i\vec{k}\cdot\vec{a}} \mathcal{F}[f](\vec{k})} \quad (8.253)$$

et pour la dilatation dans \mathbb{R}^D , $f_{\lambda}(\vec{r}) = f(\lambda\vec{r})$:

$$\boxed{\mathcal{F}[f_{\lambda}](\vec{k}) = \frac{1}{|\lambda|^D} \mathcal{F}[f]\left(\frac{\vec{k}}{\lambda}\right)} \quad (\lambda \in \mathbb{R}) \quad (8.254)$$

⁸⁸Pour une fonction $f(\vec{r})$ bien définie en $\vec{r} = 0$ (ce qui exige que la valeur $f(\vec{0})$ est indépendante du chemin suivi pour arriver à l'origine), on vérifie que la définition de $\delta(\vec{r})$: $f(\vec{0}) = \int_{\mathbb{R}^3} f(\vec{r}) \delta(\vec{r}) d^3r$ donne :

$$\delta(\vec{r}) = \frac{1}{4\pi r^2} \delta(r) . \quad (8.249)$$

De même, dans \mathbb{R}^2 on a $\delta(\vec{r}) = \frac{1}{2\pi r} \delta(r)$ et, plus généralement dans \mathbb{R}^D , $\delta(\vec{r}) = \frac{1}{S_D} \delta(r)$, où S_D est la surface de la sphère de rayon r dans \mathbb{R}^D (voir chapitre 7).

L'homogénéité permet de débusquer les omissions : la définition (8.250) dit clairement que $\delta(\vec{r})$ est homogène à L^{-3} ; si l'inadvertance faisait écrire quelque chose comme $\delta(\vec{r}) = \frac{1}{4\pi} \delta(r)$, on devrait réaliser immédiatement qu'il y a une erreur.

Si la fonction f possède certaines symétries – c'est-à-dire ne dépend pas explicitement en fait de toutes les coordonnées –, il est possible d'effectuer d'emblée certaines intégrations et de mettre sa transformée F sous la forme d'une intégrale impliquant alors moins de D variables. Par exemple, dans \mathbb{R}^3 ($D = 3$), si f est à symétrie sphérique, c'est-à-dire ne dépend que du *module* r du vecteur \vec{r} , le choix des coordonnées sphériques s'impose. On a alors ($d^3r = r^2 \sin\theta dr d\theta d\phi$) et, par définition (voir (8.246)) :

$$F(k) = \int_0^{+\infty} r^2 dr f(r) \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi e^{i\vec{k}\cdot\vec{r}} , \quad (8.255)$$

en profitant du fait que par hypothèse $f(r)$ ne dépend pas des angles. Par ailleurs, c'est bien k (et non \vec{k}) qui figure comme argument de F puisque la symétrie sphérique de f se transporte visiblement⁸⁹ sur F . Maintenant, rien n'interdit de choisir l'axe Oz le long de \vec{k} (qui est fixé⁹⁰) ; dès lors, (8.255) se réécrit comme :

$$F(k) = \int_0^{+\infty} r^2 dr f(r) \int_0^\pi \sin\theta d\theta \int_0^{2\pi} d\phi e^{ikr \cos\theta} . \quad (8.256)$$

L'intégrale sur ϕ se fait immédiatement et donne un simple facteur 2π ; celle sur θ est aussi immédiate et donne $\frac{2\sin kr}{kr}$. Au total, il ne reste que l'intégrale en r et on a :

$$F(k) = \frac{4\pi}{k} \int_0^{+\infty} r dr f(r) \sin kr . \quad (8.257)$$

Comme prévu, F ne dépend que du module de \vec{k} , ce qui est bien normal, puisque la quantité $e^{i\vec{k}\cdot\vec{r}}$ est un produit scalaire, donc invariant dans toute rotation simultanée de \vec{r} et \vec{k} (voir note 89).

De la même façon, dans \mathbb{R}^2 cette fois, on trouve :

$$F(k) = \int_0^{+\infty} r dr f(r) \int_0^{2\pi} e^{ikr \cos\theta} d\theta \equiv 2\pi \int_0^{+\infty} dr f(r) J_0(kr) \quad (8.258)$$

où J_0 est la fonction de Bessel déjà rencontrée ($J_n(x) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-in\theta} e^{ix \sin\theta} d\theta$, $n \in \mathbb{Z}$).

On rencontrera en Mécanique Quantique la relation précise entre la fonction d'onde en représentation- q , $\Psi(\vec{r}, t)$, et son image en représentation- p , $\Phi(\vec{p}, t)$; cette relation est un exemple de transformée de Fourier multidimensionnelle, qui s'écrit habituellement avec l'impulsion \vec{p} (et non pas le vecteur d'onde $\vec{k} = \hbar^{-1}\vec{p}$) ; dans \mathbb{R}^D , et avec les facteurs traditionnels dans un tel contexte :

$$\Phi(\vec{p}, t) = \frac{1}{(2\pi\hbar)^{\frac{D}{2}}} \int_{\mathbb{R}^D} e^{\frac{i}{\hbar}\vec{p}\cdot\vec{r}} \Psi(\vec{r}, t) d^D r . \quad (8.259)$$

Pour une onde plane⁹¹ $\Psi(\vec{r}, t) = L^{-\frac{D}{2}} e^{i\vec{k}_0\cdot\vec{r}}$, on trouve $\Phi(\vec{p}, t) = \left(\frac{2\pi\hbar}{L}\right)^{\frac{D}{2}} \delta(\vec{p} - \hbar\vec{k}_0)$, comme il se doit. On vérifie sans peine que Ψ et Φ ont respectivement pour dimension

⁸⁹La quantité $e^{i\vec{k}\cdot\vec{r}}$ est un produit scalaire, donc invariant dans toute rotation simultanée de \vec{r} et \vec{k} . Quand f est à symétrie sphérique, l'intégrand ne change pas par rotation, et il en est de même de F .

⁹⁰L'intégrale est un nombre qui se moque du système de coordonnées choisi pour effectuer son calcul.

⁹¹où L est la longueur du côté de la boîte cubique de confinement.

(longueur) $^{-\frac{D}{2}}$ et (impulsion) $^{-\frac{D}{2}}$, en conformité avec le fait que l'intégrale du module carré de chacune de ces fonctions dans un domaine $\mathcal{D} \subset \mathbb{R}^D$ est un nombre, plus précisément une probabilité.

Notons aussi que le caractère transverse ou longitudinal d'un champ de vecteur, $\vec{A}(\vec{r})$, se traduit très simplement à l'aide de la transformée de Fourier. En effet, soit :

$$\vec{A}(\vec{k}) = \int_{\mathbb{R}^3} d^3r \vec{A}(\vec{r}) e^{i\vec{k} \cdot \vec{r}} . \quad (8.260)$$

Pour un champ transverse, on a par définition⁹² $\text{div} \vec{A} = 0$, ce qui se traduit par :

$$\boxed{\vec{k} \cdot \vec{A}(\vec{k}) = 0 \quad (\text{champ transverse})} \quad (8.261)$$

exprimant l'orthogonalité entre \vec{A} et \vec{k} . Pour un champ longitudinal, $\text{rot} \vec{A} = 0$, c'est-à-dire que l'on a :

$$\boxed{\vec{k} \times \vec{A}(\vec{k}) = 0 \quad (\text{champ longitudinal})} \quad (8.262)$$

une relation qui dit que \vec{A} et \vec{k} sont colinéaires.

8.5 Causalité et analyticit 

La Physique est construite sur quelques principes, qui sont des affirmations  nonc es sans autre souci de d monstration que la mise   l' preuve exp rimentale. Au fil de l'histoire, certains d'entre eux sont pass s du statut de principe   celui de r sultat dont la preuve d coule de principes r put s plus fondamentaux, comme le Second Principe de la Thermodynamique : dans la formulation statistique, celui-ci est essentiellement une *cons quence* du postulat de base de la M canique statistique ( quiprobabilit  des  tats de m me  nergie pour un syst me isol ), alli    la limite thermodynamique qui assure la divergence des temps de retour⁹³.

Ici, il s'agit du Principe de causalit , qui est finalement une affirmation de bon sens : les effets ne peuvent  tre ant rieurs aux causes, – une particule au repos ne peut  tre mise en mouvement qu'*apr s* l'application d'une force, le courant dans un circuit ne peut appara tre qu'apr s avoir ferm  l'interrupteur.

⁹²En  lectromagn tisme, \vec{A}  tant le potentiel- vecteur, une telle jauge s'appelle *jauge de Coulomb*. Noter aussi que le langage de Fourier permet d'exprimer tr s simplement les relations fondamentales ; par exemple $\vec{B}(\vec{r}, t) = \vec{\nabla} \times \vec{A}$ devient $\vec{B}(\vec{k}, t) = \vec{k} \times \vec{A}(\vec{k}, t)$.

⁹³Cette affirmation, exp ditive et certes un peu trop elliptique, donne juste l'id e permettant de comprendre le fondement physique du Second Principe. En toute rigueur, certaines situations extr mes interdisent de pr tendre que l'on dispose d'une authentique d monstration de port e universelle (c'est tout le probl me de l'*ergodicit *).   l'inverse, on conna t des mod les simples permettant de bien montrer comment c'est la prise de la Limite thermodynamique qui produit la "fl che du temps".

D'un point de vue plus formel, soit $x(t)$ une grandeur dynamique d'un certain syst me soumis   une perturbation externe not e $f(t)$. Tr s souvent, l'effet de la perturbation peut se traduire par la relation suivante entre la cause, $f(t)$, et l'effet, $x(t)$:

$$x(t) = \int_{-\infty}^{+\infty} \chi(t-t')f(t') dt' . \quad (8.263)$$

L'apparition de la seule diff rence dans χ n'a rien d' vident (on attend plut t $\chi(t, t')$). La justification de ce point sera donn e plus loin (voir (8.276)), et en invoquant le th or me de convolution.

La fonction $\chi(t-t')$ s'appelle *fonction de r ponse*, ou *susceptibilit *⁹⁴, suivant le contexte. Fondamentalement, cette relation est *lin aire*, et on peut penser qu'elle appara t spontan ment si la cause est petite, en un sens   pr ciser soigneusement. Ceci n'est pas toujours vrai : d'une part une telle relation peut exister ind pendamment de l'intensit  de la perturbation (c'est le cas pour un oscillateur harmonique – voir plus loin). D'autre part, il peut arriver que m me si la cause est infiniment petite, la fonction χ n'existe pas : c'est le cas pour un syst me critique dont la susceptibilit  est infiniment grande.

Admettons que la relation (8.263) est vraie. Le Principe de causalit  exige alors que :

$$\chi(t-t') = 0 \quad \text{si } t' > t \iff \chi(t) = 0 \quad \forall t < 0 . \quad (8.264)$$

Il s'agit de montrer comment cette affirmation se traduit par des propri t s de la fonction χ , plus pr cis ment par des propri t s d'analyticit  de sa transform e de Fourier $\hat{\chi}(z)$, une fois prolong e imm diatement aux valeurs complexes de la pulsation. La relation inverse de Fourier est :

$$\chi(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{\chi}(\omega)e^{-i\omega t} d\omega . \quad (8.265)$$

L'id e est de calculer cette int grale en admettant que $\hat{\chi}(z)$ satisfait les conditions du 3^e lemme de Jordan. Dans ces conditions, on peut refermer le contour par un grand demi-cercle   l'infini. Quand $t < 0$, il faut refermer par en-haut ; selon le th or me des r sidus, il vient :

$$\chi(t) = 2i\pi \sum_{k, \Im z_k > 0} \text{Res}[\hat{\chi}(z)e^{-izt}, z_k] . \quad (8.266)$$

Supposons pour simplifier que $\hat{\chi}(z)$ est m romorphe, et que tous les p les z_k sont simples ; alors :

$$\chi(t) = 2i\pi \sum_{k, \Im z_k > 0} \Gamma_k e^{-iz_k t} , \quad \Gamma_k = \lim_{z \rightarrow z_k} (z - z_k)\hat{\chi}(z) . \quad (8.267)$$

Les fonctions $e^{-iz_k t}$  tant lin airement ind pendantes, on ne peut avoir $\chi(t) = 0 \quad \forall t < 0$ que si $\Gamma_k = 0 \quad \forall (z_k, \Im z_k > 0)$; avec l'hypoth se de p les simples, ceci signifie que la

⁹⁴ $\chi(t-t')$ mesure bien la fa on dont le syst me r agit   la sollicitation repr sent e par $f(t)$: quand la fonction χ est petite, le syst me est relativement indiff rent   la perturbation ; au contraire, quand χ est grand, le syst me "part au quart de tour" (le mot *susceptibilit * est employ  dans le m me sens que lorsque l'on dit d'une personne qu'elle est *susceptible*). Pour un syst me critique, la susceptibilit  est infinie, signifiant que le syst me "bascule"   la plus infime sollicitation. Par ailleurs, on verra par la suite (chapitre 10) que la susceptibilit  est une certaine fonction de Green, dite fonction de Green *avanc e*.

partie principale de $\hat{\chi}(z)$ autour de tout z_0 , $\Im z_0 > 0$ est en fait identiquement nulle⁹⁵. La non-violation du Principe de causalité est donc bien équivalente à l'absence de singularité de $\hat{\chi}(z)$ dans tout le demi-plan supérieur⁹⁶.

Il est utile d'illustrer ces résultats formels en traitant un exemple précis (et universel en Physique), en l'occurrence un oscillateur harmonique amorti, quitte à généraliser mentalement les résultats de portée universelle obtenus dans ce cadre restrictif.

Pour un tel oscillateur, l'équation dynamique est :

$$m\ddot{x} = \Lambda(x, \dot{x}) - kx + F(t) . \quad (8.269)$$

x est l'abscisse mesurée à partir de la position d'équilibre, m est la masse de la particule, $-kx$ la force de rappel (on posera $k = m\omega_0^2$) ; $F(t)$ est une force extérieure pour l'instant arbitraire et c'est la réponse de l'oscillateur à cette force que l'on veut étudier. Λ est la force de frottement qui introduit l'amortissement, toujours inévitable pour les systèmes réels ; pour faire simple, on choisit un frottement *fluide* où, par définition, le freinage est proportionnel à la vitesse :

$$\Lambda = -\alpha\dot{x} \quad (\alpha > 0) . \quad (8.270)$$

Avec ces choix, l'équation (8.269) devient :

$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = \frac{1}{m}F(t) \quad (\gamma = \frac{\alpha}{m} > 0) ; \quad (8.271)$$

γ a la dimension de l'inverse d'un temps.

La transformée de Fourier de $x(t)$ est :

$$X(\omega) = \int_{-\infty}^{+\infty} x(t) e^{i\omega t} dt \quad (8.272)$$

⁹⁵La généralisation au cas de pôles d'ordre supérieur à 1 ne présente pas de difficultés. Si $\hat{\chi}(z)$ a un pôle d'ordre n_0 en z_0 , la partie principale de l'intégrand se trouve à partir de :

$$\left[\frac{c_{-n_0}}{(z-z_0)^{n_0}} + \dots \right] e^{-izt} = e^{-iz_0 t} \left[\frac{c_{-n_0}}{(z-z_0)^{n_0}} + \dots \right] e^{-i(z-z_0)t} = e^{-iz_0 t} \left[\dots + \frac{(-it)^{n_0-1}}{(n_0-1)!} \frac{c_{-n_0}}{z-z_0} + \dots \right] , \quad (8.268)$$

d'où le résidu $\frac{(-it)^{n_0-1}}{(n_0-1)!} c_{-n_0} e^{-iz_0 t}$. Un tel terme ne doit apparaître puisqu'il faut $\chi(t < 0) \equiv 0$: la seule possibilité est $c_{-n_0} = 0$. Une autre façon d'énoncer le résultat est de dire que $\hat{\chi}(z)e^{-izt}$ n'a que des résidus nuls dans le demi-plan supérieur (se souvenir toutefois qu'une fonction peut avoir des pôles, mais des résidus nuls – penser à $z \rightarrow f(z) \stackrel{\text{déf}}{=} \frac{1}{z^2}$).

En fin de compte, c'est la multiplication par le facteur exponentiel e^{-izt} (qui contient toutes les puissances positives) qui assure la nullité de tous les résidus, donc l'inexistence d'une partie principale et par conséquent l'absence de toute singularité de $\hat{\chi}(z)$ dans le demi-plan supérieur.

⁹⁶Bien sûr, ce résultat dépend de la convention choisie pour la définition de la transformation de Fourier (8.106).

Il existe un autre argument montrant que $\hat{\chi}(z)$ ne peut avoir de résidus différents de zéro en tout z_k , $\Im z_k > 0$. Si tel était le cas, l'inversion de Fourier produirait des exponentielles divergentes $\propto e^{+(\Im z_k)t}$; ces solutions ne peuvent apparaître dès lors que l'on a supposé l'existence de $\mathcal{F}[x]$; il est heureux que cette raison soit compatible avec la non-violation du Principe de causalité...

et la relation inverse est :

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\omega) e^{-i\omega t} d\omega . \quad (8.273)$$

On a aussi :

$$\dot{x} \rightarrow \mathcal{F}[\dot{x}] = -i\omega \mathcal{F}[x] = -i\omega X(\omega) , \quad \ddot{x} \rightarrow \mathcal{F}[\ddot{x}] = (-i\omega)^2 \mathcal{F}[x] = -\omega^2 X(\omega) , \quad (8.274)$$

toutes  galit s qui supposent implicitement que $x(t)$ (et sa d riv e) tendent vers z ro quand $t \rightarrow \pm\infty$.

Prenons maintenant la transform e de Fourier de (8.271). En utilisant la lin arit  de cette  quation, et la lin arit  de la transformation int grale de Fourier, on trouve :

$$(-i\omega)^2 X(\omega) - i\gamma\omega X(\omega) + \omega_0^2 X(\omega) = \frac{1}{m} \Phi(\omega) , \quad (8.275)$$

o  $\Phi(\omega) = \mathcal{F}[F(t)]$. D s lors (8.275) s'inverse   vue :

$$X(\omega) = \frac{1}{m} \frac{1}{-\omega^2 - i\gamma\omega + \omega_0^2} \Phi(\omega) \equiv \hat{\chi}(\omega) \Phi(\omega) , \quad (8.276)$$

o  ceci d finit de fait la transform e de Fourier de la susceptibilit  $\hat{\chi}(\omega)$:

$$\hat{\chi}(\omega) = \frac{1}{m} \frac{1}{-\omega^2 - i\gamma\omega + \omega_0^2} \equiv |\hat{\chi}(\omega)| e^{i\phi(\omega)} . \quad (8.277)$$

En effet, par la relation inverse de Fourier (8.273), on a :

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega e^{-i\omega t} \hat{\chi}(\omega) \Phi(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega \int_{-\infty}^{+\infty} dt' e^{-i\omega t} \hat{\chi}(\omega) e^{+i\omega t'} F(t') . \quad (8.278)$$

En  changeant l'ordre des int grations, on trouve finalement :

$$x(t) = \int_{-\infty}^{+\infty} dt' \chi(t-t') F(t') , \quad \chi(t-t') = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega e^{-i\omega(t-t')} \hat{\chi}(\omega) . \quad (8.279)$$

de sorte que, selon la relation inverse (8.273), $\hat{\chi}(\omega)$ appar it bien comme $\mathcal{F}[\chi]$. On se souvient que cette fonction $x(t)$ n'est qu'une solution *particul re* de l' quation (8.271) (et d'ailleurs, elle ne contient aucune condition initiale !).

La question importante est maintenant de montrer le lien entre analyticit  et causalit , plus pr cis ment d' tablir que la fonction $\chi(t)$:

$$\chi(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega e^{-i\omega t} \hat{\chi}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega \frac{1}{m} \frac{e^{-i\omega t}}{-\omega^2 - i\gamma\omega + \omega_0^2} \quad (8.280)$$

satisfait   la r quisition du principe de causalit  $\chi(t) = 0 \forall t < 0$. On va voir que c'est bien le cas parce que $\hat{\chi}(\omega)$ est analytique dans le demi-plan sup rieur⁹⁷. En effet, les

⁹⁷Attention : r p tons que ceci est li  explicitement aux d finitions de Fourier adopt es ci-dessus. Si par exemple on prenait une autre convention en  crivant $X(\omega) = \int_{-\infty}^{+\infty} x(t) e^{-i\omega t} dt$ - c'est une autre fa on de faire tout aussi l gitime -, alors $\hat{\chi}(\omega)$ serait analytique dans le demi-plan inf rieur.

zéros du dénominateur dans l'intégrand de (8.280) sont des pôles pour $\hat{\chi}(\omega)$; ils sont donnés par :

$$z_{\pm} = -i\frac{\gamma}{2} \pm \sqrt{\omega_0^2 - \frac{\gamma^2}{4}} . \quad (8.281)$$

Il s'agit de pôles simples tant que $\gamma \neq 2\omega_0$.

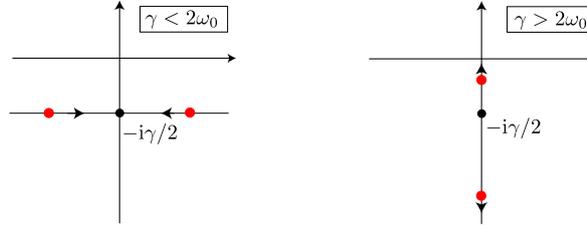


Figure 8.8: Pôles de $\hat{\chi}(z)$: à gauche, cas sous-amorti, à droite cas sur-amorti. Les flèches indiquent le mouvement des pôles quand le rapport $\frac{\omega_0}{\gamma}$ décroît à γ fixé. Dans tous les cas, les deux pôles sont dans le demi-plan inférieur, en traduction du Principe de causalité.

Pour le frottement faible (régime sous-amorti, γ petit, plus précisément $\gamma < 2\omega_0$), la racine carrée est réelle ; dans le cas contraire (frottement fort, régime suramorti), la racine carrée est imaginaire pure. On voit cependant que dans tous les cas la partie imaginaire de chaque pôle est négative (heureusement !) :

$$\Im z_{\pm} < 0 \quad \forall \gamma, \omega_0 ; \quad (8.282)$$

toutes les singularités (ici des pôles) de la susceptibilité $\hat{\chi}(\omega)$ sont dans le demi-plan inférieur : $\hat{\chi}(\omega)$ est bien analytique dans tout le demi-plan supérieur. Maintenant, quand $t < 0$, et en invoquant le lemme de Jordan, on peut calculer l'intégrale de (8.280) en refermant par un demi-cercle, situé dans le demi-plan supérieur. Par le théorème des résidus, et puisqu'il n'y a pas de pôles dans le contour ainsi formé, l'intégrale est nulle, et il vient bien $\chi(t) = 0 \quad \forall t < 0$, en conformité avec le principe de causalité.

Quand au contraire t est positif, un calcul simple de résidus (en refermant cette fois par le demi-cercle inférieur) donne, $\forall t > 0, \gamma, \omega_0$:

$$\chi(t) = \frac{1}{m} \frac{e^{-\gamma t/2}}{\sqrt{\omega_0^2 - \frac{\gamma^2}{4}}} \sin\left(\sqrt{\omega_0^2 - \frac{\gamma^2}{4}} t\right) = \frac{1}{m} \frac{e^{-\gamma t/2}}{\sqrt{\frac{\gamma^2}{4} - \omega_0^2}} \sinh\left(\sqrt{\frac{\gamma^2}{4} - \omega_0^2} t\right) . \quad (8.283)$$

Ces deux expressions sont vraies quel que soit le signe de $\omega_0 - \frac{\gamma}{2}$, mais on peut préférer l'une ou l'autre suivant que le frottement est faible ou fort : dans le premier cas, la première expression exhibe les oscillations dont l'amplitude décroît exponentiellement, dans le second, le retour à la position d'équilibre s'effectue d'une seule traite⁹⁸, mais

⁹⁸La seconde expression n'est que le prolongement analytique immédiat de la première quand la variable $\omega_0^2 - \frac{\gamma^2}{4}$ devient négative.

lentement (fluide tr s visqueux). Le cas-fronti re $\gamma = 2\omega_0$ s'obtient par passage   la limite dans les expressions ci-dessus :

$$\gamma = 2\omega_0 : \chi(t) = \frac{1}{m} t e^{-\gamma t/2} . \quad (8.284)$$

En d finitive, l' cart   l' quilibre $x(t)$, apr s extinction des transitoires – voir la Remarque ci-dessous – et pour une perturbation quelconque $F(t)$, est donn  (par exemple) par⁹⁹ :

$$x(t) = \int_{-\infty}^t dt' \frac{1}{m} \frac{e^{-\gamma(t-t')/2}}{\sqrt{\frac{\gamma^2}{4} - \omega_0^2}} \sinh\left(\sqrt{\frac{\gamma^2}{4} - \omega_0^2} (t-t')\right) F(t') . \quad (8.285)$$

Les calculs ont  t  faits explicitement pour un oscillateur amorti, mais le r sultat majeur reste vrai : pour tout syst me physique, la transform e de Fourier de toute susceptibilit , d finie avec la convention (8.272), est n cessairement une fonction analytique (holomorphe) dans le demi-plan complexe sup rieur. Dans le cas g n ral, la fonction $\hat{\chi}(\omega)$ a une forme plus compliqu e mais toutes ses singularit s sont dans le demi-plan inf rieur puisque l'on doit toujours avoir :

$$\chi(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega \hat{\chi}(\omega) e^{-i\omega t} \equiv 0 \quad \forall t < 0 . \quad (8.286)$$

Par ailleurs, quand $\gamma \ll \omega_0$, le module de $\hat{\chi}(\omega)$ a un pic tr s prononc  tout pr s de ω_0 (voir fig. 8.9). Ce r sultat a une autre port e g n rale : les *r sonances* de la susceptibilit  r v lent les fr quences propres du syst me analys . Le plus souvent, $\hat{\chi}(\omega)$ est une fraction contenant au d nominateur la fonction de transfert $Z(\omega)$; quand Z n'a que des z ros simples, $\hat{\chi}(\omega)$ peut donc se mettre sous la forme¹⁰⁰ :

$$\hat{\chi}(\omega) = \sum_{\alpha=1}^N \frac{C_\alpha}{\omega - \omega_\alpha + i\gamma_\alpha} , \quad (8.287)$$

o  N est l'ordre de l' quation diff rentielle, et o  $z_\alpha = \omega_\alpha - i\gamma_\alpha$ est l'un quelconque des p les de $\hat{\chi}(\omega) = \frac{1}{Z(\omega)}$. Le coefficient C_α n'est autre que le r sidu de $\hat{\chi}(z)$ au p le z_α .

◆ *Remarque*

  la r flexion, l'expression g n rale (8.285) est certainement en un sens incompl te : en fait, elle rel ve d'un tour de passe-passe. En effet, l' quation dynamique (8.271) est du second ordre en temps (comme toujours pour un syst me m canique), et sa

⁹⁹L'int grale s'arr te   t puisque $\chi(t-t')$ est nulle pour $t' > t$.

¹⁰⁰Il s'agit   nouveau d'un d veloppement de Mittag - L effler qui, dans le cas d'une  quation d'ordre N , n'est rien d'autre que la d composition d'une fraction rationnelle en  l ments simples.

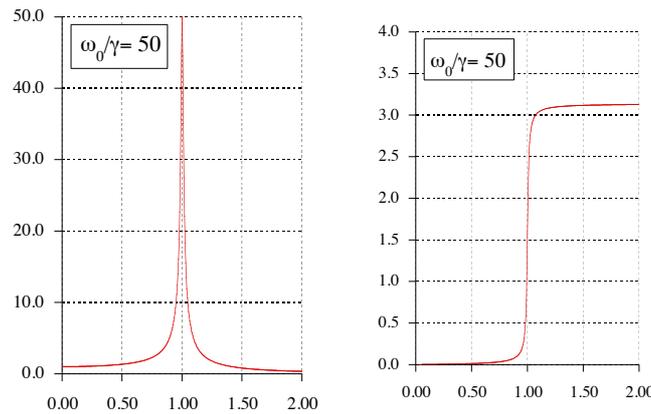


Figure 8.9: Suivant les définitions (8.277) : à gauche, module de $m\omega_0^2\hat{\chi}(\omega)$, à droite, argument $\phi(\omega)$ de $\hat{\chi}(\omega)$, dans le cas très sous-amorti ($\gamma \ll \omega_0$). L'abscisse est le rapport $\frac{\omega}{\omega_0}$.

solution la plus générale *doit* contenir des constantes d'intégration arbitraires tant que les conditions initiales, par exemple position et vitesse au départ, n'ont pas été explicitement prises en compte. Or rien de tel n'apparaît dans (8.285), qui ne peut donc représenter la solution la plus générale de (8.271).

En réalité, (8.285) représente ce que l'on appelle le *régime forcé*. Sans rentrer dans le détail, les choses se présentent en fait comme suit. À un certain instant, t_0 , où on se donne les conditions initiales x_0 et \dot{x}_0 , on branche la perturbation $F(t)$. Il en résulte que la solution générale arrive sous la forme :

$$x(t) = e^{-\gamma(t-t_0)}C(x_0, \dot{x}_0) + \int_{t_0}^t dt' \frac{1}{m} \frac{1}{\sqrt{\frac{\gamma^2}{4} - \omega_0^2}} \sinh\left(\sqrt{\frac{\gamma^2}{4} - \omega_0^2}(t-t')\right) F(t') . \quad (8.288)$$

Le premier terme – propre à l'équation homogène et qui survit sur un temps de l'ordre de γ^{-1} (c'est typiquement un régime transitoire) – exprime *l'oubli*¹⁰¹ *des conditions initiales* : suffisamment longtemps après t_0 – et à condition que le système soit amorti –, toute trace de la position et de la vitesse initiales est en pratique effacée, et seul subsiste le régime forcé obtenu auparavant. Formellement, ceci est strictement vrai quand on prend la limite $t_0 \rightarrow -\infty$ dans (8.288).

En définitive, le régime forcé, solution obtenue un peu par inadvertance, est en réalité le régime pertinent physiquement, une fois éteints les transitoires inévitables consécutifs au branchement de la perturbation, mais à la condition essentielle que

¹⁰¹Cette amnésie n'est pas présente pour les systèmes dits chaotiques qui exhibent au contraire le phénomène de *sensibilité par rapport aux conditions initiales* : deux trajectoires initialement voisines s'écartent exponentiellement l'une de l'autre au cours du temps. Ces comportements extraordinaires ne peuvent se produire que pour les solutions d'équations non-linéaires (condition nécessaire, nullement suffisante)

le système soit amorti. Sa pertinence physique est bien sûr tributaire de l'existence d'une échelle de temps τ dûment caractérisée et à la nature de la loi de relaxation, ici exponentielle. \blacklozenge

8.6 Relations de Kramers - Kronig

À la suite de l'exemple traité ci-dessus en détail, cela vaut la peine de revenir un instant sur les relations de Kramers - Kronig, déjà introduites au chapitre 6, compte tenu de leur rôle important en Physique par le lien explicite qu'elles mettent en évidence entre les parties réelle et imaginaire de la transformée de Fourier d'une susceptibilité. Avec les définitions de la section 8.5, on vient de voir que la fonction $\hat{\chi}(\omega)$ est analytique dans le demi-plan supérieur. Pour cette raison¹⁰², l'intégrale le long d'un contour sur l'axe réel contournant la singularité en ω_0 en l'excluant, et refermé par le haut, est égale à zéro (pas de pôle, pas de résidu – voir fig. 8.10) :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{\hat{\chi}(\omega)}{\omega - \omega_0} d\omega - i\pi\hat{\chi}(\omega_0) = 0 . \quad (8.289)$$

À des notations près, cette relation a déjà été obtenue au chapitre 6. En introduisant les parties réelle et imaginaire de $\hat{\chi}(\omega) = \hat{\chi}_1(\omega) + i\hat{\chi}_2(\omega)$, et en écrivant (8.289) sous la forme de deux équations réelles en séparant parties réelle et imaginaire, il vient¹⁰³ :

$$\mathcal{P} \int_{-\infty}^{+\infty} \frac{\hat{\chi}_1(\omega)}{\omega - \omega_0} d\omega = -\pi\hat{\chi}_2(\omega_0) , \quad \mathcal{P} \int_{-\infty}^{+\infty} \frac{\hat{\chi}_2(\omega)}{\omega - \omega_0} d\omega = \pi\hat{\chi}_1(\omega_0) . \quad (8.291)$$

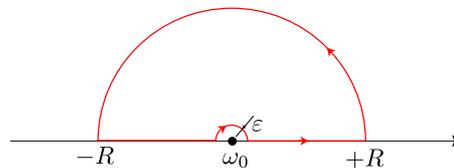


Figure 8.10: Contour d'intégration utilisé pour établir les relations de Kramers - Kronig. ($R \rightarrow +\infty$, $\varepsilon \rightarrow 0$)

Ces relations sont importantes en pratique pour la raison suivante. On peut montrer que l'absorption d'énergie d'un système excité par un champ monochromatique

¹⁰²On suppose en outre que $\lim_{|\omega| \rightarrow \infty} |\hat{\chi}(\omega)| = 0$, afin de pouvoir s'appuyer sur le premier lemme de Jordan.

¹⁰³La transformation \mathcal{H} :

$$f(x) \xrightarrow{\mathcal{H}} f_{\mathcal{H}}(x) \stackrel{\text{déf}}{=} \frac{1}{\pi} \int_{\mathcal{D}} \frac{f(x')}{x' - x} dx' \quad (8.290)$$

porte le nom de transformation de Hilbert. C'est pourquoi on dit que les parties réelle et imaginaire d'une fonction de réponse sont (essentiellement) transformées de Hilbert l'une de l'autre.

de pulsation ω_0 est proportionnelle à la partie imaginaire d'une certaine fonction de réponse¹⁰⁴ $\hat{\chi}_E$ à la pulsation du champ, soit $\hat{\chi}_{E,2}$. Il est en général facile de mesurer une telle absorption, donc de déterminer expérimentalement la fonction $\hat{\chi}_{E,2}(\omega)$. Cela étant fait, la deuxième relation (8.291) permet d'en déduire la partie réelle par transformation intégrale. Bien sûr, la qualité de la fonction $\hat{\chi}_{E,1}$ ainsi trouvée dépend crucialement de la finesse de la grille expérimentale (bonne résolution en pulsation) et de l'amplitude de l'intervalle de mesure (large bande).

Dans l'exemple traité dans la section 8.5, la susceptibilité est donnée par (8.276) et on a :

$$\hat{\chi}(\omega) = \frac{1}{m} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2} + i \frac{1}{m} \frac{\gamma \omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2} \equiv \hat{\chi}_1(\omega) + i \hat{\chi}_2(\omega) . \quad (8.292)$$

La partie imaginaire a l'allure d'une courbe en cloche, bien piquée en ω_0 si $\gamma \ll \omega_0$ (résonance marquée si frottement faible, avec une largeur $\Delta\omega \sim \gamma \ll \omega_0$). Au contraire, la partie réelle change de signe en s'annulant en $\omega = \omega_0$. Il est facile de vérifier directement (par un calcul d'intégrales réelles) que les relations de Kramers - Kronig sont satisfaites par la fonction $\hat{\chi}(\omega)$.

¹⁰⁴Ceci sera démontré au chapitre 10, à propos des fonctions de Green.

Chapitre 9

Transformation de Laplace

“Ma méthode pour surmonter une difficulté,
c’est de tourner autour.”
(George POLYÀ 1887–1985)

*Le but de ce chapitre est d’introduire la transformation de Laplace
et d’en présenter les applications les plus usuelles.*

9.1 Présentation

La transformation de Laplace est une transformation intégrale ayant un grand lien de parenté avec la transformation de Fourier, mais qui s’en démarque nettement sur plusieurs points. Notamment, on verra que la classe des fonctions admettant une transformation de Laplace est beaucoup plus vaste que celles des fonctions pour lesquelles l’intégrale de Fourier est définie.

En outre, la transformation de Laplace est particulièrement adaptée pour l’étude de la dynamique de systèmes supposés dans un état connu à un certain instant, que l’on peut toujours prendre comme origine des temps. Cette situation est très banale en Physique, qu’il s’agisse de résoudre un problème de dynamique classique, d’évolution d’un système quantique, ou de relaxation d’un système macroscopique à partir d’un état hors d’équilibre. Dans toutes ces situations, si on note f la quantité physique d’intérêt, la question est donc de résoudre l’équation décrivant l’évolution de la grandeur f et d’en obtenir l’expression, $f(t)$, à un instant t postérieur à celui où on a “déclenché le chronomètre”, après avoir préparé le système dans un état dûment caractérisé. Comme on l’a vu, au contraire, la transformation de Fourier est très commode pour l’étude des régimes forcés qui, pour un système amorti, prévalent après extinction des régimes transitoires (oubli des conditions initiales), auquel cas le régime physique s’obtient en rejetant formellement l’instant initial infiniment loin dans le passé.

