



cel-00519301, version 1 - 20 Sep 2010

Cours d'analyse

Jacques Harthong

Ce document est sous licence Creative Commons: paternité; pas d'utilisation commerciale; partage des conditions initiales à l'identique; 3.0 France

<http://creativecommons.org/licenses/by-nc-sa/3.0/deed.fr>





Table des matières

Avant-propos **v**

Un mot sur Jacques Harthong **v**

Le photocopillage **vi**

Un gros mensonge **vi**

Comment fonctionne réellement l'édition technique ou universitaire **vii**

Les publications de recherche **ix**

Conclusion **x**

I Théorèmes sur les intégrales. **1**

II Formule de Green **11**

II.1 Intégrales curvilignes **11**

II.2 Intégrales doubles **13**

II.3 Relations entre intégrales doubles et curvilignes **18**

II.4 Domaines ne vérifiant pas la condition 2 **19**

II.5 Intégrales curvilignes réductibles par quadrature **22**

II.6 Domaines à trous **24**

II.7 Homologie des lacets **26**

II.8 Intégrales curvilignes à variable complexe **28**

III Fonctions analytiques **31**

III.1 Une propriété des polynômes **31**

III.2 Fonctions analytiques **33**

III.3 Séries entières convergentes **37**

III.4 Théorie de Cauchy **41**

III.5 Fonctions multiformes 49

IV Calcul des résidus 57

IV.1 Séries de Laurent 57

IV.2 Théorème des résidus 60

IV.3 Calculs d'intégrales définies 63

IV.4 Comment calculer pratiquement les résidus 68

IV.5 Fonction d'Euler 70

IV.6 Fonctions puissance non entière 77

V Fonctions eulériennes 81

V.1 Présentation 81

V.2 Prolongements analytiques 82

V.3 Formule d'Euler 85

V.4 Dérivée de $\Gamma(z)$ 86

V.5 Développements eulériens 90

V.6 Intégrale de Hankel 92

VI Transformations conformes 97

VI.1 Transformations géométriques du plan 97

VI.2 Propriétés géométriques des fonctions analytiques 100

VI.3 Fonctions harmoniques 103

VI.4 Autres exemples 107

VI.4.1 Polynômes 107

VI.4.2 Rosettes 108

VII Transformation de Fourier 111

VII.1 Équation de la chaleur 111

VII.2 Transformation intégrale 115

VII.3 Principales propriétés de la transformation intégrale 120

VII.4 Notions de convergence 124

VII.5 Espace $\mathbb{L}_2(\mathbb{R})$ 126

VII.6 Transformation de Laplace 132

VIII Intégrales divergentes. 135

VIII.1 Calcul d'une intégrale semi-convergente 135

VIII.2 Valeur principale de Cauchy 138

VIII.3 Pseudo-fonctions de Hadamard 140

IX Théorie des distributions 153

IX.1 Modèle mathématique 153

IX.2 Définition des distributions 154

IX.3 Exemples 157

IX.4 Continuité dans l'espace $\mathcal{S}(\mathbb{R})$ 159

IX.5 Intégrales avec poids et extension 162

IX.6 À propos de l'espace $\mathcal{S}(\mathbb{R})$ 165

IX.7	Dérivation des distributions	168
IX.8	Transformation de Fourier des distributions	171
IX.8.1	Transformée de Fourier de 1	171
IX.8.2	Transformée de Fourier d'un polynôme	172
IX.8.3	Transformée de Fourier de δ et de ses dérivées	172
IX.8.4	Transformée de Fourier de $1/[ix^\alpha]_2$	173
IX.8.5	Transformée de Fourier de e^{ix^2}	173
IX.9	Limites de distributions	174

X Calculer avec les distributions. 179

X.1	Dérivées de fonctions non dérivables	179
X.2	Multiplication et convolution des distributions	186
X.3	Exemples et applications des produits et convolutions	191
X.3.1	Convolution par les distributions de Dirac	191
X.3.2	Convolution par les dérivées de δ	192
X.3.3	Régularisation	193
X.3.4	Résolution d'équations différentielles	195
X.4	Famille Y_α	197

XI Espaces de Hilbert 203

XI.1	Espaces euclidiens de dimension infinie	203
XI.2	Espaces de Hilbert	206
XI.3	Bases orthonormées	210
XI.4	Exemples de bases orthonormées	212
XI.4.1	Polynômes de Legendre	212
XI.4.2	Fonctions et polynômes d'Hermite	215
XI.4.3	Fonctions et polynômes de Laguerre	217
XI.5	Théorèmes de Weierstrass	222

XII Opérateurs 227

XII.1	Dimension infinie	227
XII.2	Opérateurs continus et opérateurs discontinus fermés	227
XII.3	Valeurs propres et spectre d'un opérateur	232
XII.4	Oscillateur quantique	237
XII.5	Opérateurs auto-adjoints et unitaires	241
XII.6	Fonctions d'opérateurs	246
XII.7	Groupes unitaires	251
XII.8	Espace $\mathbb{L}_2(S)$ et harmoniques sphériques	254
XII.9	Théorie de l'atome d'hydrogène	258

Bibliographie 263

Index 265

A photograph of a person sitting on a wooden bench, playing a harmonica. The person's hands are visible, holding the harmonica. The background is slightly blurred, showing other people and a park-like setting. The image has a warm, golden-hour lighting.

Avant-propos

Un mot sur Jacques Harthong

Jacques Harthong (1948–2005) a été mathématicien et physicien. Depuis 1972, l'année de son recrutement, il est resté à l'Université Louis Pasteur de Strasbourg jusqu'à ce que sa maladie l'emporte en 2005.

Agrégé de Mathématiques en 1971, puis docteur d'État en 1981, Jacques Harthong avait le souci de faire des *mathématiques qui servent* ; l'exemple du moiré dans son hommage à Georges Reeb témoigne de cette volonté : doivent être résolus les problèmes qui se posent et pas seulement les problèmes que l'on se pose.

Ces problèmes qui se posent, Jacques Harthong est allé les chercher en Physique (il s'agit de la Physique au sens large : la Mécanique Quantique, la Physique Statistique, l'Optique, l'Infographie etc.), ce qui traduit un certain esprit d'ouverture, esprit qu'il a gardé jusqu'à sa disparition. Lorsque quelqu'un (qui pouvait être un étudiant) venait lui poser une question sur tel ou tel sujet, il avait tout de suite la réponse ou alors, il disait qu'il ne savait pas mais qu'il allait y réfléchir. Il revenait souvent avec la bonne réponse et un commentaire sur la pertinence de la question posée. Cette curiosité scientifique agrémentée d'une réflexion approfondie est une clé pour comprendre l'homme qui a appliqué le concept d'idée intérieure (emprunté à Caspar David Friedrich : *die Stimme seines Innern*) et que l'on pourrait comprendre comme suit : « il s'agit d'un sentiment strictement personnel, qui permet au mathématicien de juger, de manière absolument subjective, ce qui est intéressant pour lui, ce qui est *sa voie* » (sic).

Cette philosophie personnelle — ce code de conduite — explique certainement la pro-

fondeur des réflexions scientifiques de Jacques Harthong. Il s'est démarqué des autres scientifiques mais avec un revers de la médaille sans doute inévitable : la marginalité. Le « sentiment strictement personnel » s'est trouvé incompatible avec le corporatisme, l'appartenance à un courant scientifique et la contrainte administrative du nombre.

Dans cette manifestation, il n'est en aucun cas question d'être panégyrique. Il s'agit de présenter la richesse des thèmes étudiés par un homme dont le but avoué était de comprendre. En quelque sorte, il s'agit d'un bref retour vers la philosophie naturelle.

Le photocopillage

Un gros mensonge

Sur la page de garde de nombreux ouvrages que vous avez consultés ou utilisés pendant vos études, vous avez certainement remarqué parfois une exhortation contre le *photocopillage* : *le photocopillage tue le livre*, accompagnée d'un plaidoyer pour la protection des auteurs et de la propriété intellectuelle. On a cherché à vous culpabiliser de photocopier un ouvrage au lieu de l'acheter en librairie.

Sachez qu'il s'agit d'une manipulation exclusivement vouée à protéger le droit au profit des éditeurs. En effet les pires ennemis de la *juste rémunération d'une création intellectuelle* ne sont pas les étudiants qui photocopient, mais les éditeurs qui le plus souvent ne versent tout simplement pas cette juste rémunération aux auteurs. Ils peuvent se le permettre grâce à la seule loi qui est, toujours et à coup sûr, appliquée : la loi du plus fort (c'est-à-dire du plus riche). C'est pourquoi cette *exhortation contre le photocopillage* que vous avez certainement remarquée plus d'une fois est vraiment le comble de l'hypocrisie.

Jusqu'à une époque récente, aucune loi n'interdisait à un étudiant de photocopier un livre pour son usage personnel. Si toute une promo le faisait individuellement (chaque étudiant photocopiant lui-même le bouquin) il n'y avait pas d'infraction. Tout différent de la manipulation signalée ci-dessus est en effet le texte (parfaitement légal, lui) qui figure aussi sur les pages de garde (alinéa premier de l'Article 40) :

« La loi du 11 mars 1957 n'autorisant, aux termes des alinéas 2 et 3 de l'Article 41, d'une part, que les copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective et, d'autre part, que les analyses et courtes citations dans un but d'exemple et d'illustration, toute représentation ou reproduction intégrale, ou partielle, faite sans consentement de l'auteur ou de ses ayant-droits ou ayant-cause, est illicite. »

La technologie ayant énormément évolué depuis trente ans, les intérêts des éditeurs se sont de plus en plus heurtés à ses progrès. Ainsi jusqu'au début des années mille neuf cent soixante-dix la reproduction d'un livre n'était pas à la portée du simple particulier ; elle l'est devenue à partir des années mille neuf cent quatre-vingt grâce à la photocopieuse de bureau omniprésente. C'est pourquoi les éditeurs ont engagé une bataille juridique et obtenu de nouvelles lois plus restrictives. Il est devenu illégal de photocopier un livre même pour un usage privé. À chacune de ces batailles juridiques ou parlementaires, les éditeurs invoquent systématiquement la protection de la création intellectuelle, c'est-à-dire la protection de la

juste rémunération d'un artiste, d'un écrivain, d'un chercheur. Plus récemment le Web à son tour a commencé à heurter les intérêts des éditeurs et de la même façon ils essaient de contraindre les États à protéger leurs sources de profit contre cette évolution. À nouveau est entonné le lamento sur les pauvres auteurs dépossédés par des pilleurs.

Je voudrais que vous sachiez que ceux qui dépossèdent les auteurs sont surtout les éditeurs et non les *pilleurs*. La protection des artistes, écrivains, ou chercheurs est le dernier souci des éditeurs. Cet argument est de la pure hypocrisie, car toutes les batailles juridiques et parlementaires menées par le lobby de l'édition visent exclusivement à la protection des profits que les éditeurs, surtout les plus gros, obtiennent grâce à l'exploitation sans scrupule des artistes, écrivains, ou chercheurs. Bien sûr un écrivain ou un chanteur de rock célèbre touche des droits d'auteur conséquents ; mais c'est surtout parce que le montant de ces droits est alors suffisamment élevé pour que le recours à des avocats soit rentable : l'auteur est payé parce qu'il est assez riche pour obtenir d'être payé. Si le coût de l'action en justice est supérieur au montant des droits d'auteur, elle devient sans intérêt et les éditeurs le savent.

Comment fonctionne réellement l'édition technique ou universitaire

Autrefois l'éditeur était un auxiliaire indispensable de l'auteur : ce dernier ne pouvait écrire ses œuvres qu'à la plume d'oie, au stylo, ou à la machine à écrire mécanique et en un seul exemplaire, qu'il devait faire très attention à ne pas perdre, car sinon il ne lui restait plus qu'à tout réécrire. Les éditeurs recevaient ce manuscrit, le faisaient typographier, relire et corriger, imprimer et diffuser. Ils avaient aussi un service juridique pour vérifier que la publication ne contient rien d'illicite (cela peut échapper à un auteur sans qu'il ait eu la moindre mauvaise intention). Depuis qu'existent les traitements de textes et les divers logiciels de mise en page, les éditeurs ne s'occupent plus du tout de la typographie ni de la correction des épreuves ; ils se contentent de changer les polices et les formats de pages (et encore : souvent ils exigent que l'auteur le fasse lui-même d'après leurs normes). Les auteurs sont sommés de remettre une disquette contenant un texte prêt à la publication (plus de service fourni par l'éditeur). Les contrats stipulent que les auteurs sont responsables du contenu de leur manuscrit (plus de service juridique). En outre, bien que les contrats stipulent expressément que les auteurs recevront des droits de l'ordre de 10% sur les ventes, ces droits ne sont, le plus souvent, même pas payés, à moins que, comme je le disais plus haut, l'auteur ne saisisse la justice. Ainsi les éditeurs ne font plus rien que d'envoyer la disquette chez un imprimeur et de diffuser l'ouvrage auprès des libraires. Au passage, ils prélèvent un pourcentage élevé. Dans une maison d'édition sérieuse ces profits servent à payer les salaires des employés, les locaux, les investissements etc. mais évidemment la pression essentielle vient comme toujours de la nécessité de maintenir des dividendes suffisamment élevés aux actionnaires. Dans les maisons moins sérieuses, on se demande où est passé tout cet argent : après le dépôt de bilan on apprend en effet avec stupeur que les secrétaires n'étaient plus payées depuis des mois, que l'imprimeur n'avait plus été payé depuis un an etc.

Bien entendu, je sais que fort peu de jeunes se laissent impressionner par les tentatives d'intimidation contre le *photocopillage*. Je vois dans les laboratoires des thésards photocopier des ouvrages entiers. Félicitations ! (c'est devenu illégal). Je voudrais quand même, en tant

qu'auteur, fournir un témoignage concret contre certaines illusions que peuvent avoir les profanes envers l'édition.

Vous êtes peut-être étudiant ou jeune chercheur, ou simplement curieux. Vous avez étudié dans des manuels ou consulté des ouvrages plus avancés de recherche. Sans doute ne vous êtes-vous pas posé beaucoup de questions sur les auteurs de ces ouvrages, mais vous avez probablement pensé que ces gens sont très compétents et connus et peut-être même que les éditeurs viennent les courtiser pour les convaincre d'écrire. Sans doute croyez vous aussi que les auteurs gagnent de l'argent avec ces ouvrages.

Cela se passe très rarement ainsi !

Certes il arrive parfois que des universitaires soient ainsi courtisés, par exemple s'ils viennent d'obtenir le prix Nobel. On viendra alors leur demander d'écrire un ouvrage pour le grand public, exposant leur philosophie générale ou expliquant de manière très simple et accessible les travaux qui leur ont valu la prestigieuse récompense. De même, s'ils sont déjà connus pour des ouvrages qui se vendent bien, on viendra leur demander d'en écrire d'autres. C'est parce qu'il y a une perspective de profit par le nombre des ventes.

Toute autre est la situation des manuels didactiques ou des monographies de recherche. Ces ouvrages sont principalement achetés par des bibliothèques. Certains étudiants achètent leurs manuels pour en disposer plus commodément ou parce qu'ils aiment les livres ; de même certains chercheurs, sachant qu'ils travailleront longtemps sur tel sujet, se procurent la bible de leur domaine afin de l'avoir en permanence sous la main. Mais la majorité des exemplaires sont achetés par les bibliothèques et cela constitue un marché subventionné par l'État, qui est assuré quasi automatiquement. Les éditeurs, pour la plupart uniquement motivés par le profit (comme le montre amplement leur mépris absolu pour le contenu des ouvrages), savent cela. Ces marchés en quelque sorte garantis par l'État ont donc l'avantage de ne pas nécessiter de stratégies commerciales risquées et coûteuses. Les ouvrages se vendent en peu d'exemplaires — de l'ordre de quelques centaines, parfois mille en langue française, parfois jusqu'à dix-mille en langue anglaise — mais les profits sont assurés grâce aux faibles coûts de production et au financement public par les bibliothèques. Le comportement des agents économiques est ici analogue à celui des ours du Yellowstone National Park qui ont trouvé plus rentable de chercher leur nourriture dans les poubelles des touristes plutôt que — comme leurs ancêtres sauvages — dans la forêt naturelle. Si ce marché était *naturel* comme le veut la théorie économique, il faudrait que les éditeurs engagent de nombreux frais pour rendre l'ouvrage plus compétitif. Au contraire de cela, dans ce type d'édition on édite sans cesse de nouveaux manuscrits qui n'ont donc pas à être rachetés ; comme je l'ai déjà dit plus haut on se contente de faire imprimer directement le texte fourni par l'auteur sur une disquette (donc on n'a pas à payer de salaires pour la typographie informatique), on ne paie rien à ces auteurs qui en plus de la conception ont même fait la typographie ⁽¹⁾.

Si par un proche ou par votre activité professionnelle vous étiez familier des milieux de l'édition, vous sauriez que la plupart des auteurs connus ont été ou sont en procès avec leur éditeur. Sachez donc que cela s'explique ainsi : les éditeurs appliquent comme une règle

(1) Les contrats prévoient le paiement de droits, mais ces droits ne sont effectivement payés que si l'auteur engage un procès qui dans la plupart des cas lui coûtera plus cher que la juste rémunération qu'il peut revendiquer.

d'or de ne payer les droits d'auteur que s'il y a menace de contrainte judiciaire. Les auteurs riches ont donc toutes les chances de toucher des royalties substantielles, les pauvres par contre n'ont aucune chance. Ces derniers sont le plus souvent des universitaires qui écrivent des ouvrages didactiques ou des monographies de recherche, car le nombre d'exemplaires vendus pour ce genre d'ouvrage est en-dessous du seuil où un recours judiciaire vaut la peine d'être tenté. Dans le domaine de la littérature de fiction, les auteurs *pauvres* sont aussi les jeunes auteurs qui essaient de percer ; inutile de dire que ceux qui percent le doivent en général bien moins à leur talent littéraire qu'à leurs talents de businessmen !

Les éditeurs savent aussi que les universitaires qui veulent publier des manuels d'enseignement ou des monographies de recherche sont davantage motivés par la renommée que l'ouvrage leur rapportera que par les très modestes droits d'auteur ; il est donc d'autant moins probable qu'ils saisissent la justice. Ainsi je ne puis accuser unilatéralement les seuls éditeurs : la vanité des universitaires est elle aussi responsable de la situation que je dénonce et les éditeurs ne font que l'exploiter. Cela n'est certainement pas un argument pour vous dissuader de photocopier !

Les publications de recherche

Un autre filon exploité sans scrupule par de grandes maisons d'édition internationales est celui des revues professionnelles où les chercheurs publient leurs travaux. Dans le Big Science System que je m'efforce de décrire par ailleurs dans ces pages, le rayonnement d'un laboratoire ou d'un chercheur individuel se mesure au nombre de publications. Pas à leur qualité, mais à leur nombre. Cela est bien connu des intéressés, qui généralement y sont résignés faute de trouver une alternative et ce n'est nié que par ceux qui profitent du système. Un laboratoire peut donc payer pour faire des publications prestigieuses et pourtant trouver cela avantageux. En outre un laboratoire ou un centre de recherches doit, pour tenir son rang, entretenir une bibliothèque riche en documentation et par conséquent maintenir des abonnements aux revues les plus consultées. Les prix de ces abonnements, essentiellement financés par les crédits de recherche, c'est-à-dire par les contribuables, augmente sans cesse sous la pression purement économique des éditeurs qui possèdent ces revues. Il s'agit de prix de monopole, qui ne sont pas déterminés par les coûts mais par la solvabilité des institutions de recherche ; cela signifie que si les principales nations (disons celles du G7) décidaient demain le doublement de tous les budgets de recherche, après-demain le prix de ces abonnements doublerait aussi car les éditeurs verraient aussitôt que les laboratoires sont devenus capables de payer le double. Je vous laisse évidemment deviner les conséquences que cela aurait pour les sept pays les plus pauvres de la planète.

Dans tout ce contexte, il est bien clair que la *juste rémunération de la création intellectuelle* n'intervient que comme slogan, alibi hypocrite destiné à couvrir les juteuses opérations décrites ci-dessus, tout comme les *droits de l'homme* n'interviennent que pour habiller les intérêts financiers.

Les récents ennuis de Napster qui distribuait gratuitement de la musique sur le Web relèvent exactement du même phénomène : au nom de la *juste rémunération des artistes et créateurs* qui sont plus souvent exploités et spoliés que rémunérés, on veut protéger un

racket. Il est inutile que je m'y étende davantage puisque suffisamment de sites s'en occupent déjà.

Je n'entrerais pas non plus dans le débat sur la justesse du système économique, qui est suffisamment (et même trop) discutée sur d'autres sites. J'ai écrit ci-dessus que la course au profit des éditeurs (comme de n'importe quelle entreprise) est essentiellement conditionnée par la nécessité d'augmenter sans cesse les dividendes des actionnaires. C'est bien connu : si cet objectif n'est pas maintenu, la cote boursière de la société baisse et une spirale s'enclenche. Ce système est-il l'horreur économique ? Où est-il *le moins mauvais possible*, en sorte que toute tentative de le corriger ne peut aboutir qu'à une aggravation de la situation générale ⁽²⁾ ? Je n'ai pas la réponse à ces questions et je n'ai pas de solution miracle à proposer. Mais je pense qu'il est toujours plus utile de contribuer à la lucidité générale en apportant, comme je l'ai fait ici, mon modeste témoignage sur les mécanismes réels du système et en dénonçant le mensonge et l'hypocrisie qui masquent ces mécanismes réels, plutôt qu'en proposant des théories et des rêves utopiques. Je rappelle en effet que nous devons les bénéfices pratiques de la science non à des gens qui auraient dit un jour « je vais chercher de quoi rendre la vie plus facile » et qui auraient en conséquence trouvé le feu, la roue, le levier, l'engrenage, le ciment, l'électricité, le transistor etc, mais à des gens qui ont cherché d'abord à comprendre le fonctionnement réel des phénomènes au lieu de croire à l'action de divers dieux, démons et esprits.