Les équations d'évolution exigent toujours la donnée additionnelle de condition(s) initiale(s). Le nombre de ces conditions dépend de la nature de l'équation dynamique. Si celle-ci est une équation différentielle d'ordre p , et si le nombre de degrés de liberté est égal à N , il faut Np conditions initiales. Si la quantité cherchée obéit à une équation aux dérivées partielles du premier ordre en temps, l'état initial est complètement défini par la donnée du champ à $t = 0$: c'est le cas de l'équation de Schrödinger¹, qui ne peut être complètement résolue pour un système quantique donné (à température nulle) que si la fonction d'onde de départ a été précisée.

Dans les problèmes d'évolution temporelle, un état initial étant donné (à un instant pris conventionnellement comme origine des temps, $t = 0$), l'objectif est d'obtenir la fonction $f(t)$ à $t > 0$. Dans ces conditions, la fonction f à $t < 0$ n'est pas définie, ou plus précisément, on se moque de savoir ce qu'elle vaut. C'est pourquoi il est usuel (et conventionnel) de la prendre *nulle*. D'ailleurs, ce choix est souvent ce que dicte la situation physique d'intérêt ; soit par exemple un oscillateur harmonique à l'équilibre (position d'équilibre, vitesse nulle). S'il reçoit instantanément un choc à un certain instant ($t = 0$), son abscisse (et sa vitesse) vont par la suite osciller dans le temps : nulles à $t < 0$, ces fonctions ne le sont plus après de l'instant du choc.

Dans toute la suite, la transformée de Laplace sera définie exclusivement pour de telles fonctions, nulles pour $t < 0$, non-nulles pour $t > 0$. Remarquons tout de suite que ce n'est pas toujours la variable temps t qui intervient : si on étudie le mouvement d'une particule confinée dans le demi-espace \mathbb{R}_+ , toute grandeur relative à cette particule sera une fonction de l'abscisse x , $\phi(x)$, identiquement nulle $\forall x < 0$. Toutefois, la variable-type utilisée ci-dessous sera notée t , afin de respecter les usages.

La transformation de Laplace est souvent associée aux calculs de circuits électriques, car c'est dans le souci de développer un calcul symbolique (on dit aussi *opérationnel*) pour ces systèmes qu'elle a été historiquement introduite. Il faut bien voir que ses propriétés, et sa puissance en tant qu'outil, ne la rende en aucune façon tributaire des circonstances de son émergence historique. La transformation de Laplace est l'une des transformations intégrales importantes de l'Analyse – tout comme les transformations de Fourier, Hilbert, Mellin,... –, et présente un grand intérêt en soi, en-dehors du fait que, dans presque tous les domaines des sciences, elle est très souvent un outil d'une extrême utilité – quand il n'est pas tout simplement irremplaçable. Tout naturellement, la trame de ce chapitre ressemblera à celle du chapitre 8.

¹La situation est la même pour l'équation classique de la diffusion, $\partial_t P = D\partial_{xx}P$ ($D =$ constante de diffusion), ou l'équation de la chaleur $\partial_t T = \kappa\partial_{xx}T$ ($\kappa =$ coefficient de diffusivité thermique). Toutes ces EDP sont de type parabolique.

9.2 Définition et formule d'inversion

Tout comme la transformation de Fourier, la transformation de Laplace est une transformation intégrale associant une fonction à une autre. On notera :

$$\boxed{F = \mathcal{L}[f] \iff f \xrightarrow{\mathcal{L}} F} \quad (9.1)$$

La correspondance précise est la suivante. Soit une certaine fonction, $f(t)$ continue par morceaux sur \mathbb{R}_+ , et ayant au plus un nombre fini de discontinuités dans tout intervalle fini de \mathbb{R} ; sa transformée de Laplace $F(z)$ est donnée par la relation intégrale suivante³ :

$$\boxed{F(z) \stackrel{\text{dét}}{=} \int_0^{+\infty} f(t)e^{-zt} dt} \quad (9.2)$$

$f(t)$ est appelé *original* et sa transformée $F(z)$ est appelée *image*. Sur cette définition, et par comparaison avec la transformation de Fourier, on voit bien que tout se passe comme si la fonction $f(t)$ était de fait nulle à tout temps négatif, avec la relation :

$$\boxed{\mathcal{L}[f(t)](z) = \mathcal{F}[Y(t)f(t)](\omega = iz)} \quad (9.3)$$

$Y(t)$ étant la fonction de Heaviside. z est la variable conjuguée de t , et cette notation affiche la couleur : dans (9.2), z est *a priori* un nombre *complexe*, de sorte que $F(z)$ est une fonction à valeurs complexes même si $f(t)$ est réelle.

La toute première question à régler est de savoir dans quelles circonstances l'intégrale (9.2) existe ; une hypothèse de base dans le présent contexte consistant à supposer que $\forall t > 0$, $f(t)$ est bornée⁴ en module par une certaine exponentielle ($|f(t)| \leq Me^{x_0 t}$). Montrons d'abord que l'intégrale $\int_0^T f(t)e^{-zt} dt$ converge (i.e. a une limite) quand T tend vers $+\infty$, ce qui est équivalent à montrer que $\lim_{T \rightarrow +\infty} \int_T^{+\infty} f(t)e^{-zt} dt = 0$; on a :

$$\left| \int_T^{+\infty} f(t)e^{-zt} dt \right| \leq \int_T^{+\infty} |f(t)| |e^{-(x+iy)t}| dt = \int_T^{+\infty} |f(t)| e^{-xt} dt, \quad (9.4)$$

où on a posé $z = x + iy$. Selon l'hypothèse sur la fonction $f(t)$, et supposant de plus $x > x_0$, on a maintenant :

$$\left| \int_T^{+\infty} f(t)e^{-zt} dt \right| \leq \int_T^{+\infty} e^{x_0 t} e^{-xt} dt = \frac{e^{-(x-x_0)T}}{x-x_0}, \quad (9.5)$$

²On peut en fait encore définir l'intégrale si $f(t)$ diverge en certains points t_n de \mathbb{R}_+ mais pas trop vite, de façon à être localement sommable ; par exemple, pour une divergence algébrique $\sim |t - t_n|^\alpha$, il faut que α (négatif puisqu'il est question de divergence) soit plus grand que -1 – voir par exemple $\mathcal{L}[t^\alpha]$, écrite en (9.18)

³Si t est un temps, z est l'inverse d'un temps, une pulsation par exemple.

⁴On peut énoncer une condition un peu moins restrictive : il faut et suffit qu'il existe $t_0 > 0$ tel que f soit sommable entre 0 et t_0 et que $\forall t > t_0$, $|f(t)| \leq M_0 e^{x_0 t}$.

montrant de toute évidence que $\lim_{T \rightarrow +\infty} \left| \int_T^{+\infty} f(t) e^{-zt} dt \right| = 0$. Noter que la convergence est *uniforme* par rapport à z dans le domaine $\Re z > x_0$: pour tout $z = x + iy$ donné, il suffit de remplacer partout x par x_1 tel que $x \geq x_1 > x_0$ pour avoir une majoration *indépendante* de z .

Une fois démontrée la convergence de l'intégrale, la même hypothèse sur la fonction f permet de trouver une borne supérieure de la limite :

$$|F(z)| \leq \int_0^{+\infty} |f(t)| e^{-xt} dt \leq M \int_0^{+\infty} e^{-(x-x_0)t} dt = \frac{M}{x-x_0} \quad (\forall z, \Re z \equiv x > x_0) . \quad (9.6)$$

Ainsi, pour toute fonction f ne croissant pas plus vite que $e^{x_0 t}$ la transformation intégrale de Laplace définit une fonction $F(z)$ bornée et bien définie dans le demi-plan $\Re z > x_0$. L'inégalité (9.6) montre aussi que :

$$\boxed{\forall F(z), \lim_{|z| \rightarrow +\infty} |F(z)| = 0 \quad (\Re z > x_0)} \quad (9.7)$$

Une fonction donnée $\Phi(z)$ qui n'a pas cette propriété ne saurait être une transformée de Laplace au sens ci-dessus⁵.

De plus, $F(z)$ est de la forme $\int_C \phi(z; t) dt$ où la dépendance en z est uniquement *via* l'exponentielle e^{-zt} , qui est une fonction holomorphe. Autrement dit, $\phi(z; t)$ est holomorphe en z pour tout t fixé ; comme l'intégrale converge uniformément en z , la fonction $F(z)$ est elle-même holomorphe⁶. Étant bornée $\forall z, \Re z > x_0$, elle est holomorphe dans tout le demi-plan à droite de la ligne verticale d'abscisse x_0 . La conclusion importante est donc, compte tenu de la forme spécifique de la transformation intégrale (9.2) :

$$\boxed{\exists x_0 \in \mathbb{R}, \exists M > 0, |f(t)| \leq M e^{x_0 t} \implies F(z) \text{ holomorphe } \forall z, \Re z > x_0} \quad (9.8)$$

Le caractère holomorphe peut aussi se démontrer comme suit. L'intégrale (9.2) est de la forme $\int_C \phi(z; t) dt$, converge uniformément en z , et la dérivée $\frac{\partial \phi}{\partial z} = -t e^{-zt} f(t)$ est bornée en module au même titre que $f(t)$; il en résulte que la dérivée $F'(z)$ s'obtient par dérivation sous le signe somme ([2], théorème 3.8 p. 56) :

$$F'(z) = \int_0^{+\infty} f(t) (-t) e^{-zt} dt ; \quad (9.9)$$

l'existence de $F'(z)$ établit le caractère holomorphe de $F(z)$. De plus, compte tenu des hypothèses sur $f(t)$, on a aussi :

$$|F'(z)| \leq \int_0^{+\infty} M t e^{-(x-x_0)t} dt = \frac{M}{(x-x_0)^2} \quad \forall z, \Re z > x_0 . \quad (9.10)$$

⁵On peut cependant donner un sens, en tant que transformées de Laplace en un sens généralisé, à des fonctions n'ayant pas cette propriété (voir p. 417). C'est ainsi que $z^n = \mathcal{L}[\delta^{(+)(n)}(t)]$, où $\delta^{(+)}(t)$ est la fonction définie en (9.39), $\delta^{(+)(n)}(t)$ désignant sa dérivée d'ordre n .

⁶Schématiquement : si l'intégrale $\int_C \phi(z; t) dt$ converge uniformément $\forall z \in \mathcal{D}$, et si les points singuliers de ϕ ne dépendent pas du paramètre z , alors $\int_\Gamma F(z) dz = 0$, où Γ est un contour fermé situé dans \mathcal{D} – une conséquence du fait que $\int_\Gamma \left[\int_C \phi(z; t) dt \right] dz = \int_C \left[\int_\Gamma \phi(z; t) dz \right] dt$ avec $\int_\Gamma \phi(z; t) dz = 0$ en vertu du théorème de Cauchy (voir [9], théorème 4, §16, p. 62, ou [7], § 5.32).

$F(z)$ a donc une dérivée bornée en module partout dans le demi-plan $\Re z > x_0$, donnée par (9.9). Le nombre x_0 est appelé *abscisse de sommabilité*.

L'abscisse de sommabilité x_0 n'est pas toujours positive, même si c'est souvent le cas. Par exemple, la fonction-porte $P(t)$:

$$P(t) \stackrel{\text{d\u00e9f}}{=} \begin{cases} 1 & \text{si } 0 < t < 1 \\ 0 & \text{autrement} \end{cases} \quad (9.11)$$

a pour transform\u00e9e de Laplace :

$$\mathcal{P}(z) = \int_0^1 e^{-zt} dt = \frac{1}{z}(1 - e^{-z}) \quad (9.12)$$

Cette fonction, ainsi d\u00e9finie par l'int\u00e9grale, est holomorphe dans \mathbb{C} (la singularit\u00e9 en $z = 0$ est clairement \u00e9liminable) ; ceci montre que pour toute fonction \u00e0 support born\u00e9⁷, l'abscisse de sommabilit\u00e9 est \u00e9gale \u00e0 $-\infty$. Tout polyn\u00f4me a une abscisse de sommabilit\u00e9 positive mais aussi petite que l'on veut puisque $e^{x_0 t}$, $x_0 > 0$ cro\u00eet \u00e0 l'infini plus vite que toute puissance de x (l'exponentielle appartient \u00e0 la classe des fonctions dites \u00e0 *d\u00e9croissance rapide*). Pour une fraction rationnelle $\frac{P(z)}{Q(z)}$ o\u00f9 P et Q sont des polyn\u00f4mes de degr\u00e9 quelconque, $x_0 > \max z_k$, o\u00f9 z_k est un z\u00e9ro r\u00e9el de $Q(z)$, ceci afin d'\u00e9carter le cas o\u00f9 une r\u00e9gularisation serait n\u00e9cessaire, par exemple par une partie principale de Cauchy. Enfin, si une fonction $f(t)$ est born\u00e9e sup\u00e9rieurement, on a $f(t) \leq e^{x_0 t}$, $\forall x_0 \geq 0$: on peut donc prendre $x_0 = 0$ si besoin est.

Il est bien \u00e9vident que, compar\u00e9e \u00e0 la transformation de Fourier des fonctions, la transformation de Laplace permet de consid\u00e9rer une classe beaucoup plus grande de fonctions : quand x_0 est strictement positif, la fonction $f(t)$ peut diverger comme $e^{x_f t}$ avec $x_f < x_0$, sa transform\u00e9e de Laplace existe quand m\u00eame. La contrainte essentielle est donc ici que l'original ne croisse pas plus vite qu'une exponentielle, alors que la transformation de Fourier, avec son exponentielle oscillante, n'a trivialement de sens que pour des fonctions absolument int\u00e9grables (rappelons toutefois qu'une fonction int\u00e9grable mais non-absolument int\u00e9grable peut n\u00e9anmoins avoir une transform\u00e9e de Fourier, par exemple⁸ $\frac{\sin x}{x}$ - voir p. 371). Inversement, il est bien clair qu'il existe des fonctions croissant trop vite \u00e0 l'infini pour avoir une transform\u00e9e de Laplace : c'est le cas de $e^{\alpha t^2}$, $\alpha > 0$, par exemple (voir Remarque p. 418) ; pour de telles fonctions, on peut dire que l'abscisse de sommabilit\u00e9 x_0 est \u00e9gale \u00e0 $+\infty$.

Quand la variable est une abscisse de position, x , la transform\u00e9e de Laplace spatiale s'\u00e9crit :

$$F(k) = \int_0^{+\infty} f(x) e^{-kx} dx ; \quad (9.13)$$

⁷Pour une fonction $f(t)$ nulle pour $t > t_0$, on a visiblement $f(t) = 0 < e^{x_0 t}$, $\forall t > t_0$, quel que soit x_0 r\u00e9el, y compris $-\infty$!

⁸Autre exemple : les deux fonctions $\cos x^2$ et $\sin x^2$ ont aussi une transform\u00e9e de Fourier, respectivement $\sqrt{\pi/2}(\cos k^2/4 + \sin k^2/4)$ et $\sqrt{\pi/2}(\cos k^2/4 - \sin k^2/4)$, obtenues en prenant les parties r\u00e9elle et imaginaire de l'int\u00e9grale gaussienne $\int_{-\infty}^{+\infty} e^{ikx} e^{ix^2} dx = \sqrt{\pi} e^{i(\pi - k^2)/4}$.

k a la dimension de l'inverse d'une longueur, c'est par exemple un vecteur d'onde. À nouveau, les grandeurs x et k sont souvent dites *conjuguées* l'une de l'autre. Ce formalisme est bien adapté pour la résolution d'équations linéaires (différentielles ou aux dérivées partielles) dans un demi-espace – par exemple, l'équation de la diffusion dans \mathbb{R}_+ avec un mur parfaitement réfléchissant en $x = 0$.

À titre d'exemple, considérons la fonction de Heaviside⁹ (échelon-unité) :

$$Y(t) \stackrel{\text{déf}}{=} \begin{cases} 0 & \forall t < 0 \\ 1 & \forall t > 0 \end{cases} \quad (9.14)$$

D'après la définition intégrale (9.2), sa transformée de Laplace est :

$$F_1(z) = \int_0^{+\infty} e^{-zt} dt = \frac{1}{z} \quad \forall z, \Re z > 0 . \quad (9.15)$$

La fonction $Y(t)$ est bornée $\forall t > 0$ par $e^{x_0 t} \forall x_0 \geq 0$. L'abscisse de sommabilité x_0 est donc nulle : $F_1(z)$ est holomorphe dans le demi-plan ouvert de droite, $\Re z > 0$; c'est bien ce que dit (9.15), la fonction $\frac{1}{z}$ l'étant dans ce demi-plan¹⁰.

Cet exemple simple permet d'illustrer une fois encore la notion de prolongement analytique, une étape indispensable pour l'utilité de la transformation de Laplace, notamment pour l'efficacité de la formule d'inversion. En effet, la fonction $F_2(z) \stackrel{\text{déf}}{=} \frac{1}{z}$ est définie dans le domaine $\mathcal{D}_2 = \mathbb{C} \setminus \{0\}$ (le plan privé de l'origine). D'un autre côté, la relation intégrale (9.15) définit une fonction $F_1(z) = \frac{1}{z}$ dans le domaine $\mathcal{D}_1 = \{z, \Re z > 0\}$. Les deux domaines se recouvrent (largement !), $\mathcal{D}_1 \subset \mathcal{D}_2$, et dans la région commune, on a bien $F_1(z) = F_2(z)$; par le théorème du prolongement analytique (p. 219), on peut donc dire que $F_2(z)$ est le prolongement analytique de $F_1(z)$ du côté $\Re z < 0$. En définitive, la transformée de Laplace de $Y(t)$ est la fonction $\mathcal{Y}(z) = \frac{1}{z}$ prolongée dans tout le plan complexe et analytique partout dans $\mathbb{C} \setminus \{0\}$:

$$\mathcal{Y}(z) \equiv \mathcal{L}[Y](z) = \frac{1}{z} \quad \forall z, z \neq 0 \quad (9.16)$$

Il convient de bien saisir la distinction entre (9.15) et (9.16).

Ainsi, et de façon générale, l'opération de prolongement étant faite, on n'est plus confiné au demi-plan $\Re z > x_0$ (ici $\Re z > 0$ puisque $x_0 = 0$) ; c'est ce qui autorisera des promenades dans \mathbb{C} affranchies de cette condition – nécessaire pour que la définition par l'intégrale (9.15) ait un sens –, tout juste contraintes, comme toujours, à contourner les singularités de $F(z)$, dûment et préalablement identifiées, lors des déformations continues des chemins d'intégration.

⁹On note qu'il n'est nullement nécessaire de préciser ce que vaut cette fonction en $x = 0$.

¹⁰Une fois prolongée, elle n'a en fait qu'une seule singularité, un pôle simple en $z = 0$.

Citons un autre exemple, également très important. Soit la fonction t^α ; sa transformée de Laplace est¹¹ :

$$\mathcal{L}[t^\alpha] = \int_0^{+\infty} t^\alpha e^{-zt} dt = \frac{\Gamma(\alpha + 1)}{z^{\alpha+1}} \quad (\Re\alpha > -1) \quad (9.18)$$

L'intégrale n'existe que si $\alpha > -1$ et $\Re z > 0$; la relation (9.18) est vraie partout dans le plan de droite, mais le prolongement en z se fait par continuité, le seul point singulier étant le point de branchement en $z = 0$, absent seulement si $\alpha \in \mathbb{N}$. Noter aussi que, avec $\Re\alpha > -1$, la transformée ainsi obtenue satisfait bien (9.7).

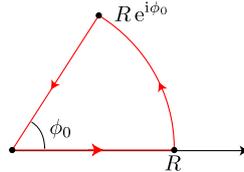


Figure 9.1: Contour en secteur circulaire illustrant la rotation de la demi-droite d'intégration pour effectuer le prolongement analytique sur l'égalité intégrale de définition de la transformée de Laplace.

Pour la transformation de Laplace, la possibilité d'effectuer le prolongement analytique peut se comprendre comme suit¹². Avec la définition intégrale (9.2) de $F(z)$, il est possible de jouer avec l'intervalle d'intégration, qui n'est autre qu'un certain contour dans le plan complexe. Montrons l'idée en considérant $Y(t)$; soit l'intégrale :

$$\int_C e^{-z\tau} d\tau \quad (9.19)$$

où le contour fermé C est formé de l'intervalle $[0, R] \subset \mathbb{R}$, de l'arc de cercle de rayon R , $0 \leq \phi \leq \phi_0 < \frac{\pi}{2}$ et du segment joignant le point $R e^{i\phi_0}$ à l'origine (voir fig. 9.1). Ce contour délimite un domaine (secteur) où $e^{-z\tau}$ est holomorphe. Le théorème de Cauchy permet donc d'écrire :

$$\int_C e^{-z\tau} d\tau = 0 . \quad (9.20)$$

En décomposant l'intégrale, on peut écrire :

$$\int_0^{+R} e^{-zt} dt + \int_0^{\phi_0} e^{-zR e^{i\phi}} iR e^{i\phi} d\phi + \int_{+R}^0 e^{-z\rho e^{i\phi_0}} e^{i\phi_0} d\rho = 0 . \quad (9.21)$$

¹¹On rappelle que :

$$\Gamma(z) = \int_0^{+\infty} x^{z-1} e^{-x} dx \quad (\Re z > 0) . \quad (9.17)$$

¹²On sait que le prolongement analytique n'est pas toujours possible.

Avec $0 < \phi \leq \phi_0 < \pi/2$, l'intégrale du milieu tend vers zéro exponentiellement avec R ; À la limite $R = +\infty$, on a donc :

$$\int_0^{+\infty} e^{-zt} dt = e^{i\phi_0} \int_0^{+\infty} e^{-zt\rho e^{i\phi_0}} d\rho . \quad (9.22)$$

Maintenant, $z = re^{i\theta}$ étant donné, l'argument de l'exponentielle de droite est $-rt\rho e^{i(\phi_0 + \theta)}$: il suffit de prendre $|\phi_0 + \theta| < \frac{\pi}{2}$ pour que l'intégrale au second membre de (9.22) converge à l'infini et donc existe ; par exemple, quand z quitte l'axe réel en montant, il suffit d'incliner la demi-droite d'intégration vers le bas afin de respecter $|\phi_0 + \theta| < \frac{\pi}{2}$. Ainsi, simplement en jouant avec le contour d'intégration, on voit d'avance que $\mathcal{L}[f](z)$ va pouvoir être prolongée bien au-delà de $\{z, \Re z > 0\}$ – en fait là où $|z| \neq 0$, c'est-à-dire partout sauf à l'origine. Bien évidemment, lorsque l'abscisse de sommabilité est égale à $-\infty$, aucun prolongement analytique n'est nécessaire ; c'est le cas pour la fonction-porte (9.11) : d'emblée, la transformation intégrale produit à elle toute seule la fonction donnée en (9.12), qui est holomorphe dans le plan entier ouvert.

Tout comme pour la transformation de Fourier, se pose naturellement la question de l'inversion de la transformation intégrale (9.2) : étant donné une fonction $F(z)$, considérée comme une image par Laplace, existe-t-il une formule donnant son original $f(t)$, et si oui, quelle est-elle :

$$\boxed{f = \mathcal{L}^{-1}[F] \iff F \xrightarrow{\mathcal{L}^{-1}} f ?} \quad (9.23)$$

La fonction $F(z)$ étant donnée, l'analyse vigilante de son expression permet l'identification complète de ses singularités ; la question est donc de trouver la formule permettant d'obtenir son original $f(t)$.

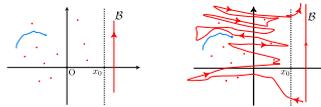


Figure 9.2: Droite de Bromwich utilisée pour l'inversion de la transformation de Laplace (voir (9.28)). Les points schématisent les singularités isolées, le tortillon une coupure, de $F(z)$, toutes situées à la gauche de cette droite.

Pour établir la formule d'inversion, on s'appuie sur ce que l'on sait de la transformation de Fourier. x_0 désignant toujours l'abscisse de sommabilité, l'introduction de la fonction de Heaviside $Y(t)$ permet de récrire (9.2) sous la forme (noter le remplacement de la borne inférieure 0 par $-\infty$) :

$$F(z = x - iy) = \int_{-\infty}^{+\infty} e^{iyt} Y(t) f(t) e^{-xt} dt \quad \forall x > x_0 . \quad (9.24)$$

Le maintien de la contrainte $x > x_0$ est indispensable puisque l'on manipule explicitement la relation intégrale de définition de $F(z)$. Cette dernière écriture montre que $F(x - iy)$ est

la transformée de Fourier de la fonction $Y(t)f(t)e^{-xt}$; en utilisant la formule d'inversion de Fourier (8.122), on a :

$$Y(t)f(t)e^{-xt} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iyt} F(x-iy) dy , \quad (9.25)$$

soit :

$$Y(t)f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{(x-iy)t} F(x-iy) dy . \quad (9.26)$$

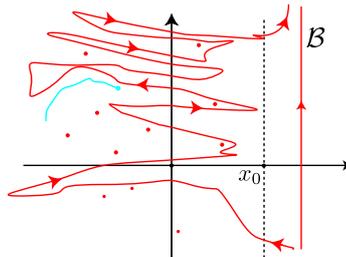


Figure 9.3: Un chemin possible pour l'inversion de Laplace, les points figurant les singularités de $F(z)$ (pôles, points de branchement, ...).

Posons maintenant $z = x-iy$ ($dy = idz$) : x étant fixé, z décrit une droite verticale quand y varie entre $\pm\infty$; notant que quand y varie de $-\infty$ à $+\infty$, z décrit sa droite de *haut en bas*, il vient $\forall t > 0$:

$$Y(t)f(t) = \frac{i}{2\pi} \int_{x+i\infty}^{x-i\infty} e^{zt} F(z) dz , \quad (9.27)$$

soit, en inversant le sens de parcours de la droite \mathcal{B} :

$$\boxed{Y(t)f(t) = \frac{1}{2i\pi} \int_{\mathcal{B}} e^{zt} F(z) dz} \quad (9.28)$$

où \mathcal{B} est une droite parallèle à l'axe imaginaire d'abscisse $x > x_0$ parcourue de *bas en haut*, appelée *droite de Bromwich*. Par construction, toutes les singularités de $F(z)$ sont à *gauche* de cette droite, puisque celle-ci est située à la *droite* de l'abscisse de sommabilité x_0 . Ceci étant acquis, toutes les déformations de \mathcal{B} sont toujours autorisées à condition de ne pas franchir de singularité de $F(z)$. Le contour en *spaghetti* de la fig. 9.3 est tout aussi légitime que \mathcal{B} , ni plus, ni moins.

Il est très instructif de voir la formule d'inversion à l'œuvre pour la fonction $Y(t)$, dont la transformée est $\frac{1}{z}$. Dans ce cas, le second membre de (9.28) est :

$$\frac{1}{2i\pi} \int_{\mathcal{B}} e^{zt} \frac{1}{z} dz . \quad (9.29)$$

Dans un premier temps, limitons la droite entre les deux points $x_1 \pm iR$ ($x_1 > x_0$), et calculons l'intégrale en refermant le contour par un grand arc de cercle de rayon R . Quand $t > 0$, la contribution du grand arc tend vers zéro (exponentiellement) à la limite $R \rightarrow +\infty$ si $\Re z < 0$: pour $t > 0$, on referme à gauche ; l'intégrand de (9.29) a un pôle simple en $z = 0$, dont le résidu vaut 1, de sorte que l'intégrale vaut $2i\pi$. Il vient donc :

$$\frac{1}{2i\pi} \int_{\mathcal{B}} e^{zt} \frac{1}{z} dz = 1, \quad t > 0. \quad (9.30)$$

À l'inverse, si $t < 0$, on referme par la droite ; alors le contour fermé ne contient aucune singularité et, par le théorème de Cauchy, l'intégrale est nulle :

$$\frac{1}{2i\pi} \int_{\mathcal{B}} e^{zt} \frac{1}{z} dz = 0, \quad t < 0. \quad (9.31)$$

Au total, on a bien :

$$\boxed{\frac{1}{2i\pi} \int_{\mathcal{B}} e^{zt} \frac{1}{z} dz = Y(t)} \quad (9.32)$$

Remarquons que si l'on fait $t = 0$ sous l'intégrale dans (9.29), on obtient une intégrale qui se calcule immédiatement puisqu'une primitive, $\ln z$, est connue :

$$\frac{1}{2i\pi} \int_{\mathcal{B}} \frac{1}{z} dz = \frac{1}{2i\pi} |\ln z|_{\mathcal{B}}. \quad (9.33)$$

La droite \mathcal{B} étant dans le demi-plan de *droite*, la variation de la partie imaginaire du logarithme, quelle que soit la détermination choisie, est $+i\pi$. Il en résulte que, pour $t = 0$, l'inversion de Laplace fournit la valeur $\frac{1}{2}$, qui est la demi-somme des valeurs à gauche et à droite de $Y(t)$ (qui n'est pas définie en $t = 0$). Au total, on peut écrire :

$$\boxed{\mathcal{L}^{-1} \left[\frac{1}{z} \right] = Y(t)} \quad (9.34)$$

où l'égalité doit être comprise *presque partout*, soit ici partout sauf en $t = 0$ (où $Y(t)$ n'a pas été définie).

Cet exemple montre bien la nécessité d'effectuer (le plus souvent mentalement et implicitement) le prolongement analytique, faute de quoi aucune incursion dans le demi-plan de gauche $\Re z \leq x_0$ ne serait autorisée, ce que l'on vient pourtant de faire pour trouver $\mathcal{L}^{-1} \left[\frac{1}{z} \right]$ lorsque $t > 0$.

Il est clair que quand l'original $f(t)$ a un (ou plusieurs) saut(s) (fini(s)), on peut en compléter la définition en convenant d'une valeur ou d'une autre au point de discontinuité¹³ : ceci ne change en aucune façon la valeur de la fonction obtenue par la transformation intégrale (on change l'original sur un ensemble de mesure nulle). D'un

¹³Souvent, on prend la valeur limite d'un côté, mais on pourrait prendre aussi n'importe quelle autre valeur – la demi-somme, un choix parmi d'autres – étant *a priori* naturel (et c'est d'ailleurs ce que donne \mathcal{L}^{-1}).

autre côté, la formule inverse fournit une valeur unique, qui n'a aucune raison de coïncider avec la valeur conventionnelle définie antérieurement. C'est pourquoi on précise usuellement que l'inversion de Laplace redonne l'original, aux valeurs aux points de discontinuité près. De fait, $\mathcal{L}^{-1}[F](t)$ et $f(t)$ ($= f(t)Y(t)$) sont *presque partout* égales. Ce phénomène, déjà rencontré à propos des séries et de l'intégrale de Fourier, ne devrait plus surprendre.

Un autre exemple ; soit $f(t) = e^{\alpha t}$, $\alpha \in \mathbb{C}$ (donc $x_0 > \Re\alpha$), dont la transformée est $\frac{1}{z-\alpha}$, fonction holomorphe dans $\mathbb{C} \setminus \{\alpha\}$. La droite de Bromwich est n'importe quelle droite¹⁴ parallèle à l'axe imaginaire et coupant l'axe réel en un point d'abscisse strictement supérieure à α . La formule d'inversion (9.28) est ici :

$$\frac{1}{2i\pi} \int_{\mathcal{B}} e^{zt} \frac{1}{z-\alpha} dz . \quad (9.35)$$

On referme le contour comme ci-dessus : par la gauche si $t > 0$, le pôle simple en $z = \alpha$ donnant au total $e^{\alpha t}$, par la droite si $t < 0$, auquel cas le contour ne contient aucune singularité et l'intégrale est nulle. En définitive :

$$\frac{1}{2i\pi} \int_{\mathcal{B}} e^{zt} \frac{1}{z-\alpha} dz = Y(t) e^{\alpha t} , \quad \forall t \quad (9.36)$$

comme il se doit. Noter aussi que, au simple vu de (9.18), on a l'important résultat :

$$\boxed{\mathcal{L}^{-1}[z^{\beta}] = \frac{1}{\Gamma(-\beta)t^{\beta+1}} Y(t) \quad (\beta < 0)} \quad (9.37)$$

Tout comme pour la transformation de Fourier, il est parfois nécessaire (ou en tout cas très utile) de définir la transformée de Laplace de "fonctions" obtenues comme limite de certaines suites. Le cas le plus utile est celui de la fonction $\delta_{\Delta t}(t)$ définie comme :

$$\delta_{\Delta t}(t) = \begin{cases} 0 & \text{si } t = 0 \\ \frac{1}{\Delta t} & \text{si } 0 < t < \Delta t \\ 0 & \text{si } t \geq \Delta t \end{cases} . \quad (9.38)$$

Si $\Delta t \ll 1$, cette fonction est très étroite et l'aire de son graphe vaut 1 : c'est un précurseur de la fonction de Dirac centrée juste à droite de $t = 0$, que l'on peut noter $\delta^{(+)}(t)$, résultant de la limite symbolique :

$$\lim_{\Delta t \rightarrow 0} \delta_{\Delta t}(t) \stackrel{\text{déf}}{=} \delta^{(+)}(t) . \quad (9.39)$$

La transformée de Laplace de $\delta_{\Delta t}(t)$ est :

$$\mathcal{L}[\delta_{\Delta t}(t)] \equiv \Delta_{\Delta t}(z) = \int_0^{\Delta t} e^{-zt} \frac{1}{\Delta t} dt = \frac{1 - e^{-z\Delta t}}{z\Delta t} . \quad (9.40)$$

¹⁴ou, bien sûr, n'importe quelle ligne laissant à sa gauche toutes les singularités de l'image.

Ceci montre que la largeur en z de $\delta_{\Delta t}(z)$ est $\frac{1}{\Delta t}$: tout comme pour la transformation de Fourier, le produit de la largeur Δt de l'original et $\Delta \Re z$ de son image – qui ici vaut strictement 1 – est d'une façon générale un nombre d'ordre unité¹⁵ :

$$\Delta t \Delta \Re z \sim 1 . \quad (9.41)$$

Maintenant, pour tout z fixé, la limite de $\Delta_{\Delta t}(z)$ quand $\Delta t \rightarrow 0$ donne :

$$|z| > 0 : \quad \Delta(z) \stackrel{\text{déf}}{=} \lim_{\Delta t \rightarrow 0} \Delta_{\Delta t}(z) = \lim_{\Delta t \rightarrow 0} \frac{1 - e^{-z\Delta t}}{z\Delta t} = 1 ; \quad (9.42)$$

ceci permet d'écrire :

$$\boxed{\mathcal{L}[\delta^{(+)}(t)] = 1} \quad (9.43)$$

Réciproquement, on a :

$$\boxed{\mathcal{L}^{-1}[1] = \delta^{(+)}(t)} \quad (9.44)$$

Ces relations sont clairement la traduction dans le langage de Laplace des égalités $\mathcal{F}[\delta(x)] = 1$ et $\mathcal{F}^{-1}[1] = \delta(x)$ obtenues dans le chapitre 8.

Notons qu'il existe plusieurs familles de fonctions (en fait, autant qu'on veut) qui, ayant toutes la même limite $\delta^{(+)}(t)$, ont également toutes $F(z) = 1$ pour transformée de Laplace. Par exemple, soit :

$$f_{\Delta t}(t) = Y(t) \frac{1}{\Delta t} e^{-\frac{t}{\Delta t}} \iff F_{\Delta t}(z) \equiv \mathcal{L}[f_{\Delta t}] = \frac{1}{\Delta t} \frac{1}{z + \frac{1}{\Delta t}} . \quad (9.45)$$

Dans la limite $\Delta t \rightarrow 0$, on trouve $F_{0+}(z) = 1$, tout comme pour la fonction $\delta^{(+)}(t)$ ci-dessus.

◆ *Remarque*

On a vu que, de la définition de la transformée de Laplace $F(z)$ d'une fonction $f(t)$, il résulte que $\lim_{|z| \rightarrow \infty} |F(z)| = 0$ (voir (9.7)). Cette propriété triviale n'est pas anodine : elle permet de dire d'emblée si une fonction $H(z)$ "tirée au hasard" peut être ou non une transformée de Laplace au sens strict.

À titre d'exemple, soit la gaussienne $G(z) = e^{\tau^2 z^2}$ ($\tau \in \mathbb{R}$), qui n'est pas bornée lorsque $|z| \rightarrow \infty$, quand z est dans le double secteur $|\text{Arg } z| < \frac{\pi}{4}$ (π). La propriété ci-dessus permet d'affirmer *a priori* qu'il n'existe pas de fonction $g(t)$ nulle à $t < 0$ dont $G(z)$ est la transformée de Laplace :

$$\nexists g(t) : \quad g(t) = \frac{1}{2i\pi} \int_{\mathcal{B}} e^{zt} e^{\tau^2 z^2} dz \quad (t > 0) , \quad g(t) = 0 \quad (t < 0) . \quad (9.46)$$

¹⁵Comme pour les couples de Fourier, ceci tient à la nature même de la transformation intégrale de définition, qui conduit immédiatement à la relation de dilatation (voir (9.70)) : plus l'original est concentré dans le temps, plus l'image est étendue en z , et inversement.