Conclusion

Le principal problème que je veux soulever n'est pas le paiement des droits d'auteurs. La plupart des universitaires ou chercheurs ont déjà un revenu confortable ; cette question du paiement des droits est pour eux secondaire. Mais justement grâce à cela, on pourrait diffuser de nombreux manuels ou monographies pour beaucoup moins cher, pour un prix qui couvrirait juste les frais d'impression et l'organisation de la diffusion auprès des libraires. Tout le coûteux travail de conception, rédaction, typographie, correction, mise en page, étant effectué gratuitement par les auteurs. Trouvez-vous acceptable que les livres soient vendus à des prix qui comptent tout ce travail, alors qu'il est effectué gratuitement mais que les éditeurs en empochent le fruit ? La plupart des auteurs de manuels ou monographies de recherche travaillent gratuitement et au lieu que cela se traduise pour l'étudiant ou la bibliothèque par des prix modiques, cela est détourné pour ce racket. Je ne propose pas de faire la révolution, mais je soutiens ce qu'on appelle l'économie solidaire, le secteur associatif, les mutuelles, les coopératives. Dans l'édition, ce secteur reste à créer (ou plus exactement à sortir de la marginalité).

J'ai juste un conseil pratique à donner à ceux qui ont quelque chose à dire et qui souhaitent l'écrire. Cela s'adresse avant tout à mes collègues qui sont jeunes chercheurs et préparent un ouvrage sur leurs travaux, mais peut-être aussi à ceux qui débutent comme auteurs de fiction et rêvent de carrière littéraire. Ce conseil est : méfiez vous comme de

(2) Voir l'expérience désastreuse entreprise par les bolchéviques. Si c'est effectivement l'horreur économique, peut-il exister un système moins mauvais, qu'il nous faut alors inventer, mais en sachant que nous ne pouvons plus guère prendre le risque d'expérimenter à nouveau comme les bolchéviques ?

la peste des éditeurs et de leurs bonnes paroles, ils ne visent qu'à vous utiliser pour faire de l'argent sur votre dos. Ne laissez pas vos droits à ces pirates en signant leurs contrats léonins. Non, pour votre juste rémunération mais pour faire baisser le prix des livres et de la documentation scientifique. C'est l'apathie ou la vanité des auteurs qui permet le racket. Faites preuve d'indépendance et publiez d'abord sur le Web. Si votre manuscrit obtient un certain succès dans le cyber-monde, il sera toujours temps de répondre à un éditeur, mais vous serez en position plus forte. Et si votre manuscrit passe totalement inaperçu (cas le plus probable, ce qui d'ailleurs ne prouve pas forcément qu'il est mauvais) vous n'aurez de toute façon rien perdu.

Pour finir, je ne vous cacherai pas ce que je souhaite : que les progrès de l'informatique entraînent la ruine des éditeurs. De plus en plus, avec les nouvelles technologies informatiques, ils deviennent de simples parasites qui n'apportent pratiquement plus aucun service réel (voir plus haut) et veulent juste maintenir leur droit au racket par la menace judiciaire. Les auteurs devraient se regrouper en associations mutualistes capables de négocier directement avec les imprimeurs et les libraires. Hélas je ne crois guère à l'extinction pure et simple des éditeurs, car ils ont de quoi payer suffisamment d'avocats pour obtenir la protection de l'État contre le mécanisme naturel du marché. Que pensez-vous d'un mouvement associatif de diffusion de livres ?

Si ces problèmes de l'édition vous intéressent, sachez qu'ils sont déjà débattus dans certains cercles, hélas trop faibles et trop marginaux. On dénonce volontiers "le Pouvoir de l'Argent", mais ce qui fait ce Pouvoir est bien plus l'apathie du plus grand nombre d'entre nous que l'Argent. Groupez-vous et *ils* ne seront plus rien !



I — Théorèmes sur les intégrales

Ce chapitre expose les conditions dans lesquelles on peut passer à la limite ou dériver sous le signe intégral. Les quatre chapitres suivants en feront un usage constant. Il est donc indispensable de l'avoir soigneusement étudié avant de passer à la suite.

Il comporte des exercices qui consistent presque tous à effectuer concrètement des démonstrations qui ne seront reprises que très rapidement aux chapitres suivants.

Lemme I.1 — Inégalité de la moyenne Soient $]a; b[$, un intervalle non nécessairement borné (c'est-à-dire qu'on peut avoir $a = -\infty$ ou $b = +\infty$), $f(t)$, fonction bornée sur $]a; b[$ et $g(t)$, une fonction telle que l'intégrale $\int_{]a; b[} |g(t)| dt$ soit convergente. Alors, l'intégrale $\int_{]a; b[} f(t)g(t) dt$ est absolument convergente et on a :

$$\left| \int_{]a; b[} f(t)g(t) dt \right| \leqslant \sup_{t \in]a; b[} \{|f(t)|\} \times \int_{]a; b[} |g(t)| dt \quad (\text{I.1})$$

Preuve L'inégalité de la moyenne est connue pour les sommes finies : si $\{a_j\}$ et $\{b_j\}$ ($j = 1, 2, 3, \dots, n$) sont deux suites finies, on aura toujours :

$$\left| \sum_{j=1}^n a_j b_j \right| \leqslant \max_{j=1, n} \{|a_j|\} \times \sum_{j=1}^n |b_j| \quad (\text{I.2})$$

Or, une intégrale est toujours une limite de sommes finies⁽¹⁾ et les inégalités larges passent à la limite. •

La difficulté, dans les diverses théories de l'intégrale, vient de l'existence ou de l'unicité de ces limites. En effet, pour avoir une notion d'intégrale cohérente, il faut que la valeur

(1) Par exemple, les sommes de Riemann dans la théorie élémentaire, mais c'est vrai aussi pour n'importe quelle théorie de l'intégrale.

limite soit indépendante de la discrétisation de la fonction. L'exemple typique pour illustrer ce problème est la fonction :

$$\chi(t) = \begin{cases} 1 & \text{si } t \text{ est rationnel} \\ 0 & \text{autrement} \end{cases} \quad (\text{I.3})$$

Si on construit une somme de Riemann en discrétisant l'intervalle par $t_j = j/N$, on obtient :

$$\int_0^1 \chi(t) dt = \lim_{N \rightarrow \infty} \sum_{j=0}^{N-1} \frac{1}{N} \chi\left(\frac{j}{N}\right) = 1 \quad (\text{I.4})$$

tandis que si on discrétise par $t_j = j\pi/N$, on obtient ⁽²⁾ :

$$\int_0^1 \chi(t) dt = \lim_{N \rightarrow \infty} \sum_{j=0}^{E(N/\pi)} \frac{\pi}{N} \chi\left(\frac{j\pi}{N}\right) = 0 \quad (\text{I.5})$$

Ainsi, la méthode de Riemann ne permet pas de définir de façon cohérente une intégrale de la fonction χ . La fonction χ est l'exemple donné par H. Lebesgue dans sa première communication sur le sujet à l'Académie des Sciences ⁽³⁾ d'une fonction qui ne peut pas être intégrée par la méthode de Riemann, mais peut l'être par la sienne.

Dans n'importe quelle approche de l'intégrale, cependant, une fois résolus les problèmes d'existence et d'unicité de la limite, la valeur de l'intégrale sera toujours une limite de sommes finies, auxquelles on peut appliquer l'inégalité (I.2). Le fait que l'intervalle $]a; b[$ soit infini, ou que la fonction à intégrer devienne infinie en a ou en b , n'y change rien non plus. Par exemple, on aura :

$$\int_0^\infty e^{-t} dt = \lim_{N \rightarrow \infty} \sum_{j=0}^N \frac{1}{N} e^{-j/N} \quad \text{ou bien} \quad \int_0^1 \frac{1}{\sqrt{t}} dt = \lim_{N \rightarrow \infty} \sum_{j=1}^N \frac{1}{N} \frac{1}{\sqrt{j/N}} \quad (\text{I.6})$$

Théorème I.1 $[a; b]$ est un intervalle borné, f_n une suite de fonctions continues sur $[a; b]$. Si f_n converge uniformément sur $[a; b]$ vers f , alors :

$$\lim_{n \rightarrow \infty} \int_a^b f_n(t) dt = \int_a^b f(t) dt \quad (\text{I.7})$$

Preuve Dire que f_n converge uniformément sur $[a; b]$ vers f signifie que la suite numérique :

$$u_n = \sup_{t \in [a; b]} |f_n(t) - f(t)| \quad (\text{I.8})$$

tend vers zéro. D'après l'inégalité de la moyenne, on a :

$$\left| \int_a^b f_n(t) dt - \int_a^b f(t) dt \right| \leq u_n \times \int_a^b 1 dt = u_n \times (b - a) \quad (\text{I.9})$$

Or, $b - a$ est fini et u_n tend vers zéro. •

(2) La sommation s'effectue de $j = 0$ à $j = E(N/\pi)$, où $E(N/\pi)$ désigne la partie entière de N/π .

(3) H. LEBESGUE, *Sur une généralisation de l'intégrale définie*, Comptes-rendus de l'Académie des Sciences, séance du 29 avril, 1901, vol. 132, p. 1025–1028

Ⓡ On peut aussi utiliser cet argument si, au lieu d'une suite de fonctions f_n , on a une famille continue f_α où α est un nombre réel qui tend vers zéro (ou vers l'infini, ou vers n'importe quoi).

Exercice I.1 Dans la démonstration du théorème III.6, on intervertit une intégration sur un lacet Γ avec la sommation d'une série géométrique ; Utiliser le théorème I.1 pour justifier cette opération. Généraliser. ■

Exercice I.2 En fin de page 80, on peut lire : “lorsqu'un tel lacet est complètement aplati sur le segment, l'intégrale devient celle de 0 à 1 des valeurs limite de $F(z)$ par dessous, plus celle de 1 à 0 des valeurs limite de $F(z)$ par dessus”.

Utiliser le théorème I.1 pour prouver la validité de ce passage à la limite. ■

Exercice I.3 Démontrer qu'une série normalement convergente de fonctions peut être intégrée terme à terme. De façon plus précise : soit une série de fonctions $\sum f_n(x)$, les f_n étant définies et continues sur $[0; A]$. On pose $a_n = \sup_{t \in [0; A]} |f_n(t)|$ et on suppose que la série numérique $\sum a_n$ est convergente. Alors, si $F_n(x) = \int_0^x f_n(t) dt$, la série des $F_n(x)$ est la primitive de la fonction $\sum f_n(x)$. ■

Voyons maintenant le cas où l'intervalle $]a; b[$ est infini, ou encore, le cas où l'intervalle est fini mais où les fonctions f_n et f peuvent devenir infinies en a ou en b , les intégrales étant cependant convergentes.

Théorème I.2 Soient $]a; b[$, un intervalle non nécessairement borné et f_n , une suite de fonctions continues sur $]a; b[$ telles que les intégrales $\int_a^b f_n(t) dt$ soient convergentes.

Si sur tout sous-intervalle borné $[A; B]$ inclus dans $]a; b[$, f_n converge uniformément sur $[A; B]$ vers f et si en outre il existe une fonction $F(t)$ telle que :

1. $\forall t \in]a; b[, \quad F(t) \geq 0$;
2. $\forall t \in]a; b[, \quad \forall n \in \mathbb{N}, \quad |f_n(t)| \leq F(t)$;
3. l'intégrale $\int_a^b F(t) dt$ est convergente.

alors :

$$\lim_{n \rightarrow \infty} \int_a^b f_n(t) dt = \int_a^b f(t) dt \quad (\text{I.10})$$

Preuve Il faut montrer que :

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N}, \quad n \geq n_0 \quad \Rightarrow \quad \left| \int_a^b f_n(t) dt - \int_a^b f(t) dt \right| \leq \varepsilon \quad (\text{I.11})$$

Soit donc $\varepsilon > 0$. Puisque l'intégrale $\int_a^b F(t) dt$ est convergente, on peut trouver A et B tels que $a < A < B < b$, ainsi que :

$$\int_a^A F(t) dt < \frac{\varepsilon}{8} \quad \text{et} \quad \int_B^b F(t) dt < \frac{\varepsilon}{8} \quad (\text{I.12})$$

Cela est valable aussi bien si $a = -\infty$ ou $b = +\infty$, que si a et b sont finis mais que $F(t)$ devient infinie en a et b . Puisque par hypothèse $|f_n(t)| \leq F(t)$, on aura aussi $|f(t)| \leq F(t)$ et par conséquent, $|f_n(t) - f(t)| \leq |f_n(t)| + |f(t)| \leq 2F(t)$, de sorte que :

$$\int_a^A |f_n(t) - f(t)| dt + \int_B^b |f_n(t) - f(t)| dt \leq 2 \int_a^A F(t) dt + 2 \int_B^b F(t) dt \leq \frac{\varepsilon}{2} \quad (\text{I.13})$$

Ainsi, pour tout ε positif, il existe A et B tels que l'intégrale de $|f_n(t) - f(t)|$ sur $]a; A[\cup]B; b[$ soit inférieure à $\frac{\varepsilon}{2}$. D'autre part, puisque sur l'intervalle $[A; B]$ (avec A et B ainsi choisis), f_n tend uniformément vers f ,

on peut dire qu'il existe un entier n_0 tel que :

$$n \geq n_0 \Rightarrow \int_A^B |f_n(t) - f(t)| dt \leq \frac{\varepsilon}{2} \quad (\text{I.14})$$

En combinant (I.13) et (I.14), on obtient ce qu'on voulait. •

On a évidemment le même résultat pour une famille continue f_α au lieu d'une suite f_n .

Exercice I.4 Utiliser le théorème I.2 pour faire une démonstration courte du théorème V.5.

Indications : prendre comme suite f_n , les fonctions :

$$f_n(t) = \begin{cases} t^{x-1} \left(1 - \frac{t}{n}\right)^n & \text{si } 0 \leq t \leq n \\ 0 & \text{si } t > n \end{cases}$$

Rappelons aussi que :

$$\lim_{n \rightarrow \infty} \left(1 - \frac{t}{n}\right)^n = e^{-t} \quad \text{et} \quad 0 \leq \left(1 - \frac{t}{n}\right)^n \leq e^{-t} \quad \blacksquare$$

Exercice I.5 Dans la phrase qui précède l'équation (V.79), justifier le passage à la limite pour $\alpha \rightarrow 0$. ■

On peut aussi énoncer une variante du théorème I.2 : on y a supposé que les f_n tendent uniformément vers f sur tout sous-intervalle $[A; B]$ et que les $|f_n|$ sont toutes majorées par $F(t)$. Pour appliquer ce théorème, il faudra donc vérifier soigneusement que *pour tout* sous-intervalle $[A; B]$, la suite numérique $u_n(A, B) = \sup_{t \in [A; B]} \{|f_n(t) - f(t)|\}$ tend vers zéro. Souvent, on peut trouver plus directement une fonction $F(t)$, positive ou nulle, dont l'intégrale sur $]a; b[$ est convergente et une suite numérique u_n qui tend vers zéro telles que :

$$\forall t \in]a; b[, \quad \forall n \in \mathbb{N}, \quad |f_n(t) - f(t)| \leq u_n F(t) \quad (\text{I.15})$$

Dans ce cas, la conclusion est la même, puisque, toujours d'après l'inégalité de la moyenne, on pourra écrire :

$$\left| \int_a^b f_n(t) dt - \int_a^b f(t) dt \right| \leq u_n \times \int_a^b F(t) dt \quad (\text{I.16})$$

On peut donc énoncer :

Théorème I.3 Soient $]a; b[$, un intervalle non nécessairement borné, f_n , une suite de fonctions continues sur $]a; b[$ telles que toutes les intégrales $\int_a^b f_n(t) dt$ soient convergentes.

S'il existe une fonction $F(t) \geq 0$ dont l'intégrale sur $]a; b[$ est convergente, ainsi qu'une suite numérique u_n qui tend vers zéro, telles que :

$$\forall t \in]a; b[, \quad \forall n \in \mathbb{N}, \quad |f_n(t) - f(t)| \leq u_n F(t) \quad (\text{I.17})$$

alors :

$$\lim_{n \rightarrow \infty} \int_a^b f_n(t) dt = \int_a^b f(t) dt \quad (\text{I.18})$$

Exercice I.6 On obtient les égalités (IV.56) par un passage à la limite sous les intégrales. Utiliser le théorème I.3 pour le justifier en détail, en majorant la différence :

$$\left| \frac{e^{-t}}{1+zt} - \frac{e^{-t}}{1-at} \right| \quad (\text{I.19})$$

Il ne faut pas oublier que sur la partie courbe des chemins Γ_1 et Γ_2 , t est complexe. ■

Les théorèmes I.1, I.2 et I.3 couvrent pratiquement tous les cas concrets qu'on peut rencontrer. Les passages à la limite dans les intégrales servent presque toujours à établir des égalités : formule d'Euler (théorème V.6 et exercice I.4), formule de Hankel (section V.6 et exercice I.5), calcul de la discontinuité de la fonction d'Euler (équation (IV.53) et exercice I.6), calculs d'intégrales diverses (sections IV.3 et IV.5) etc. Ils ont donc une utilité pratique considérable pour les calculs.

Dans tous ces cas, la condition de convergence uniforme, c'est-à-dire l'existence de la suite numérique u_n dans les théorèmes I.1, I.2 et I.3, n'est pas plus difficile à obtenir que la limite point-par-point ; c'est-à-dire qu'il ne coûte pas plus cher, dans tous ces exemples, de vérifier que $u_n = \sup_t |f_n(t) - f(t)|$ tend vers zéro, que de vérifier que pour tout t fixé, $|f_n(t) - f(t)|$ tend vers zéro ⁽⁴⁾. C'est pourquoi ces théorèmes élémentaires sont largement suffisants pour faire des calculs.

Pourtant, la condition de convergence uniforme n'est pas du tout nécessaire et H. Lebesgue a démontré ⁽⁵⁾ la version suivante du théorème I.2, où les hypothèses sont fortement affaiblies :

Théorème I.4 — Théorème de convergence dominée $]a; b[$ étant un intervalle non nécessairement borné et f_n une suite de fonctions intégrables (au sens de Lebesgue, c'est-à-dire pratiquement n'importe quoi pourvu que l'intégrale de la valeur absolue converge) sur $]a; b[$.

Si pour presque tout t ⁽⁶⁾ fixé dans $]a; b[$, la suite numérique $f_n(t)$ tend vers $f(t)$ (c'est-à-dire que f_n tend point-par-point et non uniformément, vers f) et s'il existe une fonction $F(t)$ telle que :

1. pour tout t dans $]a; b[$, $F(t) \geq 0$;
2. pour presque tout t dans $]a; b[$ et pour tout $n \in \mathbb{N}$, $|f_n(t)| \leq F(t)$;
3. la fonction F est intégrable sur $]a; b[$ (i.e. l'intégrale $\int_a^b F(t) dt$ est convergente) ;

alors :

$$\lim_{n \rightarrow \infty} \int_a^b f_n(t) dt = \int_a^b f(t) dt \quad (\text{I.20})$$

La démonstration de ce théorème est possible dans le cadre de la théorie de Lebesgue, mais utilise des propriétés fines du continuum des nombres réels, essentiellement le fait que :

(4) Il s'agit respectivement de la convergence *simple* et de la convergence *point-par-point*.

(5) Voir *Leçons sur l'intégration et la recherche de fonctions primitives*, Gauthiers-Villars, Paris, 1904.

(6) L'expression *presque partout* signifie *en dehors d'un ensemble négligeable*. Cette notion est expliquée au chapitre II, page 17.

Théorème I.5 — théorème d'Egoroff Si f_n tend point-par-point vers f , alors pour tout $\varepsilon > 0$ et tout intervalle fini $[A; B] \subset]a; b[$ il existe un ensemble $E(\varepsilon, A, B) \subset [A; B]$ de longueur totale inférieure à ε tel que la convergence de f_n vers f soit uniforme sur $[A; B] - E(\varepsilon, A, B)$

Exercice I.7 En admettant le théorème I.5 et en admettant aussi qu'on peut prendre $F(t)$ continue, prouver le théorème de convergence dominée (reprendre la démonstration du théorème I.2) et au lieu de découper l'intégrale en trois morceaux (sur $]a; A[$, $[A; B]$ et $]B; b[$), la découper en quatre (sur $]a; A[$, $[A; B] - E(\varepsilon, A, B)$, $E(\varepsilon, A, B)$ et $]B; b[$).

On considère maintenant une intégrale dépendant d'un paramètre :

$$\Phi(x) = \int_{]a; b[} f(x, t) dt \quad (\text{I.21})$$

À quelles conditions peut-on, sans risque d'erreur, dériver sous l'intégrale ? Voici d'abord un théorème facile à démontrer dans un cadre élémentaire. Comme dans les théorèmes I.3 et I.4, l'intervalle $]a; b[$ n'est pas nécessairement borné et s'il l'est, les fonctions à intégrer peuvent devenir infinies en a ou en b , mais de sorte que l'intégrale soit convergente.

Théorème I.6 — Dérivation d'une intégrale Si dans (I.21) la fonction $f(x, t)$ possède, pour tout $x \in U$ et pour tout $t \in]a; b[$, une dérivée $\partial f / \partial x$ par rapport à x , ainsi qu'une dérivée seconde $\partial^2 f / \partial x^2$ et s'il existe une fonction $F(t)$ telle que :

1. $\forall t \in]a; b[, \quad F(t) \geq 0$;
2. $\forall t \in]a; b[, \quad \forall x \in U, \quad \left| \frac{\partial^2 f}{\partial x^2}(x, t) \right| \leq F(t)$;
3. l'intégrale $\int_a^b F(t) dt$ est convergente ;

alors :

$$\frac{d\Phi}{dx}(x) = \int_a^b \frac{\partial f}{\partial x}(x, t) dt \quad (\text{I.22})$$

Preuve Toujours le même schéma. D'après la formule des accroissements finis, on a l'inégalité :

$$\forall t \in]a; b[, \quad \forall x \in U, \quad \left| f(x+h, t) - f(x, t) - h \frac{\partial f}{\partial x}(x, t) \right| \leq \frac{1}{2} h^2 \sup_{y \in U} \left\{ \left| \frac{\partial^2 f}{\partial x^2}(y, t) \right| \right\} \leq \frac{1}{2} h^2 F(t) \quad (\text{I.23})$$

d'où (par l'inégalité de la moyenne) :

$$\left| \frac{\Phi(x+h) - \Phi(x)}{h} - \int_a^b \frac{\partial f}{\partial x}(x, t) dt \right| \leq \frac{1}{2} |h| \int_a^b F(t) dt \quad (\text{I.24})$$

Puisque l'intégrale $\int_a^b F(t) dt$ est finie, le second membre de (I.24) tend vers zéro à cause du facteur h , donc :

1. $\Phi(x)$ est dérivable sur U ;
2. sa dérivée est $\int_a^b \frac{\partial f}{\partial x}(x, t) dt$ •

Cette version du théorème de dérivation est une version faible car on exige des conditions sur la dérivée seconde. En pratique toutefois, cela n'est qu'exceptionnellement une gêne. Cette version a en outre l'avantage de pouvoir être démontrée par une simple application de l'inégalité de la moyenne.

La version classique exige des hypothèses bien plus faibles, mais ne se démontre que dans le cadre de l'intégrale de Lebesgue. Le théorème qui suit est dû à H. Lebesgue, en 1904 :

Théorème I.7 On suppose seulement que pour presque tout t dans $]a; b[$ et tout x dans U , la fonction $f(x, t)$ de (I.21) possède une dérivée partielle $\partial f / \partial x$ et qu'il existe une fonction $F(t)$ telle que :

1. pour presque tout t dans $]a; b[$, $F(t) \geq 0$;
2. pour presque tout t dans $]a; b[$ et tout x dans U , $|\frac{\partial f}{\partial x}(x, t)| \leq F(t)$;
3. l'intégrale $\int_a^b F(t) dt$ est convergente ⁽⁷⁾.

La conclusion est celle du théorème I.6.

Preuve Ce théorème se démontre aisément comme corollaire du théorème de convergence dominée I.4 : soit h_n une suite numérique qui tend vers zéro (mais on suppose que $\forall n \in \mathbb{N}, h_n \neq 0$). On a :

$$\frac{\Phi(x + h_n) - \Phi(x)}{h_n} = \int_a^b \frac{f(x + h_n, t) - f(x, t)}{h_n} dt \quad (\text{I.25})$$

D'après les hypothèses, lorsque h_n tend vers zéro, la suite $[f(x + h_n, t) - f(x, t)]/h_n$ tend presque partout vers $\partial f / \partial x$; en outre d'après la formule des accroissements finis et du fait que $|\partial f / \partial x| \leq F(t)$, on a aussi pour presque tout $t \in]a; b[$, $\forall n \in \mathbb{N}$ et $\forall x \in U$:

$$\left| \frac{f(x + h_n, t) - f(x, t)}{h_n} \right| \leq F(t) \quad (\text{I.26})$$

de sorte que d'après le théorème de convergence dominée :

$$\lim_{n \rightarrow \infty} \frac{\Phi(x + h_n) - \Phi(x)}{h_n} = \int_a^b \frac{\partial f}{\partial x}(x, t) dt \quad (\text{I.27})$$

On a ainsi montré que *pour n'importe quelle* suite h_n qui tend vers zéro :

$$\frac{\Phi(x + h_n) - \Phi(x)}{h_n} \rightarrow \int_a^b \frac{\partial f}{\partial x} dt \quad (\text{I.28})$$

ce qui signifie que Φ est dérivable et que sa dérivée est :

$$\Phi'(x) = \int_a^b \frac{\partial f}{\partial x} dt \quad (\text{I.29})$$

ce qui permet de conclure. •

Exercice I.8 On trouve page 116 : “Si les conditions pour pouvoir dériver sous le signe intégral sont satisfaites. . .” Préciser ces conditions à partir des théorèmes I.6 et I.7. ■

Exercice I.9 La démonstration du théorème VII.1 utilise l'inégalité de la moyenne et cela revient à redémontrer implicitement le théorème I.6. Refaire cette démonstration en utilisant cette fois les théorèmes I.6 et I.7. ■

Exercice I.10 Soit la fonction :

$$\Phi(x) = \int_{-\infty}^{+\infty} \frac{e^{itx}}{t^2 + 1} dt = \pi e^{-|x|} \quad (\text{I.30})$$

(le calcul de l'intégrale se fait par la méthode des résidus, ce qui est un autre exercice). On voit, à cause de la valeur absolue $|x|$ dans l'expression à droite, que cette fonction n'est pas dérivable en zéro. Montrer que, justement, l'intégrale ne satisfait pas les conditions exigées pour le théorème I.7 (et encore moins pour le théorème I.6). ■

(7) Dans la théorie de H. Lebesgue, on dit que F est intégrable sur $]a; b[$ plutôt que l'intégrale $\int_a^b F(t) dt$ est convergente.

Exercice I.11 Au chapitre IV, on calcule l'intégrale (IV.27) :

$$\Psi(t) = \int_{-\infty}^{+\infty} \frac{e^{itx}}{x^4 + 1} dx = \pi \cos\left(\frac{|t|}{\sqrt{2}} - \frac{\pi}{4}\right) e^{-|t|/\sqrt{2}} \quad (\text{I.31})$$

(attention! le rôle des variables x et t est inversé par rapport à l'énoncé du théorème I.7). Vérifier directement sur l'expression à droite, que $\Psi(t)$ est deux fois dérivable en $t = 0$, mais pas trois fois. Constater par ailleurs que l'intégrale satisfait aux conditions des théorèmes I.6 et I.7 pour dériver une première fois; ensuite, que pour dériver une seconde fois, elle ne satisfait plus aux conditions du théorème I.6 mais encore à celles du théorème I.7. Enfin, que pour dériver une troisième fois, elle ne satisfait même plus aux conditions du théorème I.7. ■

Un dernier théorème très utile pour l'étude des fonctions d'une variable complexe, dont beaucoup sont définies comme des intégrales, notamment $\Gamma(z)$, $\beta(z, w)$ au chapitre V, $\text{Eu}(z)$, au chapitre IV, mais aussi les fonctions de Bessel, les transformées de Fourier ou de Laplace, au chapitre VI, les intégrales de Fresnel etc.

Théorème I.8 On considère une intégrale dépendant du paramètre complexe z :

$$\Phi(z) = \int_a^b f(z, t) dt \quad (\text{I.32})$$

L'intervalle $]a; b[$ peut, comme aux théorèmes I.2 à I.7, être infini ou, s'il est fini, les fonctions peuvent devenir infinies en a ou en b . On suppose que pour tout $t \in]a; b[$, la fonction $z \mapsto f(z, t)$ est analytique dans un domaine Ω du plan complexe.

S'il existe une fonction $F(t)$ telle que :

1. $\forall t \in]a; b[, \quad F(t) \geq 0$;
2. $\forall t \in]a; b[, \quad \forall z \in \Omega, \quad |f(z, t)| \leq F(t)$;
3. l'intégrale $\int_a^b F(t) dt$ est convergente;

alors la fonction $\Phi(z)$ est analytique dans Ω et sa dérivée analytique est :

$$\Phi'(z) = \int_a^b \frac{\partial f}{\partial z}(z, t) dt \quad (\text{I.33})$$

où $\partial f / \partial z$ est, pour t fixé, la dérivée analytique de $z \mapsto f(z, t)$.

Preuve Ainsi, puisque la dérivée seconde de $f(z, t)$ en un point $z = z_0$ de Ω est majorée par $2M_r(t)/r^2$, on peut écrire la formule des accroissements finis sous la forme

$$\left| \frac{f(z+h, t) - f(z, t)}{h} - \frac{\partial f}{\partial z}(z, t) \right| \leq 2 \frac{M_r(t)}{(r/2)^2} \quad (\text{I.34})$$

l'inégalité étant valable pour $|z - z_0| < r/2$ (on recouvre le disque de centre z_0 et de rayon r , dans lequel $|f(z, t)|$ est majoré par $M_r(t)$, par des disques de centre z et de rayon $r/2$, ce qui marche si on prend les z dans le disque de centre z_0 et de rayon $r/2$). En procédant exactement comme dans la démonstration du théorème I.6, on obtient ainsi que $\Phi(z)$ est analytique dans le disque de centre z_0 et de rayon $r/2$, pourvu que le disque de centre z_0 et de rayon r soit inclus dans Ω , ce qui est toujours possible si r est assez petit (Ω est ouvert). Comme on peut faire cela pour n'importe quel point z_0 de Ω , on a ainsi prouvé que $\Phi(z)$ est analytique au voisinage de *tout* $z_0 \in \Omega$, donc analytique dans Ω . •

Ⓡ On constate que, contrairement aux théorèmes I.6 et I.7, il n'est pas exigé que la fonction F majore la dérivée $\partial f / \partial z$, ni la dérivée seconde $\partial^2 f / \partial z^2$, mais qu'elle majore la fonction elle-même. Cela est permis par une propriété remarquable des fonctions analytiques, les *inégalités de Cauchy* d'après le

corollaire III.6.1. Les inégalités de Cauchy majorent en effet les dérivées d'une fonction analytique : si $M_r(t)$ est le maximum de $|f(z, t)|$ sur le disque $|z - z_0| \leq r$, on a, d'après l'équation (III.49) :

$$\left| \frac{\partial f}{\partial z}(z_0, t) \right| \leq \frac{M_r(t)}{r} \quad \text{et} \quad \left| \frac{\partial^2 f}{\partial z^2}(z_0, t) \right| \leq \frac{2 M_r(t)}{r^2} \quad (\text{I.35})$$

Contrairement au théorème I.6, le théorème de convergence dominée de Lebesgue n'aurait pas conduit à des hypothèses plus faibles⁽⁸⁾ : ici nous avons utilisé une majoration de la dérivée seconde comme dans le théorème I.6, alors qu'en faisant appel au théorème de convergence dominée, on aurait pu se contenter d'une majoration de la dérivée première ; mais cela n'aurait rien changé, puisqu'il y a les inégalités de Cauchy.

Concernant la formule des accroissements finis appliquée aux fonctions d'une variable complexe, elle se ramène aux fonctions d'une variable réelle de la manière suivante. Soit h un nombre complexe (celui qui devra tendre vers zéro). Si $f(z)$ est une fonction analytique de z , on peut considérer la fonction $\varphi(t) = f(z + th)$ de la variable réelle $t \in [0; 1]$ prenant évidemment des valeurs *complexes*. Sa dérivée par rapport à t est $\varphi'(t) = h f'(z + th)$, où f' désigne la dérivée analytique de f . De même pour la dérivée seconde : $\varphi''(t) = h^2 f''(z + th)$. D'autre part $\varphi(0) = f(z)$ et $\varphi(1) = f(z + h)$. La formule des accroissements finis usuelle appliquée à φ donne alors :

$$|\varphi(1) - \varphi(0) - \varphi'(0)| \leq \frac{1}{2} \sup_{t \in [0; 1]} |\varphi''(t)| \quad (\text{I.36})$$

Si on remplace les fonctions φ , φ' et φ'' par leurs expressions à partir de f , f' et f'' , cela se traduit par :

$$|f(z + h) - f(z) - h f'(z)| \leq \frac{1}{2} |h|^2 \sup_{t \in [0; 1]} |f''(z + th)| \quad (\text{I.37})$$

Si $z + th$ est, $\forall t \in [0; 1]$, contenu dans un disque de centre z_0 et de rayon r , sur lequel le maximum de $|f(z)|$ est M_r , on obtient bien l'inégalité (I.34).

Ⓡ En général, lorsque la fonction $\Phi(z)$ du théorème I.8 est analytique dans un domaine Ω , souvent infini, on ne pourra pas majorer en une seule fois la fonction $|f(z, t)|$ par $F(t)$ uniformément dans tout Ω . La plupart des exercices suivants correspondent à cette situation qui est la plus fréquente en pratique. On majorera alors $|f(z, t)|$ par une fonction $F(t)$ dans une partie seulement de Ω . Puis le domaine Ω sera obtenu comme une réunion infinie de parties U_n , $\Omega = \cup_n U_n$, sur chacune desquelles on aura une majorante F_n (mais telle que $\sup_n F_n = \infty$). Ainsi, on prouvera que $\Phi(z)$ est analytique dans chaque U_n et *par conséquent* aussi dans Ω .

Exercice I.12 Montrer, à partir des définitions V.1 et V.2, que les fonctions $\beta(x, y)$ et $\Gamma(x)$ sont analytiques dans le domaine $\{x \in \mathbb{C} \mid \Re(x) > 0\}$, $\{y \in \mathbb{C} \mid \Re(y) > 0\}$.

Appliquer le théorème I.8 aux domaines $U_n = \{x \in \mathbb{C} \mid \Re(x) > \frac{1}{n}\}$, pour tout n . ■

Exercice I.13 Montrer que la fonction d'Euler (IV.45) est analytique dans $\mathbb{C} -]-\infty; 0]$.

Appliquer le théorème I.8 aux domaines $U_n = \{z \in \mathbb{C} \mid d(z) > \frac{1}{n}\}$, où $d(z)$ est défini par la formule (V.62). ■

(8) Les conditions sur $f(z, t)$ de l'énoncé auraient pu n'être supposées que presque partout, ce qui ne se rencontre jamais dans les calculs.

Exercice I.14 Montrer que l'intégrale de Hankel (V.71) définit une fonction analytique dans tout $\mathbb{C}$. Appliquer le théorème I.8 aux domaines $U_n = \{z = x + iy \in \mathbb{C} \mid x > -n \text{ et } |y| < n\}$. ■

II — Formule de Green

II.1 Intégrales curvilignes

En mécanique ou en électricité, on est souvent amené à considérer l'intégrale d'une quantité le long d'un chemin curviligne. Soit, par exemple dans le plan, un point matériel soumis à un champ de forces $\vec{F}(x, y)$. Lorsque le point matériel effectue un déplacement rectiligne $\overrightarrow{AB}$, le travail du champ de force $\vec{F}$ est le produit scalaire $\vec{F} \cdot \overrightarrow{AB}$. Si le point matériel se déplace le long d'une courbe, il faut approcher la courbe par un polygone formé d'un grand nombre de déplacements rectilignes infinitésimaux $\overrightarrow{A_0A_1}$, $\overrightarrow{A_1A_2}$, $\overrightarrow{A_2A_3}$, $\dots$, $\overrightarrow{A_{N-1}A_N}$; le travail du champ de forces le long de la courbe est alors approché par la somme :

$$\sum_{j=0}^{N-1} \vec{F}(A_j) \cdot \overrightarrow{A_jA_{j+1}} \quad (\text{II.1})$$

dont la limite lorsque les déplacements $\overrightarrow{A_jA_{j+1}}$ tendront vers zéro et N conjointement vers l'infini, sera l'intégrale :

$$\int_{\gamma} \vec{F} \cdot d\vec{A} \quad (\text{II.2})$$

le symbole γ représentant la courbe (on précisera plus loin son sens mathématique).

Appelons $a(x, y)$ et $b(x, y)$ les composantes du vecteur $\vec{F}$ selon x et y et x_j, y_j les coordonnées du point A_j ; l'expression $\vec{F}(A_j) \cdot \overrightarrow{A_jA_{j+1}}$ s'écrira $a(x_j, y_j)(x_{j+1} - x_j) + b(x_j, y_j)(y_{j+1} - y_j)$ et la somme (II.1) devient :

$$\sum_{j=0}^{N-1} a(x_j, y_j)(x_{j+1} - x_j) + b(x_j, y_j)(y_{j+1} - y_j) \quad (\text{II.3})$$

Lorsque les déplacements $\overrightarrow{A_j A_{j+1}}$ tendent vers zéro, la somme (II.3) tend vers l'intégrale :

$$\int_{\gamma} a(x, y) dx + b(x, y) dy \quad (\text{II.4})$$

Lorsqu'on veut calculer effectivement de telles intégrales, dites curvilignes, on doit paramétrer la courbe γ ; les coordonnées (x, y) d'un point de la courbe sont alors fonction d'un paramètre t qui parcourt un intervalle, soit $(x(t), y(t))$ pour $t_0 < t < T$. Les points A_j de la courbe, de coordonnées x_j, y_j , correspondent à des valeurs discrètes de l'intervalle $[t_0; T]$: $x_j = x(t_j)$ et $y_j = y(t_j)$ et bien entendu $t_N = T$. La somme (II.1) (ou (II.3)) devient, en introduisant $x'(t) = dx/dt$ et $y'(t) = dy/dt$:

$$\sum_{j=0}^{N-1} \left(a(x(t_j), y(t_j)) x'(t_j) + b(x(t_j), y(t_j)) y'(t_j) \right) (t_{j+1} - t_j) \quad (\text{II.5})$$

puisque $(x_{j+1} - x_j) \simeq x'(t_j)(t_{j+1} - t_j)$ et $(y_{j+1} - y_j) \simeq y'(t_j)(t_{j+1} - t_j)$.

Faire tendre les déplacements $\overrightarrow{A_j A_{j+1}}$ vers zéro équivaut bien sûr à faire tendre les quantités $t_{j+1} - t_j$ vers zéro, de sorte que la somme (II.5) aura pour limite l'intégrale :

$$\int_{t_0}^T \left(a(x(t), y(t)) x'(t) + b(x(t), y(t)) y'(t) \right) dt \quad (\text{II.6})$$

Cela montre quel est le sens mathématique exact de l'intégrale (II.2) ou (II.4) : après paramétrage, elles se ramènent à des intégrales au sens usuel.

Tout le raisonnement précédent repose sur l'hypothèse que la courbe possède bien un paramétrage différentiable : on dit que la courbe est différentiable. On peut étendre le procédé à des courbes qui ne possèdent pas globalement un tel paramétrage mais qu'on peut découper en portions qui en possèdent chacun un — de telles courbes sont dites différentiables par morceaux — car l'intégrale curviligne est alors la somme des intégrales sur chaque morceau. On ne peut pas l'étendre à des courbes plus irrégulières, telles que par exemple les courbes fractales ; ces dernières exigent alors l'introduction de nouveaux concepts spécifiques.

Lorsqu'on découpe un chemin en morceaux dont chacun possède un paramétrage différentiable, on dit qu'on effectue une *décomposition* du chemin. L'opération inverse, consistant à recoller des chemins qui jusque là avaient été considérés séparément, s'appelle la *concaté-nation*. Si γ_1 et γ_2 sont deux chemins, ayant chacun le paramétrage :

$$\begin{aligned} \gamma_1 : t &\mapsto x_1(t), y_1(t) & 0 < t < S \\ \gamma_2 : t &\mapsto x_2(t), y_2(t) & 0 < t < T \end{aligned} \quad (\text{II.7})$$

alors le chemin noté γ :

$$\gamma : t \mapsto x(t), y(t) \quad 0 < t < S + T \quad (\text{II.8})$$

tel que :

$$x(t) = \begin{cases} x_1(t) & \text{si } 0 < t < S \\ x_2(t - S) & \text{si } S < t < S + T \end{cases} \quad y(t) = \begin{cases} y_1(t) & \text{si } 0 < t < S \\ y_2(t - S) & \text{si } S < t < S + T \end{cases}$$

est appelé la concaténation de γ_1 et γ_2 ; on note $\gamma = \gamma_1 + \gamma_2$. Il n'est pas nécessaire que $x_1(S) = x_2(0)$ et $y_1(S) = y_2(0)$ pour que la concaténation soit définie : un chemin peut être formé de morceaux *non connexes*, c'est-à-dire de morceaux qui ne se touchent pas.

L'intégrale est évidemment additive par rapport à la concaténation :

$$\int_{\gamma_1 + \gamma_2} a \, dx + b \, dy = \int_{\gamma_1} a \, dx + b \, dy + \int_{\gamma_2} a \, dx + b \, dy \quad (\text{II.9})$$

Étant donné un chemin γ , on désignera aussi par $-\gamma$ le chemin qui consiste à parcourir γ en sens inverse; si $x(t), y(t)$ pour $0 < t < T$ est un paramétrage de γ , $x(T-t), y(T-t)$ pour $0 < t < T$ est un paramétrage de $-\gamma$. On aura évidemment :

$$\int_{-\gamma} a \, dx + b \, dy = - \int_{\gamma} a \, dx + b \, dy \quad (\text{II.10})$$

En effet :

$$\begin{aligned} \int_{-\gamma} a(x, y) \, dx + b(x, y) \, dy &= \\ &= \int_0^T \left(a(x(T-t), y(T-t)) x'(T-t) + b(x(T-t), y(T-t)) y'(T-t) \right) dt \\ &= \int_T^0 \left(a(x(t), y(t)) x'(t) + b(x(t), y(t)) y'(t) \right) dt \end{aligned} \quad (\text{II.11})$$

II.2 Intégrales doubles

Nous ne présentons pas ici la théorie la plus générale possible, qui serait très dispendieuse en temps, mais seulement la théorie qui présente (de très loin) le meilleur rapport qualité/prix. Elle permet de couvrir tous les cas où on peut *effectivement* calculer⁽¹⁾ et est très simple.

Considérons d'abord le cas d'un domaine U du plan délimité par une courbe fermée γ , ayant les propriétés suivantes :

1. la courbe γ est différentiable par morceaux ;
2. toute droite parallèle à l'un des axes de coordonnées coupe la courbe en au plus deux points (voir figure II.1).

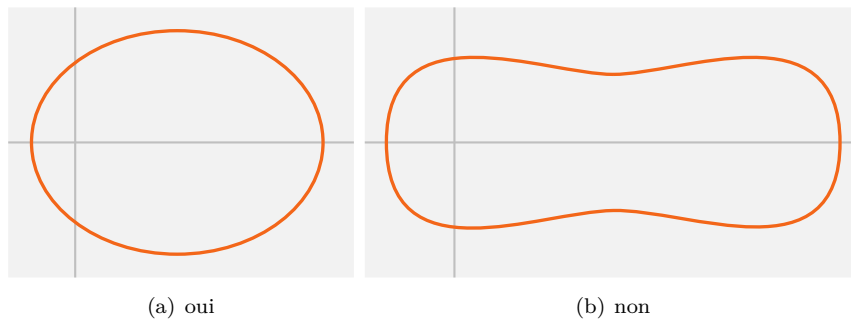


Figure II.1 – Toute droite parallèle à l'un des axes de coordonnées coupe la courbe en au plus deux points

(1) C'est-à-dire que les cas qui exigent des théories plus sophistiquées sont tous des cas où le calcul *effectif* est impossible.