D'un autre côté, rien n'interdit de considérer l'intégrale $I(t)$:

$$I(t) \stackrel{\text{déf}}{=} \frac{1}{2i\pi} \int_D e^{zt} e^{\tau^2 z^2} dz, \quad (9.47)$$

où D est une certaine droite verticale dans \mathbb{C} , d'abscisse $a \in \mathbb{R}$, parcourue de bas en haut. L'intégrale (9.47) peut s'effectuer directement en posant $z = a + iy$; ainsi :

$$I(t) = \frac{1}{2i\pi} \int_{-\infty}^{+\infty} e^{(a+iy)t} e^{\tau^2(a+iy)^2} idy = \frac{1}{2\pi} e^{at} e^{a^2\tau^2} \int_{-\infty}^{+\infty} e^{iyt} e^{-\tau^2(ay^2 - 2iaiy)} dy ; \quad (9.48)$$

utilisant $\int_{-\infty}^{+\infty} e^{-\alpha u^2 + \beta u} du = \left(\frac{\pi}{\alpha}\right)^{1/2} e^{\frac{\beta^2}{4\alpha}}$, $\forall \alpha, \Re \alpha > 0, \forall \beta \in \mathbb{C}$, on trouve :

$$I(t) = \frac{1}{2\sqrt{\pi}\tau} e^{-\frac{t^2}{4\tau^2}}. \quad (9.49)$$

Au passage, on note que l'abscisse a de la droite D disparaît ; ceci est une simple conséquence du fait que $e^{zt} e^{\tau^2 z^2}$ est une fonction entière et que l'on peut déplacer la droite (et aussi la déformer en un *spaghetti*) sans changer la valeur de $I(t)$: on ne risque pas de tomber sur des singularités qui n'existent pas.

Plus important est de noter que le résultat (9.49) est vrai quel que soit t , *positif* ou *négatif*. Ainsi, il existe bien une fonction $I(t)$ obtenue par intégration suivant la recette d'inversion de Laplace, mais en l'occurrence (9.47) ne peut être interprétée comme une relation d'inversion, en dépit des apparences : c'est juste une relation de définition d'une certaine fonction, $I(t)$ en la circonstance. Il s'avère que cette fonction n'est pas de la forme $Y(t) \times$ quelque chose, et n'est donc pas un original dont la gaussienne serait l'image par \mathcal{L} , et ce bien que l'égalité $e^{\tau^2 z^2} = \int_0^{+\infty} e^{-zt} \frac{1}{2\sqrt{\pi}\tau} e^{\frac{t^2}{4\tau^2}}$ est vraie.

On peut résumer les choses ainsi. Pour la transformation de Laplace, la succession des étapes est :

$$\text{étape (I) : } f(t) \xrightarrow{\mathcal{L}} F(z) = \int_0^{+\infty} e^{-zt} f(t) dt \equiv \mathcal{L}[f](z)$$

$$\text{étape (II) : } F(z) \xrightarrow{\mathcal{L}^{-1}} f(t) = \frac{1}{2i\pi} \int_{\mathcal{B}} e^{zt} F(z) dz \equiv \mathcal{L}^{-1}[\mathcal{L}[f](z)](t) = f(t). \quad (9.50)$$

F est l'image d'une certaine fonction f (original), obtenue suivant la prescription de Laplace, impliquant automatiquement une abscisse de sommabilité $x_0 < +\infty$, et mettant en jeu la droite \mathcal{B} située à la droite de toutes les singularités de $F(z)$. Dès lors, la fonction f "retrouvée" par la formule d'inversion est nécessairement nulle pour $t < 0$.

D'un autre côté, si on "débarque" avec une fonction $H(z)$ tirée d'un chapeau, on peut toujours définir une certaine fonction $h(t)$ (si cette écriture a un sens dans le cas particulier considéré, évidemment) :

$$h(t) = \frac{1}{2i\pi} \int_D e^{zt} H(z) dz ; \quad (9.51)$$

mais rien n'assure que la relation $\tilde{H}(z) = \int_0^{+\infty} e^{-zt} h(t) dt$ existe. En d'autres termes, la recette inverse $\frac{1}{2i\pi} \int_D e^{zt} \bullet dz$ renvoie dans un ensemble plus vaste que celui des fonctions nulles à $t < 0$.

Noter enfin que $\mathcal{L}[\delta^{(+)}]$ ne possède pas la propriété exprimée en (9.7), au contraire de la transformée de tous ses précurseurs avant la limite. \blacklozenge

9.3 Propriétés de la transformation de Laplace

À l'instar de la transformation de Fourier, la transformation de Laplace possède des propriétés importantes, qui résultent directement de la définition (9.2). On note toujours :

$$f(t) \xrightarrow{\mathcal{L}} F(z) , \quad F = \mathcal{L}[f] ; \quad F(z) \xrightarrow{\mathcal{L}^{-1}} f(t) , \quad f = \mathcal{L}^{-1}[F] . \quad (9.52)$$

■ **Linéarité** L'intégrale étant une opération linéaire, on a de toute évidence :

$$\boxed{\mathcal{L}[\lambda f + \mu g] = \lambda \mathcal{L}[f] + \mu \mathcal{L}[g]} \quad (9.53)$$

où λ et μ sont des scalaires. La transformation de Laplace est donc elle aussi bien adaptée à la résolution d'équations *linéaires*, et produira aisément les solutions admettant une transformée de Laplace, c'est-à-dire ne croissant pas plus vite qu'une certaine exponentielle.

■ **Translation** Étant donné une fonction $f(t)$, sa translatée de t_0 est la fonction $(\mathcal{T}_{t_0} f)(t) \stackrel{\text{déf}}{=} f(t - t_0)$; cette fonction est non nulle seulement pour $t > t_0$. La transformée de $\mathcal{T}_{t_0} f$ est :

$$\mathcal{L}[\mathcal{T}_{t_0} f] = \int_0^{+\infty} e^{-zt} f(t - t_0) dt = \int_{t_0}^{+\infty} e^{-zt} f(t - t_0) dt = e^{-zt_0} \int_0^{+\infty} e^{-zt'} f(t') dt' , \quad (9.54)$$

d'où :

$$\boxed{\mathcal{L}[\mathcal{T}_{t_0} f] = e^{-zt_0} \mathcal{L}[f]} \quad (9.55)$$

Dans le contexte traditionnel de la transformation de Laplace, ce résultat s'appelle parfois *théorème du retard*.

• **Exemple** À titre d'illustration élémentaire, soit l'équation :

$$f(t) = f(t - t_0) + a , \quad f(t < 0) = 0 \quad (a, t_0 \in \mathbb{R}_+) , \quad (9.56)$$

qui peut être résolue commodément par la transformation de Laplace – tout en ayant remarqué que la solution se trouve aussi à vue ; en effet, quand $0 < t < t_0$, l'équation est

$f(t) = 0 + a = a$; quand $t_0 < t < 2t_0$, l'équation est $f(t) = a + a = 2a$ et ainsi de suite. La solution nulle à $t < 0$ vaut donc à t positif :

$$f(t) = a \left[1 + E\left(\frac{t}{t_0}\right) \right] \quad (t > 0) \tag{9.57}$$

où $E(x)$ désigne la partie entière du réel x : c'est une fonction en escalier, qui vaut a entre 0 et t_0 , $2a$ entre t_0 et $2t_0$, etc.

D'une autre façon, la transformation de Laplace appliquée à (9.56) donne :

$$F(z) = e^{-zt_0} F(z) + \frac{a}{z} \iff F(z) = \frac{a}{z(1 - e^{-zt_0})} , \tag{9.58}$$

d'où, par la formule d'inversion :

$$f(t) = \frac{1}{2i\pi} \int_{\mathcal{B}} e^{zt} \frac{a}{z(1 - e^{-zt_0})} dz . \tag{9.59}$$

L'intégrale se calcule par résidus. Les pôles sont $z_0 = 0$ d'une part, les zéros de $1 - e^{-zt_0}$ soit $z_k = \frac{2ik\pi}{t_0}$, $k \in \mathbb{Z}$, d'autre part ; le pôle en 0 est double, le résidu vaut $\frac{1}{2} + \frac{t}{t_0}$, tous les autres sont simples et donnent chacun le résidu $\frac{1}{2ik\pi} e^{2ik\pi \frac{t}{t_0}}$ ($k \neq 0$). L'expression de $f(t)$ qui résulte de l'application du théorème des résidus est ainsi :

$$\frac{1}{a} f(t) = \frac{1}{2} + \frac{t}{t_0} + \sum_{k \in \mathbb{Z}^*} \frac{e^{2ik\pi \frac{t}{t_0}}}{2ik\pi} , \tag{9.60}$$

étant sous-entendu que $f(t < 0) = 0$. Le troisième terme de l'expression (9.60) est une série de Fourier, qui représente la fonction $\frac{1}{2} - \frac{t}{t_0}$ entre 0 et t_0 , t_0 -périodisée sur \mathbb{R} . Au total, l'expression au second membre de (9.60) coïncide avec (9.57).

La série de (9.60) peut aussi se calculer en passant par l'intermédiaire de la dérivée $f'(t)$, qui donne un peigne de Dirac, à un terme près, et a pour expression (toujours pour $t > 0$) :

$$\frac{1}{a} f'(t) = \frac{1}{t_0} + \frac{1}{t_0} \sum_{k \in \mathbb{Z}^*} e^{2ik\pi \frac{t}{t_0}} = \frac{2\pi}{t_0} \sum_{p \in \mathbb{Z}} \delta\left(2\pi \frac{t}{t_0} - 2p\pi\right) = \frac{1}{t_0} \sum_{p \in \mathbb{N}} \delta(t - pt_0) ; \tag{9.61}$$

l'écriture de la dernière somme (\mathbb{N} au lieu de \mathbb{Z}) anticipe le fait que f et f' sont nulles $\forall t < 0$, de sorte que les $\delta(t - pt_0)$, $p < 0$, peuvent être omis. Par intégration, et après calage de la constante d'intégration sachant que $f(0 < t < t_0) = a$, on en déduit¹⁶ :

$$f(t) = a \sum_{p \in \mathbb{N}} Y(t - pt_0) = a [Y(t) + Y(t - t_0) + Y(t - 2t_0) + \dots] \quad \forall t , \tag{9.62}$$

qui a bien les mêmes valeurs que la fonction $f(t)$ donnée en (9.57) ou en (9.60). À tout t fini, la série (9.62) ne contient qu'un nombre fini de termes non-nuls.

¹⁶Une fois connues les règles (exposées plus loin) concernant les relations entre les dérivées et les primitives d'un original et de son image, l'inversion de (9.59) se fait plus rapidement.

L'inversion peut enfin se faire comme suit. La transformée $F(z)$ dans (9.59) se développe en série entière :

$$F(z) = \frac{a}{z(1 - e^{-zt_0})} = \frac{a}{z} \sum_{n=0}^{+\infty} e^{-nzt_0} . \quad (9.63)$$

Ce développement converge uniformément $\forall z, |e^{-zt_0}| < 1$, soit $\Re z > 0$. La transformation inverse peut donc se faire en intégrant terme à terme ; se souvenant que $\mathcal{L}^{-1}\left[\frac{1}{z}\right] = Y(t)$ et compte tenu du théorème du retard donnant ici $\mathcal{L}^{-1}\left[e^{-nzt_0}\frac{1}{z}\right] = Y(t - nt_0)$:

$$\mathcal{L}^{-1}[F](t) = a \sum_{n=0}^{+\infty} \mathcal{L}^{-1}\left[e^{-nzt_0}\frac{1}{z}\right](t) = a \sum_{n=0}^{+\infty} \mathcal{T}_{nt_0} Y(t) = a \sum_{n=0}^{+\infty} Y(t - nt_0) . \quad (9.64)$$

Pour tout t fini, la série est en fait tronquée au plus grand entier n tel que $nt_0 < t$, soit $E\left(\frac{t}{t_0}\right)$ (tous les termes d'après sont nuls). En définitive, on retrouve bien l'expression (9.62), qui est partout égale à (9.57). •

Revenons sur l'important résultat pratique énoncé en (9.37), récrit sous la forme :

$$\boxed{\mathcal{L}^{-1}\left[\frac{1}{z^\alpha}\right] = \frac{t^{\alpha-1}}{\Gamma(\alpha)} Y(t) \quad (\Re \alpha > 0)} \quad (9.65)$$

Maintenant, par le théorème de translation, on en déduit :

$$\boxed{\mathcal{L}^{-1}\left[\frac{e^{-zt_0}}{z^\alpha}\right] = \mathcal{T}_{t_0} \frac{t^{\alpha-1}}{\Gamma(\alpha)} Y(t) = \frac{(t - t_0)^{\alpha-1}}{\Gamma(\alpha)} Y(t - t_0) \quad (\Re \alpha > 0)} \quad (9.66)$$

■ **Modulation** Étant donnée une fonction $f(t)$, sa modulée à ω_0 est $e^{-i\omega_0 t} f(t)$. On a :

$$\mathcal{L}[e^{-i\omega_0 t} f(t)] = \int_0^{+\infty} e^{-(z+i\omega_0)t} f(t) dt = \mathcal{L}[f](z + i\omega_0) \equiv F(z + i\omega_0) . \quad (9.67)$$

Plus généralement, on a :

$$\boxed{\mathcal{L}[e^{-\gamma t} f(t)] = F(z + \gamma)} \quad (9.68)$$

Exemple : on trouve sans peine que la fonction $f(t) = Y(t) \cos \omega t$ a pour transformée $F(z) = \frac{z}{z^2 + \omega^2}$. Soit la fonction modulée $f_{\text{mod}}(t) \stackrel{\text{déf}}{=} \sin \delta \omega t f(t)$; sa transformée de Laplace est $F_{\text{mod}}(z) = \frac{1}{2i} \left[\frac{z - i\delta\omega}{(z - i\delta\omega)^2 + \omega^2} - \frac{z + i\delta\omega}{(z + i\delta\omega)^2 + \omega^2} \right]$.

■ **Dilatation** Un changement d'échelle sur la variable d'une fonction $f(t)$ définit une nouvelle fonction f_λ telle que $f_\lambda(t) \stackrel{\text{déf}}{=} f(\lambda t)$. Supposant λ positif, on a :

$$\mathcal{L}[f_\lambda(t)] = \int_0^{+\infty} e^{-zt} f(\lambda t) dt = \frac{1}{\lambda} \int_0^{+\infty} e^{-z\frac{t'}{\lambda}} f(t') dt' , \quad (9.69)$$

d'où :

$$\mathcal{L}[f_\lambda(t)] = \frac{1}{\lambda} F\left(\frac{z}{\lambda}\right) \quad (\lambda \in \mathbb{R}_+) \quad (9.70)$$

Fondamentalement, c'est cette propriété qui est à l'origine de la "relation d'incertitude" entre Δt et $\Delta \Re z$ (voir (9.41)) : tout comme un couple de Fourier, plus f est ramassée, plus F est étendue, et réciproquement.

• **Exemple** À titre d'illustration, montrons comment la relation (9.70) permet d'obtenir astucieusement la transformée de la fonction¹⁷ $f(t) = Y(t) \ln \frac{t}{\tau}$. En effet, posant $F(z) = \mathcal{L}[\ln \frac{t}{\tau}]$ on a d'après (9.70) :

$$\mathcal{L}\left[\ln \lambda \frac{t}{\tau}\right] = \frac{1}{\lambda} F\left(\frac{z}{\lambda}\right) ; \quad (9.71)$$

En utilisant $\ln \lambda \frac{t}{\tau} = \ln \lambda + \ln \frac{t}{\tau}$, on obtient :

$$F(z) + \frac{\ln \lambda}{z} = \frac{1}{\lambda} F\left(\frac{z}{\lambda}\right) ; \quad (9.72)$$

maintenant, dérivant membre à membre par rapport à λ et faisant ensuite $\lambda = 1$, on trouve que $F(z)$ satisfait l'équation différentielle $zF'(z) + F(z) = -\frac{1}{z}$, dont la solution générale est $-\frac{1}{z} \ln \frac{z}{z_0}$, où z_0 est une constante d'intégration, que l'on peut trouver en connaissant une valeur particulière de l'intégrale de définition de $F(z)$. Si on fait $z = \tau^{-1}$, on a :

$$F(z = \tau^{-1}) = \int_0^{+\infty} e^{-\frac{t}{\tau}} \ln \frac{t}{\tau} dt = \tau \int_0^{+\infty} e^{-u} \ln u du . \quad (9.73)$$

La dernière intégrale est connue et est égale à l'opposé de la constante d'Euler¹⁸, le plus souvent notée γ :

$$\int_0^{+\infty} e^{-u} \ln u du = -\gamma = -0,577 \dots . \quad (9.75)$$

D'où, notant $\mathbf{C} = e^\gamma = 1,781 \dots$ comme le font la plupart des auteurs :

$$\mathcal{L}\left[\ln \frac{t}{\tau}\right] = -\frac{1}{z} \ln(\mathbf{C}\tau z) \quad (9.76)$$

La branche du logarithme à retenir est celle satisfaisant la relation $\ln(z^*) = (\ln z)^*$ (voir juste ci-après). L'échelle de variation de l'original au premier membre est τ , celle de l'image est $(\mathbf{C}\tau)^{-1}$, en conformité avec la relation (9.41). Noter aussi qu'il s'agit d'un exemple où l'original n'est pas borné à l'origine mais y est localement sommable. •

¹⁷On note que cette fonction n'a pas de transformée de Fourier.

¹⁸On sait plus précisément (!) que :

$$\gamma = 0,577\,215\,664\,901\,532\,860\,606\,512\,090\,082\,402\,431\,042\,159\,335\,939\,923\,598\,805\,767\,234\,884 \dots$$

Cette constante apparaît fréquemment en Analyse et en Théorie des nombres ; on peut la définir comme :

$$\gamma = \lim_{N \rightarrow +\infty} \left(\sum_{n=1}^N \frac{1}{n} - \ln N \right) , \quad (9.74)$$

égalité déjà introduite dans le chapitre 7, éq. (7.80).

■ Conjugaison

$$\mathcal{L}[f^*] = \int_0^{+\infty} e^{-zt} f^*(t) dt = \left[\int_0^{+\infty} e^{-z^*t} f(t) dt \right]^* = F^*(z^*) . \quad (9.77)$$

Cette propriété joue un rôle important lorsque $F(z)$ est une fonction multiforme, car elle permet de lever toute ambiguïté sur la détermination à considérer. Par exemple (avec $\Re\alpha > -1$) :

$$\mathcal{L}[t^\alpha] = \int_0^{+\infty} e^{-zt} t^\alpha dt = \frac{\Gamma(\alpha+1)}{z^{\alpha+1}} \quad (\Re z > 0) , \quad (9.78)$$

où, à nouveau, référence a été faite à la définition intégrale de la fonction Γ d'Euler. Lorsque α n'est pas entier, la fonction $z^{\alpha+1}$ a un point de branchement en $z = 0$, mais la coupure est ici complètement déterminée par la définition même de la transformation de Laplace. En effet, compte tenu de la relation de conjugaison (9.77), la coupure est nécessairement le demi-axe réel négatif, et on a sans ambiguïté :

$$z = r e^{i\theta} , \quad z^{\alpha+1} = r^{\alpha+1} e^{i(\alpha+1)\theta} \quad (-\pi < \theta < +\pi) . \quad (9.79)$$

■ Dérivation Un résultat particulièrement utile en pratique est le lien entre les transformées de Laplace d'une fonction et de ses dérivées. La transformée de la dérivée f' est par définition :

$$\mathcal{L}[f'] = \int_0^{+\infty} e^{-zt} f'(t) dt \quad (\forall z, \Re z > x_0) . \quad (9.80)$$

Une intégration par parties donne :

$$\mathcal{L}[f'] = [f(t)e^{-zt}]_0^{+\infty} - \int_0^{+\infty} (-z) e^{-zt} f(t) dt . \quad (9.81)$$

Le terme tout intégré vaut¹⁹ $-f(0)$, d'où :

$$\mathcal{L}[f'] = -f(0) + z\mathcal{L}[f](z) , \quad (9.82)$$

d'où l'égalité très importante :

$$\boxed{\mathcal{L}[f'] = zF(z) - f(0)} \quad (9.83)$$

De la même façon :

$$\mathcal{L}[f''] = z\mathcal{L}[f'](z) - f'(0) = z[zF(z) - f(0)] - f'(0) = z^2F(z) - zf(0) - f'(0) , \quad (9.84)$$

et plus généralement ($f^{(m)}(t) = \frac{d^m f}{dt^m}$, $m \in \mathbb{N}^*$) :

$$\begin{aligned} \mathcal{L}[f^{(m)}] &= z^m F(z) - z^{m-1}f(0) - z^{m-2}f'(0) - \dots - f^{(m-1)}(0) = \\ &= z^m F(z) - \sum_{r=0}^{m-1} z^{m-r-1} f^{(r)}(0) . \end{aligned} \quad (9.85)$$

¹⁹Il faut bien sûr comprendre que $f(0)$ désigne la limite à droite $f(0+)$. Par ailleurs, la limite $\lim_{z \rightarrow \infty} e^{-zt} f(t)$ est nulle, puisque $f(t)$ ne croît pas plus vite que $e^{x_0 t}$.

Noter qu'il existe une abscisse de sommabilité x_0 commune à $F(z)$ et à toutes les transformées des dérivées de $f(t)$ puisque si $F(z)$ est holomorphe dans $\{z : \Re z > x_0\}$, il en va de même de toutes les combinaisons (9.85).

• **Exemple** À titre d'exemple, soit l'équation :

$$f'(t) - \gamma f(t - t_0) = \phi(t) \tag{9.86}$$

où $\phi(t)$ est une source donnée, et sachant que $f(0) = f_0$ (avec toujours $f(t < 0) = 0$) ; cette équation est à la fois différentielle et aux différences finies. La transformation de Laplace donne ($\Phi = \mathcal{L}[\phi]$) :

$$zF(z) - f_0 - \gamma e^{-zt_0} F(z) = \Phi(z) \iff F(z) = \frac{\Phi(z) + f_0}{z - \gamma e^{-zt_0}} . \tag{9.87}$$

La fonction cherchée est donc :

$$f(t) = \frac{1}{2i\pi} \int_{\mathcal{B}} \frac{f_0 e^{zt}}{z - \gamma e^{-zt_0}} dz + \frac{1}{2i\pi} \int_{\mathcal{B}} \frac{\Phi(z) e^{zt}}{z - \gamma e^{-zt_0}} dz \equiv f_h(t) + f_{\text{source}}(t) . \tag{9.88}$$

Le terme contenant la source $\Phi(z)$ ne peut être explicité davantage tant que l'on ne sait pas précisément ce qu'est $\phi(t)$. En revanche, le terme homogène peut s'inverser en développant en série la fraction rationnelle :

$$f_h(t) = \mathcal{L}^{-1} \left[\frac{f_0}{z - \gamma e^{-zt_0}} \right] = f_0 \mathcal{L}^{-1} \left[\frac{1}{z} \sum_{n=0}^{+\infty} \left(\frac{\gamma}{z} e^{-zt_0} \right)^n \right] = f_0 \sum_{n=0}^{+\infty} \gamma^n \mathcal{T}_{nt_0} \mathcal{L}^{-1} \left[\frac{1}{z^{n+1}} \right] . \tag{9.89}$$

En utilisant maintenant (9.37), il vient :

$$\begin{aligned} \mathcal{L}^{-1} \left[\frac{f_0}{z - \gamma e^{-zt_0}} \right] &= f_0 \sum_{n=0}^{+\infty} \gamma^n \mathcal{T}_{nt_0} \frac{t^n Y(t)}{\Gamma(n+1)} = f_0 \sum_{n=0}^{+\infty} \frac{\gamma^n}{n!} (t - nt_0)^n Y(t - nt_0) = \\ &= f_0 \sum_{n=0}^{E(\frac{t}{t_0})} \frac{\gamma^n}{n!} (t - nt_0)^n . \end{aligned} \tag{9.90}$$

Explicitement, la fonction $f_h(t)$ définie en (9.90) est :

$$f_h(t) = \begin{cases} f_0 & \text{si } 0 < t < t_0 \\ f_0 [1 + \gamma(t - t_0)] & \text{si } t_0 < t < 2t_0 \\ f_0 [1 + \gamma(t - t_0) + \frac{\gamma^2}{2}(t - 2t_0)^2] & \text{si } 2t_0 < t < 3t_0 \\ \dots & \dots \end{cases} . \tag{9.91}$$

Cette fonction un peu bizarre a un saut de la n^e dérivée en $t = nt_0$: nulle avant, la dérivée en question vaut γ^n à $nt_0 + 0$. On ne s'en douterait pas au seul vu du graphe (voir fig.9.4) : un graphe d'apparence anodine peut se référer à une fonction un peu

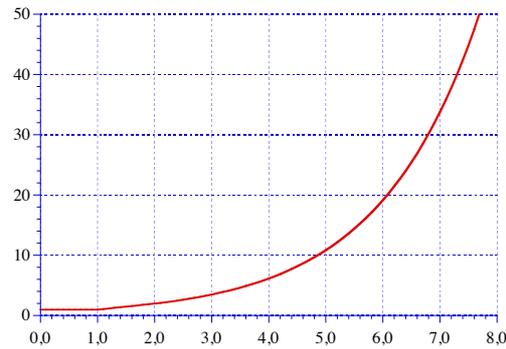


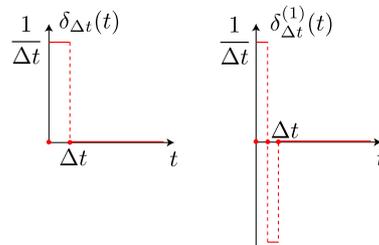
Figure 9.4: Graphe de la fonction définie en (9.91).

étrange²⁰. Sur l'expression (9.90), on voit vite que $\lim_{t_0 \rightarrow 0} f_h(t) = f_0 e^{\gamma t}$, comme il se doit.

Notons que l'usage de la transformation de Laplace n'est pas ici absolument nécessaire. En effet, la solution de l'équation homogène :

$$f'(t) - \gamma f(t - t_0) = 0 \quad (9.92)$$

peut aussi se trouver de proche en proche, en raisonnant successivement dans les intervalles $[nt_0, (n+1)t_0]$. Pour $0 < t < t_0$, l'équation est $f'(t) = 0$, soit $f(t) = f_0$; pour $t_0 < t < 2t_0$, l'équation est $f'(t) - \gamma f_0 = 0$, qui donne $f(t) = f_0 \gamma(t - t_0) + f_0$. Pour $2t_0 < t < 3t_0$, on a $f'(t) - \gamma f(t - t_0) = 0$ avec $f(t) = f_0 \gamma(t - t_0) + f_0$, soit $f'(t) - \gamma[f_0 \gamma(t - 2t_0) + f_0] = 0$, et ainsi de suite. •

Figure 9.5: Graphes des fonctions $\delta_{\Delta t}(t)$ et $\delta_{\Delta t}^{(1)}(t)$.

Montrons maintenant comment il est possible de définir les transformées de Laplace des “dérivées” de la fonction $\delta_{\Delta t}$ définie en (9.38), en acceptant de se placer sur le

²⁰Une discontinuité de la fonction saute aux yeux, tout comme une rupture de pente ; en revanche, un saut de dérivée seconde est quasi invisible, et *a fortiori* pour les dérivées d'ordre supérieur.

même terrain quelque peu symbolique. Notons qu'un précurseur gaussien de la fonction de Dirac $\delta(x)$ a pour dérivée une fonction impaire croissant très vite pour $x \lesssim 0$, et s'annule en $x=0$ et décroît très vite pour $x \gtrsim 0$. D'où l'idée de considérer la fonction²¹ :

$$\delta_{\Delta t}^{(1)}(t) = \begin{cases} +\frac{1}{\Delta t} & 0 < t < \frac{\Delta t}{2} \\ -\frac{1}{\Delta t} & \frac{\Delta t}{2} < t < \Delta t \\ 0 & \text{si } t = 0, \frac{\Delta t}{2}, \text{ et si } t \geq \Delta t \end{cases} . \quad (9.93)$$

On trouve facilement que sa transformée de Laplace est $\frac{1}{z\Delta t}(1 - e^{-\frac{1}{2}z\Delta t})^2$. Dans la limite $\Delta t \rightarrow 0$, on obtient :

$$\lim_{\Delta t \rightarrow 0} \mathcal{L}[\delta_{\Delta t}^{(1)}(t)](z) = z . \quad (9.94)$$

Revenant à la définition de $\delta^{(+)}(t)$, (9.43), et compte tenu de (9.83), on voit que l'on peut considérer $\delta_{\Delta t \rightarrow 0}^{(1)}(t)$ comme la "dérivée" de $\delta^{(+)}(t)$. Plus généralement, z^m sera la transformée de Laplace de la m^e "dérivée" $\delta^{(+)(m)}(t)$ de $\delta^{(+)}(t)$:

$$\boxed{\mathcal{L}[\delta^{(+)(m)}(t)](z) = z^m \iff \mathcal{L}^{-1}[z^m](t) = \delta^{(+)(m)}(t)} \quad (9.95)$$

Il est clair que, en raison de leur statut quelque peu symbolique, ces définitions (conventionnelles) conduisent à des images ne possédant pas la propriété (9.7) propre à toute "bonne" transformée de Laplace.

Notons enfin que (9.83) se lit aussi :

$$zF(z) = \mathcal{L}[f'](z) + f(0) , \quad (9.96)$$

de sorte que pour des "bonnes fonctions" (telles que leur transformée de Laplace est nulle à l'infini), on a :

$$\boxed{\lim_{z \rightarrow \infty} zF(z) = f(0)} \quad (9.97)$$

■ **Intégration** Soit $\int_0^t f(t')dt'$ la primitive de f qui s'annule en $t = 0$. Sa dérivée est égale à $f(t)$, dont la transformée de Laplace est $\mathcal{L}[f] = F(z)$, mais d'après (9.82), elle vaut aussi $z\mathcal{L}[\int_0^t f(t')dt'] - 0$. On en déduit :

$$\mathcal{L}\left[\int_0^t f(t')dt'\right](z) = \frac{1}{z}\mathcal{L}[f](z) = \frac{F(z)}{z} . \quad (9.98)$$

Comme pour la transformation de Fourier, la symétrie entre \mathcal{L} et \mathcal{L}^{-1} permet d'établir des propriétés duales. Par exemple, si on translate la transformée $F(z)$ d'une fonction $f(t)$, $\mathcal{T}_{z_0}F(z) = F(z - z_0)$, la transformée inverse est :

$$\mathcal{L}^{-1}[\mathcal{T}_{z_0}F] = \frac{1}{2i\pi} \int_{\mathcal{B}} e^{zt} F(z - z_0) dz = e^{z_0 t} \frac{1}{2i\pi} \int_{\mathcal{B}} e^{z't} F(z') dz' = e^{z_0 t} \mathcal{L}^{-1}[F] . \quad (9.99)$$

²¹Remarquer que l'intégrale $\int_0^t \delta_{\Delta t}^{(1)}(t') dt'$ ne donne pas $\delta_{\Delta t}(t)$ mais un triangle très pointu de base Δt et de hauteur $\frac{1}{2}$.

Ceci est en quelque sorte le théorème dual du théorème du retard. Dans le même ordre d'idée, considérons l'intégrale $\int_z^\infty F(z') dz'$, qui est l'opposé de la primitive de $F(z)$ s'annulant à l'infini. On a :

$$\int_z^\infty F(z') dz' = \int_z^\infty \left(\int_0^{+\infty} e^{-z't} f(t) dt \right) dz' \quad (9.100)$$

En prenant le chemin d'intégration en z dans le demi-plan $\Re z \geq x_1 > x_0$, l'intégrale interne est majorée comme suit ($f(t) \leq M e^{x_0 t}$) :

$$\left| \int_0^{+\infty} e^{-z't} f(t) dt \right| \leq M \int_0^{+\infty} e^{-(x_1-x_0)t} dt . \quad (9.101)$$

L'intégrale interne converge donc uniformément en z' ; intervertissant l'ordre des intégrations dans (9.100) :

$$\int_z^\infty F(z') dz' = \int_0^{+\infty} f(t) dt \left(\int_z^\infty e^{-z't} dz' \right) = \int_0^{+\infty} f(t) dt \frac{e^{-zt}}{t} \equiv \mathcal{L} \left[\frac{f(t)}{t} \right] . \quad (9.102)$$

En définitive :

$$\int_z^\infty F(z') dz' = \mathcal{L} \left[\frac{f(t)}{t} \right] \iff \frac{f(t)}{t} = \mathcal{L}^{-1} \left[\int_z^\infty F(z') dz' \right] . \quad (9.103)$$

Ceci est la propriété duale de (9.98). Par exemple²², soit à trouver $\mathcal{L} \left[\frac{e^{-at} - e^{-bt}}{t} \right]$:

$$\mathcal{L} \left[\frac{e^{-at} - e^{-bt}}{t} \right] = \int_z^\infty \mathcal{L}[e^{-at} - e^{-bt}](z') dz' = \int_z^\infty \left[\frac{1}{z'+a} - \frac{1}{z'+b} \right] dz' , \quad (9.104)$$

d'où :

$$\boxed{\mathcal{L} \left[\frac{e^{-at} - e^{-bt}}{t} \right] = \ln \frac{z+b}{z+a}} \quad (9.105)$$

On retrouve sur le logarithme les singularités (qui sont des points de branchement) prévisibles au vu de l'expression de l'original, dont l'abscisse de sommabilité x_0 est visiblement $\max(-a, -b)$.

Enfin, on peut noter que :

$$\frac{d^m F}{dz^m} = \int_0^\infty (-t)^m f(t) e^{-zt} dt \iff t^m f(t) Y(t) = (-1)^m \mathcal{L}^{-1} \left[\frac{d^m F}{dz^m} \right] . \quad (9.106)$$

■ **Convolution** La convolution de deux fonctions a été définie au chapitre 8 ; toujours notée $f * g$, elle s'obtient par l'opération suivante :

$$(f * g)(t) = \int_{-\infty}^{+\infty} f(t') g(t - t') dt' . \quad (9.107)$$

²²Ces intégrales appartiennent à l'espèce dite *intégrales de Frullani*.

Ici, toutes les fonctions sont nulles pour $t < 0$, de sorte que l'intégrale de (9.107) s'étend de fait de 0 à t seulement :

$$(f * g)(t) = \int_0^t f(t')g(t-t') dt' . \tag{9.108}$$

La transformée de $f * g$ est :

$$\mathcal{L}[f * g](z) = \int_0^{+\infty} e^{-zt} dt \int_0^t f(t')g(t-t') dt' . \tag{9.109}$$

L'interversion de l'ordre d'intégration donne :

$$\begin{aligned} \mathcal{L}[f * g](z) &= \int_0^{+\infty} f(t') dt' \int_{t'}^{+\infty} g(t-t')e^{-zt} dt = \\ &= \int_0^{+\infty} f(t')e^{-zt'} dt' \int_{t'}^{+\infty} g(t-t')e^{-z(t-t')} dt , \end{aligned} \tag{9.110}$$

soit :

$$\mathcal{L}[f * g](z) = \int_0^{+\infty} f(t')e^{-zt'} dt' \int_0^{+\infty} g(t'')e^{-zt''} dt'' , \tag{9.111}$$

d'où²³ :

$$\boxed{\mathcal{L}[f * g](z) = F(z) G(z)} \tag{9.112}$$

Enfin, on peut aussi montrer que la transformée de Laplace du produit de deux fonctions est la convolution de leurs transformées :

$$\mathcal{L}[fg](z) \stackrel{\text{déf}}{=} \int_0^{+\infty} f(t)g(t) e^{-zt} dt = \frac{1}{2i\pi} \int_{a-i\infty}^{a+i\infty} F(z')G(z-z') dz' ; \tag{9.113}$$

Si x_0 et x'_0 désignent les abscisses de sommabilité de f et g respectivement, on a $a > x_0$ et $\Re z > x'_0 + a$. $\mathcal{L}[fg](z)$ est de fait holomorphe dans le demi-plan $\Re z > x_0 + x'_0$. Schématiquement, (9.113) résulte de la suite d'égalités :

$$\begin{aligned} \mathcal{L}[fg](z) &= \int_0^{+\infty} \int_{B'} \frac{dz'}{2i\pi} e^{z't} F(z') \int_{B''} \frac{dz''}{2i\pi} e^{z''t} G(z'') e^{-zt} dt = \\ &= \frac{1}{(2i\pi)^2} \int_{B'} dz' \int_{B''} dz'' \frac{F(z')G(z'')}{z-z'-z''} , \end{aligned} \tag{9.114}$$

la dernière écriture venant de l'intégration sur la variable t . Maintenant, on effectue l'intégration en z'' en refermant par un demi-cercle à droite : F et G étant analytiques dans le demi-plan de droite, la seule singularité est le pôle simple en $z''_0 = z - z'$; le résidu est $-F(z')G(z-z')$, mais comme on tourne dans le sens négatif, c'est $(-2i\pi)$ qui apparaît dans l'expression du théorème des résidus, d'où (9.113).