Un tel domaine U est délimité par deux courbes paramétrables en x et aussi par deux courbes paramétrables en y ; en effet, d'après la propriété 2, les points du domaine U ayant pour abscisse x sont tels que $y_0(x) < y < y_1(x)$ et de même les points du domaine U ayant pour ordonnée y sont tels que $x_0(y) < x < x_1(y)$. On peut dire que le domaine U est limité en bas par la courbe paramétrée $y = y_0(x)$ et en haut par la courbe paramétrée $y = y_1(x)$, ou encore que le domaine U est limité à gauche par la courbe paramétrée $x = x_0(y)$ et à droite par la courbe paramétrée $x = x_1(y)$. Pour la suite, il suffira que ces différentes courbes soient différentiables par morceaux.

Nous allons définir une intégrale double :

$$\iint_U f(x, y) \, dx \, dy \quad (\text{II.12})$$

comme étant égale à l'intégrale simple :

$$\int_{x_0}^{x_1} g(x) \, dx \quad \text{et} \quad g(x) = \int_{y_0(x)}^{y_1(x)} f(x, y) \, dy \quad (\text{II.13})$$

où x_0 est le minimum de la fonction $y \mapsto x_0(y)$, x_1 le maximum de la fonction $y \mapsto x_1(y)$. L'intégrale double se ramène ainsi à deux intégrales simples.

Pour que cette définition présente un minimum de cohérence, il faut que l'intégrale simple :

$$\int_{y_0}^{y_1} h(y) \, dy \quad \text{avec} \quad h(y) = \int_{x_0(y)}^{x_1(y)} f(x, y) \, dx \quad (\text{II.14})$$

où y_0 est le minimum de la fonction $x \mapsto y_0(x)$, y_1 le maximum de la fonction $x \mapsto y_1(x)$, ait la même valeur que (II.13). Les fonctions $g(x)$ et $h(y)$ sont appelées les *intégrales partielles* de l'intégrale double.

Donnons un exemple où (II.13) et (II.14) ne coïncident pas, ce qui sous-entend évidemment que l'intégrale double n'est pas définie, quoique les quatre intégrales simples le soient. On prend pour U le carré $[-1; 1] \times [-1; 1]$ et pour $f(x, y)$ la fonction $(x^2 - y^2)/(x^2 + y^2)^2$. On remarque que cette fonction est égale à $\frac{\partial}{\partial y} \frac{y}{x^2 + y^2}$, donc :

$$g(x) = \int_{-1}^1 \frac{\partial}{\partial y} \frac{y}{x^2 + y^2} \, dy = \left. \frac{y}{x^2 + y^2} \right|_{-1}^{+1} = \frac{2}{x^2 + 1} \quad (\text{II.15})$$

On remarquera cependant que l'intégrale ci-dessus n'est pas convergente si $x = 0$, quoique le calcul par primitives soit appliqué formellement (et c'est là que se trouve le pot-aux-roses). On en déduit :

$$\int_{-1}^1 g(x) \, dx = 2 \arctan(x) \Big|_{-1}^{+1} = \pi \quad (\text{II.16})$$

Par ailleurs on constate aussi que $f(x, y) = \frac{\partial}{\partial x} \frac{-x}{x^2 + y^2}$, donc :

$$h(y) = \int_{-1}^1 \frac{\partial}{\partial x} \frac{-x}{x^2 + y^2} \, dx = \left. \frac{-x}{x^2 + y^2} \right|_{-1}^{+1} = \frac{-2}{1 + y^2} \quad (\text{II.17})$$

Là aussi, l'intégrale n'est pas convergente si $y = 0$; mais l'intégrale de $h(y)$ l'est :

$$\int_{-1}^1 h(y) \, dy = -2 \arctan(y) \Big|_{-1}^{+1} = -\pi \quad (\text{II.18})$$

Les deux résultats ne sont pas égaux ⁽²⁾. Cela est évidemment dû à la singularité de la fonction $f(x, y)$ au point $(0, 0)$: en fait, l'intégrale double *diverge* en ce point. On remarquera que cette divergence n'empêche pas les fonctions $g(x)$ et $h(y)$ d'être parfaitement définies et continues partout sur $[-1; 1]$. Cela montre que l'existence des intégrales (II.13) ou (II.14) n'est pas un critère suffisant pour donner un sens correct à l'intégrale double. Cette dernière se définit comme la limite d'une somme discrète, qui doit être indépendante de la discrétisation. Cela fonctionne très bien lorsque f est une fonction continue :

Théorème II.1 Soit U un domaine fermé et borné du plan vérifiant les conditions 1 et 2 ci-dessus, donc délimité d'une part par les courbes $x \mapsto y_0(x)$ (en bas) et $x \mapsto y_1(x)$ (en haut), d'autre part par les courbes $y \mapsto x_0(y)$ (à gauche) et $y \mapsto x_1(y)$ (à droite). Soit $f(x, y)$, une fonction continue sur U , alors :

$$\int_{y_0}^{y_1} h(y) dy = \int_{x_0}^{x_1} g(x) dx \quad (\text{II.19})$$

Preuve On se ramène à la définition de l'intégrale simple. Découpons les abscisses et les ordonnées selon des subdivisions discrètes x_j et y_j ⁽³⁾. Par définition de l'intégrale simple, on peut dire que :

$$g(x_j) = \int_{y_0}^{y_1} f(x_j, y) dy = \lim_y \sum_k f(x_j, y_k)(y_{k+1} - y_k) \quad (\text{II.20})$$

la somme se faisant sur les indices k pour lesquels $y_0(x_j) < y_k < y_1(x_j)$ et la limite (avec l'indice y) étant prise pour $\sup |y_{k+1} - y_k|$ tendant vers zéro. Appliquons la même définition de l'intégrale simple à (II.13) :

$$\int_{x_0}^{x_1} g(x) dx = \lim_x \sum_j g(x_j)(x_{j+1} - x_j) \quad (\text{II.21})$$

Cette fois la somme se fait sur les indices j pour lesquels $x_0(y_j) < x_j < x_1(y_j)$ et la limite (notée avec l'indice x) est prise pour $\sup |x_{j+1} - x_j|$ tendant vers zéro. En combinant (II.20) et (II.21), on obtient :

$$\int_{x_0}^{x_1} g(x) dx = \lim_x \sum_j \lim_y \sum_k f(x_j, y_k)(x_{j+1} - x_j)(y_{k+1} - y_k) \quad (\text{II.22})$$

Il est clair que la somme (finie) effectuée sur les indices j et k dans (II.22) ne dépend pas de l'ordre de sommation et qu'on peut tout aussi bien la sommer d'abord selon k et ensuite selon j ; mais il n'est pas aussi immédiatement évident qu'on peut prendre les limites dans n'importe quel ordre. Le raisonnement suivi pour parvenir à (II.22) implique qu'il faut prendre la limite (pour $\sup |y_{k+1} - y_k|$ tendant vers zéro) de la somme sur k *avant* d'effectuer la sommation sur j .

De façon analogue, on voit que :

$$\int_{y_0}^{y_1} h(y) dy = \lim_y \sum_k \lim_x \sum_j f(x_j, y_k)(x_{j+1} - x_j)(y_{k+1} - y_k) \quad (\text{II.23})$$

Cette fois, il faut prendre la limite (pour $\sup |x_{j+1} - x_j|$ tendant vers zéro) de la somme sur j *avant* d'effectuer la sommation sur k . Rien ne prouve *a priori* que les deux opérations conduisent au même résultat puisque les passages à la limite ne se font pas dans le même ordre. Toutefois, si les limites sont uniformes, c'est-à-dire si la limite, pour $\sup |x_{j+1} - x_j|$ tendant vers zéro, de la somme $\sum_j f(x_j, y_k)(x_{j+1} - x_j)(y_{k+1} - y_k)$ est uniforme par rapport aux y_k , ou si la limite, pour $\sup |y_{k+1} - y_k|$ tendant vers zéro, de la somme $\sum_j f(x_j, y_k)(x_{j+1} - x_j)(y_{k+1} - y_k)$ est uniforme par rapport aux x_j , alors le résultat ne dépend plus de l'ordre et les deux opérations conduiront au même résultat.

(2) Qu'ils soient ici égaux en valeur absolue n'est pas essentiel : si on avait pris un domaine non symétrique par rapport à l'origine il n'en serait pas ainsi.

(3) Afin d'éviter la confusion avec x_0, x_1, y_0, y_1 , déjà définis, on peut convenir que ces indices j et k prennent leurs valeurs à partir de 1 000 et non à partir de zéro ; ceci est du détail.

Il est facile, quoique fastidieux, de vérifier que si la fonction $f(x, y)$ est uniformément continue, les limites en question sont uniformes. Si le domaine U est borné, ce qui est le cas avec nos hypothèses concernant son bord, la fonction $f(x, y)$ sera automatiquement *uniformément* continue dès lors qu'elle sera simplement continue. On ne donne pas ici cette vérification de routine. On se contente d'insister sur le fait que la *raison pour laquelle le théorème est vrai* est cette affaire de limite uniforme. La démonstration s'achève donc ainsi. •

La continuité de $f(x, y)$ est une condition *suffisante* pour l'uniformité des limites mais elle n'est pas *nécessaire*. On peut affaiblir ces conditions et montrer qu'on obtient aussi l'uniformité des limites en supposant que la fonction $f(x, y)$ n'a que des discontinuités de première espèce. Cette approche de l'intégrale est connue sous le nom de *théorie de l'intégrale de Riemann*. Les mathématiciens ont cherché à la fin du siècle dernier des approches plus générales, pouvant s'appliquer à des fonctions plus irrégulières. Ainsi, H. Lebesgue a fait remarquer que la condition d'uniformité des limites n'a pas besoin d'être vraie partout, mais seulement en dehors d'un ensemble qui serait *négligeable* pour l'intégrale.

L'idée des ensembles négligeables s'explique comme suit. Considérons une suite infinie $\mathcal{F}$ d'intervalles $]a_k; b_k[$, non nécessairement disjoints et dont la réunion :

$$F = \bigcup_{k=0}^{\infty}]a_k; b_k[\quad (\text{II.24})$$

recouvre entièrement un certain ensemble donné E , c'est-à-dire que $E \subset F$. La longueur totale de ces intervalles est la somme infinie $\sum (b_k - a_k)$; elle peut être finie ou infinie, appelons la $\ell(\mathcal{F})$. Si l'ensemble E est tel qu'on puisse l'inclure dans des familles $\mathcal{F}$ d'intervalles telles que $\ell(\mathcal{F})$ puisse être rendu aussi petit qu'on veut, on dira que E est un *ensemble négligeable*. Il saute aux yeux que si par exemple E est l'intervalle $]0; 1[$, quelle que soit la manière de choisir les $]a_k; b_k[$, on ne pourra jamais faire descendre la valeur de $\ell(\mathcal{F})$ en dessous de 1. Prenons pour E , l'ensemble $\mathbb{Q}_+$ des nombres rationnels positifs. Ces derniers sont tous de la forme $r = p/q$ où p/q est une fraction irréductible (avec p et q entiers positifs). à chacune de ces fractions irréductibles $r = p/q$, on peut associer l'intervalle $[a_r; b_r] = [r - 2^{-p-q}; r + 2^{-p-q}]$. Il est clair par construction que l'ensemble E des fractions positives est contenu dans la réunion de ces intervalles, c'est-à-dire que la famille $\mathcal{F}_0$ de ces intervalles recouvre entièrement $E = \mathbb{Q}_+$; elle est indexée par les couples p, q d'entiers positifs et premiers entre eux. La longueur de l'intervalle $[a_r; b_r]$ étant 2^{1-p-q} , on peut dire alors que pour $\mathcal{F}_0$:

$$\ell(\mathcal{F}_0) = \sum_{p,q} 2^{1-p-q} \quad (\text{II.25})$$

la somme étant effectuée sur tous les couples p, q d'entiers positifs et premiers entre eux. Cette somme est forcément inférieure à ce qu'on obtiendrait en sommant sur tous les couples p, q d'entiers positifs et non nécessairement premiers entre eux. Or, ces derniers peuvent être comptés comme suit : il y en a un tel que $p + q = 2$, deux tels que $p + q = 3$, trois tels que $p + q = 4, \dots, k-1$ tels que $p + q = k, \dots$

Exercice II.1

1. Montrer l'inégalité : $\ell(\mathcal{F}_0) \leq \sum_{k=1}^{\infty} k 2^{1-k}$;

2. Conclure que $\ell(\mathcal{F}_0) \leq 2$;
3. On prend maintenant la famille $\mathcal{F}_n$ des intervalles $]r - 2^{-p-q-n}; r + 2^{-p-q-n}[$ avec p, q entiers strictement positifs et premiers entre eux. Constater que pour tout $n \geq 1$, l'ensemble E est toujours inclus dans leur réunion;
4. Vérifier que $\ell(\mathcal{F}_n) \leq 2^{-n+1}$. ■

D'après l'inégalité de la moyenne, on peut conclure que si $f(x)$ est une fonction bornée, positive et majorée par M , alors :

$$\forall n \in \mathbb{N}, \quad \int_E |f(x)| dx \leq \int_{F_n} |f(x)| dx \leq M \times \ell(F_n) \quad (\text{II.26})$$

de sorte que la contribution de l'ensemble E à l'intégrale est nulle. Ainsi, dans notre exemple, $\mathbb{Q}_+$ est un *ensemble négligeable* au sens de H. Lebesgue.

On peut, bien sûr, étendre cette notion d'ensembles négligeables au plan ou à l'espace. H. Lebesgue ⁽⁴⁾ a montré que les conditions que doivent satisfaire les fonctions pour qu'on puisse définir leur intégrale de façon logiquement correcte n'ont pas besoin d'être vérifiées dans tout le domaine d'intégration, mais seulement en dehors d'un ensemble négligeable.

Il existe beaucoup de théories mathématiques de l'intégration. La plus générale, qui sert de référence pour les mathématiciens, est la théorie de Lebesgue. La plus simple est celle de Riemann que nous utilisons ici et qui fonctionne pour les fonctions continues par morceaux. Elle consiste simplement à dire que l'intégrale $\int_a^b f(x) dx$ est la limite des sommes (dites de Riemann) $\sum_j f(x_j)(x_{j+1} - x_j)$, où x_j ($j = 0, 1, 2, 3, \dots, N$) est une subdivision discrète de l'intervalle $[a; b]$ par des points x_j tels que $\varepsilon = \sup_j |x_{j+1} - x_j|$ tende vers zéro lorsque N tend vers l'infini. On montre (facilement) que ces sommes convergent bien vers une limite *indépendante de la subdivision* lorsque la fonction f est continue (lorsqu'elle est continue par morceaux, on s'y ramène par simple découpage).

Notons bien que $\varepsilon = \sup_j |x_{j+1} - x_j|$ tend vers zéro lorsque N tend vers l'infini : elle signifie que les pas de la subdivision tendent *uniformément* vers zéro. On peut construire une théorie de l'intégration plus générale en considérant des discrétisations dont le pas ne tend pas uniformément vers zéro ⁽⁵⁾. Toutes les théories de l'intégration passent d'une façon ou d'une autre par l'intermédiaire de sommes discrètes et ne diffèrent au fond que par les conditions de convergence.

C'est pourquoi l'hypothèse que f est continue sur le domaine U (frontière comprise) entraîne que les sommes discrètes considérées (ici les sommes de Riemann) puissent être sommées dans n'importe quel ordre sans changer leur limite mais elle n'est pas *nécessaire* pour qu'il en soit ainsi.

Dans les calculs *pratiques*, il est très rare de rencontrer des cas qui nécessitent la théorie de Lebesgue ; si la fonction est discontinue, on pourra presque toujours découper le domaine en morceaux sur chacun desquels elle sera continue. Par contre on rencontrera fréquemment

(4) H. LEBESGUE, *Sur une généralisation de l'intégrale définie*. Comptes-rendus de l'Académie des Sciences, 1901, vol. 132, p. 1025–1028.

(5) J. KURZWEIL, *Generalized ordinary differential equations and continuous dependence on a parameter*. Czechoslovak Mathematical Journal, 1957, vol. 7(3), p. 418–446 — <http://dml.cz/dmlcz/100258>. Cette théorie fait tendre le pas vers zéro plus vite là où la fonction à intégrer est irrégulière que là où elle est continue ; cette astuce remplace le recours aux ensembles négligeables de Lebesgue.

le cas où la fonction f a une intégrale divergente, c'est-à-dire que la condition nécessaire pour que les sommes discrètes puissent être sommées dans n'importe quel ordre sans changer leur limite n'est pas satisfaite. Ces cas sont alors indépendants de la théorie : la différence entre (II.13) et (II.14) ne va pas disparaître si on prend l'intégrale au sens de Lebesgue ou de Kurzweil. C'est pourquoi l'approche élémentaire adoptée ici est largement suffisante pour tous les calculs pratiques. Toutefois, l'intégrale de Lebesgue est la seule qui convient pour avoir une théorie correcte des espaces $\mathbb{L}^2$ (voir les chapitres XI et XII).

II.3 Relations entre intégrales doubles et curvilignes

Supposons que $f(x, y)$ soit la dérivée partielle, soit par rapport à x , soit par rapport à y , d'une fonction connue ; c'est-à-dire que :

$$f = \frac{\partial a}{\partial y} \quad \text{ou} \quad f = \frac{\partial b}{\partial x} \quad (\text{II.27})$$

Pour pouvoir appliquer le théorème de la section précédente, on supposera ces dérivées partielles continues sur le domaine *fermé* U vérifiant les conditions 1 et 2 de la section II.2. On peut alors écrire les intégrales partielles :

– si $f = \frac{\partial a}{\partial y}$:

$$g(x) = \int_{y_0(x)}^{y_1(x)} \frac{\partial a}{\partial y}(x, y) dy = a(x, y_1(x)) - a(x, y_0(x)) \quad (\text{II.28})$$

– si $f = \frac{\partial b}{\partial x}$:

$$h(y) = \int_{x_0(y)}^{x_1(y)} \frac{\partial b}{\partial x}(x, y) dx = b(x_1(y), y) - b(x_0(y), y) \quad (\text{II.29})$$

Si on appelle γ la courbe fermée qui délimite le domaine U , on peut interpréter les fonctions $x \mapsto y_0(x)$ et $x \mapsto y_1(x)$, ou bien $y \mapsto x_0(y)$ et $y \mapsto x_1(y)$, comme des paramétrages de cette courbe. En effet, lorsque x va de x_0 à x_1 , le point de coordonnées $x, y_0(x)$ parcourt la partie inférieure de la courbe γ , puis, lorsque x revient de x_1 à x_0 , le point de coordonnées $(x, y_1(x))$ parcourt la partie supérieure de la courbe γ . Si nous paramétrons l'intégrale curviligne $\int_{\gamma} a(x, y) dx$ avec ce paramétrage, nous obtenons conformément à (II.6), l'expression :

$$\begin{aligned} \int_{\gamma} a(x, y) dx &= \int_{x_0}^{x_1} a(x, y_0(x)) dx + \int_{x_1}^{x_0} a(x, y_1(x)) dx \\ &= \int_{x_0}^{x_1} a(x, y_0(x)) dx - \int_{x_0}^{x_1} a(x, y_1(x)) dx \\ &= - \int_{x_0}^{x_1} g(x) dx = - \iint_U \frac{\partial a}{\partial y} dx dy \end{aligned} \quad (\text{II.30})$$

De la même façon, en prenant $f = \frac{\partial b}{\partial x}$, on obtiendrait en utilisant les paramétrages $y \mapsto x_0(y)$ et $y \mapsto x_1(y)$:

$$\int_{\gamma} b(x, y) dy = \iint_U \frac{\partial b}{\partial x} dx dy \quad (\text{II.31})$$

En tenant compte de l'additivité des intégrales et en regroupant ces deux résultats, on peut conclure par la relation :

$$\int_{\gamma} a(x, y) dx + b(x, y) dy = \iint_U \left[\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right] dx dy \quad (\text{II.32})$$

appelée *formule de Green*

On a vu à la section précédente que les intégrales partielles peuvent être définies et continues même si $\frac{\partial a}{\partial y}$ ou $\frac{\partial b}{\partial x}$ ont une singularité qui fait diverger l'intégrale double ; cela signifie que le membre de droite pourrait être une intégrale divergente, bien que le membre de gauche soit parfaitement défini. Quoique cela puisse faire l'objet d'une théorie à part entière, nous ne considérerons que le cas où le membre de droite est régulier, en supposant toujours pour cela que les dérivées partielles $\frac{\partial a}{\partial y}$ et $\frac{\partial b}{\partial x}$ sont continues.

II.4 Domaines ne vérifiant pas la condition 2

La démonstration donnée jusqu'ici de la formule de Green suppose que le domaine U vérifie les conditions 1 et 2 édictées en section II.2.