D'une façon générale, il existe un grand nombre de "petits" théorèmes fort utiles dans les applications. Citons à titre d'exemple celui d'Efros, qui est une généralisation du théorème de convolution ([9], § 81). Soit $F(z) = \mathcal{L}[f](z)$ et $g(t; \tau)$ dont l'image par Laplace $G(z; \tau)$ est de la forme $\Phi(z) e^{-\tau\phi(z)}$; alors, le produit $F[\phi(z)]\Phi(z)$ est la transformée de Laplace de l'intégrale $\int_0^{+\infty} f(\tau)g(t; \tau) d\tau$.

²³Ce résultat est dû à Émile Borel.

9.4 Propriétés asymptotiques

On se doute intuitivement que, tout comme pour la transformation de Fourier, il existe une relation étroite entre le comportement de $f(t)$ à l'infini et celui de $F(z)$ aux petits z . Ceci est d'ailleurs évident au vu de la définition (9.2) : si t devient très grand, l'intégrand est quasi nul sauf quand z est très petit, typiquement $|z| \lesssim t^{-1}$. Autrement dit, les grandes valeurs de t filtrent les petites valeurs de $|z|$. Montrons ceci en commençant par démontrer un résultat utile en pratique :

Si une fonction $f(t)$ admet une limite pour $t \rightarrow +\infty$, alors on a l'égalité :

$$f(+\infty) = \lim_{z \rightarrow 0} [zF(z)] , \quad (9.115)$$

où $F(z) = \mathcal{L}[f](z)$.

L'existence de la limite entraîne que la fonction $f(t)$ est bornée ; la plus petite abscisse de sommabilité est donc $x_0 = 0$, de sorte que $F(z)$ est holomorphe dans $\Re z > 0$, ainsi que $zF(z)$. L'existence de la limite assure aussi que l'intégrale $\int_0^{+\infty} f'(t) dt$ existe puisqu'elle est égale à $f(+\infty) - f(0)$. Par ailleurs, on a toujours $\int_0^{+\infty} e^{-zt} f'(t) dt = zF(z) - f(0)$, mais comme cette intégrale converge uniformément en z , on peut passer à la limite $z \rightarrow 0$ sous l'intégrale pour obtenir :

$$\lim_{z \rightarrow 0} \int_0^{+\infty} e^{-zt} f'(t) dt = \int_0^{+\infty} f'(t) dt = f(+\infty) - f(0) , \quad (9.116)$$

d'où finalement $\lim_{z \rightarrow 0} [zF(z) - f(0)] = f(+\infty) - f(0)$, et le résultat exprimé en (9.115).

D'une autre façon, plus formelle : soit une fonction $g(t)$ dont l'abscisse de sommabilité est $x_0 \leq 0$. Sa transformée de Laplace $G(z)$ est donc analytique dans $\{z : \Re z > 0\}$; l'intégrale de Laplace convergeant uniformément en z , on a :

$$\lim_{z \rightarrow 0} G(z) \stackrel{\text{déf}}{=} \lim_{z \rightarrow 0} \int_0^{+\infty} e^{-zt} g(t) dt = \int_0^{+\infty} \lim_{z \rightarrow 0} e^{-zt} g(t) dt = \int_0^{+\infty} g(t) dt , \quad (9.117)$$

ce que l'on peut écrire :

$$\lim_{z \rightarrow 0} G(z) = \int_0^{+\infty} \mathcal{L}^{-1}[G](t) dt , \quad (9.118)$$

d'une part. D'autre part, $zF(z) = \mathcal{L}[f'](z) + f(0) \iff \mathcal{L}^{-1}[zF(z)] = f'(t) + \mathcal{L}^{-1}[f(0)](t)$, d'où :

$$\lim_{z \rightarrow 0} zF(z) = \int_0^{+\infty} \mathcal{L}^{-1}[zF(z)](t) dt = \int_0^{+\infty} [f'(t) + \mathcal{L}^{-1}[f(0)](t)] dt . \quad (9.119)$$

$f(0)$ est une constante, donc d'après (9.44), $\mathcal{L}^{-1}[f(0)](t) = f(0)\delta^{(+)}(t)$. D'où :

$$\lim_{z \rightarrow 0} zF(z) = \int_0^{+\infty} [f'(t) + f(0)\delta^{(+)}(t)] dt , \quad (9.120)$$

soit finalement le résultat traduit par l'égalité (9.115). L'existence de la limite $f(+\infty)$ est manifestement une hypothèse essentielle, dont il faut toujours s'assurer *au préalable* : la fonction $F(z) \stackrel{\text{def}}{=} \mathcal{L}[\sin \omega t] = \frac{\omega}{z^2 + \omega^2}$ vérifie $\lim_{z \rightarrow 0} zF(z) = 0$, et pourtant $\sin \omega t$ n'a pas de limite à l'infini (tout au plus peut-on remarquer que $\lim_{z \rightarrow 0} zF(z)$ donne une sorte "moyenne" de $\sin \omega t$ (la plus petite abscisse de sommabilité de $\sin \omega t$ est $x_0 = 0$ puisque cette dernière fonction est bornée).

En pratique, l'usage de la transformation de Laplace est une méthode commode et puissante pour obtenir des informations précises sur une fonction $f(t)$ que l'on cherche à analyser. Le résultat ci-dessus permet de trouver la valeur de la limite quand on sait par ailleurs qu'elle existe. Il est également intéressant de savoir *comment* $f(t)$ approche sa limite (quand elle existe), ou plus généralement quel est le comportement de $f(t)$ aux grands temps, que cette fonction ait une limite ou pas. Le plus souvent, $F(z)$ est une fonction trop compliquée pour que l'on sache faire exactement l'inversion, mais il est néanmoins possible, en analysant finement $\mathcal{L}[f]$, d'obtenir le comportement asymptotique de $f(t)$.

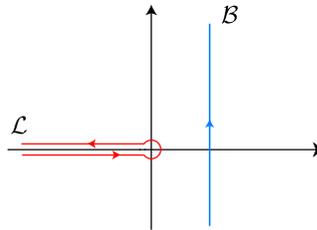


Figure 9.6: Modification de la ligne d'intégration pour exhiber le développement asymptotique de $f(t)$: on rabat la droite de Bromwich \mathcal{B} le long de la coupure, définissant le contour en lacet \mathcal{L} . Le théorème des résidus donne alors $f(t) - \mathcal{L}(t) = \sum_k \text{Res}[e^{zt}F(z), z_k]$.

La réponse est apportée par une suite de théorèmes appelés *théorèmes taubériens*, qui s'établissent tous par application de la méthode suivante ; on effectue le développement de $F(z)$ aux petits²⁴ z , et on inverse terme à terme. Par exemple, si on trouve²⁵ :

$$F(z) \simeq \sum_{\alpha > 0} a_\alpha z^\alpha, \quad z \rightarrow 0, \tag{9.121}$$

on peut alors montrer²⁶ que quand $t \rightarrow +\infty$:

$$f(t) \sim \sum_{\alpha > 0} a_\alpha \frac{t^{-\alpha-1}}{\Gamma(-\alpha)} \tag{9.122}$$

²⁴Comme toujours, si t désigne le temps, parler des "petits z " veut dire que vis-à-vis d'une autre échelle de temps τ préalablement définie, on a $|z|\tau \ll 1$.

²⁵*A priori* les α ne sont nullement tenus d'être entiers – d'ailleurs, voir juste après...

²⁶Noter aussi que (9.122) est une généralisation de (9.37) aux puissances positives de z .

Il importe de noter que le développement ainsi obtenu est en général *asymptotique* – d'où le symbole caractéristique $\sim -$, et que tout renseignement *a priori* sur les propriétés analytiques de f et/ou F est donc le bienvenu. En essence, le résultat (9.122) s'obtient en rabattant la droite de Bromwich le long de la coupure \mathbb{R}_- (voir fig. 9.6) et en remarquant que les éventuels pôles situés à gauche de l'axe imaginaire donnent des termes exponentiellement petits $\sim e^{-\gamma_k t}$, $\gamma_k > 0$, alors que le "lacet" \mathcal{L} associé aux deux bords de la coupure donnent des lois en puissance $\sim t^{-\nu}$. Une fois la droite \mathcal{B} rabattue en \mathcal{L} , chaque terme z^α donne la contribution écrite sous forme d'intégrales réelles ($\varepsilon > 0$) :

$$\int_{+\infty}^{-\varepsilon} d(\rho e^{-i\pi}) e^{-\rho t} \rho^\alpha e^{-i\alpha\pi} d\rho + \int_{+\varepsilon}^{+\infty} d(\rho e^{+i\pi}) e^{-\rho t} \rho^\alpha e^{+i\alpha\pi} d\rho + I_\varepsilon, \quad (9.123)$$

où I_ε est l'intégrale sur le cercle coupé de rayon ε autour de l'origine. Après $\lim_{\varepsilon \rightarrow 0}$ et réarrangement des termes, on trouve $-2i\pi \frac{\sin \alpha\pi}{\pi} \frac{\Gamma(\alpha+1)}{t^{\alpha+1}}$; compte tenu du facteur global $\frac{1}{2i\pi}$ dans la formule d'inversion, et de la formule des compléments (7.48), on obtient le développement (9.122). Noter que pour $t \gg (\inf_k \gamma_k)^{-1}$, les lois-puissance dominent les exponentielles, de sorte que les éventuels résidus (laissés sous le tapis) sont effectivement négligeables.

Par ailleurs, on note que ceci est *symboliquement* le résultat établi en (9.65), prolongé aux valeurs quelconques de α . Remarquer enfin que si le développement (9.121) ne contient que des puissances entières, le second membre de (9.122) est vide de sens puisque la fonction Γ n'est pas définie pour $n \in -\mathbb{N}$; ceci n'est pas surprenant puisque $\mathcal{L}^{-1}[z^m]$, $m \in \mathbb{N}$ donne la dérivée m^e de la "fonction" de Dirac...

La mise en œuvre de ce type d'analyse exige un peu de savoir-faire et de vigilance, comme le montre l'exemple suivant. Supposons que l'on obtienne d'une façon ou d'une autre la transformée de Laplace suivante :

$$F(z) = \frac{1}{1 + \Phi(z)}, \quad (9.124)$$

où $\Phi(z)$ est une fonction compliquée annihilant tout espoir de pouvoir effectuer exactement l'inversion de Laplace. En revanche, on suppose connu le comportement de Φ aux petits z ; admettons qu'il soit de la forme²⁷ :

$$\Phi(z) \simeq C z^\alpha \quad (\alpha > 0 \text{ non-entier}), \quad (9.125)$$

ce qui autorise à écrire :

$$F(z) \simeq \frac{1}{1 + C z^\alpha} \equiv F_{\text{ap}}(z), \quad z \rightarrow 0. \quad (9.126)$$

L'idée est d'approximer $f(t)$ aux grands temps par $\mathcal{L}^{-1}[F_{\text{ap}}] \stackrel{\text{déf}}{=} f_{\text{ap}}(t)$.

L'erreur consisterait à dire que, quand z est petit et si $\alpha > 0$, z^α peut être négligé devant 1, d'où l'on déduirait que $f(t) \sim \mathcal{L}^{-1}[1] = \delta^{(+)}(t)$... La règle *absolue*

²⁷Ne pas écrire $F(z) \simeq 1 - C z^\alpha$ au motif que z^α est petit ! Une telle écriture conduirait à des objets ressemblant à des dérivées fractionnaires de $\delta^{(+)}(t)$ et ne serait vraiment ici d'aucune utilité.

est de conserver *le terme multiforme dominant*. On s'en convainc en écrivant en détail l'expression de $f_{\text{ap}}(t) = \mathcal{L}^{-1}[F_{\text{ap}}]$ résultant d'une déformation continue de la droite de Bromwich²⁸ pour lui faire entourer la coupure de z^α , choisie le long du demi-axe réel négatif, de façon à écrire l'inversion de Laplace sous la forme d'une intégration portant sur une variable réelle, une écriture évitant l'apparition de canulars. Après un petit calcul ($z = xe^{\pm i\pi}$ sur les bords supérieure et inférieure de la coupure, $x > 0$), on trouve finalement que $f_{\text{ap}}(t)$ se met en effet sous la forme de l'intégrale *réelle* :

$$f_{\text{ap}}(t) = \frac{C}{\pi} \sin \alpha\pi \int_0^{+\infty} e^{-xt} \frac{x^\alpha}{1 + 2Cx^\alpha \cos \alpha\pi + C^2x^{2\alpha}} dx . \quad (9.127)$$

? Établir ce résultat.

Sur cette expression, on peut maintenant sans risque exploiter le fait que t est très grand : quand t est grand, l'exponentielle e^{-xt} "casse" l'intégrand pour $x \gtrsim t^{-1}$ (on dit que e^{-xt} effectue une "coupure" de l'intégrand), rendant *a priori* secondaire ce qui se passe loin de l'origine²⁹. On peut ainsi maintenant en toute sécurité faire un développement limité de l'intégrand pour les petits x et écrire :

$$f(t) \sim f_{\text{ap}}(t) \sim \frac{C}{\pi} \sin \alpha\pi \int_0^{+\infty} x^\alpha e^{-xt} [1 + \mathcal{O}(x^\alpha)] dx . \quad (9.128)$$

d'où :

$$f(t) \sim \frac{C}{\pi} \sin \alpha\pi \frac{\Gamma(\alpha + 1)}{t^{\alpha+1}} . \quad (9.129)$$

En utilisant la formule des compléments $\Gamma(-\alpha)\Gamma(1+\alpha) = -\frac{\pi}{\sin \alpha\pi}$, on obtient finalement :

$$\boxed{f(t) \sim \frac{C}{\Gamma(-\alpha)} t^{-(\alpha+1)} \quad (\alpha > 0, t \rightarrow +\infty)} \quad (9.130)$$

On retiendra que ce scénario est typique de l'apparition d'une relaxation en *loi-puissance* ; noter aussi que ce terme est bien nul pour les valeurs entières de α , auquel cas la fonction $F_{\text{ap}}(z)$ n'a pas de coupure et prend les mêmes valeurs de part et d'autre de l'axe réel négatif : l'expression (éventuellement *asymptotique*) de $f(t)$ dépend alors d'autres singularités de $F(z)$ non envisagées dans le calcul précédent, donnant précisément des contributions sous-dominantes quand α n'est pas entier.

Par ailleurs, quand $|z|$ augmente, il faut se poser tôt ou tard la question des termes omis dans l'expression approchée (9.125) de $\Phi(z)$: on pourrait tout à fait avoir quelque chose comme $\Phi(z) \simeq Cz^\alpha + Dz^2$ auquel cas l'analyse ci-dessus peut sembler

²⁸Ce faisant, il faut bien sûr contourner les autres éventuelles singularités, d'où l'importance d'en savoir le plus possible sur les propriétés analytiques de la vraie transformée $F(z)$! On suppose ici que $F(z)$ n'a pas d'autres singularités que la coupure – ou que tous ses pôles ont une partie réelle négative, auquel cas la conclusion qui suit n'est pas altérée, une fois éteints les transitoires correspondants.

²⁹Il convient toutefois de rester prudent, et de vérifier que le dénominateur dans (9.127) ne peut pas s'annuler sur \mathbb{R}_+ . Dans le cas contraire, l'intégrale exige une régularisation – se souvenir de la mésaventure de Vlasov (voir p. 337) !

incorrecte si $\alpha > 2$. À nouveau, on est tenté de penser que si $\alpha > 2$ le terme multiforme peut être omis car infiniment petit devant le terme en z^2 (au motif que $|z|^\alpha \ll |z|^2$ si z tend vers zéro et si $\alpha > 2$). C'est faux : le même type de calcul montre que les termes non-multiformes se compensent et que la puissance de t donnant le comportement asymptotique est inchangée ! Ce constat valide l'affirmation suivant laquelle il faut conserver le terme multiforme *dominant* (ce qui veut dire que l'on peut oublier z^β comparé à z^α si $\beta > \alpha$, α et β non-entiers), même si sa puissance (non-entière) est *supérieure* à d'autres termes en puissance entière (voir la sous-section 9.5.4 pour un exemple).

Le résultat intéressant est le suivant : la présence d'une puissance *quelconque* (i.e. non-entière) dans $F_{\text{ap}}(z)$ donne à $f(t)$ un déclin en loi-puissance avec un exposant non entier. Depuis une bonne vingtaine d'années, les comportements en loi-puissance sont massivement apparus en Physique (notamment à propos des systèmes fractals caractérisés par une certaine invariance d'échelle, dans les phénomènes de diffusion anormale, pour décrire les relaxations lentes, ...). Un déclin en loi-puissance donne en général une physique qualitativement différente d'un déclin exponentiel banal : pour ce dernier, on peut sans ambiguïté définir un temps de relaxation, ce qui n'est pas possible pour une relaxation en loi-puissance quand l'exposant est petit ; à la limite, cet exposant peut être nul : c'est le cas lorsque $f(t) \propto \frac{1}{\ln(t/\tau)}$, la fonction $\frac{1}{\ln x}$ décroissant plus lentement que $x^{-\alpha}$ quel que soit $\alpha > 0$.

9.5 Quelques applications de la transformation de Laplace

9.5.1 Équations différentielles linéaires à coefficients constants

Il s'agit sans doute de l'application la plus élémentaire (et c'est d'ailleurs pour résoudre de telles équations que la transformation de Laplace – ou ses avatars – a été inventée). Soit l'équation linéaire :

$$\sum_{p=0}^n a_p \frac{d^p f}{dt^p} = \phi(t) , \quad f(t < 0) = 0 , \quad (9.131)$$

complétée par la donnée d'autant de conditions initiales que nécessaire, par exemple les n nombres $f(0), f'(0), \dots, f^{(n-1)}(0)$. En prenant la transformée de Laplace, compte tenu de (9.85), il vient :

$$a_0 F(z) + \sum_{p=1}^n a_p \left[z^p F(z) - \sum_{r=0}^{p-1} z^{p-r-1} f^{(r)}(0) \right] = \Phi(z) , \quad (9.132)$$

d'où :

$$F(z) = \frac{1}{Z(z)} \left[\Phi(z) + \sum_{p=1}^n a_p \sum_{r=0}^{p-1} z^{p-r-1} f^{(r)}(0) \right] , \quad Z(z) \stackrel{\text{déf}}{=} \sum_{p'=0}^n a_{p'} z^{p'} . \quad (9.133)$$

Sur cette expression, on retrouve bien le fait que la solution s'obtient comme la somme d'une combinaison linéaire (dépendant des n conditions initiales) représentant la solution générale de l'équation homogène, et d'une solution particulière de l'équation complète qui contient la source $\phi(t)$. L'expression formelle de $f(t)$ est donc :

$$f(t) = \sum_{r=0}^{n-1} f^{(r)}(0) \sum_{p=r+1}^n a_p \frac{1}{2i\pi} \int_{\mathcal{B}} \frac{z^{p-r-1}}{Z(z)} e^{zt} dz + \frac{1}{2i\pi} \int_{\mathcal{B}} \frac{\Phi(z)}{Z(z)} e^{zt} dz \equiv f_h(t) + f_{\text{source}}(t) . \quad (9.134)$$

La fonction au dénominateur, $Z(z)$, s'appelle³⁰ *fonction de transfert* ; sauf coïncidence accidentelle³¹, ses n zéros z_k donnent des pôles pour $F(z)$. Par le seul fait de la présence de $Z(z)$ – et sans préjuger d'autres contributions venant de singularités du crochet dans (9.133) –, l'expression de $f(t)$ contiendra donc une combinaison linéaire d'exponentielles $e^{z_k t}$. Si toutes les parties réelles des z_k sont négatives, ces termes tendent vers zéro quand t tend vers l'infini et représentent les régimes transitoires qui disparaissent rapidement. Les autres résidus (ceux venant des singularités de $\Phi(z)$) décrivent le régime forcé. Il est clair que le théorème de Rouché (chapitre 6, p. 279) est d'une grande utilité pour une telle analyse.

Il est possible d'expliciter le terme homogène $f_h(t)$ de la solution (9.134). Si $Z(z)$ n'a que des zéros simples, notés $z_k = \gamma_k + i\omega_k$, on peut écrire (par le théorème fondamental de l'algèbre) $Z(z) = a_n \prod_{k=1}^n (z - z_k)$, où les n zéros z_k sont tous différents – et apparaissent par paires conjuguées si les coefficients a_p sont tous réels. Alors, le théorème des résidus³² permet d'écrire :

$$f_h(t) = \sum_{k=1}^n \left[\sum_{r=0}^{n-1} f^{(r)}(0) \sum_{p=r+1}^n a_p \frac{z_k^{p-r-1}}{a_n \prod_{l=1, l \neq k}^n (z_k - z_l)} e^{z_k t} \right] = \sum_{k=1}^n \frac{e^{z_k t}}{Z'(z_k)} \sum_{r=0}^{n-1} f^{(r)}(0) \sum_{p=r+1}^n a_p z_k^{p-r-1} . \quad (9.136)$$

Au total, $f_h(t)$ est de la forme³³ :

$$f_h(t) = \sum_{k=1}^n A_k \frac{e^{z_k t}}{Z'(z_k)} , \quad (9.137)$$

où les A_k incorporent les n valeurs initiales $\{f^{(m)}(0)\}_{0 \leq m \leq n-1}$. On retrouve bien que la solution homogène est une combinaison linéaire d'exponentielles associées aux zéros du polynôme caractéristique, paramétrée par les conditions initiales.

³⁰Attention, la terminologie est variable d'un auteur à l'autre : c'est parfois $\frac{1}{Z(z)}$ qui porte ce nom.

³¹avec des zéros des numérateurs. Notons à ce propos un effet intéressant : si la source possède un zéro commun z_0 avec la fonction de transfert, z_0 devient une singularité éliminable dans le terme de source. Un choix judicieux de la source permet ainsi d'éliminer des réponses indésirables.

³²ou l'application systématique de :

$$\mathcal{L}^{-1} \left[\frac{1}{z - z_k} \right] = e^{z_k t} . \quad (9.135)$$

Le dénominateur dans (9.136) n'est autre que la dérivée de $Z(z)$ calculée au pôle, soit $Z'(z_k)$.

³³Toujours dans l'hypothèse où tous les zéros de $Z(z)$ sont simples.

Pour un système physique *stable*³⁴, on a $\Re z_k = \gamma_k \leq 0$. Si tous les γ_k sont strictement négatifs, le terme homogène disparaît de fait pour $t \gg \frac{1}{\min_k |\gamma_k|}$; au-delà de cet instant, l'oubli des conditions initiales est quasi-total et $f(t)$ se réduit en pratique à $f_{\text{source}}(t)$: c'est le régime forcé.

Si un z_k est imaginaire pur³⁵, par exemple $z_{k_0} = i\omega_{k_0}$, alors le terme homogène ne tend pas vers zéro ; après extinction des autres termes (transitoires, eux, puisqu'ayant $\gamma_k \neq 0$), et pour des $a_p \in \mathbb{R} \forall p$:

$$f_h(t) \simeq 2\Re \frac{e^{i\omega_{k_0}t}}{Z'(i\omega_{k_0}t)} \sum_{r=0}^{n-1} f^{(r)}(0) \sum_{p=r+1}^n a_p (i\omega_{k_0}t)^{p-r-1}, \quad t \gg \frac{1}{\min_{k \neq k_0} |\gamma_k|}. \quad (9.138)$$

Ainsi, seules subsistent aux grands temps des oscillations non-amorties relatives au(x) pôle(s) imaginaire(s) pur(s) de $\frac{1}{Z(z)}$.

Clairement, la fonction $\hat{\chi}(z) = \frac{1}{Z(z)}$ joue un rôle essentiel ; toutes les autres quantités apparaissant dans l'expression (9.134) de la solution dépendent de la *préparation* du système (les conditions initiales) et de la source externe $\phi(t)$ qui lui est imposée : on peut donc dire que $\hat{\chi}$ décrit la dynamique propre irréductible du système étudié³⁶.

Le terme de source est de la forme :

$$f_{\text{source}}(t) = \frac{1}{2i\pi} \int_{\mathcal{B}} \hat{\chi}(z) \Phi(z) e^{zt} dz, \quad \hat{\chi}(z) = \frac{1}{Z(z)}. \quad (9.139)$$

Cette expression montre que $\mathcal{L}[f_{\text{source}}](z) = \hat{\chi}(z) \Phi(z)$, donc que $f_{\text{source}}(t)$ est la convolution $(\mathcal{L}^{-1}[\hat{\chi}]) * \phi$. On peut donc écrire :

$$f_{\text{source}}(t) = \int_0^t \chi(t-t') \phi(t') dt'. \quad (9.140)$$

où la fonction $\chi = \mathcal{L}^{-1}[\hat{\chi}]$ est de ce fait égale à :

$$\chi(t) = \mathcal{L}^{-1} \left[\frac{1}{\sum_{p=0}^n a_p z^p} \right] \equiv \mathcal{L}^{-1} \left[\frac{1}{Z(z)} \right]. \quad (9.141)$$

Faisant à nouveau l'hypothèse que les zéros z_k de $Z(z)$ sont tous différents les uns des autres (zéros simples), il vient :

$$\chi(t) = Y(t) \sum_{k=1}^n \frac{e^{z_k t}}{Z'(z_k)} \equiv Y(t) \sum_{k=1}^n \Gamma_k e^{z_k t}, \quad Z(z_k) = 0. \quad (9.142)$$

³⁴Par là, on entend ici qu'aucune divergence de nature exponentielle ne peut se produire. Alors l'abscisse de sommabilité est non-positive ; tous les pôles de la fonction méromorphe $\frac{1}{Z(z)}$ ont une partie réelle négative ou nulle.

³⁵Si les coefficients a_p sont tous réels, les zéros imaginaires purs arrivent aussi par paires complexes conjuguées.

³⁶Pour l'oscillateur harmonique $\ddot{x} + \gamma\dot{x} + \omega_0^2 = \frac{1}{m}f(t)$, on a $X(z) = [Z(z)]^{-1} [zx(0) + \dot{x}(0) + \frac{1}{m}\Phi(z)]$ avec $Z(z) = z^2 + \gamma z + \omega_0^2$; $[Z(z = -i\omega)]^{-1}$ est - au facteur $\frac{1}{m}$ près - la susceptibilité définie en (8.276).

◆ *Remarque*

Si deux z_k coïncident (on parle alors de *dégénérescence*), $z_{k_1} = z_{k_2}$, il existe un pôle double. Plutôt que de faire le calcul idoïne du résidu, le plus simple est prendre la limite dans la forme générale (9.142), en isolant deux z_{k_i} quelconques et en imaginant que l'un tend vers l'autre, par exemple $z_{k_2} \rightarrow z_{k_1}$. La somme au second membre de (9.142) contient donc les deux termes à analyser plus particulièrement :

$$\frac{e^{z_{k_1}t}}{(z_{k_1} - z_{k_2})Z'(z_{k_1})} + \frac{e^{z_{k_2}t}}{(z_{k_2} - z_{k_1})Z'(z_{k_2})}, \quad (9.143)$$

soit :

$$\frac{1}{z_{k_1} - z_{k_2}} \left[\frac{e^{z_{k_1}t}}{Z'(z_{k_1})} - \frac{e^{z_{k_2}t}}{Z'(z_{k_2})} \right], \quad (9.144)$$

Les deux dérivées Z' dans les dénominateurs deviennent égales quand $z_{k_2} \rightarrow z_{k_1}$; les autres termes se combinent en restituant formellement la dérivée de e^{zt} par rapport à z ; à la limite, on trouve pour ces deux termes :

$$\frac{e^{z_{k_1}t}}{Z'(z_{k_1})} \left[t - \frac{Z''(z_{k_1})}{Z'(z_{k_1})} \right]. \quad (9.145)$$

Un pôle double produit donc typiquement un terme en $t e^{z_k t}$; plus généralement, un pôle d'ordre n en z_k fournira des contributions $\propto t^{n-1} e^{z_k t}$. ◆

L'expression (9.140) est caractéristique des systèmes *linéaires* et exprime le fait que la *réponse*³⁷ du système à une source extérieure est caractérisée par une susceptibilité $\chi(t)$, laquelle ne dépend clairement que du système (c'est donc un attribut intrinsèque de ce système). On note également que cette expression respecte le Principe de causalité : la réponse (effet) à l'instant t ne dépend que des valeurs de la source (cause) aux instants *antérieurs*, puisque l'intégrale n'implique que les instants $t' \leq t$.

Pour un système dont tous les modes sont amortis ($\Re z_k = \gamma_k < 0 \ \forall k$), il est par ailleurs possible de définir la transformée de Fourier de $\chi(t)$, $\hat{\chi}_F$:

$$\hat{\chi}_F(\omega) = \int_{-\infty}^{+\infty} e^{i\omega t} \chi(t) dt \equiv \hat{\chi}(z = -i\omega). \quad (9.146)$$

D'après (9.142), on a :

$$\hat{\chi}_F(\omega) = \sum_{k=1}^n \frac{i\Gamma_k}{\omega - iz_k} = \sum_{k=1}^n \frac{i\Gamma_k}{\omega - (i\gamma_k - \omega_k)}. \quad (9.147)$$

Toutes les singularités de $\hat{\chi}_F(\omega)$ sont des pôles en $-\omega_k - i|\gamma_k|$ et sont donc toutes situées dans le demi-plan *inférieur*. L'expression de $\chi(t)$ par la transformation de Fourier inverse, $\chi(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\omega t} \hat{\chi}_F(\omega) d\omega$, montre que $\chi(t)$ est bien nulle $\forall t < 0$: en pareil cas, on

³⁷après extinction des transitoires.

referme le contour par un grand demi-cercle du côté $\Im\omega > 0$ et on ne ramasse aucune singularité ; par le théorème de Cauchy, l'intégrale donnant $\chi(t)$ est alors nulle³⁸. Remarque que pour une équation différentielle d'ordre fini n , la transformée de Fourier $\hat{\chi}_F(\omega)$ se décompose en au plus n éléments simples ; formellement, une équation différentielle d'ordre infini introduirait naturellement le développement de Mittag - Leffler de la fonction méromorphe $\hat{\chi}_F(\omega)$.

9.5.2 Équations aux différences finies

Une équation aux différences finies est une équation où figurent les valeurs d'une fonction inconnue en plusieurs points séparés d'une distance finie, alors qu'une équation différentielle ordinaire, par les dérivées qu'elle contient, n'implique que des valeurs distinctes infiniment proches d'une même fonction³⁹. Notamment, une relation de récurrence entre des quantités $f_m, f_{m+1}, f_{m+2}, \dots$ (m entier) relève de cette catégorie : en définissant une fonction $f(t)$ telle que $f(t) = f_m$ quand $m \leq t < m + 1$, la relation de récurrence implique les valeurs $f(t), f(t + 1), f(t + 2), \dots$ de la fonction en différents points.

De telles équations surviennent souvent en Physique, où il est parfois utile de discrétiser une variable continue. L'exemple-type est celui où on préfère travailler sur un réseau isomorphe à \mathbb{Z}^D , plutôt que de raisonner dans l'espace continu \mathbb{R}^D , quitte à revenir vers le *continuum* en bout de course, quand on analyse le problème à grande échelle ; c'est ainsi que la marche de l'ivrogne sur réseau (avec des déplacements vers les sites premiers voisins) reproduit l'équation classique de la diffusion quand on regarde le réseau de loin⁴⁰. Par ailleurs, la discrétisation intervient forcément dès que l'on doit résoudre un problème numériquement : les machines ne connaissent pas les infiniment petits !

Soit donc une relation de récurrence linéaire d'ordre n , impliquant $n + 1$ termes consécutifs $\forall m \geq 0$:

$$\sum_{p=0}^n a_p f_{m+p} = 0, \quad m = 0, 1, 2, \dots \quad (9.148)$$

d'une suite de quantités f_m définies pour $m \geq 0$, qu'il s'agit de trouver ; les coefficients a_p sont supposés indépendants de m , au même titre que les coefficients de (9.131) étaient

³⁸Ces considérations pourraient faire croire qu'il existe un lien direct entre stabilité et causalité : il n'en est rien. En fait, tous les systèmes physiques, qu'ils soient stables ou instables, obéissent au Principe de causalité. Le seul fait, pour l'argument, d'utiliser la transformation de Fourier élimine *de facto* les systèmes instables (e^{at} , $a > 0$, n'a pas de transformée de Fourier !) et ne sont retenus pour la suite que les systèmes qui à la fois sont stables et satisfont la causalité. Le lien apparent entre causalité et stabilité est donc illusoire et artificiel.

³⁹De ce point de vue, l'équation (9.86) est une équation qui est à la fois différentielle et aux différences finies.

⁴⁰Ce que l'on réalise en prenant formellement la limite nulle pour le pas du réseau.

supposés indépendants⁴¹ de la variable t . On voit tout de suite que l'on ne pourra résoudre cette équation que moyennant la donnée supplémentaire de n valeurs "initiales", par exemple les n valeurs f_q , $q = 0, 1, 2, \dots, n-1$: alors la valeur suivante f_n est déterminée par l'équation (9.148), et de proche en proche toutes les valeurs $f_{m>n}$ le sont aussi. Définissons maintenant la fonction $f(t)$:

$$f(t) \stackrel{\text{d\u00e9f}}{=} f_m \quad m \leq t < m+1 ; \quad (9.149)$$

alors, l'équation (9.148) devient :

$$\sum_{p=0}^n a_p f(t+p) = 0 \quad \forall t > 0 . \quad (9.150)$$

Pour transformer par Laplace une telle équation, il faut connaître $\mathcal{L}[f(t+p)]$ en fonction de $F(z) = \mathcal{L}[F](z)$. On a par exemple :

$$\begin{aligned} \mathcal{L}[f(t+1)](z) &= \int_0^{+\infty} e^{-zt} f(t+1) dt = e^z \int_0^{+\infty} e^{-z(t+1)} f(t+1) dt = \\ &= e^z \left[F(z) - \int_0^1 e^{-zt} f(t) dt \right] . \end{aligned} \quad (9.151)$$

On sait, par les conditions initiales, que $f(t) = f_0$ quand $0 \leq t < 1$; il vient donc :

$$\begin{aligned} \mathcal{L}[f(t+1)](z) &= e^z \left[F(z) - f_0 \int_0^1 e^{-zt} dt \right] = e^z \left[F(z) - f_0 \frac{1 - e^{-z}}{z} \right] = \\ &= e^z F(z) - f_0 \frac{e^z - 1}{z} . \end{aligned} \quad (9.152)$$

De la même façon, on trouve (f_1 est la valeur de $f(t)$ pour $1 \leq t < 2$, qu'il faut aussi se donner) :

$$\mathcal{L}[f(t+2)](z) = e^{2z} F(z) - f_0 \frac{e^z(e^z - 1)}{z} - f_1 \frac{e^z - 1}{z} , \quad (9.153)$$

et ainsi de suite. Au total, la transformée de l'équation (9.148) est de la forme :

$$\left[\sum_{p=0}^n a_p e^{pz} \right] F(z) = \Lambda(f_0, f_1, \dots, f_{n-1}; z) , \quad (9.154)$$

où $\Lambda(f_0, f_1, \dots, f_{n-1}; z)$ est une forme linéaire des n conditions "initiales" – les valeurs de $f(t)$ à la borne de gauche de chaque intervalle –, paramétrée par la variable z , que l'on sait trouver dans chaque cas particulier. La solution est donc :

$$F(z) = \frac{\Lambda(f_0, f_1, \dots, f_{n-1}; z)}{\sum_{p=0}^n a_p e^{pz}} , \quad (9.155)$$

⁴¹Dans un cas comme dans l'autre, le problème est d'une tout autre difficulté si les coefficients sont variables – en général, on ne sait pas alors trouver la solution sous une forme exploitable...

où $Z(e^z) = \sum_{p=0}^n a_p e^{pz}$ est un polynôme de degré n dans la variable e^z et joue le rôle d'une fonction de transfert discrète.

L'analyse de l'expression (9.155) procède suivant les mêmes lignes que précédemment, à propos des équations différentielles. Elle montre un résultat général intéressant : $f(m \leq t < m+1) \equiv f_m$ est une combinaison linéaire des $e^{z_k m} \equiv (e^{z_k})^m$, où $Z(e^{z_k}) = 0$: chaque f_m est donc de la forme $\sum_k c_k X_k^m$ où $Z(X_k) = 0$, les constantes c_k étant trouvées à partir des conditions initiales. Ce résultat est l'équivalent, pour les équations aux différences finies, du résultat retrouvé dans la sous-section 9.5.1 : pour une équation telle que (9.131), les solutions sont des combinaisons linéaires des $e^{z_k t}$ où les z_k sont les zéros⁴² de $Z(X) = \sum_{p=0}^n a_p X^p$.

Bien sûr, la méthode fonctionne aussi pour une équation inhomogène, du genre :

$$\sum_{p=0}^n a_p f_{m+p} = \phi_m, \quad m = 0, 1, 2, \dots, \quad (9.156)$$

en définissant la fonction $\phi(t) = \phi_m$ si $m \leq t < m+1$.

• **Exemple 1** Cette technique, appliquée à l'équation du second ordre ($n = 2$) :

$$f_{m+2} - 5f_{m+1} + 6f_m = 0, \quad f_0 = 0, f_1 = 1, \quad (9.157)$$

donne la solution $f_m = 3^m - 2^m$ (qui est bien une combinaison linéaire des zéros de $Z(X) = X^2 - 5X + 6$). Pour l'équation inhomogène :

$$f_{m+2} - 5f_{m+1} + 6f_m = 4^m, \quad (9.158)$$

on trouve $f_m = \frac{1}{2}(4^m - 2^m)$. •

• **Exemple 2** Traitons un autre exemple simple, montrant le lien entre équations aux différences finies et équations différentielles, et donnant l'occasion d'attirer l'attention sur une difficulté propre aux équations non-linéaires. Soit l'équation linéaire :

$$f_{m+1} = \alpha f_m, \quad f_0 \text{ donné}. \quad (9.159)$$

Bien évidemment, on voit immédiatement que la solution est $f_m = \alpha^m f_0$. L'exercice consiste à retrouver ce résultat en utilisant la transformation de Laplace.

Pour la clarté des grandeurs introduites, t désigne le temps ; on définit maintenant la fonction $f(t)$:

$$f(t) = f_m \quad m\Delta t \leq t < (m+1)\Delta t \quad (9.160)$$

où Δt est un intervalle de temps donné, ce qui permet de récrire (9.159) comme suit :

$$f(t + \Delta t) = \alpha f(t), \quad f_0 \text{ donné}. \quad (9.161)$$

⁴²On dit parfois que $Z(X) = 0$ est l'équation caractéristique de l'équation différentielle.

On a :

$$\begin{aligned} \mathcal{L}[f(t + \Delta t)](z) &= \int_0^{+\infty} e^{-zt} f(t + \Delta t) dt = e^{z\Delta t} \int_0^{+\infty} e^{-z(t+\Delta t)} f(t + \Delta t) dt = \\ &= e^{z\Delta t} \left[F(z) - \int_0^{\Delta t} e^{-zt} f(t) dt \right], \end{aligned} \quad (9.162)$$

soit :

$$\begin{aligned} \mathcal{L}[f(t + \Delta t)](z) &= e^{z\Delta t} \left[F(z) - f_0 \int_0^{\Delta t} e^{-zt} dt \right] = \\ &= e^{z\Delta t} \left[F(z) - f_0 \frac{1 - e^{-z\Delta t}}{z} \right] = e^{z\Delta t} F(z) - f_0 \frac{e^{z\Delta t} - 1}{z}. \end{aligned} \quad (9.163)$$

Au total, la transformation de Laplace sur (9.161) produit :

$$e^{z\Delta t} F(z) - f_0 \frac{e^{z\Delta t} - 1}{z} = \alpha F(z) \iff F(z) = f_0 \frac{1 - e^{-z\Delta t}}{z(1 - \alpha e^{-z\Delta t})}, \quad (9.164)$$

d'où :

$$f(t) = \mathcal{L}^{-1} \left[f_0 \frac{1 - e^{-z\Delta t}}{z(1 - \alpha e^{-z\Delta t})} \right]. \quad (9.165)$$

En développant en série la fraction rationnelle en $e^{-z\Delta t}$, il vient :

$$f(t) = f_0 \mathcal{L}^{-1} \left[\frac{1 - e^{-z\Delta t}}{z} \sum_{n=0}^{+\infty} \alpha^n e^{-nz\Delta t} \right]. \quad (9.166)$$

Utilisant $\mathcal{L}^{-1} \left[\frac{e^{-nz\Delta t}}{z} \right] = Y(t - n\Delta t)$, on trouve :

$$f(t) = f_0 \sum_{n=0}^{+\infty} \alpha^n [Y(t - n\Delta t) - Y(t - (n+1)\Delta t)]. \quad (9.167)$$

Ceci est une fonction en escalier, dont la marche située entre $n\Delta t$ et $(n+1)\Delta t$ est à la hauteur $\alpha^n f_0$. On peut donc aussi l'écrire (E désigne toujours la partie entière) :

$$f(t) = f_0 \alpha^{E\left(\frac{t}{\Delta t}\right)}. \quad (9.168)$$

En quelque sorte, $f(t)$ est une exponentielle discrétisée, décroissante si $\alpha < 1$, croissante dans le cas contraire.

Ceci peut aussi se voir en envisageant la limite continue en temps de l'équation aux différences, qui devient différentielle dans cette limite. Clairement, pour que ceci soit possible, il faut simultanément faire tendre Δt vers zéro et α vers 1, puisque si Δt devient très petit, $f(t + \Delta t)$ tend vers⁴³ $f(t)$. D'ailleurs, il est évident que la limite différentielle n'existe que si α est de la forme :

$$\alpha \rightarrow 1 + \gamma \Delta t \quad \Delta t \rightarrow 0; \quad (9.169)$$

⁴³On fait bien sûr l'hypothèse que f est une fonction continue.

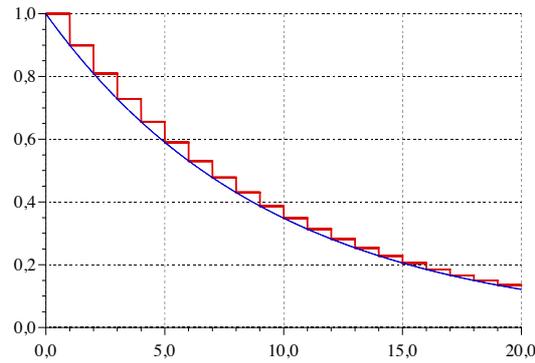


Figure 9.7: Variations comparées de la fonction (9.168) et de sa limite en temps continu (l'abscisse est $t/\Delta t$, $f_0 = 1$).

ceci assure que pour les petits Δt , (9.161) prend l'aspect d'une équation "pré-différentielle".

Effectuons donc la limite suivant cette procédure dans la solution (9.164). En développant l'exponentielle $e^{z\Delta t} \simeq 1 + z\Delta t + \mathcal{O}(t^{\geq 2})$ et tenant compte de (9.169), il vient :

$$F(z) \rightarrow f_0 \frac{z\Delta t}{z[z\Delta t - \gamma\Delta t]} = f_0 \frac{1}{z - \gamma} \equiv F_{\text{cont}}(z) . \quad (9.170)$$

La fonction $f_{\text{cont}}(t) = \mathcal{L}^{-1}[F_{\text{cont}}]$ est $f_0 e^{\gamma t}$, c'est bien la solution de la limite $\Delta t \rightarrow 0$ de l'équation aux différences (9.161), qui devient alors $f'(t) = \alpha f$. C'est aussi la limite de $\alpha^n = e^{n \ln(1 + \gamma\Delta t + \dots)}$ (voir (9.168)) dans la limite $n \rightarrow +\infty$, $n\Delta t = C^{\text{ste}} = t$, avec $\lim \alpha^n = \lim_{n \rightarrow +\infty} \left(1 + \gamma \frac{t}{n}\right)^n = e^{\gamma t}$. γ est ≥ 0 selon que $\alpha \geq 1$. La relation formelle étroite entre équations différentielles et équations aux différences se généralise aisément au-delà du premier ordre (voir chapitre 10, section 10.5).

On voit ainsi que les deux équations cousines, aux différences et différentielles, donnent essentiellement le même résultat, quand on regarde la solution de la première à grande échelle (vu de loin, le Δt est petit). Cette coïncidence est très rassurante, mais elle ne tient que pour les équations linéaires. Il est possible de fabriquer des équations non-linéaires très simples, dont les deux versions discrète et continue n'ont pas du tout les mêmes solutions (dans la version discrète, on assiste souvent à la génération spontanée de points dits singuliers [5], qui changent radicalement la nature des solutions⁴⁴) ; l'inverse peut aussi se produire : il arrive aussi que la solution discrète soit régulière partout, alors

⁴⁴On rencontrera des exemples au chapitre 10, section 10.5, mais on peut déjà en donner un. L'équation $f' = f^2 - f$, $f(0) = 2$, a pour solution $f(x) = \frac{2}{2 - e^x}$, qui diverge en $x_0 = \ln 2$. D'un autre côté, l'équation aux différences $a_{n+1} = a_n^2$ s'écrit aussi $a_{n+1} - a_n = a_n^2 - a_n$ et partage un cousinage évident avec l'équation différentielle. Pourtant, ses solutions sont $a_n = 2^{2^n}$ et sont finies quel que soit n !

que sa cousine continue a des comportements singuliers (voir l'exemple de la note 44). Ce fait peut légitimement déclencher quelque inquiétude, car les équations non-linéaires sont le plus souvent insolubles analytiquement : l'obligation de recourir à un calcul sur ordinateur (pour lequel la discrétisation est une nécessité incontournable) permet de suspecter d'avance de grandes difficultés pour ne pas dénaturer le problème posé et en trouver bel et bien les solutions. En la matière, un immense savoir-faire est clairement de rigueur, et il est essentiel de procéder à des analyses locales des solutions qui permettent de trancher dans le vif si la machine fournit des solutions délirantes. •

9.5.3 Équations aux dérivées partielles

Une équation aux dérivées partielles est une relation entre une fonction de plusieurs variables et ses dérivées. Un chapitre ultérieur du cours sera consacré à de telles équations ; il s'agit ici, à propos d'un exemple particulier mais très important, de montrer comment les transformations intégrales de Laplace et de Fourier permettent d'obtenir rapidement la solution.

L'exemple choisi est celui du propagateur d'une particule libre en Mécanique quantique, c'est-à-dire de l'objet mathématique qui permet de construire la fonction d'onde à tout instant, connaissant la fonction d'onde au départ, à l'instant t_0 pris comme origine des temps.

On a vu dans le cours de Mécanique quantique que la fonction d'onde $\Psi(x, t)$ pour une particule libre de masse m se déplaçant dans \mathbb{R} obéit à l'équation de Schrödinger⁴⁵ :

$$i\hbar \frac{\partial}{\partial t} \Psi(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi(x, t) \quad (x \in \mathbb{R}) , \quad (9.171)$$

où \hbar est la célébrissime constante de Planck divisée par 2π . C'est une équation (parabolique) du premier ordre en temps (et du second ordre en espace). Autrement dit, pour trouver $\Psi(x, t)$ à un instant quelconque, il faut se donner un état initial $\Psi(x, t_0)$. Cette équation est en un sens infiniment plus compliquée qu'une équation différentielle : on peut dire qu'elle nécessite en fait une *infinité* de conditions initiales, toutes les valeurs du champ Ψ au départ, contenues dans la donnée de la *fonction* $\Psi(x, t_0)$.

Il se révèle commode d'introduire le *propagateur*⁴⁶ $U(x, t; x_0, t_0)$ qui relie l'état à l'instant t à l'état initial selon :

$$\Psi(x, t) = \int U(x, t; x_0, t_0) \Psi(x_0, t_0) dx_0 , \quad (9.172)$$

⁴⁵La particule étant libre, il n'y a pas d'énergie potentielle $V(x)$ dans (9.171).

⁴⁶Une fois connu le propagateur pour un problème donné, le calcul de $\Psi(x, t)$ se réduit à une simple intégration avec l'état initial donné, conformément à (9.172).

La terminologie est fluctuante : le propagateur s'appelle aussi fonction de Green ; par ailleurs, le terme *propagateur* est aussi utilisé aussi dans une acception plus vaste : dans tous les cas, il s'agit d'une représentation ou d'une autre de ce qui permet de relier explicitement un "ancêtre" à sa descendance. Il doit être clair que la notion de propagateur n'est pas spécifiquement quantique : d'une façon générale, le propagateur est l'objet qui fait évoluer le système dans l'espace-temps.

cette relation tenant quel que soit $\Psi(x, t_0)$; le noyau $U(x, t; x_0, t_0)$ est en quelque sorte une “matrice continue”. Compte tenu de (9.171), le propagateur satisfait l'équation :

$$i\hbar \frac{\partial}{\partial t} U(x, t; x_0, t_0) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} U(x, t; x_0, t_0) \quad (9.173)$$

avec la condition initiale :

$$U(x, t = t_0; x_0, t_0) = \delta(x - x_0) , \quad (9.174)$$

exprimant que $U(x, t_0; x_0, t_0)$ est le noyau de l'opérateur identité $\mathbf{1}$ (entre t_0 et $t = t_0$, le système n'évolue et reste ce qu'il est). Physiquement, $U(x, t; x_0, t_0)$ ne dépend que de la *différence* des temps : on peut donc toujours prendre $t_0 = 0$; dans la suite, on pose :

$$U(x, t; x_0, t_0 = 0) \equiv U(x, t; x_0) . \quad (9.175)$$

Il est facile de voir directement à partir de (9.171) que l'intégrale de $|\Psi(x, t)|^2$ est une constante du mouvement (la probabilité totale se conserve, la particule ne “disparaît” pas, elle est toujours forcément quelque part), ce qui est assuré⁴⁷ par le fait que U est un opérateur unitaire satisfaisant :

$$U^{-1}(x, t; x_0) = U^*(x, t; x_0) = U(x, -t; x_0) . \quad (9.177)$$

Il suffit donc de déterminer une fonction U_+ égale à U si $t > 0$ et nulle si $t < 0$; le noyau U à $t < 0$ s'en déduit par (9.177). U_+ est appelé propagateur causal, ou avancé.

L'équation (9.173), écrite pour U_+ , se résout aisément en effectuant successivement la transformation de Laplace en t , puis celle de Fourier en x . On commence par poser⁴⁸ :

$$K(x, z; x_0) = \int_0^{+\infty} U_+(x, t; x_0) e^{-zt} dt \quad (\Re z > 0) . \quad (9.178)$$

La restriction $\Re z > 0$ est suffisante ; il est possible de s'en convaincre intuitivement en remarquant que U_+ étant un opérateur unitaire, toutes ses valeurs propres sont de module égal à 1 et sont bornées. L'intégrale (9.178) converge dès que z a une partie réelle finie, aussi petite soit-elle ; $K(x, z; x_0)$ n'a donc aucune singularité dans le demi-plan de droite $\Re z > 0$.

La transformée de Laplace de l'équation (9.173) pour $U_+(x, t; x_0)$ se construit suivant la règle (9.82) et s'écrit :

$$i\hbar [z K(x, z; x_0) - \delta(x - x_0)] = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} K(x, z; x_0) , \quad (9.179)$$

⁴⁷La conservation de la probabilité totale est assurée par le fait que de l'équation fondamentale (9.171) on peut *déduire* une équation de conservation locale entre une densité ρ et un courant \vec{j} :

$$\frac{\partial \rho}{\partial t} + \text{div} \vec{j} = 0 , \quad (9.176)$$

où $\rho = \Psi^* \Psi \equiv |\Psi|^2$ et $\vec{j} = \frac{\hbar}{2im} [\Psi^* \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^*]$.

⁴⁸Pour être tout à fait explicite, il conviendrait de noter K_+ cette transformée de Laplace. On se souviendra qu'elle est reliée au propagateur avancé U_+ .

de sorte que l'équation pour K reste différentielle en x . Réarrangeant les termes :

$$z K(x, z; x_0) - \frac{i\hbar}{2m} \frac{\partial^2}{\partial x^2} K(x, z; x_0) = \delta(x - x_0) . \quad (9.180)$$

Pour résoudre (9.180), il est maintenant commode de faire une transformation de Fourier en x et de poser :

$$K(x, z; x_0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathcal{U}(k, z; x_0) e^{-ikx} dk , \quad (9.181)$$

d'où, selon la formule d'inversion de Fourier :

$$\mathcal{U}(k, z; x_0) = \int_{-\infty}^{+\infty} K(x, z; x_0) e^{+ikx} dx . \quad (9.182)$$

La transformée de Fourier de $\delta(x - x_0)$ est e^{ikx_0} ; la transformée de Fourier de l'équation (9.180) est donc :

$$z \mathcal{U}(k, z; x_0) + \frac{i\hbar}{2m} k^2 \mathcal{U}(k, z; x_0) = e^{ikx_0} , \quad (9.183)$$

d'où l'on déduit immédiatement :

$$\mathcal{U}(k, z; x_0) = \frac{e^{ikx_0}}{z + \frac{i\hbar k^2}{2m}} . \quad (9.184)$$

Soit maintenant $\mathcal{K}(k, t; x_0)$ la transformée de Laplace inverse de $\mathcal{U}(k, z; x_0)$; comme $\mathcal{L}^{-1}(z + a)^{-1} = e^{-at}$:

$$\mathcal{K}(k, t; x_0) = e^{ikx_0} e^{-i \frac{\hbar k^2}{2m} t} . \quad (9.185)$$

On note au passage la relation :

$$[\mathcal{K}(k, t; x_0)]^* = \mathcal{K}(-k, -t; x_0) , \quad (9.186)$$

qui exprime la symétrie dans le renversement du temps : en renversant le temps et la vitesse ($\hbar k$ étant l'impulsion mv de la particule, inverser la vitesse c'est changer k en son opposé), on obtient le complexe conjugué de \mathcal{K} . Ceci est en accord avec le fait que si $\Psi(x, t)$ est solution de l'équation de Schrödinger, $\Psi(x, -t)^*$ l'est aussi. Pour avoir enfin $U_+(x, t; x_0)$ il suffit maintenant de prendre la transformée de Fourier inverse de \mathcal{K} ; compte tenu de (9.185) :

$$U_+(x, t; x_0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk e^{-ikx} e^{ikx_0} e^{-i \frac{\hbar k^2}{2m} t} . \quad (9.187)$$

Cette intégrale gaussienne se calcule aisément, par un changement de variable élémentaire. On trouve ainsi finalement :

$$U_+(x, t; x_0) = \left(\frac{m}{2i\pi\hbar t} \right)^{1/2} e^{i \frac{m(x-x_0)^2}{2t}} \equiv U(x, t; x_0) \quad \forall t > 0 \quad (9.188)$$

qui est donc le propagateur U pour $t > 0$. On reconnaît dans l'exponentielle donnant U_+ l'action classique $S_{\text{class}} = \frac{m(x-x_0)^2}{2t}$ de la particule libre partie de x_0 à $t = 0$ et arrivée en x à l'instant t – un fait pour le moins surprenant à première vue, qui est le point de départ de la formulation de Feynman de la Mécanique quantique dans le formalisme de l'intégrale de chemins (voir aussi la remarque de la note 50 p. 446).

Pour avoir U pour $t < 0$, il suffit d'utiliser l'unitarité de U (voir (9.177)). Il vient ainsi :

$$U(x, t < 0; x_0) = [U_+(x, -t; x_0)]^* = \left[\left(\frac{m}{2i\pi\hbar t} \right)^{1/2} e^{\frac{i}{\hbar} \frac{m(x-x_0)^2}{2t}} \right]^* . \quad (9.189)$$

La branche de la racine carrée est celle qui prend des valeurs réelles positives sur l'axe réel positif ; il en résulte que⁴⁹ :

$$(z^{1/2})^* = (z^*)^{1/2} \quad (9.190)$$

et finalement :

$$U(x, t < 0; x_0) = \left(\frac{m}{2(-i)\pi\hbar(-t)} \right)^{1/2} e^{-\frac{i}{\hbar} \frac{m(x-x_0)^2}{2(-t)}} = \left(\frac{m}{2i\pi\hbar t} \right)^{1/2} e^{\frac{i}{\hbar} \frac{m(x-x_0)^2}{2t}} . \quad (9.191)$$

Ainsi, l'expression de U est la même, que t soit positif ou négatif⁵⁰ ; rétablissant t_0 ⁵¹ :

$$U(x, t; x_0, t_0) = \left(\frac{m}{2i\pi\hbar(t-t_0)} \right)^{\frac{1}{2}} e^{\frac{i}{\hbar} \frac{m(x-x_0)^2}{2\hbar(t-t_0)}} \equiv \left(\frac{m}{2i\pi\hbar(t-t_0)} \right)^{\frac{1}{2}} e^{\frac{i}{\hbar} S_{\text{class}}(x, x_0; t, t_0)} \quad (9.193)$$

cette expression étant valable quel que soit t , positif ou négatif. Il n'est pas difficile – en se souvenant de l'expression du précurseur gaussien de la fonction de Dirac (et en effectuant le prolongement analytique) –, de se convaincre que $\lim_{t \rightarrow t_0} U(x, t; x_0, t_0) = \delta(x - x_0)$, comme il se doit (un bon exemple d'ailleurs où la méthode de la phase stationnaire se confond avec la méthode du col).

Ce résultat fournit aussi le propagateur pour l'équation de diffusion libre d'une particule *classique* ($\hbar = 0$). En effet, en conséquence de la loi de Fick affirmant que le

⁴⁹C'est un nouvel exemple montrant comment une prescription physique (ici l'unitarité de U , elle-même imposée par la nécessaire conservation de la norme de Ψ) verrouille la coupure d'une fonction multiforme.

⁵⁰Indépendamment du préfacteur contenant la constante de Planck, il est remarquable que la fonction action *classique* S_{class} apparaisse aussi simplement dans l'expression du propagateur. Il n'en va pas toujours ainsi, mais c'est vrai pour tous les Hamiltoniens au plus quadratiques en \vec{p} et \vec{r} (par exemple : oscillateur harmonique, particule accélérée par un champ constant). Ce fait, très surprenant de prime abord, est le point de départ de la formulation de Feynman en intégrale fonctionnelle de la Mécanique quantique.

⁵¹La généralisation à trois dimensions d'espace est immédiate :

$$U(\vec{r}, t; \vec{r}_0, t_0) = \left(\frac{m}{2i\pi\hbar(t-t_0)} \right)^{\frac{3}{2}} e^{\frac{i}{\hbar} \frac{m(\vec{r}-\vec{r}_0)^2}{2\hbar(t-t_0)}} . \quad (9.192)$$

courant est proportionnel au gradient de la densité, $j = -D \frac{\partial P}{\partial x}$, l'équation de conservation $\frac{\partial P}{\partial t} + \text{div } j = 0$ donne l'équation de diffusion (en une dimension d'espace⁵²) :

$$\frac{\partial}{\partial t} P(x, t) = D \frac{\partial^2}{\partial x^2} P(x, t) , \quad (9.194)$$

où $P(x, t)$ et la densité de probabilité (c'est cette fois une quantité positive ou nulle, alors que $\Psi(x, t)$, *essentiellement* à valeurs complexes, est une *amplitude de probabilité*). Les deux équations (9.171) et (9.194) sont formellement identiques, on passe de l'une à l'autre par la substitution :

$$\frac{i\hbar}{2m} \rightarrow D \iff \frac{m}{\hbar} \rightarrow \frac{i}{2D} . \quad (9.195)$$

Il en résulte que le propagateur pour l'équation de la diffusion, usuellement noté W , s'obtient de l'expression (9.193) en y faisant la substitution ci-dessus et vaut donc :

$$W(x, t; x_0, t_0) = \frac{1}{\sqrt{4\pi D(t-t_0)}} e^{-\frac{(x-x_0)^2}{4D(t-t_0)}} . \quad (9.196)$$

Il est intéressant de noter que les deux équations (9.171) et (9.194), quoique formellement identiques, produisent des comportements en temps radicalement différents. L'équation de Schrödinger est invariante par renversement du temps⁵³, alors que l'équation de la diffusion donne une évolution irréversible (la tache d'encre à la surface de l'eau s'étale toujours !). La différence tient à l'absence ou à la présence du nombre imaginaire fondamental i , un détail (!) qui a donc son importance...

Noter que $W(x, t; x_0, t_0)$ est plus précisément une (densité de) probabilité *conditionnelle* ; comme on a choisi la condition initiale $\delta(x-x_0)$, $W(x, t; x_0, t_0)$ est la (densité de) probabilité de trouver la particule en x à l'instant t sachant qu'elle était en x_0 à $t = t_0$ avec probabilité 1 (certitude).

9.5.4 Équations intégro-différentielles

La transformation de Laplace s'impose pour les équations contenant à la fois des opérateurs différentiels et intégraux, quand ceux-ci se présentent sous forme de convolutions. Typiquement, ces équations sont de la forme :

$$\mathbf{L}[f(t)] = \phi(t) + \int_{\mathbb{R}} K(t-t')f(t') dt' , \quad (9.197)$$

⁵²À nouveau, la généralisation à \mathbb{R}^3 est immédiate : la loi de Fick s'écrit alors $\vec{j} = -\vec{\nabla}P$ et l'équation de conservation est $\frac{\partial P}{\partial t} + \text{div } \vec{j} = 0$. L'équation de la diffusion qui en résulte est alors $\frac{\partial}{\partial t} P(\vec{r}, t) = D\Delta P(\vec{r}, t)$ où Δ est le Laplacien.

⁵³On sait qu'un *paquet d'ondes* non confiné (pas de murs, pas de potentiel attractif) s'étale toujours, mais c'est simplement parce que la relation de dispersion quantique n'est pas linéaire (au contraire du champ électromagnétique), de sorte que les différentes composantes n'avancent pas à la même vitesse (la vitesse de groupe n'est pas égale à la vitesse de phase).

où \mathbf{L} est un opérateur différentiel linéaire à coefficients constants, $f(t)$ la fonction cherchée et $\phi(t)$ une source. L'image par Laplace du premier membre produit un polynôme en z qui multiplie $F(z) \equiv \mathcal{L}[f]$ et une combinaison linéaire des conditions initiales. Quant au second membre, par le théorème de convolution (9.112), sa transformée de Laplace est simplement $\Phi(z) + \mathcal{K}(z)F(z)$, dans des notations évidentes ; au total, la solution formelle de (9.197) s'écrit à vue, et le travail s'achève en effectuant la transformation inverse.

À titre d'exemple physique, considérons le mouvement d'une grosse particule⁵⁴ dans un fluide (mouvement Brownien) : en raison des multiples impacts incessants des petites particules du fluide sur la grosse particule, celle-ci effectue un mouvement en tous sens, que l'on peut comprendre comme le résultat de l'action sur la grosse particule d'une force $\mathcal{F}(t)$ contenant une composante aléatoire. En réalité, le fluide sert à la fois de moteur et de frein : s'il est le siège de la force erratique qui pousse la grosse particule tantôt dans un sens, tantôt dans l'autre, le mouvement de celle-ci est freiné par le fluide lui-même. Au total, désignant par $v_p(t)$ la vitesse de la grosse particule, une bonne équation (phénoménologique) simulant bien l'ensemble des effets physiques indissociables est :

$$\boxed{m \frac{dv_p}{dt} = -\alpha v_p + F(t)} \quad (9.198)$$

où le premier terme au second membre est la force de freinage dans l'hypothèse du frottement fluide et où $F(t)$ est la force dite *de Langevin*, variant aléatoirement très vite (comparée à l'échelle de temps τ_R introduite plus loin), et de moyenne nulle. (9.198) porte le nom d'*équation de Langevin*. Si on prend la moyenne de (9.198) sur un ensemble statistique, on obtient :

$$m \frac{dv}{dt} = -\alpha v \quad (v \stackrel{\text{déf}}{=} \langle v_p \rangle) , \quad (9.199)$$

les $\langle \dots \rangle$ représentant la moyenne statistique. Cette équation s'intègre à vue et donne $v(t) = v(0)e^{-\gamma t}$, $\gamma = \frac{\alpha}{m}$: si on donne une impulsion au départ à la particule, son mouvement (en moyenne) s'amortit au bout du temps⁵⁵ $\tau_R = \frac{m}{\alpha}$.

Le terme de frottement dans (9.198), $-\alpha v$, est un terme instantané. Il est parfois nécessaire d'introduire un terme de frottement *retardé*, par exemple pour traduire certains effets collectifs qui induisent une *mémoire*, ou pour incorporer la réaction de la grosse particule sur le bain. Alors, l'équation de Langevin (9.198) se généralise en ($\gamma = \frac{\alpha}{m}$) :

$$m \frac{dv_p}{dt} = -\alpha \int_{t_0}^t \omega(t-t')v_p(t')dt' + F(t) \iff \frac{dv_p}{dt} = -\gamma \int_{t_0}^t \omega(t-t')v_p(t')dt' + \frac{1}{m}F(t) ; \quad (9.200)$$

la mémoire est décrite par $\omega(t-t')$ (*noyau*), qui tend vers zéro à l'infini (amnésie totale aux temps infinis). t_0 est le temps de départ du mouvement (on peut toujours le choisir comme origine, $t_0 = 0$). L'intégrale n'implique que des instants $t' < t$, en respect du Principe de causalité.

⁵⁴Tout est relatif : la grosse particule a typiquement un diamètre de l'ordre du micron, qui est gigantesque relativement à la taille (atomique) des particules du fluide.

⁵⁵Le temps de relaxation τ_R est d'autant plus long que le frottement est faible et la masse élevée.

Si $\omega(t-t')$ est une fonction très piquée (mémoire courte), aucun changement qualitatif important n'est à attendre par rapport au frottement instantané. Techniquement, il suffit d'ailleurs d'utiliser la formule de la moyenne pour écrire l'intégrale sous la forme $v_p(t + \delta t) \int_{t_0}^t \omega(t-t') dt' \propto v_p(t + \delta t)$; si le noyau $\omega(t-t')$ coupe très vite, δt est de fait très petit, et on retrouve la forme avec freinage local en temps à peu de chose près. En revanche, quand $\omega(t)$ décroît lentement, des changements *qualitatifs* sont possibles, surtout quand $\omega(t)$ est positive à tout temps (ce que l'on supposera) : alors, si la vitesse devient négative, le second membre peut être positif, donnant une accélération *positive* (effet de sillage⁵⁶ dans le bain).

La moyenne statistique de (9.200) donne :

$$\frac{dv}{dt} = -\gamma \int_{t_0}^t \omega(t-t') v(t') dt' , \quad (9.201)$$

dont la transformée par Laplace est :

$$zV(z) - v_0 = -\gamma \Omega(z)V(z) , \quad (9.202)$$

d'où la transformée de Laplace de la solution $V(z) \equiv \mathcal{L}[v(t)]$:

$$V(z) = \frac{v_0}{z + \gamma \Omega(z)} . \quad (9.203)$$

La solution $v(t)$ s'obtient par $v(t) = \mathcal{L}^{-1}[V(z)]$, et tout dépend de la variation de Ω en z , c'est-à-dire, par dualité, de la forme précise du noyau de mémoire $\omega(t-t')$.

Supposons d'abord que ω est une fonction positive décroissant vite, sur une échelle de temps bien caractérisée τ_m . Alors, par la relation (9.41), $\Omega(z)$ est une fonction lentement variable sur l'échelle $\frac{1}{\tau_m}$, et $v(t)$ a essentiellement un comportement exponentiel aux grands temps : pour $|z|\tau_m \ll 1$, on a $V(z) \simeq \frac{v_0}{z + \gamma \Omega(0)}$, soit $v(t) \simeq e^{-\Omega(0)\gamma t}$ pour $t \gg \tau_m$. Précisons les choses en prenant :

$$\omega(t) = \frac{1}{\tau_m} e^{-\frac{t}{\tau_m}} ; \quad (9.204)$$

alors $\Omega(z) = \frac{1}{\tau_m} \frac{1}{z + \tau_m^{-1}}$, et :

$$V(z) = \frac{v_0}{z + \gamma \frac{1}{z\tau_m + 1}} . \quad (9.205)$$

Par inversion \mathcal{L}^{-1} , il est facile d'obtenir l'expression exacte de $v(t)$ (laissé à titre d'exercice. On voit sans calcul que pour $|z|\tau_m \ll 1$, $V(z) \simeq \frac{v_0}{z + \gamma}$, de sorte que pour les temps $t \gg \tau_m$, et pour $\gamma\tau_m \ll 1$, on a $v(t) \simeq v_0 e^{-\gamma t}$: la mémoire finie n'introduit de changement notable que pour $t \lesssim \tau_m$. Noter que si l'on prend $\tau_m = 0+$, $\omega(t) = \delta^{(+)}(t)$, et on a exactement $v(t) = v_0 e^{-\gamma t}$ à tout temps. On remarque aussi que, même avec un noyau exponentiel, il apparaît un changement qualitatif si τ_m devient assez grand :

⁵⁶Tout comme le *sillage* d'un électron dans un cristal polarisable peut conduire à une interaction *attractive* entre électrons – le fondement de la théorie BCS pour la supraconductivité.

si $\gamma\tau_m > \frac{1}{2}$ les pôles de $V(z)$ deviennent complexes et la vitesse moyenne ... *oscille* ! L'examen intuitif de l'équation elle-même confirme cette possibilité : si ω décroît avec une échelle de temps assez longue, la dérivée de la vitesse prend de grandes valeurs négatives au début du mouvement, et la vitesse devient alors négative. Comme on s'attend physiquement à ce que la vitesse moyenne $v(t)$ tende vers zéro, on voit qu'un régime *oscillatoire* (amorti) va survenir, qui peut être compris comme un effet de *sillage*. Tout se passe comme si le fluide, polarisé par la particule, opposait une force antagoniste pour supprimer les lacunes de densité créées en son sein antérieurement par le mouvement de la grosse particule.

Prenons maintenant au contraire une fonction $\omega(t - t')$ décroissant lentement, ce qui signifie physiquement que la particule a une grande mémoire du passé (mémoire à longue portée). Une modélisation possible est :

$$\omega(t) = \frac{1}{\tau_m} \frac{1}{1 + (\frac{t}{\tau_m})^2} ; \quad (9.206)$$

avec ce choix, $\Omega(z)$ ne s'exprime pas à l'aide des fonctions élémentaires : on rencontre les fonctions spéciales dites *cosinus intégral*, $\text{Ci}(z)$, et *sinus intégral*, $\text{Si}(z)$, qui sont répertoriées, et dont on connaît notamment le développement pour $|z| \ll 1$. On trouve ainsi :

$$\Omega(z) = \text{Ci}(Z) \sin Z + [\frac{\pi}{2} - \text{Si}(Z)] \cos Z \quad (Z = z\tau_m) \quad (9.207)$$

où :

$$\text{Ci}(Z) \stackrel{\text{déf}}{=} \gamma_E + \ln Z + \int_0^Z \frac{1}{t} (\cos t - 1) dt , \quad \text{Si}(Z) \stackrel{\text{déf}}{=} \int_0^Z \frac{1}{t} \sin t dt , \quad (9.208)$$

γ_E étant la constante d'Euler. En reprenant alors la méthode décrite p. 432, il est possible d'obtenir des informations précises sur le régime aux grands temps de la vitesse. Pour la phase finale ($t \gg \tau_m$), il faut connaître le développement de $\Omega(z)$ pour $z\tau_m \ll 1$; le calcul donne :

$$\Omega(z) \simeq \frac{\pi}{2} + (\gamma_E - 1)z\tau_m + z\tau_m \ln z\tau_m , \quad (9.209)$$

En utilisant les techniques de la section 9.4 (le terme multiforme est ici un logarithme), on trouve le régime aux grands temps :

$$\boxed{v(t) \simeq \frac{4\tau_m}{\pi^2\gamma t^2} v_0 \quad (t \gg \tau_m)} \quad (9.210)$$

La vitesse moyenne relaxe cette fois vers zéro suivant une loi-puissance $\sim t^{-2}$, et non plus essentiellement suivant une banale exponentielle comme dans le cas de la mémoire courte.

9.5.5 Réduction de l'ordre d'une équation différentielle

Afin de montrer l'utilité de la transformation de Laplace pour ce propos, le mieux est de traiter un exemple⁵⁷. Soit l'équation différentielle du second ordre :

$$-\psi'' + \frac{1}{a^3}x\psi(x) = k_0^2\psi, \quad \psi(x < 0) = 0. \quad (9.211)$$

Cette équation apparaît en Mécanique quantique : c'est l'équation aux valeurs propres donnant les états propres $\psi(x)$ d'une particule chargée négativement, soumise à un champ électrique constant dirigé vers les x croissants, et confinée sur le demi-axe \mathbb{R}_+ ; a et k_0 sont des réels. Physiquement, il est évident que seuls des états liés peuvent exister⁵⁸, d'énergie positive (c'est pourquoi k_0 est réel⁵⁹), et on ne s'intéresse donc qu'aux solutions $\psi(x)$ qui tendent vers zéro quand $x \rightarrow +\infty$. Plus généralement, cette méthode peut se révéler fort utile pour le champ central en Mécanique quantique : une fois les variables séparées en $\psi(\vec{r}) = \frac{1}{r}u(r)Y_{lm}(\theta, \phi)$, l'équation aux valeurs propres pour la fonction radiale $u(r)$ est une équation sur \mathbb{R}_+ , avec une barrière parfaitement réfléchissante à l'origine.