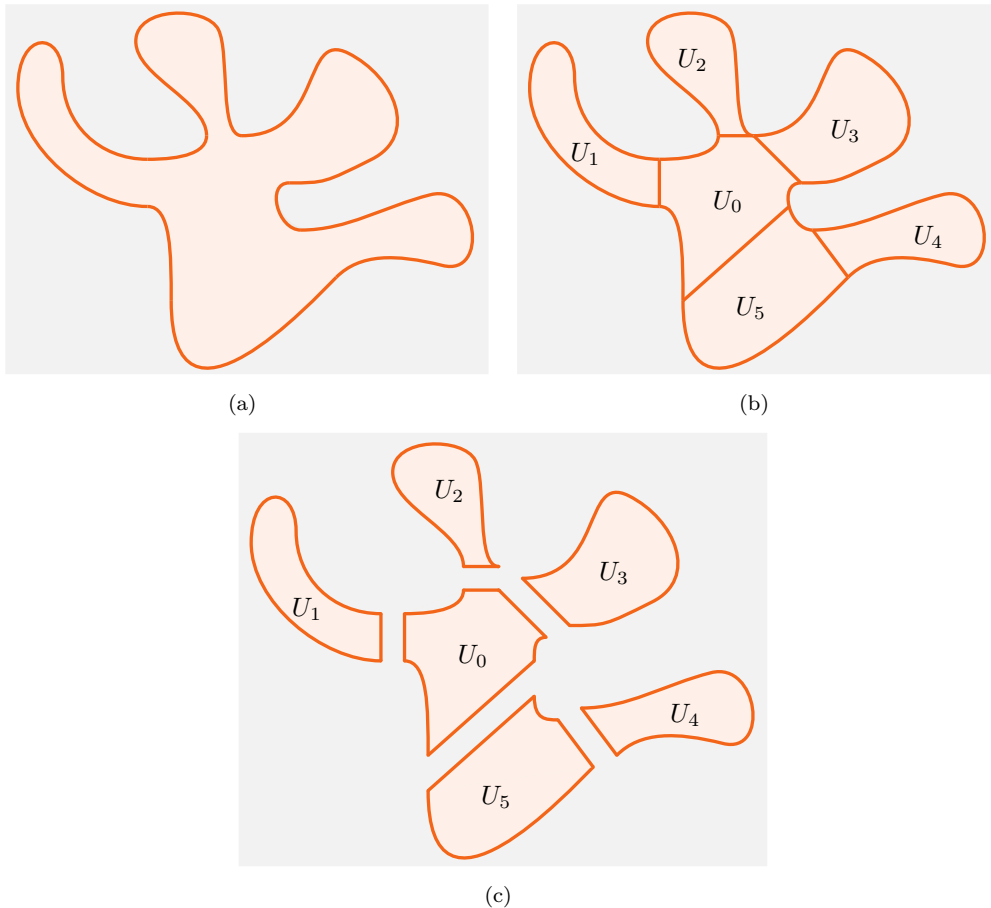


Figure II.2 – Méthode du pontage

Il est possible maintenant de montrer qu'elle reste vraie pour des domaines ne vérifiant que la condition 1, c'est-à-dire des domaines de forme quelconque. Pour effectuer cette généralisation, on recourt à la méthode du *pontage* que nous décrivons maintenant.

Il est en effet toujours possible de découper n'importe quel domaine U du plan ⁽⁶⁾ en régions $U_0, U_1, U_2, \dots$ qui vérifient la condition 2, comme le montre la figure II.2. Les régions U_j sont délimitées par des courbes fermées γ_j , différentiables par morceaux : ces courbes sont formées de fragments de la frontière γ de U et de *ponts* — qui sont rectilignes sur la figure II.2, mais rien n'impose qu'ils soient rectilignes. L'intégrale double $\iint_U$ ne peut pas être définie par la réduction à deux intégrations successives comme en II.2 car le domaine U ne vérifie pas la condition 2, mais on posera évidemment qu'elle est la somme des $\iint_{U_j}$, chacune de ces dernières étant réductible à deux intégrations successives ; en outre, la formule de Green démontrée précédemment s'applique séparément à chacun des domaines U_j , de sorte que :

$$\iint_U \left[\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right] dx dy = \sum_j \iint_{U_j} f(x, y) dx dy = \sum_j \int_{\gamma_j} a(x, y) dx + b(x, y) dy \quad (\text{II.33})$$

Lorsqu'on paramètre les chemins γ_j sur les morceaux qui forment les ponts, on constate que l'intégrale correspondante est toujours annulée par celle qui provient du domaine contigu, où le même pont est parcouru en sens inverse. En sommant les $\int_{\gamma_j}$, les contributions des ponts disparaîtront globalement et seuls subsisteront les contributions des morceaux issus du chemin initial γ . De sorte que :

$$\begin{aligned} \iint_U \left[\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right] dx dy &= \sum_j \iint_{U_j} f(x, y) dx dy = \sum_j \int_{\gamma_j} a(x, y) dx + b(x, y) dy \\ &= \int_{\gamma} a(x, y) dx + b(x, y) dy \end{aligned} \quad (\text{II.34})$$

Ainsi la formule de Green reste valable pour des domaines de forme quelconque et la condition 2 n'était qu'un intermédiaire technique.

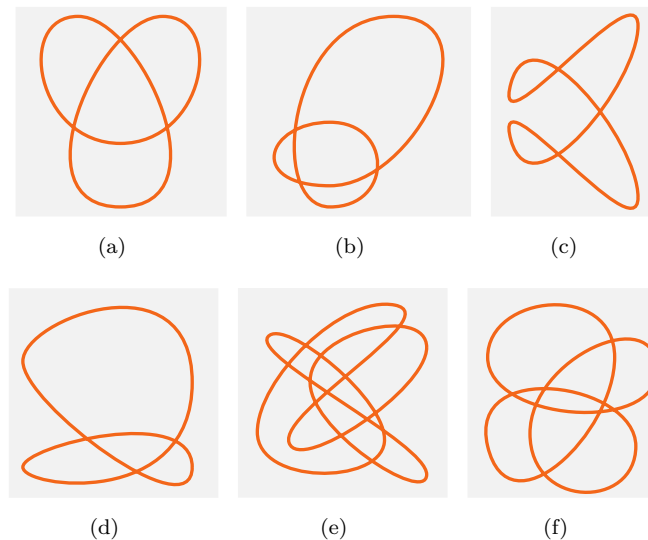


Figure II.3 – Lacets enchevêtrés

Un cas encore plus général est celui où le chemin γ , quoique fermé, n'est pas le bord d'un domaine. Des exemples de tels chemins sont donnés sur la figure II.3. Nous appellerons

⁽⁶⁾ Nous supposons toujours que le domaine est délimité par une courbe différentiable par morceaux.

lacet un chemin fermé différentiable par morceaux et nous dirons qu'un lacet est *simple* s'il est le bord d'un domaine du plan, qu'il est *multiple*, *enchevêtré* ou *entrelacé* dans le cas contraire. Les lacets de la figure II.3 sont donc des lacets *enchevêtrés*.

La formule de Green peut être étendue à des lacets enchevêtrés, à condition d'y introduire les modifications adéquates.

La figure II.4 montre comment procéder : elle montre en II.4(a) un chemin entrelacé γ qui ne peut être la frontière d'un domaine. Considérons le chemin *éclaté* représenté en II.4(b). Ce chemin est formé de quatre chemins γ_j qui eux constituent chacun la frontière d'un domaine U_j ($j = 0, 1, 2, 3$). On peut donc écrire :

$$\sum_{j=0}^{j=3} \int_{\gamma_j} a(x, y) dx + b(x, y) dy = \sum_{j=0}^{j=3} \int_{U_j} \left[\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right] dx dy \quad (\text{II.35})$$

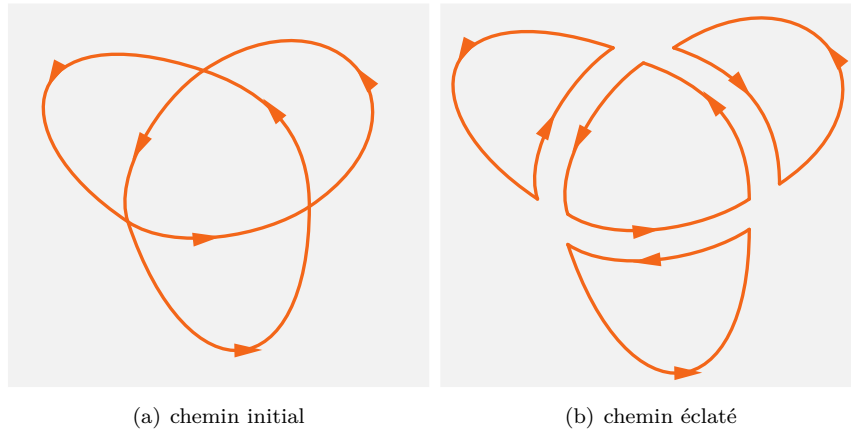


Figure II.4 – Formule de Green – elle s'applique à chacun des morceaux

Cependant, dans la somme (II.35), les intégrales s'annulent deux à deux sur les parties dédoublées, alors qu'il n'en est pas ainsi dans l'intégrale sur le chemin initial de la figure II.4. Si on fait le bilan de tous les morceaux, on obtient :

$$\int_{\gamma} a(x, y) dx + b(x, y) dy = \sum_{j=1}^{j=3} \int_{\gamma_j} a(x, y) dx + b(x, y) dy + 2 \int_{\gamma_0} a(x, y) dx + b(x, y) dy \quad (\text{II.36})$$

Le coefficient 2 devant le dernier terme du membre de droite compense les annulations mutuelles : en effet, si on fait simplement la somme sans coefficient, la contribution correspondant à γ_0 disparaît complètement ; il faut donc la rajouter une seconde fois.

En appliquant la formule de Green à chacun des termes de la somme, on obtient alors :

$$\int_{\gamma} a(x, y) dx + b(x, y) dy = 2 \int_{U_0} \left[\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right] dx dy + \sum_{j=1}^{j=3} \int_{U_j} \left[\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right] dx dy \quad (\text{II.37})$$

Il est donc possible d'utiliser la formule de Green même pour des lacets multiples, à condition de faire le bilan des annulations mutuelles.

Pour conclure cette section, signalons encore que l'hypothèse que les dérivées partielles $\frac{\partial a}{\partial y}$ et $\frac{\partial b}{\partial x}$ soient continues sur U peut aisément être affaiblie en les supposant seulement continues par morceaux, en précisant comme suit le sens de cette expression :

Définition II.1 Une fonction $f(x, y)$ est continue par morceaux sur le domaine fermé U s'il existe une famille de domaines fermés U_j délimités par des courbes différentiables et de fonctions f_j continues sur U_j telle que :

- U est la réunion des U_j ;
- les intérieurs des U_j sont disjoints (leurs frontières peuvent se chevaucher) ;
- sur tout l'intérieur de U_j , f coïncide avec f_j ;

Ce qu'il faut bien comprendre de cette définition, c'est que f peut être discontinue sur la frontière des U_j mais doit avoir de chaque côté un prolongement par continuité ; une fonction f qui tend vers l'infini en des points situés sur les frontières des U_j peut être continue à l'intérieur des U_j , mais ne sera pas continue par morceaux sur U .

On peut alors dire que la formule de Green s'applique aussi au cas où les dérivées partielles $\frac{\partial a}{\partial y}$ et $\frac{\partial b}{\partial x}$ sont continues par morceaux sur U . En effet, il suffit de l'appliquer d'abord à chacun des morceaux U_j ; en faisant ensuite la somme, les intégrales curvilignes sur les frontières intérieures s'annuleront deux à deux, car les fonctions $a(x, y)$ et $b(x, y)$ seront, elles, continues.

II.5 Intégrales curvilignes réductibles par quadrature

Lorsque l'expression $a(x, y) dx + b(x, y) dy$ qui figure sous le signe $\int_\gamma$ d'une intégrale curviligne est la différentielle d'une fonction connue $f(x, y)$, c'est-à-dire si :

$$a(x, y) dx + b(x, y) dy = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = df \quad (\text{II.38})$$

autrement dit si $a(x, y)$ et $b(x, y)$ sont les dérivées partielles d'une même fonction, alors l'intégrale curviligne est aisément intégrable, sans même recourir à un paramétrage de la courbe. En effet, soit $(x(t), y(t))$ avec $t_0 < t < t_1$, un paramétrage de γ ; on aura par définition de l'intégrale curviligne :

$$\int_\gamma a(x, y) dx + b(x, y) dy = \int_{t_0}^{t_1} \left[\frac{\partial f}{\partial x} x'(t) + \frac{\partial f}{\partial y} y'(t) \right] dt \quad (\text{II.39})$$

L'expression entre crochets est la dérivée de la fonction composée $t \mapsto f(x(t), y(t))$, d'où :

$$\int_\gamma a(x, y) dx + b(x, y) dy = f(x(t_1), y(t_1)) - f(x(t_0), y(t_0)) \quad (\text{II.40})$$

Les points M_0 de coordonnées $(x(t_0), y(t_0))$ et M_1 de coordonnées $(x(t_1), y(t_1))$ sont respectivement l'origine et l'extrémité du chemin, supposé ici être d'un seul tenant. Si le chemin est fermé (entrelacé ou non), on a évidemment $M_0 = M_1$ et l'intégrale curviligne est forcément nulle. Cela recoupe la formule de Green, puisque dans le cas présentement considéré $\frac{\partial b}{\partial x} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial a}{\partial y}$ et par conséquent l'expression sous l'intégrale double est nulle.

Le problème qui a longtemps intéressé les mathématiciens est la réciproque : étant donnée une expression de la forme $a(x, y) dx + b(x, y) dy$, peut-on trouver des conditions simples sur les fonctions a et b pour que cette expression soit la différentielle d'une fonction ?

Nous venons de voir que si $a(x, y) dx + b(x, y) dy$ est la différentielle d'une fonction $f(x, y)$ sur un domaine Ω du plan, alors son intégrale est nulle sur tout lacet contenu dans Ω . La réciproque de cet énoncé est vraie :

Si pour tout chemin fermé γ sur un domaine U du plan (x, y) l'intégrale curviligne :

$$\int_{\gamma} a(x, y) dx + b(x, y) dy$$

est nulle, alors il existe une fonction $f(x, y)$, définie et différentiable sur Ω telle que $a = \frac{\partial f}{\partial x}$ et $b = \frac{\partial f}{\partial y}$. $\square$

Il suffit en effet de poser :

$$f(x_1, y_1) = \int_{\gamma} a(x, y) dx + b(x, y) dy \quad (\text{II.41})$$

où γ est n'importe quel chemin, contenu dans Ω , d'origine (x_0, y_0) et d'extrémité (x_1, y_1) . L'intégrale ne dépend pas du chemin, car si on prend deux chemins différents ayant tous deux (x_0, y_0) pour origine et (x_1, y_1) pour extrémité, alors le chemin formé par la concaténation du premier et du second parcouru en sens inverse est fermé. Pour calculer les dérivées partielles de $f(x_1, y_1)$, on considérera séparément les deux fonctions $x \mapsto f(x, y_1)$ et $y \mapsto f(x_1, y)$. Soit un chemin γ_h allant de (x_1, y_1) à $(x_1 + h, y_1)$, paramétré par t avec $0 < t < h$. Alors :

$$f(x_1 + h, y_1) - f(x_1, y_1) = \int_{\gamma_h} a(x, y) dx + b(x, y) dy = \int_0^h a(x_1 + t, y_1) dt \quad (\text{II.42})$$

ce qui signifie bien que $a = \frac{\partial f}{\partial x}$. On procède de même pour la fonction $y \mapsto f(x_1, y)$.

Une autre réciproque est possible. Si $a(x, y) dx + b(x, y) dy$ est la différentielle d'une fonction, alors $\frac{\partial b}{\partial x} = \frac{\partial a}{\partial y}$, car on sait que $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$.

La réciproque serait donc :

Si $a(x, y)$ et $b(x, y)$ sont deux fonctions différentiables sur un domaine Ω telles que $\frac{\partial b}{\partial x} = \frac{\partial a}{\partial y}$, alors il existe une fonction $f(x, y)$, définie et différentiable sur Ω qui satisfait les conditions $a = \frac{\partial f}{\partial x}$ et $b = \frac{\partial f}{\partial y}$. $\square$

Elle n'est pas toujours vraie : si on se réfère à la formule de Green, on voit que la condition $\frac{\partial b}{\partial x} = \frac{\partial a}{\partial y}$ entraîne que sur tout sous-domaine U de Ω délimité par le chemin fermé γ , on aura $\int_{\gamma} a(x, y) dx + b(x, y) dy = 0$ mais cela ne signifie pas que l'intégrale est nulle pour tout chemin fermé : la formule de Green assure seulement qu'elle est nulle pour tout chemin fermé qui est contenu, *ainsi que le domaine U qu'il délimite*, dans Ω . La figure II.5 illustre la nuance.

Sur cette même figure, bien que l'expression $\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}$ soit nulle sur le domaine blanc, on ne peut pas déduire de la formule de Green que $\int_{\Gamma_1} a(x, y) dx + b(x, y) dy = 0$, car il faudrait pour cela que $\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}$ soit nulle *partout* à l'intérieur de la région délimitée par le chemin Γ_1 . Par contre, on a bien :

$$\int_{\Gamma_2} a(x, y) dx + b(x, y) dy = 0 \quad \text{et} \quad \int_{\Gamma_3} a(x, y) dx + b(x, y) dy = 0 \quad (\text{II.43})$$

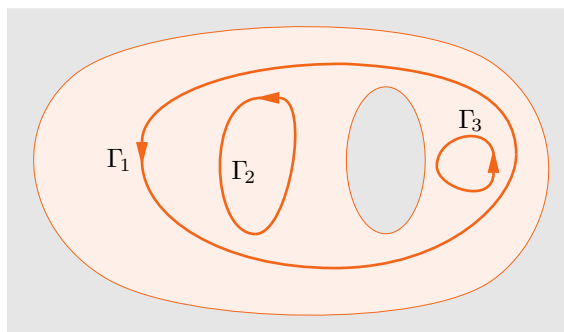


Figure II.5 – Exemple de domaine à trous

Terminologie

Lorsque les fonctions $a(x, y)$ et $b(x, y)$ satisfont à la condition $\frac{\partial b}{\partial x} = \frac{\partial a}{\partial y}$, on dit que la forme différentielle $a(x, y) dx + b(x, y) dy$ est *fermée* ; lorsqu'il existe une fonction $f(x, y)$ telle que $a = \frac{\partial f}{\partial x}$ et $b = \frac{\partial f}{\partial y}$, on dit que la forme différentielle $a(x, y) dx + b(x, y) dy$ est *exacte* et que f est la fonction *primitive* de la forme différentielle. La réciproque ci-dessus peut donc s'énoncer : *toute forme différentielle fermée est exacte*. On voit qu'elle est fautive dans le cas de la figure II.5.

II.6 Domaines à trous

Cette réciproque est donc parfois fautive, pour des domaines *à trous* — nous préciserons le sens de cette expression plus loin. Un exemple de domaine à trou est le plan privé de l'origine : $\mathbb{R}^2 - (0, 0)$: c'est l'exemple le plus simple. Sur ce domaine, le contre-exemple le plus simple à la réciproque mentionnée ci-dessus est celui de l'expression différentielle :

$$-\frac{y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy \quad (\text{II.44})$$

Nous verrons dans la suite de ce cours des exemples variés mais celui-ci est l'exemple-type. Sur la figure II.6, l'expression $\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}$ étant nulle sur le domaine blanc, on peut déduire de

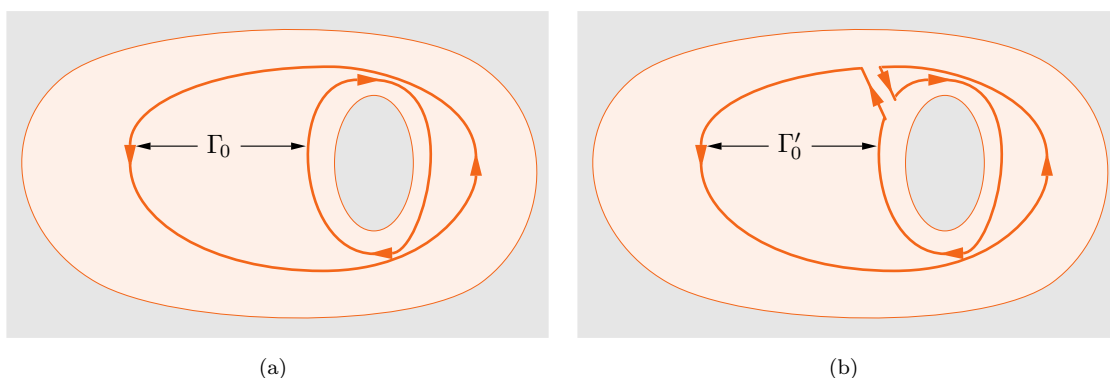


Figure II.6 – Chemin à deux composantes

la formule de Green que :

$$\int_{\Gamma_0'} a(x, y) dx + b(x, y) dy = 0 \quad (\text{II.45})$$

car le chemin Γ_0 entoure une région entièrement contenue dans le domaine blanc. Par conséquent, on aura aussi :

$$\int_{\Gamma_0} a(x, y) dx + b(x, y) dy = 0 \quad (\text{II.46})$$

On peut aussi exprimer ce résultat en disant que $\int a(x, y) dx + b(x, y) dy$ a la même valeur sur les deux composantes (orientées dans le même sens) du chemin Γ_0 . Le calcul des dérivées donne :

$$\frac{\partial a}{\partial y} = \frac{-x^2 + y^2}{(x^2 + y^2)^2} \quad \frac{\partial b}{\partial x} = \frac{-x^2 + y^2}{(x^2 + y^2)^2} \quad (\text{II.47})$$

et on constate que $\frac{\partial a}{\partial y} = \frac{\partial b}{\partial x}$. Cette égalité a lieu seulement dans le domaine troué $\mathbb{R}^2 - (0, 0)$, puisque au point $(0, 0)$ lui-même les fonctions a et b sont singulières ; en particulier, on ne peut pas considérer une intégrale du type $\iint_U [\frac{\partial a}{\partial y} - \frac{\partial b}{\partial x}] dx dy$ sur un domaine U qui contiendrait l'origine, puisque cette intégrale serait divergente. Par contre il n'y a aucun problème pour un domaine U qui ne contient pas l'origine. Cela implique que la formule de Green n'est applicable que pour des chemins qui ne contournent pas l'origine et on ne peut donc pas en déduire que $\int_{\gamma} a dx + b dy = 0$ si γ est un chemin qui entoure l'origine. Si on veut connaître la valeur de cette intégrale, il faut la calculer à la main.