L'exemple qui suit est donné à titre d'illustration, car on sait résoudre directement l'équation⁶⁰ (9.211). Les solutions qui s'annulent à l'origine et à l'infini sont les fonctions $\psi_n(x)$ définies comme suit⁶¹ :

$$\psi_n(x) = C_n \operatorname{Ai}\left(\frac{x}{a} - (k_{0,n}a)^2\right) \quad (n \in \mathbb{N}^*), \quad x > 0, \quad (9.213)$$

où $\operatorname{Ai}(x)$ est une fonction spéciale répertoriée (c'est l'une des deux fonctions conventionnellement appelées fonctions d'Airy). $\operatorname{Ai}(x)$ tend vers zéro quand $x \rightarrow +\infty$ et oscille autour de zéro pour x négatif (voir fig. 9.8). Géométriquement, les graphes des $\psi_n(x)$ s'obtiennent donc à partir de celui de la seule et unique fonction $\operatorname{Ai}(x)$, en le translatant vers la droite de sorte que son premier zéro, puis son deuxième zéro, \dots , coïncide avec l'origine, et en effaçant toute la partie située à gauche. Les fonctions $\psi_n(x)$ sont orthogonales au sens où⁶² :

$$\int_0^{+\infty} \psi_n^*(x)\psi_{n'}(x) dx \equiv C_n C_{n'} \int_0^{+\infty} \operatorname{Ai}\left(\frac{x}{a} - (k_{0,n}a)^2\right) \operatorname{Ai}\left(\frac{x}{a} - (k_{0,n'}a)^2\right) dx = \delta_{nn'}. \quad (9.214)$$

⁵⁷Une variante de cette méthode permet de résoudre généralement les équations différentielles dont les coefficients sont au plus des fonctions linéaires de la variable. Sur cette méthode (dite aussi *Méthode de Laplace*), voir par exemple [51], Appendice p. 691.

⁵⁸La particule est soumise à deux effets antagonistes : le mur en $x = 0$ sur lequel elle ne peut que se réfléchir et la renvoie vers $x > 0$, et le champ qui la tire vers l'origine.

⁵⁹L'énergie E de la particule est $E = \frac{\hbar^2 k_0^2}{2m}$. Elle est forcément supérieure au minimum d'énergie potentielle, ici égal à zéro.

⁶⁰En effet, en posant $\psi(x) = f(X)$ avec $X = \frac{x}{a} - (k_0 a)^2$, l'équation (9.211) s'écrit :

$$f''(X) = Xf(X), \quad (9.212)$$

qui est précisément l'équation dite d'Airy. La condition aux limites $\psi(x = 0) = 0$ se traduit par $f(-(k_0 a)^2) = 0$: les $-(k_0 a)^2$ sont donc bien les zéros de la bonne fonction d'Airy, celle qui, en outre, s'annule à l'infini (on cherche des états liés) – d'où la solution (9.213).

⁶¹ C_n est une constante de normalisation.

⁶²Comme pour tout état lié à une dimension, les fonctions propres sont en fait réelles, à une phase constante près, qui n'a aucun sens physique.

Ceci n'est pas facile à démontrer, mais résulte de l'orthogonalité de deux fonctions propres associées à deux valeurs propres distinctes.

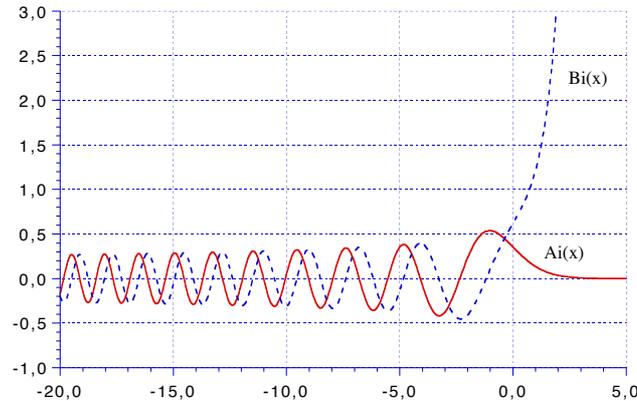


Figure 9.8: Les deux solutions conventionnelles linéairement indépendantes de l'équation d'Airy (9.212).

Le décalage de l'argument dans l'expression (9.213) assure que $\psi_n(0) = 0$ car les $-(k_{0,n}a)^2$ sont précisément les abscisses des zéros de $\text{Ai}(x)$ (c'est comme toujours une condition aux limites, ici $\psi(x=0) = 0$, qui produit la quantification) ; les trois premières valeurs sont $(k_{0,1}a)^2 = 2.338\,107\,43\dots$, $(k_{0,2}a)^2 = 4.087\,949\,445\dots$, $(k_{0,3}a)^2 = 5.520\,559\,819\dots$ et les fonctions correspondantes sont représentées sur la fig. 9.9. On note que si $\psi_n(x)$ est continue en $x = 0$ (la fonction d'onde est *toujours* continue), sa dérivée ne l'est pas : elle est nulle en 0^- , finie en 0^+ , en conséquence du mur infranchissable vers les x négatifs, traduit par le saut infini de potentiel en $x = 0$. Les énergies de la particule sont donc les quantités discrètes $E_n = \frac{\hbar^2 k_{0,n}^2}{2m}$.

L'équation (9.211) permet de trouver directement le comportement des solutions à l'infini⁶³ ; en effet, pour x "grand" (à préciser le cas échéant), on peut oublier $k_0^2\psi$ comparé à $x\psi$; l'équation (9.211) est alors approximativement :

$$-\psi'' + \frac{1}{a^3}x\psi(x) \simeq 0 . \quad (9.215)$$

Il s'agit de trouver le comportement dominant de $\psi(x)$ à l'infini. En allant du plus simple au plus compliqué, on constate qu'un comportement du genre $x^{-\alpha}$ ne convient pas : on ne peut pas équilibrer les termes dominants. En revanche, un comportement du genre exponentielle étirée convient :

$$\psi \simeq e^{-\lambda x^\alpha} \implies \psi''(x) \simeq [-\lambda\alpha(\alpha-1)x^{\alpha-2} + \lambda^2\alpha^2x^{2\alpha-2}]e^{-\lambda x^\alpha} . \quad (9.216)$$

⁶³On sait aussi le faire sur la définition conventionnelle de $\text{Ai}(x)$, mais c'est assez délicat. L'autre fonction d'Airy, notée $\text{Bi}(x)$, diverge exponentiellement quand $x \rightarrow +\infty$ (voir fig. 9.8).

Les termes dominants satisfont (9.215) à condition de prendre $2\alpha - 2 = 1$, soit $\alpha = \frac{3}{2}$ et $\lambda = \pm \frac{2}{3}a^{-3/2}$. Les solutions cherchées se comportent donc comme :

$$\psi(x) \simeq Ce^{\pm \frac{2}{3}(\frac{x}{a})^{3/2}} \equiv \psi_{\text{as}}(x) , \quad (9.217)$$

où C est une constante indéterminée, visiblement proportionnelle⁶⁴ à $\psi'(0)$. Cherchant des états liés, seules les solutions $e^{-\frac{2}{3}(\frac{x}{a})^{3/2}}$ sont à retenir. C'est cette information qui, convenablement traduite dans le langage de Laplace, permettra d'achever la résolution du problème.

Montrons maintenant l'utilité de la transformation de Laplace ; posant :

$$\Psi(k) = \mathcal{L}[\psi](k) = \int_0^{+\infty} e^{-kx} \psi(x) dx , \quad (9.218)$$

l'image de Laplace de l'équation (9.211) est⁶⁵ :

$$-[k^2\Psi(k) - \psi'(0)] + \frac{1}{a^3}\mathcal{L}[x\psi(x)](k) = k_0^2\Psi(k) . \quad (9.219)$$

Or on a :

$$\mathcal{L}[x\psi(x)](k) = \int_0^{+\infty} e^{-kx} [x\psi(x)] dx = -\frac{d}{dk} \int_0^{+\infty} e^{-kx} \psi(x) dx = -\frac{d}{dk}\Psi(k) . \quad (9.220)$$

L'équation (9.219) se transforme en :

$$\Psi'(k) + a^3(k^2 + k_0^2)\Psi(k) = a^3\psi'(0) . \quad (9.221)$$

Il s'agit d'une équation différentielle du *premier* ordre : la transformation de Laplace réduit d'une unité l'ordre de l'équation à résoudre. La solution de (9.221) est :

$$\Psi(k) = a^3\psi'(0)e^{-a^3(\frac{k^3}{3} + k_0^2k)} \int_K^k e^{a^3(\frac{k'^3}{3} + k_0^2k')} dk' , \quad (9.222)$$

où K est une constante d'intégration⁶⁶.

Pour trouver K , on peut utiliser le raisonnement suivant. Les propriétés de $\psi(x)$ à l'infini sont reliées au comportement de $\Psi(k)$ pour $k \sim 0$. Par conséquent, la transformée

⁶⁴De toute façon, l'équation (9.211) est homogène : elle ne permet donc de trouver $\psi(x)$ qu'à un facteur près. Tout facteur arbitraire se reporte à la fois sur C et $\psi'(0)$, ce qui convainc que seul le rapport $\psi'(0)/C$ est pertinent.

⁶⁵La fonction d'onde est toujours continue ; nulle à gauche, elle est donc nulle en $x = 0$, c'est pourquoi le terme $-k\psi(0)$ venant de $\mathcal{L}[\psi']$ est omis. Tant que le potentiel n'a que des sauts finis, la dérivée de la fonction d'onde est elle aussi toujours continue ; elle ne présente un saut (fini) que lorsque le potentiel a un saut infini (exemple : le puits infiniment profond).

⁶⁶On peut se demander pourquoi une seule constante d'intégration apparaît alors que l'équation de départ (9.211) est du *second ordre*. Ceci est visiblement lié à la procédure de réduction de l'ordre, qui élimine les solutions dénuées de transformées de Laplace. Si on ne perd ici aucune solution acceptable, c'est juste que de telles solutions sont physiquement à rejeter.

de Laplace Ψ_{as} de la forme asymptotique (9.217) se comporte près de $k = 0$ comme la vraie transformée de Laplace $\Psi(k)$. On a :

$$\Psi_{\text{as}}(k) = \int_0^{+\infty} e^{-\frac{2}{3}\left(\frac{x}{a}\right)^{3/2}} e^{-kx} dx = \int_0^{+\infty} e^{-\frac{2}{3}\left(\frac{x}{a}\right)^{3/2}} (1 - kx + \dots) dx . \quad (9.223)$$

En particulier, on sait calculer l'intégrale⁶⁷ pour $k = 0$:

$$\Psi_{\text{as}}(k = 0) = \int_0^{+\infty} e^{-\frac{2}{3}\left(\frac{x}{a}\right)^{3/2}} dx = a \left(\frac{2}{3}\right)^{\frac{1}{3}} \Gamma\left(\frac{2}{3}\right) . \quad (9.224)$$

Ceci est aussi la valeur de $\Psi(k = 0)$. Il en résulte que l'égalité suivante doit être satisfaite (voir (9.222)) :

$$a^2 \psi'(0) \int_K^0 e^{a^3\left(\frac{k'^3}{3} + k_0^2 k'\right)} dk' = \left(\frac{2}{3}\right)^{\frac{1}{3}} \Gamma\left(\frac{2}{3}\right) , \quad (9.225)$$

ce qui permet d'écrire la solution (9.222) sous la forme :

$$\Psi(k) = a^3 \psi'(0) e^{-a^3\left(\frac{k^3}{3} + k_0^2 k\right)} \left[\int_0^k e^{a^3\left(\frac{k'^3}{3} + k_0^2 k'\right)} dk' + C \right] \equiv \psi'(0) \Phi(k) , \quad (9.226)$$

où C est une nouvelle constante à déterminer. Pour la trouver, on raisonne comme suit. Ici, on a simplement $\mathcal{L}[\psi''] = k^2 \Psi(k) - \psi'(0)$ puisque $\psi(x = 0) = 0$; d'un autre côté, $\mathcal{L}[\psi'']$ tend vers zéro quand $k \rightarrow +\infty$, d'où :

$$\lim_{k \rightarrow +\infty} [k^2 \Psi(k) - \psi'(0)] = 0 \iff \psi'(0) = \lim_{k \rightarrow +\infty} k^2 \Psi(k) . \quad (9.227)$$

Compte tenu de (9.226), on en déduit que :

$$\lim_{k \rightarrow +\infty} \left[k^2 a^3 e^{-a^3\left(\frac{k^3}{3} + k_0^2 k\right)} \left(\int_0^k e^{a^3\left(\frac{k'^3}{3} + k_0^2 k'\right)} dk' + C \right) \right] = 1 . \quad (9.228)$$

Ceci permet en principe de trouver la constante C , ce qui achève de déterminer $\Psi(k)$.

Le retour vers $\psi(x)$ exige l'inversion de Laplace de $\Psi(k)$, qui n'est pas simple. On peut toutefois toucher du doigt le phénomène important de quantification de l'énergie de la particule. En effet, puisque $\psi'(x) = \mathcal{L}^{-1}[kF(k)](x)$, on trouve finalement (voir (9.226)) :

$$\psi'(x) = \psi'(0) \mathcal{L}^{-1}[k\Phi(k)](x) . \quad (9.229)$$

Il en résulte que l'on doit avoir :

$$\psi'(0) = \psi'(0) \mathcal{L}^{-1}[k\Phi(k)](x = 0) , \quad (9.230)$$

soit :

$$\mathcal{L}^{-1}[k\Phi(k)](0) - 1 = 0 . \quad (9.231)$$

⁶⁷Faire le changement de variable $\frac{2}{3}\left(\frac{x}{a}\right)^{3/2} = X$.

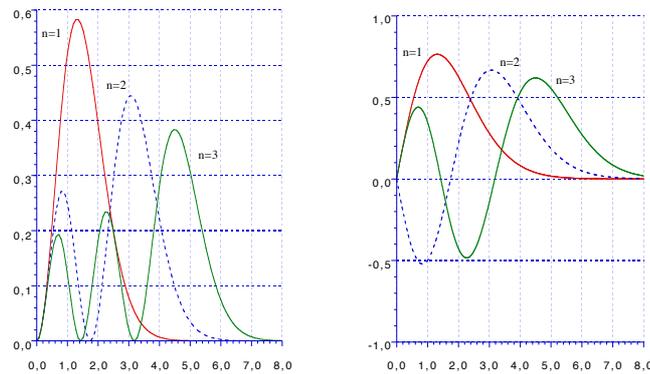


Figure 9.9: À droite : premiers états propres normalisés $\psi_n(x)$ ($\int_0^{+\infty} |\psi_n(x)|^2 dx = 1$) de la particule (généralement donnés en (9.213)) ; l'état fondamental ne s'annule pas pour $x > 0$, le premier état excité s'annule une fois, le second deux fois, etc. À gauche les densités $|\psi_n(x)|^2$ correspondantes. L'abscisse est $\frac{x}{a}$. Plus n augmente, plus le *maximum maximorum* de la densité se décale vers la droite : plus l'état est excité, plus la position moyenne de la particule s'éloigne de l'origine.

Le premier membre de (9.231) est une certaine fonction du seul paramètre sans dimension du problème à savoir $k_0 a$. Cette équation signifie donc que seules les valeurs de ce paramètre qui annulent le premier membre satisfont toutes les conditions requises⁶⁸. a est une donnée liée au champ électrique, c'est donc k_0 qui est contraint de prendre certaines valeurs discrètes $k_{0,n}$: l'énergie de la particule est quantifiée, toutes les énergies (positives) ne sont pas accessibles. Pour une particule quantique liée, il existe un ensemble fini ou isomorphe à \mathbb{N} (c'est ce cas ici) d'énergies possibles, au contraire de ce qu'envisage implicitement la Mécanique classique, où l'énergie peut prendre n'importe quelle valeur.

9.5.6 Une application du théorème de convolution : le pendule isochrone de Huyghens

Le théorème de convolution permet aussi de résoudre simplement le problème du pendule isochrone de Huyghens (période indépendante de l'amplitude d'oscillation). Plus

⁶⁸Les conditions essentielles sont d'imposer aux solutions de tendre vers zéro à l'infini, donc de représenter des états liés (seuls les états liés ont une énergie quantifiée), et de s'annuler à l'origine pour tenir compte du mur infranchissable : il faut retenir que ce sont les conditions imposées à la fonction d'onde qui produisent spontanément le miracle de la quantification de l'énergie. D'ailleurs, un phénomène analogue de quantification se produit pour une corde vibrante dont les extrémités sont maintenues fixes : la longueur d'onde λ ne peut pas avoir n'importe quelle valeur, seules les valeurs entières $\times \frac{\lambda}{2} = L$ sont possibles ($L =$ distance entre les extrémités fixes).

précisément, il s'agit de trouver la courbe C d'équation cartésienne $y = f(x)$, supposée inversible en $x = f^{-1}(y) \stackrel{\text{d'ef}}{=} h(y)$ ayant la propriété suivante : on lâche sans vitesse initiale, dans la pesanteur terrestre (g), un point matériel de masse m qui décrit cette courbe ; alors, le temps de parcours T entre une altitude $y_0 > 0$ et l'altitude nulle ($y = 0$) est *indépendant* de la valeur de l'altitude de départ.

On écrit, s étant l'abscisse curviligne le long de la courbe :

$$mgy_0 = \frac{1}{2}mv^2 + mgy, \quad dt = \frac{ds}{v} = \frac{ds}{\sqrt{2g(y_0 - y)}}. \quad (9.232)$$

Le temps d'arrivée en $y = 0$ en partant de y_0 a donc l'expression intégrale :

$$T = \frac{1}{\sqrt{2g}} \int_0^{y_0} \frac{ds}{\sqrt{y_0 - y}}. \quad (9.233)$$

On a $ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + h'^2(y)} dy \equiv \phi(y)dy$; d'où :

$$T = \frac{1}{\sqrt{2g}} \int_0^{y_0} \frac{\phi(y)}{\sqrt{y_0 - y}} dy. \quad (9.234)$$

Soit $\Phi(k)$ la transformée de Laplace de $\phi(y)$; en prenant la transformée de Laplace de (9.234), sachant que⁶⁹ $\mathcal{L}(y^{-\frac{1}{2}}) = \frac{\Gamma(\frac{1}{2})}{k^{\frac{1}{2}}}$ et en utilisant le théorème de convolution, il vient :

$$\frac{T}{k} = \frac{1}{\sqrt{2g}} \Phi(k) \frac{\sqrt{\pi}}{k^{\frac{1}{2}}} \iff \Phi(k) = \sqrt{\frac{2g}{\pi}} \frac{T}{k^{\frac{1}{2}}} \iff \phi(y) = \frac{\sqrt{2g}}{\pi} \frac{T}{y^{\frac{1}{2}}}. \quad (9.235)$$

Par définition de $\phi(y)$, on a donc :

$$1 + h'^2(y) = \frac{2g}{\pi^2} \frac{T^2}{y} \iff \frac{dx}{dy} = \sqrt{\frac{a}{y} - 1}, \quad a = \frac{2}{\pi^2} gT^2. \quad (9.236)$$

D'où (en utilisant le fait que x est nul si $y = 0$) :

$$x = \int_0^y du \sqrt{\frac{a}{u} - 1}. \quad (9.237)$$

Cette intégrale se calcule aisément en posant $u = a \sin^2 \frac{\theta}{2}$; on trouve :

$$x = \frac{a}{2}(\theta + \sin \theta), \quad y = \frac{a}{2}(1 - \cos \theta). \quad (9.238)$$

Il s'agit bien d'une *cycloïde*⁷⁰. Les oscillations d'un pendule cycloïdal sont donc indépendantes de l'amplitude, un fait qui n'est (approximativement) vrai pour un pendule simple que dans le cas de très petites oscillations.

⁶⁹ $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

⁷⁰En retournant verticalement la courbe (9.238), on obtient la trajectoire décrite par une mouche agrippée à une roue de bicyclette de rayon a roulant sur le sol sans glisser.

Chapitre 9

Équations différentielles. Fonctions de Green

Il s'agit d'une part de rappeler des résultats concernant les équations différentielles, d'autre part de profiter d'un contexte familier pour introduire la notion de Fonction de Green

Il ne s'agit pas ici de refaire un cours sur les équations différentielles, mais de rappeler des résultats essentiels et de profiter d'un contexte en principe assez familier pour introduire, à propos des équations *linéaires*, la notion de *fonction de Green*, dont l'usage est très fréquent et très utile en Physique. Par ailleurs, la brève discussion autour des équations différentielles est aussi l'occasion de dire quelques mots sur leurs pendants discrets, les équations aux différences, qui apparaissent si souvent en Physique.

9.1 Généralités et définitions

Soit f une fonction de $\mathcal{D} \subseteq \mathbb{C}$ dans $f(\mathcal{D}) \subseteq \mathbb{C}$, munies de toutes les dérivées $f', f'', \dots, f^{(m)}$ souhaitables¹. On appelle *équation différentielle* une relation entre f et certaines de ses dérivées satisfaite pour toutes les valeurs de la variable z (réelle ou complexe) où la fonction et ses dérivées sont définies. Formellement, une telle relation s'écrit toujours :

$$\Phi(z, f, f', \dots, f^{(m)}, \dots) = 0 \quad \forall z \in \mathcal{D} , \quad (9.1)$$

où $f^{(m)} = \frac{d^m f}{dz^m}$. Ainsi, les équations suivantes :

$$z^5 f'' + \frac{1}{z} f' + f = \pi , \quad z f' - 3f^2 = 0 , \quad f^{(4)} f^2 + 3 \sin z f = 0 , \quad f' + e^{f(z)} = 0 \dots \quad (9.2)$$

sont des équations différentielles.

La fonction f et la variable z sont *a priori* des quantités complexes, ce qui suppose implicitement que f est une fonction analytique ; en effet, dans le cas contraire, on ne saurait pas ce que signifie $f'(z)$! Il est tout à fait loisible de se restreindre au champ réel, en ne considérant qu'une variable réelle x , et une fonction $f(x)$ à valeurs réelles. Ce faisant, en un sens, on peut paradoxalement considérer des équations plus générales, dans la mesure où on ne sait pas toujours les prolonger à vue dans \mathbb{C} . Par exemple, l'équation différentielle réelle $f'(x) + |x| = 0$ ne peut recevoir de sens si on choisit x dans \mathbb{C} ; en effet, la fonction $|z|$ n'est pas holomorphe, donc la "dérivée" f' , égale à $-|x|$ d'après l'équation considérée, n'est pas définie. D'ailleurs, on sait que les fonctions analytiques sont des objets remarquables et robustes : forcément, les équations différentielles mettant en jeu de

¹ f peut aussi être à valeurs réelles. Si f est une fonction analytique, on sait qu'alors elle est infiniment dérivable.

telles fonctions sont elles-mêmes remarquables et ne peuvent prétendre à une forme d'universalité. Dans la suite, sauf exception dûment mentionnée, les équations différentielles étudiées porteront sur des fonctions analytiques dans un certain domaine \mathcal{D} . Les exemples donnés en (9.2) sont bien de cette espèce – notamment, toutes les fonctions jouant le rôle de coefficients sont analytiques, sauf en leurs points singuliers (isolés).

L'ordre d'une équation différentielle est l'ordre de la plus haute dérivée y figurant explicitement. Les équations différentielles (9.2) sont respectivement d'ordre deux, un, quatre et un.

Il existe une variété infinie d'équations différentielles, ce qui incite à effectuer une classification permettant d'énoncer des théorèmes dont la validité repose sur certains traits caractéristiques de ces équations. La classification fondamentale distingue deux grandes espèces :

1. les équations *linéaires*, caractérisées par le fait que la fonction inconnue et ses dérivées apparaissent au plus à la puissance un (en particulier, une équation linéaire ne contient pas de termes du genre f^p , ff' , $f^{(n)}f^{(m)}$, ni de fonctions $\alpha(f)$, etc).

En outre, il est d'usage de distinguer les équations *homogènes*, où f et ses dérivées figurent dans tous les termes à la puissance un, les équations *inhomogènes* où figure un terme additif indépendant de f et de ses dérivées. Ainsi, les équations :

$$f'' + \frac{1}{z^2}f' + f = 0, \quad f^{(3)} + f = 0 \tag{9.3}$$

sont des équations linéaires et homogènes, alors que :

$$f^{(6)} + e^z f' + f - \ln z = 0 \tag{9.4}$$

est une équation linéaire inhomogène. Souvent, on place au second membre le terme indépendant de f , justifiant les appellations traditionnelles *équation avec ou sans second membre*, parfois commodes mais au fond sans grande utilité. Dans le cas homogène, si f est une solution, alors λf est encore solution, λ étant un nombre quelconque.

Au total, une équation différentielle linéaire d'ordre N est toujours plus précisément de la forme :

$$\sum_{n=0}^N a_n(z)f^{(n)}(z) = \phi(z) \quad (a_N \neq 0); \tag{9.5}$$

les coefficients a_n sont *a priori* des fonctions de z , le cas extrême étant celui où ce sont de simples constantes (on parle alors d'équation à coefficients constants) ; la fonction ϕ est donnée (on l'appelle souvent *source*) ; l'équation devient homogène si $\phi = 0$. Pour la référence ultérieure, notons que (9.5) peut s'écrire formellement comme suit :

$$\hat{L}f = \phi \tag{9.6}$$

où \hat{L} désigne l'opérateur différentiel :

$$\hat{L}(s) = \sum_{n=0}^N a_n(z) \frac{d^n}{dz^n} . \tag{9.7}$$

2. les équations *non-linéaires*, caractérisées par le fait que la fonction inconnue et ses dérivées apparaissent dans des monômes $(f^{(n)})^\lambda (f^{(m)})^\mu \dots$, ou dans des fonctions, sans aucune restriction. Les équations suivantes sont non-linéaires :

$$z^5(f'')^2 + \frac{1}{z}f' + f = \pi, \quad \ln(f') - 5f = 0, \quad f'f^{(4)} + \frac{3}{f^2} \cosh z = z . \tag{9.8}$$

Il est bien clair que rien ne bride l'imagination quand il s'agit de donner un exemple d'équation non-linéaire...

La distinction linéaire – non-linéaire est absolument fondamentale. Pour des équations linéaires, on sait que l'ensemble des solutions peut être muni d'une structure d'espace vectoriel, ce qui permet d'énoncer de nombreux théorèmes concernant notamment l'existence et l'unicité des solutions. Pour les équations non-linéaires, la situation est nettement plus difficile, moins confortable, parfois même un peu acrobatique.

Ces deux types d'équations se rattachent à des mondes très différents. On peut dire que les équations linéaires décrivent des phénomènes très banals, ne conduisant à aucune vraie surprise. Au contraire, les équations non-linéaires contiennent une richesse incommensurable, et engendrent parfois des solutions exotiques, présentant de surcroît une extrême variabilité par rapport à des changements *a priori* anodins. On rencontrera quelques exemples dans la suite, mais il est utile de marquer la différence spectaculaire sur un premier exemple très simple.

L'équation linéaire :

$$f' + \frac{f}{z-1} = 0 \quad (9.9)$$

a pour solution² $f(z) = \frac{f(0)}{1-z}$, où $f(0)$ est la valeur (prescrite à l'avance – voir plus loin, section 9.2). $f(0)$ est un simple facteur qui ne change en rien la forme de la solution, et notamment n'affecte pas ses singularités, ici un unique pôle simple en $z = 1$ – la simple vision de l'équation (9.9) permet d'ailleurs de deviner que $z = 1$ est un point particulier, où il se passe quelque chose de remarquable.

Par contraste, soit l'équation non-linéaire :

$$f' = f^2 \quad (9.10)$$

Si on choisit $f(0) = 1$, la solution est :

$$f(z) = \frac{1}{1-z} . \quad (9.11)$$

Ainsi, en prenant $f(0) = 1$ dans les deux cas, les deux équations (9.9) et (9.10) ont exactement la même solution – et pourtant, la considération de (9.10) ne permet nullement de soupçonner que $z = 1$ est un point remarquable : un pôle apparaît spontanément dans la solution, que l'on aurait pas deviné en regardant l'équation (c'est pourquoi on parle de *singularité spontanée*). Maintenant, choisissons une autre condition initiale, par exemple $f(0) = 2$; la solution est alors $\frac{2}{1-2z}$: elle a encore un pôle, mais il est maintenant en $z = \frac{1}{2}$! Ce simple changement de condition initiale a profondément modifié la solution...

D'une façon générale, pour une équation linéaire du premier ordre $f' = \Phi(f, z)$, avec la condition $f(a) = A$, la solution existe et est une fonction analytique dans le voisinage de a pourvu que $\Phi(Z, z)$ soit une fonction analytique vis-à-vis de chacun de ses arguments en $Z = A$ et $z = a$. Rien de tel ne peut être affirmé pour une équation non-linéaire, pour laquelle le domaine d'analyticit  et le rayon de convergence des solutions sont en g n ral impr visibles en raison pr cis ment de la possibilit  d'apparition spontan e de singularit s.

Une autre diff rence spectaculaire entre  quations lin aires et non-lin aires tient aux comportements compar s de leurs solutions avec celles de leurs  quations aux diff rences  quivalentes (voir section 9.5).

9.2 Conditions initiales. Conditions aux limites

Comme une  quation diff rentielle implique des d riv es, sa r solution passe par des op rations d'int gration, qui introduisent in vitablement des constantes d'int gration. De ce fait, la seule donn e d'une  quation diff rentielle permet, si on sait la faire, de trouver sa solution *g n rale*, qui forme un ensemble de cardinal infini en g n ral.

En Physique, l'arriv e d'une  quation diff rentielle est le r sultat de la construction d'un mod le qui, pour une situation physique donn e (et un probl me bien pos ), ne peut donner qu'une et une seule solution. En d'autres termes, il ne suffit pas de disposer d'une  quation diff rentielle, il convient de la compl ter par des conditions sugg r es par l'exp rience que l'on veut d crire th oriquement, ou impos es par des principes

²obtenue par int gration imm diate !

physiques (celui de causalité par exemple), ou par le sens physique de la théorie en cours (par exemple, la nécessité pour la fonction d'onde d'un état lié d'être normalisable). Ainsi, s'agissant de trouver la dynamique d'une particule classique, il faut connaître sa vitesse et sa position initiales³ ; faute de quoi, on est dans l'impossibilité de trouver toutes les constantes d'intégration. On parle de problème *mal posé* si l'on ne dispose pas d'autant de conditions que nécessaire : un problème bien posé a une et une seule solution.

Il est d'usage de parler de *conditions initiales*, lorsque la variable est le temps. Par exemple, pour un oscillateur harmonique (bille ponctuelle de masse m attachée à un ressort parfait de constante de raideur k , à l'équilibre au point d'abscisse x_e), l'équation masse \times accélération = force s'écrit⁴ :

$$m\ddot{x}(t) = -k(x - x_e) \iff \ddot{x} = -\omega^2(x - x_e) \quad (k = m\omega^2) . \quad (9.12)$$

La solution générale est :

$$x(t) = x_e + A \cos \omega t + B \sin \omega t . \quad (9.13)$$

Si maintenant on dit que la position initiale est x_0 et la vitesse initiale v_0 , on peut écrire le système :

$$(x_e + A \cos \omega t + B \sin \omega t)_{t=0} = x_0 , \quad (-\omega A \sin \omega t + \omega B \cos \omega t)_{t=0} = v_0 \quad (9.14)$$

d'où l'on tire les expressions des deux constantes d'intégration A et B . De la sorte, la seule et unique solution est :

$$x = x_e + (x_0 - x_e) \cos \omega t + \frac{v_0}{\omega} \sin \omega t ; \quad (9.15)$$

C'est bien l'écriture explicite des conditions initiales qui permet d'extraire l'unique solution physique associée à ces conditions de départ.

En fait, prescrire des valeurs données pour la solution (et ses dérivées) en un certain point n'est pas la seule façon de déterminer de façon unique la solution du problème (bien) posé. Par exemple, soit l'équation homogène du second ordre :

$$f''(x) - k^2 f(x) = 0 \quad (k \in \mathbb{R}_+) . \quad (9.16)$$

Sa solution générale est :

$$f(x) = A e^{+kx} + B e^{-kx} . \quad (9.17)$$

Supposons maintenant que, pour des raisons imposées par le problème examiné, la fonction f soit astreinte à être bornée et continue⁵. Alors, la solution acceptable physiquement est la seule et unique fonction :

$$f(x) = f(0) e^{-k|x|} , \quad (9.18)$$

obtenue en supprimant dans l'expression générale (9.17) l'exponentielle divergente selon le signe de x . Dans ce cas, il reste à dire ce que vaut $f(0)$, par d'autres considérations⁶.

Enfin, il faut savoir que le nombre de paramètres dont dépend au total la solution la plus générale peut être *supérieure* à l'ordre de l'équation. Par exemple, la solution générale de l'équation réelle non linéaire du premier ordre $f' = f^{\frac{1}{3}}$ est telle que $f^{2/3} = \frac{2}{3}(x + C)$, où C est quelconque. L'extraction de la racine donne $f(x) = \alpha [\frac{2}{3}(x + C)]^{\frac{3}{2}}$, avec $\alpha = \pm 1$; dans ce cas précis, on peut aussi dire qu'il y a deux familles de solutions, associées aux deux branches distinctes solutions de $(e^{2ik\pi})^{3/2}$, chaque branche contenant une constante d'intégration C .

³Ainsi, il faut deux conditions initiales par degré de liberté, puisque l'équation fondamentale de la dynamique est du deuxième ordre en temps.

⁴ x_e désigne l'abscisse du point d'équilibre.

⁵C'est exactement le cas pour les équations aux valeurs et fonctions propres rencontrées en Mécanique quantique.

⁶Les experts en Mécanique quantique auront reconnu en (9.16) l'équation aux valeurs propres pour une particule liée par un potentiel attractif $-g\delta(x)$ ($g > 0$), équation écrite pour $x \neq 0$. Le saut de la dérivée f' en $x = 0$ est pilotée par l'intensité g de δ , représentant la "profondeur" du puits. Dans ce contexte, la valeur $f(0)$ s'obtient en écrivant la condition de normalisation de la fonction d'onde.

9.3 Équations différentielles linéaires à coefficients constants

9.3.1 Rappel de quelques résultats

La forme générale d'une équation différentielle linéaire homogène d'ordre N à coefficients constants est :

$$\sum_{n=0}^N a_n f^{(n)}(z) = 0 \quad (a_N \neq 0) \iff \hat{L}f(z) = 0 . \quad (9.19)$$

Pour obtenir sa solution, il suffit de remarquer que tout repose d'une part sur la fonction exponentielle, douée de sa propriété caractéristique (la dérivée est proportionnelle à la fonction elle-même – de proche en proche, c'est vrai pour toute dérivée d'ordre quelconque), d'autre part sur le fait que l'équation étant linéaire, toute combinaison linéaire de solutions est encore solution.

La fonction exponentielle est en effet la seule à posséder la propriété suivante⁷ :

$$\frac{d^m}{dz^m} e^{\lambda z} = \lambda^m e^{\lambda z} \quad (m \in \mathbb{N}) \iff [f^{(m)}(z) = \lambda^m f(z) \iff f(z) = e^{\lambda z}] . \quad (9.20)$$

Il en résulte immédiatement que si l'on injecte la forme $f(z) = e^{\lambda z}$ dans (9.19), on trouve :

$$\left(\sum_{n=0}^N a_n \lambda^n \right) e^{\lambda z} = 0 \quad \forall z . \quad (9.21)$$

Cette équation n'est satisfaite que si λ est l'une des racines du polynôme de degré N , $P_N(\lambda)$ construit avec les coefficients a_n de l'équation différentielle et appelé *polynôme caractéristique* :

$$\sum_{n=0}^N a_n \lambda^n = 0 \iff P_N(\lambda) = 0 \iff \lambda \in \{\lambda_k\}_{1 \leq k \leq N} . \quad (9.22)$$

Maintenant, par construction, chaque fonction $f_k = e^{\lambda_k z}$ est solution de l'équation :

$$\left(\sum_{n=0}^N a_n \lambda_k^n \right) e^{\lambda_k z} = 0 \iff P_N(\lambda_k) e^{\lambda_k z} = 0 \iff \hat{L}f_k(z) = 0 ; \quad (9.23)$$

Par ailleurs, \hat{L} est un opérateur linéaire, puisque l'opération de dérivation est une opération linéaire. De cette propriété découle l'égalité :

$$\hat{L}(\alpha\phi(z) + \beta\psi(z)) = \alpha\hat{L}\phi(z) + \beta\hat{L}\psi(z) , \quad (9.24)$$

où les constantes α et β sont quelconques, tout comme les fonctions ϕ et ψ . Il en résulte immédiatement que toute combinaison linéaire des solutions particulières $e^{\lambda_k z}$:

$$f(z) = \sum_{k=1}^N C_k e^{\lambda_k z} \quad (9.25)$$

est solution de (9.19), *quels que soient les coefficients*⁸ C_k . En effet :

$$\hat{L} \left(\sum_{k=1}^N C_k e^{\lambda_k z} \right) = \sum_{k=1}^N C_k \hat{L} e^{\lambda_k z} ; \quad (9.26)$$

comme $\hat{L}e^{\lambda_k z} = 0$ par construction, la dernière somme est nulle, quels que soient les C_k , et l'équation différentielle est bien satisfaite par l'expression (9.25).

⁷D'où, en particulier, le développement de Taylor infini $e^{\lambda z} = \sum_{n=0}^{+\infty} \frac{z^n}{n!} (\lambda^n e^{\lambda z})_{z=0} = \sum_{n=0}^{+\infty} \frac{z^n}{n!} \lambda^n$.

⁸On devine que cette propriété permet de doter l'ensemble des solutions d'une structure d'espace vectoriel de dimension N , égale à l'ordre de l'équation différentielle.

La fonction d'expression (9.25) est de fait la solution générale de l'équation : on peut en effet montrer que *toutes* les solutions de l'équation homogène (9.19) sont de cette forme.

Par le théorème fondamental de l'algèbre, le polynôme $P_N(\lambda)$ défini en (9.22) possède N zéros complexes λ_k et s'écrit après factorisation :

$$P_N(\lambda) = a_N \prod_{k=1}^N (\lambda - \lambda_k) ; \quad (9.27)$$

si tous les λ_k sont différents, les N fonctions $e^{\lambda_k z}$ sont linéairement indépendantes, et constituent une base de l'espace vectoriel des solutions.

Si plusieurs λ_k sont égaux entre eux, c'est-à-dire si l'équation $P_N(\lambda) = 0$ a des racines multiples (on dit aussi qu'il y a *dégénérescence*), alors le nombre de fonctions $e^{\lambda_k z}$ linéairement indépendantes est strictement inférieur à N . Par exemple, s'il n'existe qu'une racine multiple, λ_{k_0} , supposée d'ordre r , alors le polynôme P_N s'écrit $P_N(\lambda) = a_N (\lambda - \lambda_{k_0})^r \prod_{k=1}^{N-r} (\lambda - \lambda_k)$. Il est facile de vérifier que toutes les fonctions $z^p e^{\lambda_{k_0} z}$, $p = 0, 1, \dots, r-1$ sont solutions de l'équation, et elles sont visiblement linéairement indépendantes. Il s'ensuit que dans ce cas, la solution générale s'écrit :

$$f(z) = \left(\sum_{p=0}^{r-1} c_p z^p \right) e^{\lambda_{k_0} z} + \sum_{\lambda_k \neq \lambda_{k_0}} C_k e^{\lambda_k z} \quad (r \geq 1) . \quad (9.28)$$

Lorsque l'équation est inhomogène :

$$\sum_{n=0}^N a_n f^{(n)}(z) = \phi(z) \quad (a_N \neq 0) \iff \hat{L} f(z) = \phi(z) , \quad (9.29)$$

la solution la plus générale s'obtient en faisant la somme d'une solution particulière f_{part} et de la solution générale de l'équation homogène associée :

$$f(z) = f_{\text{part}}(z) + \sum_{k=1}^N C_k e^{\lambda_k z} . \quad (9.30)$$

En effet, la différence $f(z) - f_0(z)$ satisfait :

$$\sum_{n=0}^N a_n f^{(n)}(z) = 0 \quad (9.31)$$

dont la solution la plus générale est bien $\sum_{k=1}^N C_k e^{\lambda_k z}$. On verra que c'est dans ce cas que la méthode des fonctions de Green révèle toute son utilité (section 9.6) – tant qu'il n'est question que d'équations différentielles.

9.3.2 Résolution à l'aide de transformations intégrales

Il s'agit juste d'un rappel, pour la complétude. On a vu au ch. 7 que la transformation de Laplace permet de résoudre systématiquement les équations du genre (la variable est ici t) :

$$\sum_{p=0}^N a_p \frac{d^p f}{dt^p} = \phi(t) . \quad (9.32)$$

Si $F = \mathcal{L}[f]$, $\Phi = \mathcal{L}[\phi]$, la solution est :

$$F(z) = \frac{1}{Z(z)} \left[\Phi(z) + \sum_{p=1}^N a_p \sum_{r=0}^{p-1} z^{p-r-1} f^{(r)}(0) \right] \quad (9.33)$$

où $Z(z) = \sum_{p=0}^N a_p z^p$ est souvent appelé *fonction de transfert* – mais c’est aussi le polynôme caractéristique $P_N(z)$ introduit plus haut (voir (9.22)). Il s’agit bien de la solution générale, prenant en compte les N constantes d’intégration $f^{(r)}(0)$, $r = 0, 1, \dots, N-1$. L’inversion de Laplace fait apparaître les $e^{\lambda_k t}$, les λ_k étant les zéros de $Z(z)$ et donnant des pôles pour $F(z)$.

Notons que toutes les solutions d’une équation homogène à coefficients constants admettent une transformée de Laplace, puisqu’elles sont des combinaisons linéaires d’exponentielles $e^{\lambda_k t}$. Il en résulte que le formalisme ci-dessus est opérationnel dès que la source ϕ a également une transformée de Laplace.

On peut aussi utiliser la transformation de Fourier, dans l’hypothèse où toutes les transformées existent ; avec $\tilde{F}(\omega) = \mathcal{F}[f]$, $\tilde{\Phi}(\omega) = \mathcal{F}[\phi]$, on trouve :

$$\tilde{F}(\omega) = \frac{\tilde{\Phi}(\omega)}{\sum_{p=0}^N a_p (-i\omega)^p} \equiv \frac{1}{Z(-i\omega)} \tilde{\Phi}(\omega) . \quad (9.34)$$

On sait que (9.34) ne représente pas la solution générale⁹(il n’y a pas de constantes d’intégration !), mais le *régime forcé*, effectivement indépendant des conditions initiales, et réalisé physiquement après extinction des transitoires à condition que tous les zéros λ_k de $Z(z)$ aient une partie réelle strictement négative¹⁰ – c’est bien ce fait qui assure l’amnésie vis-à-vis des valeurs initiales. C’est bien aussi ce que dit l’expression (9.33) : avec la condition ci-dessus pour tous les λ_k , la somme contenant les conditions initiales disparaît de fait dès que $t \gg (\min_k |\Re \lambda_k|)^{-1}$.

9.4 Équations différentielles linéaires à coefficients variables

Dans le cas homogène, une équation d’ordre N est de la forme :

$$\sum_{n=0}^N a_n(z) f^{(n)}(z) = 0 \quad (a_N(z) \neq 0) ; \quad (9.35)$$

la dépendance effective des coefficients par rapport à la variable z modifie du tout au tout la difficulté du problème. En effet, une exponentielle du genre $e^{\lambda z}$ n’est plus solution puisque le report dans (9.35) donne l’équation $(\sum_{n=0}^N a_n(z) \lambda^n) e^{\lambda z} = 0$, qui ne peut visiblement pas être satisfaite avec des $a_n(z)$ variables.

Toutefois, grâce au caractère linéaire de l’équation, on peut démontrer que la solution générale de cette équation est de la forme :

$$f(z) = \sum_{k=1}^N C_k E_k(z) , \quad (9.36)$$

où les $E_k(z)$ sont N solutions particulières linéairement indépendantes. Ainsi, grâce à la linéarité, l’ensemble des solutions peut à nouveau être muni d’une structure d’espace vectoriel. La grosse difficulté, quand les coefficients sont variables, consiste à trouver effectivement ces N solutions $E_k(z)$ – dans le cas où les coefficients sont constants, c’est au contraire en principe un jeu d’enfant de trouver les N solutions $e^{\lambda_k z}$, les λ_k étant les zéros du polynôme caractéristique $P_N(\lambda)$.

La méthode générale de résolution repose sur des théorèmes démontrés par Fuchs, précisant les conditions qui permettent de chercher les solutions sous la forme¹¹ $(z - z_0)^\alpha S(z)$, où α est un exposant (pas forcément entier) à déterminer, et où $S(z)$ est une série entière en $(z - z_0)$ dont les coefficients sont trouvés par une relation de récurrence¹² ; cette dernière s’obtient en reportant dans (9.35) la forme $f(z) = (z - z_0)^\alpha S(z)$ et en identifiant

⁹Mais, en exécutant les acrobaties décrites dans le ch 6, on peut cependant retrouver la solution générale en faisant réapparaître les constantes liées aux conditions initiales, et retomber bel et bien sur ses pieds.

¹⁰Dans le cas contraire, la solution diverge quand $t \rightarrow +\infty$ et ne saurait avoir une transformée de Fourier.

¹¹C’est possible quand z_0 est un point ordinaire ou un point singulier régulier.

¹²Le produit $(z - z_0)^\alpha S(z)$ porte le nom de série de Frobenius.

les termes de même puissance en $(z - z_0)$. Une telle relation dépend des valeurs “initiales” des c_n , fixées par l’application de conditions aux limites, ou par l’exigence que la fonction soit bornée en certains points.

C’est cette méthode qui est mise en œuvre, par exemple, en Mécanique quantique pour trouver les fonctions d’onde stationnaires de l’équation aux valeurs propres dans le cas d’une énergie potentielle V variant gentiment dans l’espace¹³. Pour mémoire, citons quelques équations apparaissant fréquemment en Physique :

- équation d’Airy¹⁴ : $f''(z) = zf(z)$
- équation de Weber - Hermite¹⁵ : $f''(z) + (\nu + \frac{1}{2} - \frac{1}{4}z^2)f(z) = 0$
- équation de Bessel¹⁶ : $f''(z) + \frac{1}{z}f'(z) + \left(1 - \frac{\nu^2}{z^2}\right)f(z) = 0$.

9.5 Équations différentielles et équations aux différences

Montrons d’abord comment on peut mettre en parallèle une équation différentielle et une équation aux différences. Des équations de ce dernier type apparaissent souvent en Physique, soit directement, soit parce que le problème examiné, s’il s’exprime naturellement en termes de grandeurs continues, se résout plus agréablement avec des variables discrètes – ou inversement¹⁷. Les équations aux différences sont clairement d’usage universel dans les méthodes de résolution numérique d’équations que l’on ne sait pas traiter analytiquement, lorsqu’un algorithme de proche en proche ou itératif est utilisé.

Soit une suite de nombres f_n ($n \in \mathbb{N}$) obéissant à une certaine relation de récurrence. Par exemple :

$$f_{n+1} = \alpha f_n \quad (\alpha \in \mathbb{R}_+) . \tag{9.37}$$

Ceci se réécrit trivialement :

$$f_{n+1} - f_n = (\alpha - 1)f_n \tag{9.38}$$

et se traduit en bon français en disant que la variation de f entre n et $n + 1$ est proportionnelle à la valeur de f au point n – clairement, on peut dire que le premier membre est une *dérivée discrète* : c’est le taux de variation de f_n pour une variation unité de la variable, l’entier n . Une valeur de départ f_0 étant donnée (à nouveau, outre l’équation, il faut se donner des conditions supplémentaires, ici une seule suffit), la solution de (9.37) est manifestement :

$$f_n = \alpha^n f_0 . \tag{9.39}$$

D’un autre point de vue, l’écriture (9.38) évoque une équation différentielle du genre :

$$f'(x) = kf(x) \tag{9.40}$$

en posant que la valeur de la fonction $f(x)$ aux points $x = n\Delta x$ est précisément égale à f_n . La solution de (9.40) est $f(x) = f(0)e^{kx}$. L’identification va jusqu’au bout en notant que :

$$f_n = \alpha^n f_0 = f_0 e^{n \ln \alpha} = f_0 e^{(\frac{\ln \alpha}{\Delta x})(n\Delta x)} \equiv f(x = n\Delta x) . \tag{9.41}$$

¹³Pour $V(x)$ constant par morceaux, on évite la méthode de Fuchs en travaillant intervalle par intervalle, et en recollant les morceaux conformément aux prescriptions résultant du sens physique de la fonction d’onde (continuité de la fonction et de sa dérivée tant que $V(x)$ a des sauts *finis*).

¹⁴On la trouve souvent dans les problèmes d’optique, et elle arrive pour une particule quantique chargée soumise à un champ constant (par exemple, un électron soumis à un champ électrique indépendant du temps et uniforme dans l’espace – voir l’exemple d’application de la transformation de Laplace donné à la fin du ch. 7).

¹⁵rencontrée notamment à propos de l’oscillateur harmonique quantique.

¹⁶rencontrée partout en Physique. Bessel était astronome et a introduit ces fonctions qui portent depuis son nom en accomplissant l’analyse harmonique du problème de Kepler.

¹⁷De façon caricaturale, on pourrait dire que le théoricien face à un problème continu s’empresse de le discrétiser, et devant un problème discret ne peut s’empêcher de le “continuer”... En pratique, la continuation doit se faire avec discernement ; par exemple, une limite continue en temps doit respecter le Principe de causalité : l’apparition de fonctions $\delta(t)$ – en tant que limite de fonctions à mémoire courte – peut conduire à des difficultés, tant que l’on a pas réalisé que c’est plutôt $\delta^{(+)}(t)$ qui, de fait, est la bonne fonction à considérer. En outre, toutes les subtilités liés à l’alternative discret \leftrightarrow continu (voir notamment section 9.5) doivent rester présentes à l’esprit.

et à condition de poser $k = \frac{\ln \alpha}{\Delta x}$ (et bien sûr $f_0 = f(0)$).

Ceci étant fait, il est alors légitime d'affirmer que l'équation aux différences (9.37) (ou (9.38)) est la version discrète de l'équation différentielle (9.40). Ces deux écritures représentent finalement la même réalité : on peut dire que les f_n sont des points expérimentaux relevés quand on dispose de la résolution définie par Δx , et constituent la mesure de la grandeur théorique continue $f(x)$. En tout cas, les deux écritures, discrète et continue, se fondent l'une dans l'autre (il suffit d'imaginer que la résolution est de meilleure en meilleure, c'est-à-dire que Δx est rendu de plus en plus petit).

Cette quasi-coïncidence n'a pas hélas l'universalité que l'on pourrait espérer – d'ailleurs, l'identification complète exige de poser $k = \frac{\ln \alpha}{\Delta x}$, d'où l'importance du signe de α . Ce point s'éclaire en réalisant que l'assimilation d'une fonction $f(x)$ avec ses valeurs ponctuelles $f(n\Delta x)$ exige que, à l'échelle Δx , $f(x)$ soit une fonction *lentement variable*. Et de fait, si on prend $\alpha < 0$ dans (9.37), alors les différents f_n alternent en signe : la fonction $f(x)$ correspondante n'est sûrement pas à variation lente et on peut même dire intuitivement que, dans la limite $\Delta x \rightarrow 0$, cette fonction n'est certainement ni continue, ni dérivable.

Une question très importante est donc d'élucider les relations précises entre les solutions d'une équation aux différences et celles de son équation différentielle associée¹⁸. Les quelques exemples suivants donnent une idée de la difficulté du sujet, qui doit inciter à la prudence.

À titre de première illustration, changeons simplement le signe de α dans (9.37) en prenant maintenant :

$$f_{n+1} = -\alpha f_n \quad (\alpha \in \mathbb{R}_+) , \quad (9.42)$$

dont la solution est $f_n = (-1)^{n+1} \alpha^{n+1} f_1$. L'équation différentielle correspondante se trouve en récrivant (9.42) sous la forme $f_{n+1} - f_n = -(\alpha + 1)f_n$, soit :

$$f'(x) = -(\alpha + 1)f(x) \iff f(x) = f(0) e^{-(\alpha+1)x} . \quad (9.43)$$

Sans aucun doute, les deux versions donnent des fonctions complètement différentes ; par exemple, si $\alpha > 1$, la solution discrète diverge en oscillant, alors que la version continue tend vers zéro exponentiellement vite ! La compréhension en profondeur de ces points assez subtils repose sur la considération des singularités des équations considérées¹⁹. Clairement, le traitement discret, inévitable sur ordinateur, des équations différentielles (même linéaires) exige un certain savoir-faire, et doit toujours être conduit avec discernement et vigilance.

De surcroît, la situation, en pratique, est rarement aussi simple : en général, une récurrence n'implique pas seulement deux termes consécutifs. Tout en maintenant l'aspect linéaire, on peut avoir des relations du genre :

$$f_{n+2} - \alpha f_{n+1} + \beta f_n = 0 ; \quad (9.45)$$

une telle récurrence est dite du *second ordre* puisque chaque f_n dépend des deux f_n précédents. (9.45) s'écrit aussi :

$$f_{n+2} - 2f_{n+1} + f_n = (\alpha - 2)f_{n+1} + (1 - \beta)f_n ; \quad (9.46)$$

Tout comme $f_{n+1} - f_n$ est la dérivée (première) discrète, $f_{n+2} - 2f_{n+1} + f_n \equiv (f_{n+2} - f_{n+1}) - (f_{n+1} - f_n)$ peut être considérée comme une dérivée seconde discrète.

Cette dernière remarque se généralise d'ailleurs comme suit. En définissant $Df_n = f_{n+1} - f_n$, on a :

$$Df_n = f_{n+1} - f_n , \quad D^2 f_n = D(Df_n) = f_{n+2} - 2f_{n+1} + f_n , \quad \dots , \quad D^k f_n = \sum_{p=0}^k (-1)^p C_k^p f_{n+k-p} . \quad (9.47)$$

¹⁸ Une question particulièrement pertinente si un traitement numérique de l'équation différentielle s'impose. . .

¹⁹ C'est ici le cas parce que le point à l'infini est un *point singulier irrégulier*. De fait, la compréhension en profondeur de telles anomalies passe par l'analyse des points singuliers de l'équation différentielle. Très brièvement : on classe les singularités d'une équation différentielle linéaire en examinant les singularités (au sens de la théorie des fonctions analytiques) des coefficients apparaissant dans (9.5), réécrite de façon canonique (après division par $a_N(x)$) :

$$f^{(N)}(x) + p_{n-1}(x)f^{(n-1)}(x) + \dots + p_1(x)f'(x) + p_0(x)f(x) = 0 ; \quad (9.44)$$

On dit que x_0 est un point singulier *régulier* si toutes les quantités $(x - x_0)^k p_{n-k}(x)$ sont analytiques dans un voisinage de x_0 , alors que certains coefficients $p_m(x)$ ont une singularité en x_0 . La définition s'étend au cas du point à l'infini en effectuant le changement de variable $X = \frac{1}{x}$. Un point qui n'est pas régulier est dit *irrégulier*, c'est le cas de $x = \infty$ dans l'exemple traité ci-dessus. Pour plus de détails, voir par exemple l'ouvrage de BENDER et ORSZAG, ch. 3.

En adoptant maintenant les correspondances :

$$f_n \leftrightarrow f(x = n\Delta x) , \quad D^k \leftrightarrow f^{(k)}(x) . \quad (9.48)$$

on est en mesure d'associer à une équation aux différences une et une seule équation différentielle, et réciproquement. Par exemple, en faisant apparaître $f_{n+1} - f_n$ au second membre de (9.46), l'équation différentielle associée à (9.45) par les règles (9.47) est :

$$f'' = (\alpha - 2)f' + (\alpha - \beta - 1)f . \quad (9.49)$$

Remarques

1. Les relations (9.48) peuvent s'écrire de façon plus formelle comme suit. Soit l'opérateur de translation²⁰ élémentaire \mathcal{T} défini comme :

$$\mathcal{T} f_n \stackrel{\text{déf}}{=} f_{n+1} . \quad (9.50)$$

Alors, $Df_n = (\mathcal{T} - 1)f_n$, $D^2f_n = (\mathcal{T} - 1)^2f_n, \dots, D^k f_n = (\mathcal{T} - 1)^k f_n$. Le développement de $(\mathcal{T} - 1)^k$ fait bien apparaître les coefficients du binôme comme en (9.47)

2. On pourrait aussi définir la dérivée discrète comme $Df_n = f_n - f_{n-1}$, ou $Df_n = \frac{1}{2}(f_{n+1} + f_n) - \frac{1}{2}(f_n + f_{n-1})$. Il ne s'agit en général que de détails conventionnels. Toutefois, la formulation de Feynman de la Mécanique quantique (intégrale de chemin) exige parfois un choix particulier, notamment en présence d'un champ magnétique.

À titre d'exemple d'une récurrence d'ordre deux²¹, soit :

$$f_{n+2} = f_n \iff f_{n+2} - 2f_{n+1} + f_n = -2(f_{n+1} - f_n) \quad (9.51)$$

dont, selon (9.47), l'équivalent continu est :

$$f'' = -2f' . \quad (9.52)$$

La solution générale de l'équation discrète est visiblement :

$$f_n = A + (-1)^n B \quad (A \text{ et } B \text{ constantes arbitraires}) , \quad (9.53)$$

cependant que la solution de l'équation différentielle est :

$$f(x) = C + De^{-2x} . \quad (9.54)$$

À nouveau, ces dernières solutions n'ont *a priori* aucun rapport avec les solutions (9.53) de l'équation discrète ! Notamment, le comportement à l'infini ($x \rightarrow \infty$) de l'expression (9.54) n'a rien à voir avec celui ($n \rightarrow \infty$) de (9.53). Tout au plus peut-on observer que les solutions lentement variables (en fait triviales puisqu'elles sont constantes) obtenues en prenant $B = D = 0$ coïncident ; en revanche, les solutions variables ($A = C = 0$ - ce sont en général les plus intéressantes !) n'ont strictement rien à voir l'une avec l'autre²².

Répetons que, même pour une équation linéaire, la discrétisation, ou l'opération inverse ("continuation") peuvent ne pas être des opérations innocentes. Pour être aussi complet que possible dans un tour d'horizon, on doit aussi faire remarquer que la résolution d'une version peut être un jeu d'enfant, alors que la résolution de l'autre est un tour de force ; exemple : $f_{n+1} - f_n = \frac{1}{f_n}$, dont la version continue est $f' = \frac{1}{f}$ - qui n'est pas vraiment

²⁰Le mot *translation* a un sens profond, indépendamment du choix évident de la terminologie (\mathcal{T} fait passer de n à $n + 1$. En effet, la dérivée est au cœur des opérateurs infinitésimaux de translation dans l'espace, que l'on rencontre par exemple en Mécanique quantique.

²¹Ce type de récurrence se résout en posant $f_n = r^n$, ce qui conduit à une équation algébrique de degré N pour l'inconnue r . On sait que l'on peut aussi résoudre en effectuant une transformation de Laplace (ch 7).

²²Autre exemple : la récurrence $f_{n+1} = nf_n$ et l'équation différentielle $f' = (x - 1)f$ sont équivalentes, selon la correspondance (9.47). Pourtant, les solutions respectives, $f_n = (n - 1)!f_1$ et $f(x) = f(0)e^{\frac{x^2}{2} - x}$ ont des comportements radicalement différents à l'infini : par la formule de Stirling, on a $f_n \sim n^n$, qui diverge infiniment plus vite que e^{x^2} .

difficile à résoudre... Autre exemple célèbre : la diabolique application logistique $f_{n+1} = rf_n(1 - f_n)$, $r > 0$, dont la version continue est réellement triviale²³.

A fortiori, une équation différentielle *non-linéaire* et sa version discrétisée peuvent avoir des solutions très différentes. Plus précisément, il peut arriver que la version discrète ait des solutions insoupçonnables au vu de la version continue : à nouveau, le diable est à l'œuvre et peut faire des intrusions spectaculaires en jouant le scénario des singularités spontanées (une telle catastrophe n'est pas toujours sûre, heureusement ou malheureusement, c'est selon).

Par exemple, soit l'équation différentielle :

$$f'(x) = f^2(x) - f(x) \quad (9.55)$$

que l'on cherche à résoudre avec la condition $f(0) = 2$. La solution est²⁴ :

$$f(x) = \frac{2}{2 - e^x} \quad (9.56)$$

et présente une divergence en $x = \ln 2$. L'équation aux différences correspondante est :

$$f_{n+1} - f_n = f_n^2 - f_n, \quad (9.57)$$

dont la solution²⁵, pour $f_0 = 2$, est $f_n = 2^{2^n}$, qui est finie pour tout n fini. Cette distinction est générale : contrairement à leurs équivalentes non-linéaires continues, les équations aux différences non-linéaires ne sont pas sujettes au phénomène d'apparition de singularités spontanées.

Il convient donc d'être toujours extrêmement prudent quand il s'agit de prévoir le comportement des solutions d'une équation différentielle au vu de celui des solutions de l'équation aux différences, ou inversement.

Ces résultats sont une source de perplexité, et en tout cas doivent inciter à la prudence. En effet, dès qu'il faut résoudre numériquement une équation, force est de recourir à la discrétisation : l'ordinateur ne connaît ni le continu, ni l'infini. La résolution numérique d'une équation différentielle linéaire ou non-linéaire exige donc un grand savoir-faire, et doit toujours être conduite avec vigilance. La règle méthodologique de base est d'avoir le constant souci de procéder à de multiples vérifications de la solution obtenue numériquement (symétries, stabilité vis-à-vis d'un bruit numérique forcé, comportements limites, ...).

Un canular numérique^a

Pour illustrer les remarques précédentes, citons un canular numérique. Soit la relation de récurrence :

$$x_{n+1} = 111 - \frac{1130}{x_n} + \frac{3000}{x_n x_{n-1}}, \quad (9.58)$$

commençant avec $x_0 = 2$ et $x_1 = -4$. La relation de récurrence est du second ordre (elle met en jeu trois termes consécutifs) et non-linéaire : tout pour plaire.

On peut montrer que, avec ces valeurs de départ, la suite des nombres x_n a pour limite $x_\infty = 6$. Essayez de mettre cette limite en évidence en programmant l'itération sur une machine ne connaissant que les nombres^b...

Comme le dit l'auteur de l'article "*on observe parfois en machine une bonne et rapide convergence vers un résultat totalement faux*".

^aVoir l'article de J.-M. MULLER dans *La Recherche*, numéro spécial sur la théorie des nombres, juillet-août 1995.

^bEn revanche, les algorithmes formels du genre *Mathematica* se tirent bien d'affaire, à la condition expresse de n'effectuer aucune opération *numérique* pendant l'itération.

²³ On obtient l'équation différentielle $f' = (r - 1)f - f^2$.

²⁴ Poser $g = \frac{1}{f}$ conduit à une équation linéaire pour g .

²⁵ On trouve cette solution en posant $g_n = \log_2 f_n$, ce qui donne la récurrence linéaire $g_{n+1} = 2g_n$.

9.6 Fonctions de Green

Les fonctions de Green ne sont pas des fonctions spéciales (au sens où on parle de fonctions de Bessel, fonctions de Weber, fonctions d’Airy *etc*) : cette terminologie désigne en réalité un objet mathématique sur lequel est fondée une méthode de résolution des équations *linéaires*, qu’elles soient différentielles ou aux dérivées partielles (voir ch. 10). Dans cette section, sauf mention contraire, les variables seront supposées réelles.

S’agissant des équations différentielles, l’intérêt des fonctions de Green réside principalement en ceci : pour résoudre une équation inhomogène, il faut d’une part déterminer la solution générale de l’équation homogène qui lui est associée, d’autre part trouver une solution particulière de l’équation complète, puis faire la somme des deux et enfin caler les constantes d’intégration avec les données supplémentaires indispensables. La fonction de Green est un moyen systématique de trouver précisément cette solution particulière²⁶. En outre la méthode des fonctions de Green se prête bien à la rediscussion des liens importants entre conditions aux limites et analyticit  (par exemple : causalit  et analyticit ).

Par ailleurs, s’il est certain que la m thode de Green r v le pleinement son utilit    propos des  quations *aux d riv es partielles*, les id es de base peuvent d j  s’ noncer dans un contexte beaucoup plus simple, celui des  quations diff rentielles : c’est ce qui est fait ci-dessous.

9.6.1 Pr liminaires

La m thode des fonctions de Green fait un usage intensif de la fonction de Dirac, dont la r gle op rationnelle fondamentale est, pour m moire :

$$\int_{-\infty}^{+\infty} f(x)\delta(x-x_0) dx = f(x_0) \tag{9.59}$$

o  la fonction f est continue en x_0 . En prenant $f(x) = 1 \forall x$, on a :

$$\int_{-\infty}^{+\infty} \delta(x-x_0) dx = 1 . \tag{9.60}$$

Dans les notations de la Th orie des distributions, et pour $x_0 = 0$, (9.59) s’ crit :

$$\langle \delta, f \rangle = f(0) ; \tag{9.61}$$

pour des fonctions s’annulant   l’infini, une int gration par parties permet de d finir la “d riv e” de la fonction δ ; on trouve ainsi :

$$\int_{-\infty}^{+\infty} f(x)\delta'(x-x_0) dx = [\delta(x-x_0)f(x)]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} f'(x)\delta(x-x_0) dx = -f'(x_0) ; \tag{9.62}$$

en prenant   nouveau $x_0 = 0$, ceci s’ crit formellement :

$$\langle \delta', f \rangle = -f'(0) . \tag{9.63}$$

$\delta(x-x_0)$ est une id alisation de fonctions $\delta_\varepsilon(x-x_0)$, r guli res et tr s pointues dont l’aire vaut 1 (pr curseurs de δ). Si leur largeur autour de x_0 est ε , alors leur hauteur au maximum est d’ordre $\frac{1}{\varepsilon}$. Pour ces pr curseurs, on a les relations approch es :

$$\int_a^b f(x)\delta_\varepsilon(x-x_0) dx \simeq f(x_0) , \quad \int_a^b \delta_\varepsilon(x-x_0) dx \simeq 1 \quad (a < x_0 < b, x_0 - a, b - x_0 \gg \varepsilon) , \tag{9.64}$$

  la condition expresse que $f(x)$ soit lentement variable   l’ chelle ε . En raisonnant intuitivement avec les δ_ε et en passant   la limite $\varepsilon \rightarrow 0+$, on peut retrouver tr s rapidement tous les r sultats utiles concernant les r gles op rationnelles relatives   δ et   ses d riv es.

²⁶On connait  galement un autre moyen pour trouver une telle solution, c’est la m thode dite de *variation des constantes*.

En particulier, l'intégrale d'une δ_ε , $\int_a^x \delta_\varepsilon(x-x_0)dx$ ($a \ll x_0$), est une fonction de x , qui passe rapidement, sur une échelle d'ordre ε , de la valeur zéro à la valeur un (par exemple, la fonction $\delta_{\text{th},\varepsilon}(x-x_0) = \frac{1}{2}(1+\tanh \frac{x-x_0}{\varepsilon})$, avec $\varepsilon \ll 1$). À la limite $\varepsilon = 0+$, la primitive de δ est la fonction échelon-unité, nulle si $x < x_0$, égale à 1 si $x > x_0$ (sans qu'il soit nécessaire à ce stade de la définir en $x = x_0$). Si toutefois on utilise comme précurseur $\frac{1}{\pi\varepsilon} \sin \frac{x-x_0}{\varepsilon}$, le résultat de Dirichlet permet d'écrire très précisément²⁷ :

$$\int_{-\infty}^x \delta(x' - x_0) dx' = \begin{cases} 0 & \forall x < x_0 \\ \frac{1}{2} & \text{si } x = x_0 \\ 1 & \forall x > x_0 \end{cases} \equiv \theta(x - x_0) ; \quad (9.65)$$

(il est en effet parfois nécessaire, pour des raisons physiques, de préciser la valeur en x_0) ; il en va d'ailleurs ainsi quand on prend $\delta_{\text{th},\varepsilon}(x - x_0)$. Quoi qu'il en soit dans les détails, le résultat à retenir est :

$$\theta(x - x_0) = \int_{-\infty}^x \delta(x' - x_0) dx' , \quad \frac{d}{dx} \theta(x - x_0) = \delta(x - x_0) . \quad (9.66)$$

On note à nouveau que l'opération d'intégration adoucit les irrégularités²⁸ : δ a deux sauts d'amplitude infinie au même point (l'un dans un sens, l'autre dans le sens contraire), alors que sa primitive n'a qu'un saut fini, d'amplitude unité d'ailleurs.

Ce phénomène se confirme maintenant en prenant la primitive de la fonction échelon :

$$\int_{a < x_0}^x \theta(x - x_0) dx = \begin{cases} 0 & \forall x < x_0 \\ x - x_0 & \forall x > x_0 \end{cases} , \quad (9.67)$$

qui est une fonction continue, y compris en x_0 . Ainsi, toute primitive non-triviale de θ , fonction discontinue, est une fonction continue. On peut aussi retenir les résultats inverses : si une fonction continue présente une rupture de pente en un point x_0 , sa dérivée a un saut fini, sa dérivée seconde contient un terme proportionnel à δ , etc :

$$f'(x) = f'_{\text{reg}}(x) + [f(x_0+0) - f(x_0-0)] \theta(x - x_0) , \quad f''(x) = f''_{\text{reg}}(x) + [f(x_0+0) - f(x_0-0)] \delta(x - x_0) , \quad (9.68)$$

où f'_{reg} et f''_{reg} désignent les dérivées obtenues par les moyens usuels.

Ainsi, la solution de l'équation²⁹ $f'' - g\delta(x) = 0$ est de la forme $f(x) = Ax + B + C|x|$, celle de $f' + g\delta(x - x_0) = 0$ est $f(x) = A + B\theta(x - x_0)$, etc. Dans tous les cas, le saut fini se produit dans la dérivée $f^{(N-1)}$ si l'équation est d'ordre N (si $N \geq 2$, la fonction elle-même est donc continue). Pour trouver ce saut, il suffit d'intégrer de part et d'autre du point de concentration de la fonction δ . Par exemple, pour la première équation, on trouve ainsi $f'(+0) - f'(-0) = g$, ce qui donne $2C = g$, d'où la solution générale $f(x) = Ax + B + \frac{g}{2}|x|$; pour la deuxième équation, on a $f(x_0+0) - f(x_0-0) = -g$, d'où $B = -g$ et la solution générale $f(x) = A - g\theta(x - x_0)$.

9.6.2 Définition des fonctions de Green

Soit une équation différentielle linéaire inhomogène :

$$\sum_{n=0}^N a_n(x) f^{(n)}(x) = \phi(x) \iff \hat{L}(x) f(x) = \phi(x) , \quad \hat{L}(x) \stackrel{\text{d\u00e9f}}{=} \sum_{n=0}^N a_n(x) \frac{d^n}{dx^n} . \quad (9.69)$$

²⁷La fonction θ définie en (9.65) diffère donc de la fonction de Heaviside $Y(t)$ introduite dans le ch. 7 ; elle diffère également de la fonction Θ introduite dans le ch. 8 à propos des fonctions de répartition.

²⁸Comme la musique, l'intégration adoucit les mœurs.

²⁹L'équation de Poisson $\Delta V = \frac{\rho}{\varepsilon_0}$ prend la forme $V'' = \frac{q\delta(x)}{\varepsilon_0}$ sur \mathbb{R} et pour une charge ponctuelle située à l'origine. C'est pourquoi, à une dimension d'espace, le potentiel coulombien, défini comme satisfaisant l'équation de Poisson, varie comme $|x|$, et non pas comme $\frac{1}{|x|}$ ($V(x) = -\frac{q}{2\varepsilon_0}|x| + C^{\text{ste}}$).

La (les) fonction(s) de Green de cette équation sont les fonctions satisfaisant l'équation où la source ϕ a été remplacée³⁰ par $\delta(x - x')$. Si on note $G(x, x')$ la fonction de Green, alors par définition de G :

$$\sum_{n=0}^N a_n(x) \frac{\partial^n}{\partial x^n} G(x, x') \stackrel{\text{d}\acute{\text{e}}\text{f}}{=} \delta(x - x') \iff \hat{L}(x) G(x, x') \stackrel{\text{d}\acute{\text{e}}\text{f}}{=} \delta(x - x') . \quad (9.70)$$

Ceci étant, une solution de (9.69) s'exprime alors comme :

$$f(x) = \int_{-\infty}^{+\infty} G(x, x') \phi(x') dx' . \quad (9.71)$$

En effet, faisant passer $\hat{L}(x)$ sous l'intégrale en x' (commutant intégration et dérivations), on a selon (9.70) :

$$\hat{L}(x) f(x) = \int_{-\infty}^{+\infty} \hat{L}(x) G(x, x') \phi(x') dx' = \int_{-\infty}^{+\infty} \delta(x - x') \phi(x') dx' = \phi(x) , \quad (9.72)$$

qui reconstitue bien l'équation satisfaite par la fonction f . La fonction ainsi construite est donc *une* solution (particulière) de l'équation (9.69), dont la solution générale s'écrit :

$$f(x) = f_h(x) + \int_{-\infty}^{+\infty} G(x, x') \phi(x') dx' , \quad (9.73)$$

où $f_h(x)$ désigne la solution générale de l'équation homogène associée $\hat{L}f = 0$.

Autrement dit, une fois que l'on connaît la fonction de Green d'une équation différentielle, une solution particulière s'obtient par une simple intégrale, conformément à (9.71). Tout le travail consiste ainsi³¹ à déterminer la fonction de Green ; celui-ci fait, le passage d'une source à une autre se fait sans labour supplémentaire, ou presque³² ; c'est là l'un des avantages de la méthode de Green. En outre, comme le premier membre de (9.69) est caractéristique du système étudié (il en contient toute la dynamique interne), alors que le second membre (la source) représente une action extérieure sur ce système, la fonction de Green contient toute l'information sur la dynamique interne de ce système.

La difficulté est ainsi exclusivement reportée sur la détermination de la *bonne* (voir plus loin) fonction de Green du problème physique posé. Évidemment, pour une équation dont les coefficients a_n sont constants, il n'y a pas de difficulté majeure ; au contraire, et comme pour les équations différentielles, la dépendance éventuelle $a_n(x)$ augmente considérablement la difficulté de trouver effectivement $G(x, x')$.

Les écritures ci-dessus supposent toutes les grandeurs sans dimension, les fonctions et les variables. À l'inverse, si x est une longueur, alors on doit avoir $[a_n L^{-n}][G] = [L]^{-1}$ quel que soit n , de sorte que $[G] = [a_n^{-1} L^{n-1}]$.