Effectuons ce calcul dans le cas où γ est un cercle de centre l'origine et de rayon R . Le paramétrage est alors $x(t) = R \cos t, y(t) = R \sin t$ avec $0 < t < 2\pi$ d'où :

$$\int_{\gamma} a(x, y) dx + b(x, y) dy = \int_0^{2\pi} \frac{(-R \sin t)(-R \sin t) + (R \cos t)(R \cos t)}{(R \cos t)^2 + (R \sin t)^2} dt = 2\pi \quad (\text{II.48})$$

On constate en effet que cette valeur n'est pas nulle. Si on considère maintenant l'intégrale $\int_{\gamma'} a(x, y) dx + b(x, y) dy$ sur un chemin γ' *quelconque* entourant l'origine, alors le chemin obtenu par la concaténation de γ' et du cercle γ parcouru en sens inverse aurait été un chemin fermé entourant un domaine U en forme de couronne et ne contenant pas l'origine — figure II.6, auquel on peut donc appliquer la formule de Green, ce qui montre que $\int_{\gamma'} = \int_{\gamma} = 2\pi$.

Ce contre-exemple est là pour montrer que si γ entoure l'origine, l'intégrale sur γ d'une expression différentielle non définie en $(0, 0)$ n'est pas *forcément* nulle ; mais il peut quand même arriver qu'elle soit nulle. Par exemple, prenons :

$$a = \frac{x}{\sqrt{x^2 + y^2}} \quad \text{et} \quad b = \frac{y}{\sqrt{x^2 + y^2}} \quad (\text{II.49})$$

La situation est la même que dans l'exemple précédent. La quantité $\frac{\partial a}{\partial y} - \frac{\partial b}{\partial x}$ est nulle en dehors de $(0, 0)$ et non définie en $(0, 0)$ alors que sur le cercle γ :

$$\int_{\gamma} a(x, y) dx + b(x, y) dy = \int_0^{2\pi} \frac{(R \cos t)(-R \sin t) + (R \sin t)(R \cos t)}{\sqrt{(R \cos t)^2 + (R \sin t)^2}} dt = 0 \quad (\text{II.50})$$

L'intégrale est nulle même sur des lacets qui entourent l'origine, mais cela ne se déduit pas de la formule de Green et n'est pas une vérité générale.

II.7 Homologie des lacets

Dans la section II.4, nous avons examiné ce que devenait la formule de Green pour des lacets multiples tels que celui de la figure II.4. Nous avons pu constater que le lacet multiple pouvait être décomposé en lacets simples, de telle sorte que l'intégrale curviligne sur le lacet multiple soit une combinaison linéaire, à coefficients entiers⁽⁷⁾, des intégrales sur chaque élément simple (voir (II.36)).

Considérons par exemple l'intégrale de l'expression différentielle (II.44) sur le lacet de la figure II.4. En combinant (II.36) et (II.48) on voit que l'intégrale vaudra 4π si l'origine des coordonnées est dans la partie U_0 du domaine $\Omega = \mathbb{R}^2 - (0, 0)$, 2π si l'origine est dans l'une des parties U_1 , U_2 , ou U_3 et zéro si l'origine est à l'extérieur du lacet (le cas où l'origine serait exactement sur le lacet est exclu, car l'intégrale serait alors divergente). En décomposant n'importe quel lacet enchevêtré (par exemple ceux de la figure II.3) en éléments simples, on comprend aisément que l'intégrale de l'expression différentielle (II.44) sur un tel lacet sera toujours un multiple entier algébrique de 2π , la valeur du coefficient entier dépendant de la position relative de l'origine des coordonnées par rapport au lacet.

Lorsqu'on considère des intégrales curvilignes d'expressions différentielles (satisfaisant sur un domaine Ω à la condition $\frac{\partial b}{\partial x} = \frac{\partial a}{\partial y}$) sur des lacets enchevêtrés, la question importante qui se pose est de savoir dans quelle mesure l'intégrale dépend du lacet. Si le domaine Ω dans lequel on considère les lacets est dépourvu de trous, la question est vite résolue, puisque l'intégrale est nulle quel que soit le lacet. Si le domaine comporte des trous, il y aura des expressions différentielles telles que (II.44) pour lesquelles les intégrales curvilignes seront non nulles. On donnera la définition suivante :

Définition II.2 Deux lacets γ_1 et γ_2 sont dits *homologiquement équivalents* sur un domaine Ω si le lacet $\gamma = \gamma_1 - \gamma_2$ (concaténation de γ_1 et de γ_2 parcouru en sens inverse) est le bord d'un domaine U entièrement contenu dans Ω .

On comprend alors que, si l'égalité $\frac{\partial b}{\partial x} = \frac{\partial a}{\partial y}$ a lieu partout dans Ω , elle aura lieu dans U et que la formule de Green donne alors forcément :

$$\int_{\Gamma_1} a(x, y) dx + b(x, y) dy = \int_{\Gamma_2} a(x, y) dx + b(x, y) dy \quad (\text{II.51})$$

puisque $\gamma_1 - \gamma_2$ est le bord de U . La décomposition d'un lacet enchevêtré en lacets simples (figure II.4) est la réalisation pratique de cette définition. On voit que l'homologie est une propriété purement géométrique du domaine Ω .

Ⓡ Deux lacets peuvent être homologiquement équivalents pour un domaine et pas pour un autre ; par exemple si on supprime un trou.

La condition $\frac{\partial b}{\partial x} = \frac{\partial a}{\partial y}$ est évidemment essentielle pour que l'homologie des lacets entraîne l'égalité des intégrales. Si cette condition n'était pas satisfaite dans le domaine, la moindre déformation du chemin changerait la valeur de l'intégrale.

On voit facilement que l'homologie est une relation d'équivalence. étant donné que les chemins (et donc les lacets) peuvent se concaténer et que l'intégrale est alors additive, il est

(7) Le fait que pour n'importe quel lacet on aura toujours une *combinaison linéaire, à coefficients entiers* n'a pas été démontré de façon générale, mais seulement dans le cas particulier du lacet de la figure II.4. Il est toutefois aisé d'imaginer comment le procédé utilisé pourrait se généraliser.

clair que l'ensemble des classes d'équivalences de l'homologie forme un groupe commutatif, appelé *groupe d'homologie* du domaine. Pour des domaines du plan, on peut déterminer facilement ce groupe d'homologie : il est réduit à l'élément neutre zéro pour des domaines sans trous ; il est égal au groupe $\mathbb{Z}$ des entiers algébriques pour les domaines à un seul trou : en effet, pour tout lacet, un seul de ses éléments simples au plus peut contenir l'unique trou ; la classe d'homologie du lacet est alors simplement déterminée par le nombre de fois qu'on compte cet élément simple pour compenser les annulations mutuelles sur les ponts et il y a autant de possibilités que d'entiers algébriques.

Plus généralement, le groupe est égal à $\mathbb{Z}^n$ pour les domaines à n trous. Cela se comprend sans difficulté. Parmi les éléments simples issus de la décomposition d'un lacet, certains entourent un trou, d'autres aucun, d'autres plusieurs à la fois. Dans ce dernier cas il est toujours possible en introduisant des ponts convenablement choisis, de se ramener à des éléments encore plus simples qui chacun n'entourent qu'un seul trou, la contribution des ponts dans l'intégrale disparaissant par annulation mutuelle. On peut dire que n'importe quel lacet est homologiquement équivalent à une concaténation de lacets dits élémentaires⁽⁸⁾ qui n'entourent qu'un seul trou à la fois. Cela est illustré sur les figures II.7(a) et II.7(b). On associe à chaque trou le nombre, compté algébriquement, de lacets élémentaires qui l'entourent : ainsi, si le trou $N^{\circ}j$ est entouré de p lacets élémentaires orientés dans le sens positif et q lacets élémentaires orientés dans le sens négatif, le j^{e} nombre sera $p - q$. Pour un domaine à n trous, on obtient une famille ordonnée de n entiers algébriques, c'est-à-dire (dans le langage de la théorie des ensembles) un élément de $\mathbb{Z}^n$, qui caractérise donc la classe d'équivalence ou classe d'homologie du lacet.

La notion d'homologie est quelque peu aplatie dans le cas des domaines du plan, car leurs caractéristiques géométriques sont trop particulières ; elle devient plus intéressante pour des surfaces courbes de dimension supérieure à deux. L'homologie est entièrement déterminée par la décomposition des lacets en éléments simples et le pontage : elle caractérise donc les propriétés des domaines (et des surfaces) relatives au découpage en régions élémentaires, indépendamment des dimensions métriques. Ces propriétés se conservent si on déforme les surfaces sans rien déchirer ; par exemple, le nombre de trous est conservé dans les déformations continues : pour créer un nouveau trou ou en réunir deux en un seul, il faut déchirer. La décomposition d'un lacet et les liaisons par des ponts conservent leur structure combinatoire, même si les longueurs subissent d'énormes distorsions. Par conséquent, le groupe d'homologie d'une surface condense toute l'information sur la structure purement topologique de la surface.

Dans le cas des domaines du plan, l'information serait tout aussi bien exprimée par le nombre de trous et le recours au groupe d'homologie est inutile. C'est pourquoi nous n'en parlerons plus dans cet ouvrage consacré au plan. En revanche, l'homologie devient un outil mathématique très puissant pour étudier des hypersurfaces de dimension quelconque (surtout supérieure à deux), car l'intuition géométrique, limitée à l'espace à trois dimensions, disparaît alors complètement. La discipline mathématique consacrée à l'étude de l'homologie

(8) Ne pas confondre *lacet élémentaire* qui n'entoure qu'un seul trou à la fois, et *lacet simple* qui délimite une région du plan.

logie des hypersurfaces s'appelle la *topologie algébrique*. L'idée remonte à B. Riemann et H. Poincaré. Elle a été développée systématiquement au vingtième siècle, notamment par H. Cartan.

II.8 Intégrales curvilignes à variable complexe

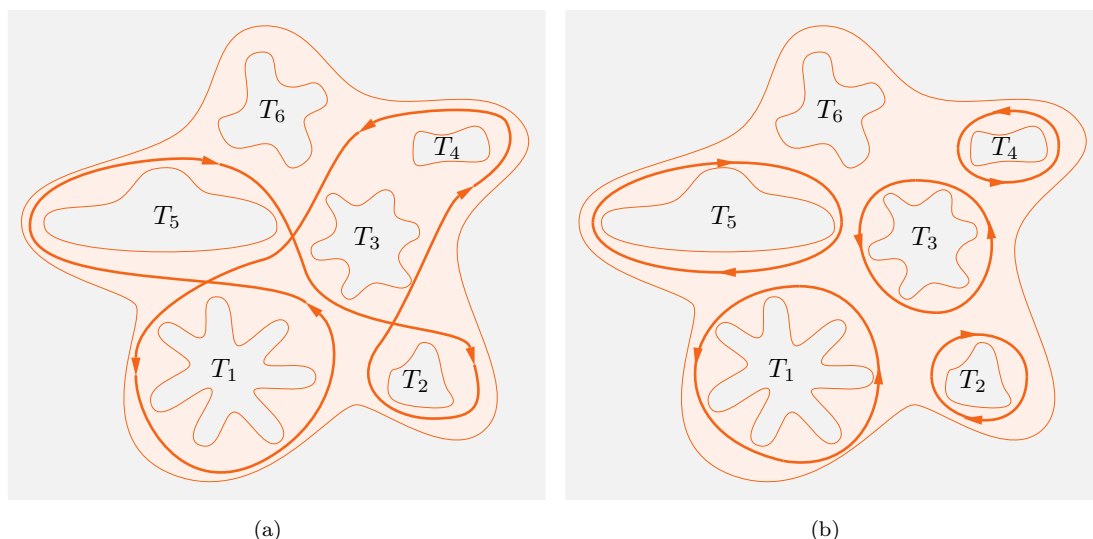


Figure II.7 – Le lacet enchevêtré sur la figure (a) est homologiquement équivalent à la concaténation des lacets simples de (b) : cette décomposition met en évidence la classe d'homologie du chemin : il tourne une fois en sens positif autour des trous N^{os} 1, 3 et 4, une fois en sens rétrograde autour des trous N^{os} 2 et 5 et pas du tout autour du trou N^o 6. On peut donc représenter la classe d'homologie du lacet par un élément du groupe $\mathbb{Z}^6$, l'élément $(1, -1, 1, 1, -1, 0)$

Les nombres complexes ont toujours été un moyen commode de représenter un point du plan, ou, ce qui revient au même, un couple de nombres réels ; au lieu d'écrire (x, y) , on écrit $x + iy$.

Par conséquent tout ce qui a été fait dans les sections précédentes peut être traduit dans le langage des nombres complexes. N'importe quelle fonction $f(x, y)$ peut être interprétée comme une fonction de la variable complexe $z = x + iy$. Toutefois, le recours aux nombres complexes n'a d'intérêt que dans la mesure où il apporte une simplification des opérations algébriques, ou du moins un allègement sensible des notations.

Voyons comment on pourrait traduire les expressions différentielles du type $a(x, y) dx + b(x, y) dy$. Il est naturel de poser $dz = dx + i dy$; mais on peut constater après quelques manipulations que cela ne permet pas d'écrire n'importe quelle expression différentielle. Par exemple $z^2 dz = ((x^2 - y^2) dx - 2xy dy) + i(2xy dx + (x^2 - y^2) dy)$. Si on prend l'expression à peine modifiée $((x^2 - y^2) dx - xy dy) + i(2xy dx + (x^2 - y^2) dy)$ (on a simplement remplacé dans la partie réelle $2xy$ par xy), on essaiera vainement de l'écrire sous la forme $f(z) dz$: en effet la seconde expression est égale à $z^2 dz + xy dx$, donc il faudrait pouvoir écrire $xy dx$ sous la forme $(A + iB)(dx + i dy) = (A dx - B dy) + i(B dx + A dy)$, ce qui exigerait que $B dx + A dy$ soit identiquement nul (et donc que A et B soient elles-mêmes nulles). Le problème vient de ce que dans ces calculs algébriques où dx , dy et dz sont traités comme des variables,

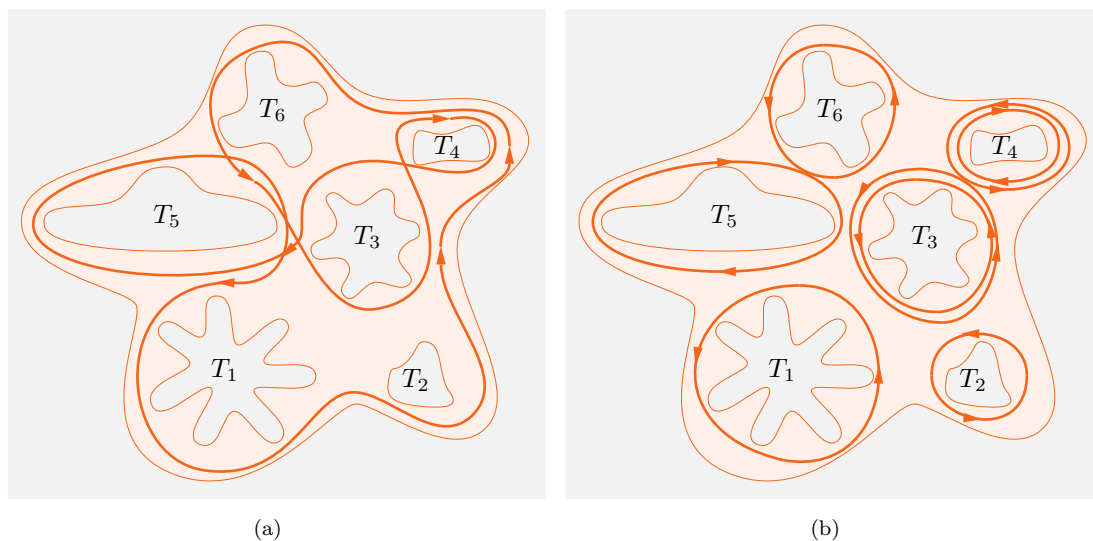


Figure II.8 – Le lacet enchevêtré sur la figure (a) est homologiquement équivalent à la concaténation des lacets simples de (b) : cette décomposition met en évidence la classe d'homologie du chemin : il tourne une fois en sens positif autour des trous N°s 1, 2 et 6, une fois en sens rétrograde autour du trou N° 5, deux fois en sens positif autour du trou N° 3 et une fois dans chacun des deux sens (ce qui fait un bilan nul) autour du trou N° 4. On peut donc représenter la classe d'homologie du lacet par un élément du groupe $\mathbb{Z}^6$, l'élément $(1, 1, 2, 0, -1, 1)$

il n'est pas possible d'exprimer dx seul ou dy seul en fonction de dz . Pour y parvenir, il faut introduire aussi $d\bar{z} = dx - i dy$. Alors on peut formellement écrire $dx = \frac{1}{2}(dz + d\bar{z})$ et $dy = \frac{1}{2i}(dz - d\bar{z})$. Par exemple $xy dx$ s'écrira $\frac{1}{2}(xy dz + xy d\bar{z})$.

Bien entendu, il ne faut pas perdre de vue au cours des manipulations algébriques formelles de dx , dy , dz et $d\bar{z}$, que les expressions différentielles $a(x, y) dx + b(x, y) dy$ ne représentent pas des nombres et ne prennent leur sens que dans l'intégrale : ce qui représente un nombre est en réalité $a(x(t), y(t)) x'(t) + b(x(t), y(t)) y'(t)$. De même $f(z) dz + g(z) d\bar{z}$ est en réalité une abréviation pour $f(z(t)) z'(t) + g(z(t)) \bar{z}'(t)$, avec $z(t) = x(t) + iy(t)$, $z'(t) = x'(t) + iy'(t)$ et $\bar{z}'(t) = x'(t) - iy'(t)$. Les manipulations algébriques sur dx et dy ne sont qu'un allègement d'écriture ; les calculs effectués sur les expressions $a(x(t), y(t)) x'(t) + b(x(t), y(t)) y'(t)$ seraient directement légitimés par le sens numérique, mais alourdis par l'écriture et pourtant *formellement* semblables.

Ayant ainsi précisé le sens des écritures en nombres complexes, voyons comment se traduisent les résultats de la section II.5. Il s'agit de savoir à quelle condition doit satisfaire une expression du type $f(z) dz + g(z) d\bar{z}$ pour être la différentielle d'une fonction (il est toujours sous-entendu que les dérivées partielles de f et g sont continues). Posons $a(x, y) = \Re(f(x + iy))$, $b(x, y) = \Im(f(x + iy))$, $u(x, y) = \Re(g(x + iy))$, $v(x, y) = \Im(g(x + iy))$. Alors :

$$\begin{aligned} f(z) dz + g(z) d\bar{z} &= (a + ib)(dx + i dy) + (u + iv)(dx - i dy) \\ &= ((a + u) dx + (-b + v) dy) + i((b + v) dx + (a - u) dy) \end{aligned} \quad (\text{II.52})$$

Pour que cela soit la différentielle d'une fonction, il faut que la partie réelle et la partie

imaginaire soient la différentielle d'une fonction. Nous avons vu que pour cela il faut que :

$$\frac{\partial(-b+v)}{\partial x} = \frac{\partial(a+u)}{\partial y} \quad \text{et} \quad \frac{\partial(b+v)}{\partial y} = \frac{\partial(a-u)}{\partial x} \quad (\text{II.53})$$

En utilisant les nombres complexes, on peut écrire cela sous la forme :

$$\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(u+iv) = \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)(a+ib) \quad (\text{II.54})$$

Afin de condenser encore plus les notations, posons :

$$\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) = 2\frac{\partial}{\partial z} \quad \text{et} \quad \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) = 2\frac{\partial}{\partial \bar{z}} \quad (\text{II.55})$$

Le facteur 2 qui apparaît en coefficient des seconds membres semble ici purement conventionnel, mais nous en verrons la justification plus tard (en III.2). Alors (II.54) s'écrit sous la forme très simple :

$$\frac{\partial g}{\partial z} = \frac{\partial f}{\partial \bar{z}} \quad (\text{II.56})$$

Si l'expression différentielle considérée se réduit au premier terme $f(z)dz$, cette condition devient $\partial f/\partial \bar{z} = 0$ (car alors $g = 0$) ; si elle se réduit au second terme $g(z)d\bar{z}$, la condition devient $\partial g/\partial z = 0$ (car alors $f = 0$). Nous verrons, au chapitre III, que les fonctions qui vérifient $\partial f/\partial \bar{z} = 0$ sont appelées *fonctions analytiques* et nous étudierons l'ensemble de leurs propriétés. Il résulte de (II.56) que sur un domaine sans trou, $f(z)$ est analytique si et seulement si l'intégrale curviligne de $f(z)dz$ sur tout lacet est nulle.

Bien entendu, il n'y a là aucun résultat nouveau par rapport à ce qui a été vu dans les sections II.5, II.6 et II.7 ; il s'agit des *mêmes* résultats, exprimés dans des notations mieux adaptées aux nombres complexes.

L'opérateur $\partial f/\partial z$ peut être interprété comme une véritable dérivation par rapport à z lorsque f est analytique, comme nous le verrons en section III.2. Son sens n'est plus du tout aussi clair si f n'est pas analytique ; il ne faut y voir alors qu'une écriture commode.