9.6.3 Exemples

Il s'agit maintenant d'illustrer les idées précédentes dans des cas très simples, où d'ailleurs la solution peut être obtenue en quelques lignes, sans l'usage des fonctions de Green³³. Ce qui suit doit donc être considéré comme un simple exercice d'illustration des idées énoncées, où le risque est nul d'être aveuglé par une difficulté technique qui serait ici d'intérêt très secondaire.

En outre, cette illustration sera l'occasion de retrouver des idées importantes, comme par exemple le lien étroit entre causalité et analyticit . Enfin, ces cas tr s simples permettent de pr ciser le contenu physique

³⁰La fonction de Green satisfait donc la m me  quation que la fonction f inconnue, la source  tant remplac e par une perturbation de type percussion centr e en x' . En d'autres termes, si dans (9.69) on choisit $\phi(x) = \delta(x)$, la solution de cette  quation avec une source impulsion-unit  est $f(x) = G(x, x' = 0)$.

³¹supposant connue la solution g n rale de l' quation homog ne, mais d'ailleurs... (voir plus loin)

³²On trouve une situation comparable en M canique quantique, o  les  quations de Heisenberg permettent de trouver la dynamique des observables. La consid ration d'un  tat initial ou d'un autre, laborieuse dans la description de Schr dinger, est quasi-imm diate dans la vision de Heisenberg.

³³On pourrait objecter que l'on fait donner l'artillerie pour tuer une mouche, mais ce serait hors de propos.

de la fonction de Green, dont l'universalité dépasse très largement le cadre simpliste des exemples traités. Par exemple, on verra que quand la variable est le temps ($x \rightarrow t$), la partie imaginaire de la transformée de Fourier de la fonction de Green décrit, pour un système non purement mécanique (dissipatif), l'absorption d'énergie par ce système.

L'oscillateur harmonique isolé

Comme premier exemple, prenons l'équation de l'oscillateur harmonique (bille quasi-ponctuelle de masse m attachée à un ressort parfait de constante de raideur k), supposé *isolé* c'est-à-dire en l'absence de toute interaction avec un milieu produisant de l'amortissement (il s'agit donc d'un système purement mécanique, à peu de degrés de liberté – en fait un seul –, dont le mouvement s'effectue à énergie constante) ; dans les variables physiques $x =$ écart à l'équilibre, $t =$ temps, $\phi =$ force par unité de masse, l'équation avec une force (excitation) quelconque est³⁴ :

$$m\ddot{x} = -kx + F(t) \iff \ddot{x} + \omega_0^2 x = \phi(t), \quad k = m\omega_0^2, \quad \phi(t) = \frac{F(t)}{m}; \quad (9.74)$$

dans la suite, la quantité $\phi(t)$ sera appelée *source*. Ceci étant posé, l'équation satisfaite par la fonction de Green associée est :

$$\frac{\partial^2}{\partial t^2} G(t, t') + \omega_0^2 G(t, t') = \delta(t - t'); \quad (9.75)$$

ici, G a la dimension d'un temps. Dans les notations en cours, (9.73) se transcrit comme suit :

$$x(t) = x_h(t) + \int_{-\infty}^{+\infty} G(t, t')\phi(t')dt', \quad x_h(t) = a \cos \omega_0 t + b \sin \omega_0 t. \quad (9.76)$$

Compte tenu des remarques préliminaires (sous-section 9.6.1), le terme le plus singulier est celui de plus haute dérivée, ici une dérivée seconde, et c'est elle qui contient le saut infini imposé par la présence de δ au second membre. Il en résulte, selon (9.66), que la primitive de cette dernière (la dérivée première, donc) contient une fonction θ (un saut fini) et que la primitive de la dérivée (soit G) est continue en $t' = t$, conformément à (9.67).

Ce rappel étant fait, pour $t \neq t'$, (9.75) se réduit à :

$$\frac{\partial^2}{\partial t^2} G(t, t') + \omega_0^2 G(t, t') = 0 \quad \forall t \neq t'. \quad (9.77)$$

On peut donc écrire³⁵ :

$$G(t, t') = \begin{cases} C \cos \omega_0 t + D \sin \omega_0 t & \forall t < t' \\ A \cos \omega_0 t + B \sin \omega_0 t & \forall t > t' \end{cases}; \quad (9.78)$$

les quantités A, B, C et D sont les constantes d'intégration vis-à-vis de l'intégration en t , et sont donc *a priori* des fonctions de t' . La question est maintenant de les déterminer.

Le point à bien comprendre est que les deux expressions apparaissant dans (9.78) sont les expressions d'une seule et même fonction, $G(t, t')$, chacune étant valide dans une région donnée, $t < t'$ ou $t > t'$. Pour trouver les "constantes" A, B, C et D – ou en tout cas en déterminer certaines, il suffit d'exploiter ce que l'on sait de G et de sa dérivée. L'idée est d'utiliser les deux expressions (9.78) et, partant de chaque région, de venir vers leur frontière commune (ici, le point $t = t'$, *l'instant t' étant fixé*) pour y écrire les conditions sur G et sa dérivée. C'est typiquement ce que l'on appelle des conditions³⁶ de *raccordement*. Par exemple, on sait que G

³⁴Dans ces notations, la source est homogène à une accélération. Par ailleurs, dans (9.74), l'origine est prise au point d'équilibre ($x_e = 0$).

³⁵On note que, par définition, $G(t, t')$ a de part et d'autre de $t = t'$ l'allure d'une solution générale de l'équation homogène – c'est pourquoi la détermination de G est tributaire de la connaissance de cette solution générale, et c'est bien dans un tel contexte que G est efficace pour trouver une solution particulière.

³⁶Cette procédure est systématique en Mécanique quantique quand l'énergie potentielle $V(x)$ est continue par morceaux : en chaque point de discontinuité *finie* de V , la fonction propre $\psi(x)$ et sa dérivée sont continues. C'est seulement lorsque V a un saut infini (par exemple $V(x) \propto \delta(x - x_0)$, ou puits carré infini) que ψ' a un saut, ψ restant *toujours* continue puisque l'équation aux valeurs et fonctions propres $H\psi = E\psi$ est du second ordre d'espace (ou impulsion).

est continue, y compris en $t = t'$. On part donc de l'expression $C \cos \omega_0 t + D \sin \omega_0 t$, vraie à gauche de t' , et on fait tendre t vers t' par valeurs inférieures ; ceci donne la valeur à gauche :

$$G(t = t' - 0, t') = C \cos \omega_0 t' + D \sin \omega_0 t' ; \quad (9.79)$$

De même, la valeur à droite en $t = t'$ est $G(t = t' + 0, t') = A \cos \omega_0 t' + B \sin \omega_0 t'$; la continuité de G donne une première relation³⁷ entre les "constantes" A, B, C et D :

$$C \cos \omega_0 t' + D \sin \omega_0 t' = A \cos \omega_0 t' + B \sin \omega_0 t' ; \quad (9.80)$$

En revanche, on sait que la dérivée $\frac{\partial G}{\partial t}$ a un saut, en $t = t'$, dont on trouve l'amplitude en intégrant en t (9.75) de part et d'autre de t' . L'intégration de la dérivée seconde donne la variation de la dérivée première de part et d'autre du point t' ; par ailleurs, comme G est une fonction bornée continue, le terme en $\omega_0^2 G$ donne zéro (intégration sur un intervalle de mesure nulle) – quant au terme $\delta(t - t')$ du second membre, il donne 1 par définition de δ . Au total on obtient ainsi l'égalité³⁸ :

$$\left(\frac{\partial}{\partial t} G(t, t') \right)_{t=t'+0} - \left(\frac{\partial}{\partial t} G(t, t') \right)_{t=t'-0} = 1 . \quad (9.81)$$

En allant chercher à nouveau les deux expressions (9.78), on trouve la deuxième équation de raccordement :

$$-\omega_0 A \sin \omega_0 t' + \omega_0 B \cos \omega_0 t' - [-\omega_0 C \sin \omega_0 t' + \omega_0 D \cos \omega_0 t'] = 1 ; \quad (9.82)$$

Les deux équations (9.80) et (9.82) produisent finalement le système linéaire inhomogène :

$$\begin{cases} (A - C) \cos \omega_0 t' + (B - D) \sin \omega_0 t' = 0 \\ -(A - C) \sin \omega_0 t' + (B - D) \cos \omega_0 t' = \frac{1}{\omega_0} \end{cases} , \quad (9.83)$$

dont la solution est :

$$(A - C) = -\frac{1}{\omega_0} \sin \omega_0 t' , \quad (B - D) = \frac{1}{\omega_0} \cos \omega_0 t' . \quad (9.84)$$

Selon (9.78), on a donc :

$$G(t, t') = \begin{cases} C(t') \cos \omega_0 t + D(t') \sin \omega_0 t & \forall t < t' \\ \left[C(t') - \frac{1}{\omega_0} \sin \omega_0 t' \right] \cos \omega_0 t + \left[D(t') + \frac{1}{\omega_0} \cos \omega_0 t' \right] \sin \omega_0 t & \forall t > t' \end{cases} ; \quad (9.85)$$

ceci peut s'écrire d'une seule façon :

$$G(t, t') = \left[C(t') - \frac{1}{\omega_0} \theta(t - t') \sin \omega_0 t' \right] \cos \omega_0 t + \left[D(t') + \frac{1}{\omega_0} \theta(t - t') \cos \omega_0 t' \right] \sin \omega_0 t , \quad (9.86)$$

ou encore³⁹ :

$$G(t, t') = C(t') \cos \omega_0 t + D(t') \sin \omega_0 t + \frac{1}{\omega_0} \theta(t - t') \sin \omega_0(t - t') . \quad (9.87)$$

On vérifie aisément que cette fonction est bien solution de (9.75), quelles que soient les fonctions $C(t')$ et $D(t')$, et en constitue donc la solution générale. En effet, on a :

$$\frac{\partial^2}{\partial t^2} [\theta(t - t') \sin \omega_0(t - t')] = \delta'(t - t') \sin \omega_0(t - t') + 2\omega_0 \delta(t - t') \cos \omega_0(t - t') - \omega_0^2 \theta(t - t') \sin \omega_0(t - t') . \quad (9.88)$$

Par ailleurs $\langle \delta', f \rangle = -f'(0)$, $\langle \delta, f \rangle = f(0)$, d'où :

$$\delta'(t - t') \sin \omega_0(t - t') = -\omega_0 \delta(t - t') \cos \omega_0(t - t') = -\omega_0 \delta(t - t') , \quad \delta(t - t') \cos \omega_0(t - t') = \delta(t - t') . \quad (9.89)$$

³⁷ Attention : il faut se souvenir que t' est fixé ; ne pas déduire de cette équation $A = C$ et $B = D$ au motif que les sin et les cos sont des fonctions linéairement indépendantes !

³⁸ Même type de remarque que dans la note 37.

³⁹ On note que, à ce stade, $G(t, t')$ n'est pas une fonction de la seule différence $t - t'$.

Au total, il vient :

$$\frac{\partial^2}{\partial t^2} [\theta(t-t') \sin \omega_0(t-t')] = \omega_0 \delta(t-t') - \omega_0^2 \theta(t-t') \sin \omega_0(t-t') , \quad (9.90)$$

de sorte que :

$$\begin{aligned} \frac{\partial^2}{\partial t^2} G(t, t') &= -\omega_0^2 [C(t') \cos \omega_0 t + D(t') \sin \omega_0 t] - \omega_0 \theta(t-t') \sin \omega_0(t-t') + \delta(t-t') \\ &= -\omega_0^2 G(t, t') + \delta(t-t') \quad \forall C, D \quad \text{CQFD} . \end{aligned} \quad (9.91)$$

Dans l'expression (9.87), il reste deux "constantes" indéterminées, $C(t')$ et $D(t')$, et c'est bien normal puisque l'équation (9.75) satisfaite par $G(t, t')$ est une équation aux dérivées partielles du second ordre, et qu'aucune condition supplémentaire n'a encore été imposée à $G(t, t')$. Pour l'instant, la solution générale de l'équation (9.74) a pour expression (9.76) où $G(t, t')$ est donnée en (9.87). Il reste ainsi 4 "constantes" encore indéterminées pour l'instant : les deux (vraies) constantes a et b dans (9.76) et les deux fonctions $C(t)$ et $D(t)$. Les conditions initiales spécifiant, par exemple, la position et la vitesse à un instant de référence donneront deux relations supplémentaires, mais il faut visiblement invoquer d'autres contraintes pour tout trouver.

Comme toujours, s'agissant d'un problème physique bien posé, ce sont des considérations physiques qui vont permettre de lever toute ambiguïté ou indétermination. Ici, c'est le principe de causalité que l'on invoque : il est certain que, en vertu de la causalité et selon (9.76), on doit avoir $G(t, t') = 0$ si $t < t'$, car la réponse x à l'instant t ne saurait dépendre de la valeur de la source à un instant *postérieur*. Revenant à l'expression (9.87) de la solution générale, l'exigence du principe de causalité conduit immédiatement à $C = D \equiv 0$.

Physiquement, la *bonne* fonction de Green est donc, quels que soient t et t' , donnée par (9.87) avec $C = D = 0$, soit :

$$G_{\text{av}}(t, t') = \frac{1}{\omega_0} \theta(t-t') \sin \omega_0(t-t') ; \quad (9.92)$$

l'indice av (pour fonction de Green *avancée*, on dit aussi fonction de Green *causale*)⁴⁰ rappelle qu'il s'agit de la fonction nulle pour $t < t'$. Le facteur $\theta(t-t')$, qui apparaît *spontanément* dans le calcul est là pour pour ça et fait ce qu'il faut.

Il est important de souligner ce fait essentiel : c'est l'application de conditions particulières, dictées par des considérations physiques, qui permet de lever toute ambiguïté pour la réponse attendue. Ici, c'est le Principe de causalité qui achève la détermination de la fonction de Green, imposant la condition aux limites⁴¹ $G(t, t') = 0$ si $t < t'$, *i.e.* $t - t' < 0$. En outre, on note que G est maintenant une fonction de la seule différence des temps $t - t'$: une fois les considérations physiques prises en compte, le mouvement physique de la bille ne dépend que de l'intervalle *écoulé* à partir d'un certain instant⁴².

Une chose doit aussi être remarquée : la fonction de Green (9.92) n'a pas ici de transformée de Fourier au sens usuel. Par exemple, pour la fonction avancée, l'écriture ($\mathcal{G}_{\text{av}} = \mathcal{F}[G_{\text{av}}]$) :

$$\mathcal{G}_{\text{av}}(\omega) = \int_{-\infty}^{+\infty} dt \frac{1}{\omega_0} \theta(t-t') \sin \omega_0(t-t') e^{i\omega(t-t')} = \frac{1}{\omega_0} \int_0^{+\infty} dt'' \sin \omega_0 t'' e^{i\omega t''} \quad (???) \quad (9.94)$$

⁴⁰On définit de même la fonction de Green *retardée* (ou anticausale), $G_{\text{ret}}(t, t')$ nulle, elle, pour $t > t'$, en prenant $C = \frac{1}{\omega_0} \sin \omega_0 t'$ et $D = -\frac{1}{\omega_0} \cos \omega_0 t'$, ce qui donne :

$$G_{\text{ret}}(t, t') = -\frac{1}{\omega_0} \theta(t'-t) \sin \omega_0(t-t') \equiv G_{\text{av}}(t', t) . \quad (9.93)$$

En quelque sorte, $G_{\text{av}}(t, t')$ propage le système du présent vers le futur, $G_{\text{ret}}(t, t')$ propage du présent vers le passé. La dernière égalité traduit la symétrie par renversement du temps pour un système purement mécanique (non-amorti, c'est-à-dire non dissipatif).

⁴¹De même, en Mécanique quantique, ce sont les conditions imposées aux fonctions propres – conditions *mathématiques* imposées en vertu du sens *physique* attribué à ces fonctions – qui produisent la quantification spontanée de l'énergie des états liés. Le même type de quantification se produit pour une corde vibrante dont les extrémités sont fixées.

⁴²On peut voir là la signature de l'uniformité du temps : que l'on démarre une *manip* à midi ou au goûter ne change pas le résultat observé au bout d'un quart d'heure.

n'a pas de sens, et doit donc être régularisée. Une façon de faire converger l'intégrand à l'infini est d'introduire "à la main" un facteur exponentiel convergent $e^{-\varepsilon t}$ ($\varepsilon > 0$!) et de prendre la limite $\varepsilon \rightarrow 0$ après intégration. On redéfinit ainsi :

$$\mathcal{G}_{av}(\omega) \stackrel{\text{d\'ef}}{=} \lim_{\varepsilon \rightarrow +0} \frac{1}{\omega_0} \int_0^{+\infty} dt \sin \omega_0 t e^{i\omega t} e^{-\varepsilon t} = \lim_{\varepsilon \rightarrow +0} \frac{1}{2\omega_0} \left(\frac{1}{\omega + \omega_0 + i\varepsilon} - \frac{1}{\omega - \omega_0 + i\varepsilon} \right). \quad (9.95)$$

Chacune des fractions rationnelles donne la partie principale de Cauchy, \mathcal{P} , et une fonction de Dirac ; au total, ainsi régularisée, la transformée de Fourier de la fonction de Green avancée est :

$$\mathcal{G}_{av}(\omega) = \frac{1}{2\omega_0} \left(\mathcal{P} \frac{1}{\omega + \omega_0} - i\pi\delta(\omega + \omega_0) - \mathcal{P} \frac{1}{\omega - \omega_0} + i\pi\delta(\omega - \omega_0) \right). \quad (9.96)$$

La justification physique d'une telle régularisation – *a priori* assez artificielle – apparaîtra clairement dans l'étude de l'oscillateur amorti (voir ci-dessous) – dont la fonction de Green, au contraire, possède tout naturellement une transformée de Fourier.

Quoi qu'il en soit, admettant pour l'instant cette régularisation pour les transformées de Fourier et par le théorème de convolution appliqué à (9.71), la transformée $X(\omega) = \mathcal{F}[x(t)](\omega)$ est donnée par⁴³ :

$$X(\omega) = \frac{1}{2}(a - ib)\delta(\omega + \omega_0) + \frac{1}{2}(a + ib)\delta(\omega - \omega_0) + \mathcal{G}_{av}(\omega)\Phi(\omega), \quad (9.97)$$

où $\Phi = \mathcal{F}[\phi]$, et où a et b sont les constantes apparaissant en (9.76).

En définitive, la solution la plus générale et physiquement acceptable du mouvement de l'oscillateur mécanique en présence de la source $\phi(t)$ est⁴⁴ :

$$x(t) = x_h(t) + \frac{1}{\omega_0} \int_{-\infty}^t \sin \omega_0(t - t')\phi(t') dt', \quad (9.98)$$

où la borne supérieure de l'intégrale prend en compte le fait que $G_{av}(t, t')$ est nulle si $t' > t$. On peut maintenant fixer les constantes a et b contenues dans $x_h(t)$. Par exemple, si on se dit qu'à un certain instant t_0 , position et vitesse valent respectivement x_0 et v_0 , on peut écrire :

$$x_0 = a \cos \omega_0 t_0 + b \sin \omega_0 t_0 + \frac{1}{\omega_0} \int_{-\infty}^{t_0} \sin \omega_0(t_0 - t')\phi(t') dt', \quad (9.99)$$

et :

$$v_0 = -a\omega_0 \sin \omega_0 t_0 + b\omega_0 \cos \omega_0 t_0 + \int_{-\infty}^{t_0} \cos \omega_0(t_0 - t')\phi(t') dt', \quad (9.100)$$

ce qui fournit deux équations⁴⁵ pour les deux constantes a et b .

Ainsi, la fonction de Green étant trouvée, la solution relative à une source donnée s'obtient par une simple intégration ; un changement de source n'exige que le calcul d'une seule nouvelle intégrale : le travail est fait une fois pour toutes par la détermination explicite de la bonne fonction de Green associée à l'équation considérée.

Pour illustrer ceci, supposons que l'oscillateur, au repos jusqu'à un certain instant conventionnellement pris comme origine $t = 0$, est soumis à partir de $t = 0$ à une perturbation $\phi(t)$ non nulle (*i.e.* $\phi(t) = Y(t)\tilde{\phi}(t)$). Dans ces conditions, les deux constantes a et b sont nulles, comme on le voit en écrivant les deux équations (9.99) et (9.99) avec $t_0 = 0$: comme $\phi(t) = 0$ pour $t < 0$, les deux intégrales sont encore nulles à $t = 0$, et le système pour a et b n'a que la solution triviale $a = b = 0$, d'où il résulte $x_h(t) \equiv 0$. Dans ces conditions précises, l'expression de la solution se réduit à :

$$x(t) = \int_{-\infty}^{+\infty} G_{av}(t, t')\phi(t') dt' = \frac{1}{\omega_0} \int_0^t \sin \omega_0(t - t')\phi(t') dt' \quad (t > 0), \quad (9.101)$$

⁴³Se souvenir que $\mathcal{F}[e^{-i\omega_0 t}] = \delta(\omega - \omega_0)$.

⁴⁴La source étant homogène à une accélération, le second membre de (9.98) est bien une longueur.

⁴⁵Le système (9.99) et (9.100) a un déterminant égal à 1, et a donc toujours une solution et une seule.

où la borne inférieure résulte du fait que $\phi(t)$ est maintenant supposée nulle à $t < 0$.

Par exemple, si on applique de l'extérieur une force constante⁴⁶ $\frac{F_0}{m}$ à partir de $t = 0$, alors la source est $\phi(t) = \frac{F_0}{m}Y(t)$; l'oscillateur étant toujours initialement au repos (avant l'arrivée de la perturbation), la solution est :

$$x(t) = \frac{F_0}{m\omega_0} \int_0^t \sin \omega_0(t-t') dt' = \frac{F_0}{m\omega_0^2}(1 - \cos \omega_0 t) . \quad (9.102)$$

L'oscillation est toujours harmonique, mais s'effectue maintenant de part et d'autre d'une nouvelle position d'équilibre d'abscisse $x_{\text{eq}} = \frac{F_0}{m\omega_0^2} \equiv \frac{F_0}{k}$, correspondant au point d'équilibre entre la force de rappel vers $x = 0$ et la force constante F_0 appliquée en plus de l'extérieur ($m\omega_0^2 x_{\text{eq}} \equiv kx_{\text{eq}}$); (9.102) se réécrit comme :

$$x(t) = x_{\text{eq}}(1 - \cos \omega_0 t) . \quad (9.103)$$

De fait, l'application soudaine d'une force constante ne fait que déplacer le point d'équilibre autour duquel s'effectue l'oscillation, dont la fréquence est inchangée⁴⁷. Après application de la force, l'énergie (nulle avant) est constante et égale à $\frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega_0^2(x - x_{\text{eq}})^2 = \frac{F_0^2}{2m\omega_0^2} \equiv \frac{1}{2}m\omega_0^2 x_{\text{eq}}^2$: c'est l'énergie potentielle de la bille en $x = 0$ quand le point d'équilibre est x_{eq} .

L'oscillateur harmonique amorti

Il vaut la peine de regarder le cas de l'oscillateur en présence de frottement, afin d'exhiber les différences qualitatives entre un système *réel* (en pratique, il y a toujours de l'amortissement) et l'oscillateur isolé, qui est toujours un peu une vue de l'esprit : physiquement on doit considérer le cas non-amorti comme l'idéalisation d'un oscillateur réel observé sur une échelle de temps courte par rapport à l'amortissement (soit $t \ll \gamma^{-1}$ dans la notation qui va être introduite). Pour s'en tenir à la description la plus simple, on choisit un frottement proportionnel à la vitesse (frottement fluide). Dans ces conditions, l'équation fondamentale en présence d'une source $F(t)$ est :

$$m\ddot{x} = -\alpha\dot{x} - kx + F(t) \iff \ddot{x} + \gamma\dot{x} + \omega_0^2 x = \phi(t) \quad (\gamma > 0, F = m\phi) . \quad (9.104)$$

γ^{-1} a la dimension d'un temps, c'est une échelle de temps supplémentaire par rapport à ω_0^{-1} qui (à un facteur 2π près) donne la période d'oscillation. Il existe ainsi désormais un paramètre sans dimension⁴⁸, $Q \equiv \frac{\omega_0}{\gamma}$, dont les valeurs grandes ou petites devant l'unité donneront deux types de comportements qualitativement différents (respectivement : régime sous-amorti et régime sur-amorti), le cas limite $Q = +\infty$ correspondant à l'oscillateur idéal sans amortissement (isolé).

Ceci étant, l'équation satisfaite par la fonction de Green associée à (9.104) est alors :

$$\frac{\partial^2}{\partial t^2} G(t, t') + \gamma \frac{\partial}{\partial t} G(t, t') + \omega_0^2 G(t, t') = \delta(t - t') . \quad (9.105)$$

La procédure suivie maintenant est parallèle à celle entreprise précédemment. Pour $t \neq t'$, (9.105) est :

$$\frac{\partial^2}{\partial t^2} G(t, t') + \gamma \frac{\partial}{\partial t} G(t, t') + \omega_0^2 G(t, t') = 0 \quad \forall t \neq t' . \quad (9.106)$$

Afin de simplifier les écritures, on écrit la solution de cette équation en combinaison linéaire d'exponentielles complexes (et non plus en sin et cos) :

$$G(t, t') = \begin{cases} C_+ e^{i\omega_+ t} + C_- e^{i\omega_- t} & \forall t < t' \\ A_+ e^{i\omega_+ t} + A_- e^{i\omega_- t} & \forall t > t' \end{cases} , \quad (9.107)$$

⁴⁶Exemple : une particule chargée liée harmoniquement, que l'on soumet à $t \geq 0$ à un champ électrique constant.

⁴⁷Le même phénomène se produit en Mécanique quantique.

⁴⁸ Q mesure ce qu'il est convenu d'appeler *facteur de qualité* de l'oscillateur : plus Q est grand, plus la résonance est fine.

où ω_{\pm} désigne les racines de l'équation $-\omega^2 + i\gamma\omega + \omega_0^2 = 0$, soit $\omega_{\pm} = \pm\sqrt{\omega_0^2 - \frac{\gamma^2}{4}} + i\frac{\gamma}{2} \equiv \pm\tilde{\omega}_0 + i\frac{\gamma}{2}$: maintenant, en conséquence de l'amortissement, les fréquences caractéristiques ont une partie imaginaire finie (mais toujours positive, quels que soient⁴⁹ ω_0 et γ).

Les équations de raccordement s'écrivent ici :

$$C_+e^{i\omega_+t'} + C_-e^{i\omega_-t'} = A_+e^{i\omega_+t'} + A_-e^{i\omega_-t'} , \quad (9.108)$$

$$i\omega_+A_+e^{i\omega_+t'} + i\omega_-A_-e^{i\omega_-t'} - [i\omega_+C_+e^{i\omega_+t'} + i\omega_-C_-e^{i\omega_-t'}] = 1 . \quad (9.109)$$

Afin d'aller plus vite, cherchons d'emblée la fonction de Green avancée, ce qui revient à prendre $C_{\pm} = 0$. On trouve alors, après calcul⁵⁰ :

$$G_{av}(t, t') = \frac{1}{\tilde{\omega}_0} \theta(t - t') e^{-\frac{\gamma}{2}(t-t')} \sin \tilde{\omega}_0(t - t') . \quad (9.111)$$

Ici apparaît une première différence qualitative avec le cas non amorti – la transformée de Fourier de G_{av} existe au sens usuel :

$$\mathcal{G}_{av}(\omega) = \frac{1}{\tilde{\omega}_0} \int_0^{+\infty} dt e^{-\frac{\gamma}{2}t} \sin \tilde{\omega}_0 t e^{i\omega t} = \frac{1}{2\tilde{\omega}_0} \left(\frac{1}{\omega + \tilde{\omega}_0 + i\frac{\gamma}{2}} - \frac{1}{\omega - \tilde{\omega}_0 + i\frac{\gamma}{2}} \right) = \frac{1}{\tilde{\omega}_0^2 - (\omega + i\frac{\gamma}{2})^2} ; \quad (9.112)$$

Cette expression montre que le paramètre de frottement γ (donnant l'amortissement) se substitue tout naturellement au paramètre "fictif" ε introduit en (9.95) pour effectuer la régularisation. Comme tous les systèmes physiques sont amortis (même "infinitement peu"), la procédure de régularisation (9.95), somme toute *a priori* assez artificielle, s'en trouve justifiée – au moins dans l'hypothèse du frottement fluide. Notons que, en tant que transformée de Fourier d'une fonction à valeurs réelles, \mathcal{G}_{av} possède la symétrie :

$$\mathcal{G}_{av}(-\omega) = [\mathcal{G}_{av}(\omega)]^* . \quad (9.113)$$

Ceci signifie que la partie réelle de $\mathcal{G}_{av}(\omega)$ est paire, que sa partie imaginaire est impaire⁵¹ :

$$\mathcal{G}_{av}(\omega) = \mathcal{G}_{av1}(\omega) + i\mathcal{G}_{av2}(\omega) : \quad \mathcal{G}_{av1}(-\omega) = \mathcal{G}_{av1}(\omega) , \quad \mathcal{G}_{av2}(-\omega) = -\mathcal{G}_{av2}(\omega) . \quad (9.114)$$

Par ailleurs, l'expression intégrale de $G_{av}(t, t')$:

$$G_{av}(t, t') = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\omega(t-t')} \mathcal{G}_{av}(\omega) d\omega \quad (9.115)$$

reflète bien une fois encore le lien entre analyticit  et causalit . En effet, pour $t > t'$, on referme le contour par le bas, et on ramasse les deux r sidents de $\mathcal{G}_{av}(z)$ en $z_{\pm} = -i\frac{\gamma}{2} \pm \tilde{\omega}_0$. Au contraire, pour $t < t'$, on ferme par en haut, mais comme il n'y a aucune singularit , l'int grale est nulle comme il se doit. Avec les d finitions en cours⁵², le Principe de causalit  se traduit par le fait que $\mathcal{G}_{av}(\omega)$ est analytique dans le demi-plan sup rieur. Quand l'amortissement est nul, les p les se trouvent sur la droite r elle, et il faut bien r gulariser, d'une fa on ou d'une autre, puisque le chemin d'int gration bute sur les singularit s, d'o  la partie principale de Cauchy et la fonction de Dirac.

⁴⁹Si $\gamma < 2\omega_0$, les ω_{\pm} ont une partie r elle finie (r gime sous-amorti, frottement faible). Dans le cas contraire, $\gamma > 2\omega_0$, on a $\omega_{\pm} = i\left(\frac{\gamma}{2} \pm \sqrt{\frac{\gamma^2}{4} - \omega_0^2}\right)$: les ω_{\pm} sont alors imaginaires pures (r gime sur-amorti, frottement fort, pas d'oscillation pendant relaxation), leurs parties imaginaires  tant toutes deux positives.

⁵⁰De la m me fa on, la fonction retard e est :

$$G_{ret}(t, t') = -\frac{1}{\tilde{\omega}_0} \theta(t' - t) e^{-\gamma(t-t')} \sin \tilde{\omega}_0(t - t') \neq G_{av}(t', t) . \quad (9.110)$$

La sym trie de renversement du temps est bris e par l'amortissement, comparer avec la note 40.

⁵¹Attention : l'usage est fr quent de noter \mathcal{G}' et \mathcal{G}'' respectivement les parties r elle et imaginaire de \mathcal{G} , mais il ne s'agit pas de ses d riv es !

⁵²Si on inverse la convention de d finition de la transformation de Fourier, les deux demi-plans s' changent.

Il a été affirmé plus haut que la fonction de Green contient toute la dynamique interne du système. Ceci se voit bien sur la transformée de Fourier $\mathcal{G}_{\text{av}}(\omega)$. Selon (9.112), et compte tenu de l'expression de $\tilde{\omega}_0$, on a :

$$\mathcal{G}_{\text{av}}(\omega) = \frac{1}{\omega_0^2 - \omega^2 - i\gamma\omega} = \frac{\omega_0^2 - \omega^2}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2} + i \frac{\gamma\omega}{(\omega^2 - \omega_0^2)^2 + \gamma^2\omega^2} . \quad (9.116)$$

Notamment, le module de $\mathcal{G}_{\text{av}}(\omega)$ présente un pic très marqué lorsque $\gamma \ll \omega_0$ ($Q \gg 1$), ce qui traduit le phénomène banal de *résonance* lorsqu'on excite un oscillateur (sous-amorti) à une pulsation ω voisine de sa pulsation propre ω_0 .

L'importance de la fonction de Green tient aussi au fait qu'elle décrit l'absorption d'énergie par l'oscillateur. En effet, $m\phi(t) = F(t)$ est la force *extérieure* ; par conséquent, le travail de cette force est par définition l'énergie *reçue* par l'oscillateur : $dW = Fdx = F\dot{x}dt$. L'énergie totale fournie à (absorbée par) l'oscillateur est donc⁵³ :

$$\Delta E = \int_{-\infty}^{+\infty} F\dot{x}dt \equiv \left[\int_{-\infty}^{+\infty} F\dot{x}e^{i\omega t} dt \right]_{\omega=0} = \mathcal{F}[F\dot{x}](\omega=0) . \quad (9.117)$$

On a $\mathcal{F}[\dot{x}] = -i\omega\mathcal{F}[x]$; en supposant l'oscillateur au repos avant l'arrivée de la perturbation ϕ , la solution homogène est identiquement nulle et il reste seulement $\mathcal{F}[x] = \mathcal{G}(\omega)\Phi(\omega)$ (voir (9.98)). Par le théorème de convolution, il vient alors :

$$\mathcal{F}[F\dot{x}](\omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} m\Phi(\omega - \omega')[-i\omega'\mathcal{G}_{\text{av}}(\omega')\Phi(\omega')] d\omega' , \quad (9.118)$$

de sorte que l'énergie absorbée est donnée par :

$$\Delta E = \frac{m}{2i\pi} \int_{-\infty}^{+\infty} \Phi(-\omega')\omega'\mathcal{G}_{\text{av}}(\omega')\Phi(\omega') d\omega' . \quad (9.119)$$

En outre, comme Φ est la transformée de Fourier d'une fonction réelle, $\Phi(-\omega) = \Phi(\omega)^*$; d'où :

$$\Delta E = \frac{m}{2i\pi} \int_{-\infty}^{+\infty} \omega'\mathcal{G}_{\text{av}}(\omega')|\Phi(\omega')|^2 d\omega' , \quad (9.120)$$

et, en vertu des symétries exprimées par (9.114), seule la partie imaginaire contribue à l'intégrale :

$$\Delta E = \frac{m}{2\pi} \int_{-\infty}^{+\infty} \omega' |\Phi(\omega')|^2 \Im\mathcal{G}_{\text{av}}(\omega') d\omega' = \frac{m}{\pi} \int_0^{+\infty} \omega' |\Phi(\omega')|^2 \Im\mathcal{G}_{\text{av}}(\omega') d\omega' . \quad (9.121)$$

Ainsi, la partie imaginaire de la fonction de Green avancée donne-t-elle directement l'énergie absorbée par l'oscillateur. Dans le cas d'une source quasi-monochromatique à ω_e , de durée $T \gg \omega_e^{-1}$, $\Phi(\omega)$ est un précurseur $\delta_{1/T}(\omega \pm \omega_e)$ de $\delta(\omega \pm \omega_e)$, qui va filtrer l'intégrand ; au total, on trouve que l'énergie absorbée est proportionnelle à $T\Im\mathcal{G}(\omega_e)$: l'énergie moyenne absorbée par unité de temps, $\frac{\Delta E}{T}$, varie comme $\Im\mathcal{G}(\omega_e)$, elle est d'autant plus grande que la pulsation de la source est voisine de la pulsation propre ω_0 de l'oscillateur. En outre, dans la limite $\gamma = 0+$, $\Delta E = 0$: si l'oscillateur est non-amorti (isolé), son énergie est constante... comme il se doit.

Tous ces résultats s'appliquent à un oscillateur harmonique amorti, quelle que soit sa nature physique précise. Par exemple, pour un circuit RLC aux bornes duquel on applique la tension $v(t)$, l'équation pour l'intensité $i(t)$ est :

$$L \frac{d^2 i}{dt^2} + R \frac{di}{dt} + \frac{1}{C} i = \dot{v}(t) . \quad (9.122)$$

Par rapport à l'oscillateur matériel de référence, le dictionnaire des paramètres est le suivant :

$$m \leftrightarrow L , \quad R \leftrightarrow \gamma , \quad \frac{1}{C} \leftrightarrow \omega_0^2 . \quad (9.123)$$

L'inductance joue bien un rôle inertiel, en s'opposant aux variations brusques de courant (elle arrondit les sauts), la résistance est bien le terme dissipatif, quant à C^{-1} , c'est bien ce qui rappelle l'intensité vers la valeur zéro. La condition de résonance du circuit est $LC\omega^2 = 1$, en conformité avec $\omega = \omega_0$ pour l'oscillateur matériel.

⁵³Ici, pas de terme de chaleur $\dot{d}Q$!