En conclusion :

$f(z)$ et $g(z)$ étant deux fonctions de la variable complexe $z = x + iy$ et continûment différentiables par rapport à x et y , pour que l'expression différentielle $f(z)dz + g(z)d\bar{z}$ soit la différentielle d'une fonction définie sur un domaine sans trous Ω , il faut et il suffit que soit satisfaite la relation :

$$\frac{\partial g}{\partial z} = \frac{\partial f}{\partial \bar{z}}$$

□



III — Fonctions analytiques

III.1 Une propriété des polynômes

Pour une fonction arbitraire, la dérivée est définie comme une limite ; par contre pour les polynômes, la dérivée peut être calculée de manière purement algébrique. Par exemple, la dérivée de X^k est kX^{k-1} . On peut exprimer cela de manière algorithmique : *on obtient la dérivée de X^k en diminuant l'exposant d'une unité et en multipliant le tout par l'ancien exposant*. Le résultat de cette transformation purement algébrique coïncide avec la limite de $((X+t)^k - X^k)/t$ lorsque t tend vers zéro. On peut trouver d'autres fonctions (autres que les polynômes) ayant une telle propriété : par exemple la fonction exponentielle, où l'opération algébrique est encore plus simple puisque la dérivée est identique à la fonction. Si on considère l'ensemble de toutes les fonctions possibles, on a l'impression que celles qui jouissent d'une telle propriété sont exceptionnelles et très rares. Inversement, si on considère les fonctions usuelles des mathématiques, qui sont définies par les opérations arithmétiques addition, multiplication, division, exponentiation, puis combinées par composition, on a au contraire l'impression qu'elles constituent la règle et non l'exception. D'ailleurs, ce qui fait la puissance du *calcul infinitésimal* est justement la possibilité de calculer les dérivées algébriquement, grâce à des formules telles que $(fg)' = f'g + fg'$, $(f/g)' = (f'g - fg')/g^2$, $(f \circ g)' = (f' \circ g) g'$ etc.

Dans le calcul de la dérivée d'une fonction par la limite du taux d'accroissement, il était essentiel que la variable soit réelle ; si la variable est complexe, la notion même d'accroissement perd son sens puisque les nombres complexes ne sont pas ordonnés. Par contre, dans le calcul purement algébrique de la dérivée, il importait peu que la variable soit réelle ou complexe, parce que les opérations $+$, $-$, $\times$, $/$ s'étendent aux nombres complexes. Ainsi

avec $z = x + iy$, la fonction z^k a toujours un sens clair et sa dérivée aussi : rien n'empêche de dire *par définition* que kz^{k-1} est la dérivée de z^k . Il est indifférent que la variable soit réelle ou complexe, pour calculer la dérivée ou la primitive d'un polynôme. Par exemple, le polynôme :

$$P(x) = x^7 - 3x^6 + x^5 + 4x^4 - 7x^3 + 2x^2 - x - 2 \quad (\text{III.1})$$

a pour dérivée :

$$P'(x) = 7x^6 - 18x^5 + 5x^4 + 16x^3 - 21x^2 + 4x - 1 \quad (\text{III.2})$$

et pour primitive (nulle en zéro) :

$$Q(x) = \frac{x^8}{8} - \frac{3x^7}{7} + \frac{x^6}{6} + \frac{4x^5}{5} - \frac{7x^4}{4} + \frac{2x^3}{3} - \frac{x^2}{2} - 2x \quad (\text{III.3})$$

Ces opérations, consistant à passer à la dérivée ou à la primitive, restent formellement valables si on remplace x par un nombre complexe. Le problème est évidemment de comprendre le sens de ces opérations lorsque la variable est complexe.

On peut constater directement que, tout comme lorsque x est réel, kz^{k-1} est aussi la limite de $((z+w)^k - z^k)/w$ lorsque w tend vers zéro, mais avec la différence essentielle que voici : un nombre réel t peut tendre vers zéro à gauche, à droite, ou de façon alternée (en oscillant) mais un nombre complexe (parce qu'il varie selon deux dimensions) peut tendre vers zéro selon n'importe quelle direction ou même en suivant une spirale. Or l'expression $((z+w)^k - z^k)/w$ tend vers kz^{k-1} *quelle que soit la manière* dont w tend vers zéro. C'est donc une contrainte nettement plus forte que dans le domaine réel.

Pour bien comprendre cela, voyons un exemple simple. On va considérer d'une part la fonction $f(z) = z^2 = x^2 - y^2 + i2xy$, qui est donc polynomiale et d'autre part, la fonction $g(z) = x^2 - y^2 + ixy$, qui est aussi une fonction de la variable complexe z et elle-même à valeurs complexes, tout comme $f(z)$, mais qui ne peut pas s'exprimer à partir de z uniquement comme produit ou puissance : pour exprimer $g(z)$, on est obligé de séparer x et y , on ne peut l'écrire comme polynôme en z .

Cherchons alors les limites des rapports :

$$\frac{f(z+w) - f(z)}{w} \quad \text{et} \quad \frac{g(z+w) - g(z)}{w} \quad (\text{III.4})$$

On voit que :

$$\frac{f(z+w) - f(z)}{w} = \frac{z^2 + 2zw + w^2 - z^2}{w} = 2z + w \quad (\text{III.5})$$

Le calcul est très simple car on utilise la multiplication et la division des nombres complexes. Lorsque w tend vers zéro (et peu importe selon quelle direction), l'expression (III.5) tend vers $2z$. Considérons la deuxième situation. Posons $w = u + iv = r(\cos \theta + i \sin \theta)$:

$$\begin{aligned} \frac{g(z+w) - g(z)}{w} &= \frac{(x+u)^2 - (y+v)^2 + i(x+u)(y+v) - x^2 + y^2 - ixy}{u + iv} \\ &= \left[(2x \cos \theta + r \cos^2 \theta - 2y \sin \theta - r \sin^2 \theta) + \right. \\ &\quad \left. i(x \sin \theta + y \cos \theta + r \cos \theta \sin \theta) \right] (\cos \theta - i \sin \theta) \end{aligned} \quad (\text{III.6})$$

Lorsque r tend vers zéro (θ restant fixe), cela tend vers une limite *qui dépend de θ* :

$$\lim_{r \rightarrow 0} \frac{g(z+w) - g(z)}{w} = (x(1 + \cos^2 \theta) - y \sin \theta \cos \theta) - i(x \sin \theta \cos \theta - y(1 + \sin^2 \theta)) \quad (\text{III.7})$$

Dans le cas de la fonction $g(z)$, il n'existe donc pas de limite unique lorsque w tend vers zéro ; la limite dépend de la direction θ selon laquelle w tend vers zéro. Bien entendu, la fonction g , considérée comme fonction des deux variables réelles x et y , est dérivable et possède des dérivées partielles selon x et selon y . On remarque que *si z et w sont réels*, alors $\theta = 0$ ou π et on trouve dans ces deux cas la même limite ; c'est seulement lorsqu'on sort de l'axe réel (θ prenant alors des valeurs autres que 0 ou π) que les limites prennent des valeurs incompatibles.

Le calcul de la limite (III.7) montre qu'il ne peut pas y avoir de *dérivation algébrique* pour $g(z)$; s'il y en avait une, elle serait aussi valable pour les nombres complexes que pour les nombres réels, car les opérations telles que multiplication ou division sont formellement identiques dans les deux cas.

On appelle *fonctions analytiques* les fonctions qui, comme les polynômes ou l'exponentielle, peuvent être dérivées ou intégrées de façon purement algébrique. Toutefois ceci ne constitue pas encore une définition rigoureuse car, si nous pouvons donner très simplement une formule algébrique explicite de dérivation ou d'intégration pour les polynômes, ou pour la fonction exponentielle, nous ne savons pas encore ce que peut être une telle dérivation ou intégration algébrique, en général.

Nous allons présenter trois approches équivalentes :

1. la dérivabilité complexe ;
2. l'intégration complexe ;
3. le développement en série entière.

III.2 Fonctions analytiques

La dérivabilité complexe est la propriété évoquée dans la section précédente, où nous avons vu que les polynômes la possèdent. On peut proposer la définition :

Définition III.1 — Fonction analytique On dira qu'une fonction $f(z)$ de la variable complexe z , à valeurs également complexes et dont les dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont continues dans D est *analytique* dans un domaine D si en tout point z de D le rapport $(f(z+w) - f(z))/w$ tend vers une limite unique lorsque w tend vers zéro.

Ⓡ Dans la définition III.1, nous avons exigé que les dérivées partielles soient continues, afin de pouvoir appliquer tranquillement la formule de Green, démontrée, rappelons-le, dans le cas où les dérivées partielles sont continues ; mais en réalité, il n'existe aucune fonction f telle que $(f(z+w) - f(z))/w$ tende vers une limite unique lorsque w tend vers zéro et dont les dérivées partielles soient discontinues.

Les mathématiciens s'efforcent généralement de faire le minimum d'hypothèses et tout particulièrement d'éviter les hypothèses non absolument nécessaires. Cependant, dans les cas pratiques qu'on rencontrera, la continuité des dérivées partielles sera toujours évidente, de sorte que l'hypothèse inutile de la définition (III.1) ne sera *jamaïs* gênante.

Selon cette définition, les polynômes sont analytiques, mais la fonction $g(x + iy) = x^2 - y^2 + ixy$ ne l'est pas.

Les polynômes ne sont pas seulement algébriquement dérivables, ils sont aussi algébriquement intégrables. Cela implique qu'une expression différentielle $P(z) dz$ est intégrable par quadrature. En effet, soit l'intégrale curviligne $\int_{\gamma} P(z) dz$ le long d'un chemin γ paramétré par $t \mapsto z(t)$ ($0 < t < T$). L'intégrale vaut alors :

$$\int_0^T P(z(t)) z'(t) dt \quad (\text{III.8})$$

et la fonction $t \mapsto P(z(t)) z'(t)$ est bien la dérivée de la fonction composée $t \mapsto Q(z(t))$ (où Q est la primitive de P), de sorte que l'intégrale se calcule par quadrature :

$$\int_{\gamma} P(z) dz = Q(z_1) - Q(z_0) \quad (\text{III.9})$$

avec $z_1 = z(T)$ et $z_0 = z(0)$. D'après les relations (II.53), (II.54), ou (II.56), on devrait donc avoir $\frac{\partial P}{\partial \bar{z}} = 0$, pour tout polynôme.

Il est facile de le vérifier directement ; il suffit d'ailleurs de le vérifier pour les fonctions $z \mapsto z^n$:

$$\frac{\partial(x + iy)^n}{\partial x} = n(x + iy)^{n-1} \quad \text{et} \quad \frac{\partial(x + iy)^n}{\partial y} = n(x + iy)^{n-1} i \quad (\text{III.10})$$

ce qui montre bien que :

$$\frac{\partial(x + iy)^n}{\partial y} = i \frac{\partial(x + iy)^n}{\partial x} \Rightarrow \frac{\partial z^n}{\partial \bar{z}} = 0 \quad (\text{III.11})$$

D'après la section II.1, cette relation doit être nécessairement satisfaite non seulement pour les polynômes, mais pour toute fonction $f(z)$ de la variable complexe $z = x + iy$ dont l'intégrale $\int f(z) dz$ sur tout lacet est nulle.

Si on sépare la partie réelle de la partie imaginaire, $f(x + iy) = u(x, y) + i v(x, y)$, cette relation s'écrit sous la forme équivalente :

$$\frac{\partial f}{\partial \bar{z}} = 0 \Rightarrow \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{et} \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} \quad (\text{III.12})$$

Ces relations sont appelées *relations de Cauchy et Riemann*. Les relations (II.53), (II.54) et (II.56) sont plus générales car elles concernent des expressions différentielles de la forme $f(z) dz + g(z) d\bar{z}$; mais dans le cas particulier où $g = 0$, elles se réduisent à (III.12). En reprenant la section II.1, on peut résumer :

Une fonction $f(z) = u(x, y) + i v(x, y)$ de la variable $z = x + iy$ ayant la propriété que $\int f(z) dz$ est nulle sur tout chemin fermé du domaine Ω , vérifie nécessairement dans Ω les relations de Cauchy et Riemann :

$$\begin{bmatrix} \frac{\partial u}{\partial x} & = & \frac{\partial v}{\partial y} \\ \frac{\partial v}{\partial x} & = & -\frac{\partial u}{\partial y} \end{bmatrix}$$

Si Ω est un domaine sans trous, ou plus généralement si le chemin fermé entoure une région sans trou de Ω , on a aussi la réciproque. □

Rappelons une dernière fois que *dans la logique du présent exposé*, la réciproque mentionnée ci-dessus ne serait vraie en toute rigueur que sous l'hypothèse (en fait inutile) que les dérivées partielles $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, $\frac{\partial v}{\partial y}$ sont continues.

Voyons maintenant à quelle condition une fonction $f(z)$ de la variable complexe est dérivable d'une manière analogue à la dérivation purement algébrique des polynômes. Nous avons vu, dans la section III.1, que la dérivation algébrique des polynômes entraînait que la limite du rapport $(f(z+h) - f(z))/h$ était la même quelle que soit la direction selon laquelle h tend vers zéro. La propriété que nous recherchons est donc celle-là. Supposons données les dérivées partielles $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$ et $\frac{\partial v}{\partial y}$. Leur existence est évidemment préalable au problème posé. Un développement limité de $u(x+k, y+l)$ et de $v(x+k, y+l)$ nous donne :

$$\begin{aligned} u(x+k, y+l) - u(x, y) &= \frac{\partial u}{\partial x} k + \frac{\partial u}{\partial y} l + \text{reste}_u \\ v(x+k, y+l) - v(x, y) &= \frac{\partial v}{\partial x} k + \frac{\partial v}{\partial y} l + \text{reste}_v \end{aligned} \quad (\text{III.13})$$

les restes étant de la forme $(k+il)\varepsilon(k, l)$ où $\varepsilon(k, l)$ tend vers zéro lorsque $k+il$ tend vers zéro. Si on pose $h = k+il$:

$$\begin{aligned} \frac{f(z+h) - f(z)}{h} &= \frac{(u(x+k, y+l) - u(x, y)) + i(v(x+k, y+l) - v(x, y))}{k+il} \\ &= \frac{\left(\frac{\partial u}{\partial x} k + \frac{\partial u}{\partial y} l\right) + i\left(\frac{\partial v}{\partial x} k + \frac{\partial v}{\partial y} l\right)}{k+il + \varepsilon(h)} \end{aligned} \quad (\text{III.14})$$

La partie provenant des restes, $\varepsilon(k, l)$, tend vers zéro ; par conséquent le terme principal déterminera la limite de notre expression lorsque h tend vers zéro. Exprimons le en coordonnées trigonométriques r, θ telles que $h = re^{i\theta}$:

$$\begin{aligned} &= \frac{\left[\frac{\partial u}{\partial x} r \cos \theta + \frac{\partial u}{\partial y} r \sin \theta\right] + i\left[\frac{\partial v}{\partial x} r \cos \theta + \frac{\partial v}{\partial y} r \sin \theta\right]}{r} \cdot (\cos \theta - i \sin \theta) \\ &= \left[\frac{\partial u}{\partial x} \cos^2 \theta + \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) \cos \theta \sin \theta + \frac{\partial v}{\partial y} \sin^2 \theta\right] + \\ &\quad + i\left[\frac{\partial v}{\partial x} \cos^2 \theta + \left(\frac{\partial v}{\partial y} - \frac{\partial u}{\partial x}\right) \sin \theta \cos \theta - \frac{\partial u}{\partial y} \sin^2 \theta\right] \end{aligned} \quad (\text{III.15})$$

On peut encore le transformer en utilisant les relations trigonométriques $\cos^2 \theta = (1 + \cos 2\theta)/2$, $\sin \theta \cos \theta = 1/2 \sin 2\theta$ et $\sin^2 \theta = (1 - \cos 2\theta)/2$:

$$\begin{aligned} &= \frac{1}{2} \left[\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) \cos 2\theta + \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right) \sin 2\theta \right] + \\ &\quad + \frac{i}{2} \left[\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) + \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) \cos 2\theta + \left(\frac{\partial v}{\partial y} - \frac{\partial u}{\partial x}\right) \sin 2\theta \right] \end{aligned} \quad (\text{III.16})$$

Cette expression est la limite, lorsque θ reste fixe et que r tend vers zéro, du rapport $(f(z+h) - f(z))/h$; pour qu'elle soit indépendante de θ , il faut et il suffit que les coefficients de $\sin 2\theta$ et $\cos 2\theta$ soient nuls, ce qui se traduit par

$$\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = 0 \quad \text{et} \quad \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = 0 \quad (\text{III.17})$$

où on reconnaît les relations de Cauchy et Riemann déjà rencontrées.

On voit par la même occasion que *reciproquement*, si les relations de Cauchy et Riemann sont satisfaites, alors les coefficients de $\sin 2\theta$ et $\cos 2\theta$ dans (III.16) seront nuls, donc la limite sera toujours :

$$\begin{aligned} \frac{1}{2} \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) + \frac{i}{2} \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) &= \frac{\partial f}{\partial z} = \frac{\partial u}{\partial x} - i \frac{\partial u}{\partial y} = \frac{\partial v}{\partial y} + i \frac{\partial v}{\partial x} \\ &= \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y} = \frac{\partial f}{\partial x} = -i \frac{\partial f}{\partial y} \end{aligned} \quad (\text{III.18})$$

quel que soit le comportement de θ : si h tend vers zéro selon une trajectoire absolument arbitraire, θ pourra varier arbitrairement sans que cela change quoi que ce soit à la limite ci-dessus, puisque, lorsque les relations de Cauchy et Riemann sont satisfaites, l'expression (III.16) ne dépend pas de θ .

On comprend maintenant aussi la nécessité du facteur 2 dans (II.55) : il a été introduit pour que l'expression purement conventionnelle $\frac{\partial f}{\partial z}$ coïncide avec la limite de $(f(z+h) - f(z))/h$.

Ainsi, dans un domaine Ω , les conditions sur $f(z)$ pour que d'une part, le rapport $(f(z+h) - f(z))/h$ ait en tout point z de Ω une limite quand h tend vers zéro dans le plan complexe et d'autre part que l'intégrale $\int_{\gamma} f(z) dz$ soit nulle sur tout lacet de Ω (du moins à tout lacet qui n'entoure pas de trou) sont identiques. Il est donc équivalent, pour une fonction $f(z)$ de la variable complexe définie sur Ω :

- que $(f(z+h) - f(z))/h$ ait en tout point z de Ω une limite quand h tend vers zéro dans le plan complexe, qu'on appelle la dérivée de f et qu'on note $f'(z)$;
- que $\int_{\gamma} f(z) dz = 0$ sur tout lacet qui n'entoure pas un trou, c'est-à-dire une région où la fonction f présente des discontinuités .

On appelle *analytiques* ou *holomorphe*, les fonctions qui satisfont ces conditions.

Il est souvent utile d'avoir les relations de Cauchy et Riemann en coordonnées polaires. Pour les trouver, partons de la formule générale du changement de coordonnées :

$$\frac{\partial}{\partial r} = \frac{\partial x}{\partial r} \frac{\partial}{\partial x} + \frac{\partial y}{\partial r} \frac{\partial}{\partial y} \quad \text{et} \quad \frac{\partial}{\partial \theta} = \frac{\partial x}{\partial \theta} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta} \frac{\partial}{\partial y} \quad (\text{III.19})$$

Puisque $x = r \cos \theta$ et $y = r \sin \theta$, on obtient :

$$\frac{\partial}{\partial r} = \cos \theta \frac{\partial}{\partial x} + \sin \theta \frac{\partial}{\partial y} \quad \text{et} \quad \frac{\partial}{r \partial \theta} = -\sin \theta \frac{\partial}{\partial x} + \cos \theta \frac{\partial}{\partial y} \quad (\text{III.20})$$

Si on tient compte de (III.12), on voit que :

$$\frac{\partial u}{\partial r} = \frac{\partial v}{r \partial \theta} \quad \text{et} \quad \frac{\partial v}{\partial r} = -\frac{\partial u}{r \partial \theta} \quad (\text{III.21})$$

On a aussi :

$$\begin{aligned} \frac{\partial}{\partial r} + i \frac{\partial}{r \partial \theta} &= \left[\cos \theta \frac{\partial}{\partial x} + \sin \theta \frac{\partial}{\partial y} \right] + i \left[-\sin \theta \frac{\partial}{\partial x} + \cos \theta \frac{\partial}{\partial y} \right] = e^{-i\theta} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \\ \frac{\partial}{\partial r} - i \frac{\partial}{r \partial \theta} &= \left[\cos \theta \frac{\partial}{\partial x} + \sin \theta \frac{\partial}{\partial y} \right] - i \left[-\sin \theta \frac{\partial}{\partial x} + \cos \theta \frac{\partial}{\partial y} \right] = e^{+i\theta} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) \end{aligned} \quad (\text{III.22})$$

ce qui donne l'expression en coordonnées polaires des opérateurs $\frac{\partial}{\partial z}$ et $\frac{\partial}{\partial \bar{z}}$:

$$e^{-i\theta} \frac{\partial}{\partial \bar{z}} = \frac{1}{2} \cdot \left(\frac{\partial}{\partial r} + i \frac{\partial}{r \partial \theta} \right) \quad \text{et} \quad e^{+i\theta} \frac{\partial}{\partial z} = \frac{1}{2} \cdot \left(\frac{\partial}{\partial r} - i \frac{\partial}{r \partial \theta} \right) \quad (\text{III.23})$$

Lorsque la fonction f satisfait les conditions de Cauchy et Riemann, on obtient des expressions plus simples :

$$f'(z) = \frac{\partial f}{\partial z} = e^{-i\theta} \frac{\partial f}{\partial r} = -ie^{-i\theta} \frac{\partial f}{r\partial\theta} \quad (\text{III.24})$$

Pour résumer, une condition nécessaire et suffisante pour que le rapport $(f(z+h) - f(z))/h$ ait une limite lorsque h tend vers zéro dans le plan complexe est que la fonction f satisfasse les relations de Cauchy et Riemann.

Lorsqu'il en est ainsi, on note $f'(z)$ cette limite ; celle-ci est alors donnée par l'une des expressions suivantes :

$$f'(z) = \frac{\partial f}{\partial z} = \frac{\partial f}{\partial x} = -i \frac{\partial f}{\partial y} = e^{-i\theta} \frac{\partial f}{\partial r} = -ie^{-i\theta} \frac{\partial f}{r\partial\theta} \quad (\text{III.25})$$

Comme annoncé plus haut, nous nous proposons maintenant de montrer que ces conditions sont encore équivalentes à une troisième, à savoir que la fonction $f(z)$ est égale à la somme d'une série entière convergente $\sum_{n=0}^{\infty} a_n (z - z_0)^n$, les a_n étant des nombres complexes. Nous avons vu que les polynômes, ainsi que la fonction $\exp(z)$ sont analytiques et cela résultait de la possibilité d'une dérivation algébrique. Or, le cas le plus général où une dérivation algébrique est possible est celui de la série entière, qui comme les polynômes est une somme (convergente) de fonctions du type $(z - z_0)^n$. Pour cela il nous faut étudier préalablement les propriétés de ces séries entières.

III.3 Séries entières convergentes

La forme la plus générale de série entière est :

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n \quad (\text{III.26})$$

Les a_n sont des nombres complexes appelés *coefficients* de la série et z_0 un nombre complexe appelé le *centre* de la série.

Les séries entières complexes ont la propriété remarquable de toujours converger dans un disque ; sous forme mathématique précise, cette propriété s'énonce comme suit, d'après N. Abel :

Théorème III.1 Pour une série entière (III.26), il existe toujours un nombre réel $R \geq 0$ (qui dépend de la suite a_n) tel que :

- si $|z - z_0| < R$, la série converge ;
- si $|z - z_0| > R$, la série diverge ;
- si $|z - z_0| = R$, on ne peut rien dire.

Preuve Soit w un nombre tel que $\lim_{n \rightarrow \infty} a_n w^n = 0$. Alors, si $|z - z_0| < |w|$, la série converge (l'hypothèse $|z - z_0| < |w|$ devient vide si $w = 0$). En effet, si $\lim_{n \rightarrow \infty} a_n w^n = 0$, il existe n_0 tel que pour $n \geq n_0$, $|a_n w^n| \leq 1$. D'autre part, si $|z - z_0| < |w|$, le rapport $r = |z - z_0|/|w|$ est < 1 . Par conséquent, pour $n \geq n_0$, $|a_n (z - z_0)^n| = |a_n w^n| r^n \leq r^n$. Or, la série $\sum r^n$ est convergente et majore en module la série (III.26).

Considérons maintenant les nombres réels positifs t tels que $\sum_{n \geq 0} |a_n| t^n$ converge. Ces nombres forment un intervalle, car si t_0 est l'un d'eux, tout $t \leq t_0$ en fait partie aussi d'après ce qui précède. Cet intervalle

possède une borne supérieure R , éventuellement infinie, qui est le nombre R annoncé par le théorème. En effet, si $|z - z_0| < R$ la série converge d'après la première partie de la démonstration, car il suffit de prendre pour w un nombre tel que $|z - z_0| < |w| < R$: puisque $|w| < R$, $|w|$ fait partie de l'intervalle des t tels que $\sum_{n \geq 0} |a_n| t^n$ converge, donc d'après la première partie de la démonstration $\sum_{n \geq 0} a_n (z - z_0)^n$ converge aussi. Inversement, supposons que $|z - z_0| > R$. Si $\sum_{n \geq 0} a_n (z - z_0)^n$ convergerait, cela entraînerait d'après la première partie de la démonstration, que si $|w| < |z - z_0|$, $\sum_{n \geq 0} a_n |w|^n$ converge ; or si on prend w tel que $R < |w| < |z - z_0|$, on aurait un nombre $t = |w|$ pour lequel la série converge et qui serait supérieur à la borne supérieure R , ce qui est absurde. •

Ce théorème signifie que le domaine de convergence d'une série entière est forcément un disque, avec éventuellement des lacunes sur la frontière. Cela implique par exemple, que si une série entière converge sur un carré, ou sur un triangle, elle convergera forcément sur un domaine plus gros (au moins le disque circonscrit).

Le disque de centre z_0 et de rayon R est le plus gros disque à l'intérieur duquel la série $\sum_{n \geq 0} a_n (z - z_0)^n$ converge ; on l'appelle le *disque de convergence* et le nombre R est appelé le *rayon de convergence*.

Il n'existe aucune règle générale concernant la convergence de la série sur la frontière de ce disque. On donne traditionnellement les exemples suivants, qui ont tous pour disque de convergence le disque de centre 0 et de rayon 1 ; ils ne diffèrent que pour ce qui se passe sur le cercle $|z| = 1$.

Exemple III.1 La série $\sum_{n \geq 0} z^n$ diverge en tout point du cercle frontière. □

Exemple III.2 La série $\sum_{n \geq 0} \frac{z^n}{n!}$ converge en tout point du cercle frontière ; elle est même *absolument* convergente en tout point du cercle. □

Exemple III.3 La série $\sum_{n \geq 0} \frac{(-1)^{[\sqrt{n}]} z^n}{n}$, où $[\sqrt{n}]$ désigne la partie entière de la racine carrée de n , est semi-convergente en tout point du cercle frontière, mais n'est *absolument* convergente en *aucun* point de ce cercle. □

Exemple III.4 La série $\sum_{n \geq 0} \frac{z^n}{n}$ est semi-convergente en tout point du cercle frontière excepté $z = 1$ où elle diverge. □

Exemple III.5 La série $\sum_{n \geq 0} \frac{d_n z^n}{n}$, où d_n désigne le nombre d'entiers dont la factorielle est un diviseur de n , diverge en tout point d'argument rationnel (en degrés) du cercle frontière et converge en certains points irrationnels connus. □

La tradition a accumulé d'innombrables exemples (plus compliqués) illustrant les différents phénomènes qui peuvent se produire sur le bord du disque de convergence. Il faut toutefois éviter le malentendu suivant : *le fait que la série diverge sur le cercle ne signifie pas que la fonction qu'elle définit n'est pas prolongeable à travers le cercle et au-delà* ; ainsi dans l'exemple III.1, la somme de la série $\sum_{n \geq 0} z^n$ est $1/(1 - z)$ à l'intérieur du disque et bien que la série diverge sur le cercle frontière, cette fonction s'y prolonge (excepté en $z = 1$). Nous reparlerons de cela plus loin quand nous aborderons les problèmes de prolongements.

Les séries entières sont en quelque sorte des polynômes de degré infini. Elles conservent un bon nombre de propriétés des polynômes, parfois sous une forme un peu dégradée. Par exemple, un polynôme est forcément convergent (et donc défini) dans tout le plan, alors qu'une série ne converge que dans un disque ; toutefois, il peut arriver que le rayon de

convergence soit infini : c'est le cas pour la fonction $\exp(z)$.

Nous donnons ci-après les principales propriétés des séries entières, sous la forme d'une petite liste de théorèmes.

Théorème III.2 Le rayon de convergence R d'une série entière $\sum_{n \geq 0} a_n (z - z_0)^n$ est donné par :

$$\frac{1}{R} = \limsup_{n \rightarrow \infty} |a_n|^{1/n} \quad (\text{III.27})$$

Ce théorème a été proposé par J. Hadamard en 1888.

Rappelons rapidement la définition de $\limsup$. Si u_n est une suite absolument quelconque de nombres réels, alors la suite $v_n = \sup_{k \geq n} u_k$ est décroissante (et la suite $w_n = \inf_{k \geq n} u_k$ est croissante). Si v_n est minorée, elle converge vers une limite appelée par définition $\limsup u_n$; sinon on pose $\limsup u_n = +\infty$. Le cas de $\liminf$ est analogue.

Ⓡ Si u_n est croissante, v_n sera constante :

1. $u_n = \sqrt[n]{0.5} \Rightarrow v_n = 1 \Rightarrow \limsup_{n \rightarrow \infty} u_n = 1$;
2. $u_n = n \Rightarrow v_n = +\infty \Rightarrow \limsup_{n \rightarrow \infty} u_n = +\infty$.

Preuve Pour alléger les notations, travaillons dans le cas où $z_0 = 0$. La limite supérieure de $u_n = |a_n|^{1/n}$ est la limite tout court de la suite décroissante et minorée $v_n = \sup_{k \geq n} u_k$. Autrement dit, pour tout $\varepsilon > 0$, il existe un entier n_ε tel que $n \geq n_\varepsilon \Rightarrow 1/R - \varepsilon < v_n < 1/R + \varepsilon$. Or v_n est la borne supérieure des $u_k = |a_k|^{1/k}$ pour $k \geq n$, ce qui veut dire :

- d'une part, que $\forall k \geq n$, $u_k \leq v_n < 1/R + \varepsilon$; ainsi $\forall k \geq n_\varepsilon$, $u_k < 1/R + \varepsilon$;
- d'autre part, que parmi les u_k de rang $k \geq n$, il y en aura toujours au moins un aussi proche qu'on voudra de v_n et cela quelque soit n ; c'est-à-dire $\forall n, \exists k \geq n, u_k \geq v_n - \varepsilon > 1/R - 2\varepsilon$; on peut encore exprimer cela en disant que pour tout ε , il y aura une infinité de u_k pour lesquels $u_k > 1/R - 2\varepsilon$.

En regroupant ces deux aspects de la limite, on voit que pour tout ε :

1. il existera n_ε tel que $\forall k \geq n_\varepsilon$, $u_k < 1/R + \varepsilon$;
2. pour une infinité de k on aura $u_k > 1/R - 2\varepsilon$.

De 1, on déduit que $|a_n z^n| < (|z|/R + \varepsilon|z|)^n < (|z|/R + \varepsilon R)^n$. Puisque $|z|/R < 1$ et que ε peut être pris aussi petit qu'on veut, prenons ε tel que $\alpha = |z|/R + \varepsilon R < 1$. Tous les termes de la série à partir du rang n_ε sont donc majorés (en module) par α^n , avec $\alpha < 1$. Cela prouve déjà que $\sum |a_n z^n|$ converge. Ainsi :

$$|z| < R \Rightarrow \sum_{n \geq 0} |a_n z^n| \text{ converge} \quad (\text{III.28})$$

On a seulement prouvé que R est inférieur ou égal au rayon de convergence. Pour s'assurer qu'il est bien égal au rayon de convergence, il reste à vérifier que la série diverge si $|z| > R$; cela se vérifie de façon analogue, en utilisant cette fois 2 : pour tout $\varepsilon > 0$ il existe une infinité de n tels que $|a_n|^{1/n} > 1/R - 2\varepsilon$. Par conséquent, pour une infinité de n , $|a_n z^n| > (|z|/R - \varepsilon|z|)^n > (|z|/R - 2\varepsilon R)^n$. Puisque $|z|/R > 1$ et que ε peut être pris aussi petit qu'on veut, prenons ε tel que $\beta = |z|/R - 2\varepsilon R > 1$. Il y a donc une infinité de termes de la série qui sont minorés (en module) par β^n avec $\beta > 1$, ce qui signifie bien qu'elle diverge. •

Théorème III.3 Sur un disque fermé Ω inclus dans l'intérieur du disque de convergence, la série entière est normalement convergente, c'est-à-dire uniformément majorée par une série numérique convergente.

Preuve Il suffit de remarquer qu'il existe un nombre A tel que $0 < A < R$ et tel que Ω soit contenu dans le disque de centre z_0 et de rayon A . Alors tout z de Ω vérifie aussi $|z - z_0| < A$ et la série $\sum a_n (z - z_0)^n$ est donc uniformément majorée dans Ω par la série numérique $\sum |a_n| A^n$, qui évidemment converge puisque $A < R$. •

Théorème III.4 On peut intégrer et dériver terme par terme une série entière : si $f(z) = \sum_{n \geq 0} a_n (z - z_0)^n$ (définie à l'intérieur du disque de convergence), alors f est analytique dans ce disque et $f'(z) = \sum_{n \geq 1} n a_n (z - z_0)^{n-1}$. La série dérivée a le même rayon de convergence et on peut dériver à nouveau :

$$\begin{aligned} f''(z) &= \sum_{n \geq 2} n(n-1) a_n (z - z_0)^{n-2} \\ f'''(z) &= \sum_{n \geq 3} n(n-1)(n-2) a_n (z - z_0)^{n-3}, \quad \text{etc.} \end{aligned} \quad (\text{III.29})$$

Cette propriété exprime l'existence d'une dérivation purement algébrique, analogue à celle des polynômes ; ainsi la somme d'une série entière donne, dans le domaine constitué par l'intérieur du disque de convergence, une fonction analytique.

Preuve L'intégration terme par terme ne pose pas de problème en vertu du théorème III.3 : une série normalement convergente peut toujours être intégrée terme par terme, du fait de l'inégalité de la moyenne. C'est la dérivation terme par terme qui est une propriété remarquable des séries entières et que les autres sortes de séries ne possèdent pas : essayer par exemple avec la série $\sum_{n \geq 1} \frac{1}{n^2} \cos(n^2 x)$.

En utilisant le théorème III.2, il est facile de vérifier que la série dérivée a le même rayon de convergence que la série initiale :

$$|n a_n|^{1/n} = n^{1/n} |a_n|^{1/n} \quad \text{et} \quad n^{1/n} \rightarrow 1 \quad (\text{III.30})$$

Prenons toujours, pour alléger l'écriture, $z_0 = 0$. On peut dire jusqu'à présent que la série initiale $f(z) = \sum a_n z^n$ et la série dérivée $g(z) = \sum n a_n z^{n-1}$ définissent chacune une fonction analytique dans le disque de convergence. Cependant, il n'est pas encore prouvé que $g = f'$, c'est-à-dire que $g(z) = \lim_{h \rightarrow 0} (f(z+h) - f(z))/h$. Cela peut se faire directement comme suit, quoique généralement, les auteurs le déduisent de la formule intégrale de Cauchy que nous verrons au paragraphe suivant ; une démonstration directe, à la main, a toutefois l'avantage de mieux montrer les raisons très simples qui font que ça marche.

On peut écrire :

$$\frac{f(z+h) - f(z)}{h} - g(z) = \sum_{n \geq 0} a_n \left(\frac{(z+h)^n - z^n}{h} - n z^{n-1} \right) \quad (\text{III.31})$$

Comme dans le cas des polynômes, le problème se ramène à des fonctions z^n . Pour analyser les expressions correspondantes, on utilise la formule du binôme : on sait que $(z+h)^n = \sum_{k=0}^n C_n^k z^k h^{n-k}$, d'où :

$$(z+h)^n - z^n - n h z^{n-1} = \sum_{k=0}^{n-2} C_n^k z^k h^{n-k} \quad (\text{III.32})$$

Comme la somme dans le membre de droite ne contient pas de puissance de h inférieure à 2, on peut mettre h^2 en facteur, ce qui donne :

$$(z+h)^n - z^n - n h z^{n-1} = h^2 \sum_{k=0}^{n-2} C_n^k z^k h^{n-2-k} \quad (\text{III.33})$$

En utilisant l'inégalité du triangle, on peut majorer le module du second membre par :

$$|h|^2 \sum_{k=0}^{n-2} C_n^k |z|^k |h|^{n-2-k} \quad (\text{III.34})$$

Enfin, on remarque la relation :

$$C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)}{(n-k)(n-1-k)} \cdot \frac{(n-2)!}{k!(n-2-k)!} = \frac{n(n-1)}{(n-k)(n-1-k)} \cdot C_{n-2}^k \quad (\text{III.35})$$

Dans la somme (III.34), l'indice k varie entre 0 et $n-2$, donc le dénominateur $(n-k)(n-1-k)$ dans (III.35) reste toujours ≥ 1 . Par conséquent il résulte de (III.35) que :

$$C_n^k \leq n(n-1) C_{n-2}^k \quad (\text{III.36})$$

Si on applique cette inégalité à (III.34) (qui majorait (III.34)), on obtient l'inégalité :

$$\left| \frac{(z+h)^n - z^n}{h} - n z^{n-1} \right| \leq |h| \sum_{k=0}^{n-2} n(n-1) C_{n-2}^k |z|^k |h|^{n-2-k} = |h| n(n-1) (|z| + |h|)^{n-2} \quad (\text{III.37})$$

Enfin, reportant cela dans (III.31) :

$$\left| \frac{f(z+h) - f(z)}{h} - g(z) \right| \leq |h| \sum_{n \geq 0} n(n-1) |a_n| (|z| + |h|)^{n-2} \quad (\text{III.38})$$

On peut maintenant conclure : puisqu'on fait tendre h vers zéro, on peut le prendre suffisamment petit pour que $b = |z| + |h| < R$ (si $|z| < R$ on peut toujours insérer $|h|$ entre $|z|$ et R). Alors la série de (III.38), $\sum n(n-1) |a_n| (|z| + |h|)^{n-2}$, est uniformément majorée par la série numérique $\sum n(n-1) |a_n| b^{n-2}$ qui converge et le facteur $|h|$ fait tendre le second membre de l'inégalité (III.38) vers zéro, ce qui montre que la limite de $[f(z+h) - f(z)]/h$ est bien $g(z)$. •

Ⓡ Il est peut-être nécessaire ici de rafraîchir encore quelque peu la mémoire du lecteur à propos des $\limsup$ et $\liminf$. La proposition suivante est vraie : soient deux suites u_n et v_n , dont l'une (disons u_n) a une limite, alors :

$$\limsup_{n \rightarrow \infty} u_n \cdot v_n = \lim_{n \rightarrow \infty} u_n \cdot \limsup_{n \rightarrow \infty} v_n \quad \text{et} \quad \liminf_{n \rightarrow \infty} u_n \cdot v_n = \lim_{n \rightarrow \infty} u_n \cdot \liminf_{n \rightarrow \infty} v_n \quad (\text{III.39})$$

On la démontre aisément en recourant directement aux définitions. Par contre, lorsqu'aucune des deux suites n'a de limite, l'égalité $\limsup u_n \cdot v_n = \limsup u_n \cdot \limsup v_n$ est généralement fautive, comme le montre ce contre-exemple :

$$u_n = \begin{cases} 0 & \text{si } n \text{ pair} \\ 1 & \text{si } n \text{ impair} \end{cases} \quad \text{et} \quad v_n = \begin{cases} 1 & \text{si } n \text{ pair} \\ 0 & \text{si } n \text{ impair} \end{cases} \quad (\text{III.40})$$

Ⓡ Cette démonstration est plus longue que celles qui utilisent des résultats ultérieurs mais elle est plus transparente : elle montre bien qu'avec les séries entières, tout se passe comme pour les polynômes ; le choix fait ici de recourir à la formule du binôme montre que c'est bien le caractère purement algébrique de la dérivation d'une série entière qui rend tout cela possible. Toutefois, la définition III.1 ne fait pas appel aux opérations algébriques (addition, multiplication, élévation à une puissance) qui s'étendent à l'identique des réels aux complexes. Elle passe par la notion de limite qui est très différente dans le plan complexe, du fait de ses deux dimensions. C'est pourquoi la question essentielle reste encore sans réponse : une fonction qui vérifie simplement la propriété que $(f(z+h) - f(z))/h$ possède une limite lorsque h (complexe) tend vers zéro a-t-elle forcément toujours une dérivée algébrique analogue aux polynômes ? La réponse à cette question est *oui* : pour le prouver il suffit d'établir que toute fonction ayant cette propriété est nécessairement la somme d'une série entière, qui se dérive alors algébriquement. C'est ce que nous allons étudier dans le paragraphe suivant.

III.4 Théorie de Cauchy

La présente section expose les résultats les plus fondamentaux de la théorie des fonctions analytiques. La tradition l'appelle *théorie de Cauchy* en hommage à celui qui a eu la première idée ⁽¹⁾ et qui ensuite en a établi les résultats essentiels. La théorie transmise par la tradition et que nous résumons ici, est très différente de l'approche initiale de Cauchy et est le produit d'une lente maturation qui s'étend sur un siècle. La notion d'intégrale curviligne est absente

(1) A.-L. CAUCHY, *Mémoire sur les intégrales définies*, lu à l'Institut le 22 août 1814, imprimé en 1827. On peut le trouver dans *Œuvres complètes de Cauchy*, tome I, 1^{re} série, Gauthier-Villars, Paris, 1882, p. 329.

dans le mémoire de 1814, seules y apparaissent des intégrales simples sur un intervalle réel ou doubles (par contre le passage par les valeurs complexes y est essentiel).

Toutefois, la clé de la théorie de 1814 est bien la relation entre intégrale double sur un domaine et intégrale curviligne sur le bord de ce domaine. La formule de Green (II.32) n'était pas encore connue et la lecture de son mémoire montre bien que la relation *générale* entre intégrale double sur un domaine et intégrale curviligne sur le bord de ce domaine n'était pas perçue — c'est une idée qui vient de la théorie électromagnétique développée par Ampère et Green un peu plus tard. Pourtant, Cauchy l'a bien perçue indirectement, sous une forme plus spéciale : lorsque le domaine de l'intégrale double est un rectangle parallèle aux axes.

Dans ce premier mémoire, il ne parle pas encore d'intégrale de contour sur le périmètre du rectangle mais il utilise de manière essentielle, quoique dans un cas particulier, la relation exprimée par la formule de Green, en traduisant l'intégrale sur le périmètre du rectangle par quatre intégrales simples (une pour chacun des quatre côtés du rectangle).

Ajoutons que le contenu du mémoire est présenté comme une méthode nouvelle pour calculer des intégrales définies et non comme l'exposé d'une théorie des fonctions de la variable complexe ; l'idée générale de fonction analytique n'y apparaît pas. Nous avons vu, en section II.2, que l'intégrale double de la fonction $(x^2 - y^2)/(x^2 + y^2)^2$ sur le carré $[-1; 1] \times [-1; 1]$ ne donnait pas la même valeur selon qu'on intègre d'abord en x puis en y ou inversement ; la différence entre les deux valeurs est 2π , qui est justement l'intégrale sur le bord du carré de l'expression différentielle $(-y dx + x dy)/(x^2 + y^2)$; cette dernière expression se trouve être la partie imaginaire de $dz/z = (dx + i dy)/(x + iy)$. Dans le mémoire de 1814, Cauchy cherchait un moyen de prédire la différence entre les deux manières de calculer une intégrale double lorsque la fonction à intégrer devient infinie et a trouvé que, pour un rectangle, cette différence était ce que Green a appelé, plus tard en 1827, l'intégrale curviligne sur le bord. Cauchy a par la suite approfondi ce travail dans un livre ⁽²⁾ puis dans un article ⁽³⁾ ⁽⁴⁾. C'est donc en cherchant un moyen de prédire la différence entre les deux manières de calculer une intégrale double singulière que Cauchy a peu à peu découvert les propriétés des fonctions analytiques et la théorie qui porte son nom.

La présentation traditionnelle de la théorie de Cauchy, que nous suivons ici, est due à la postérité, notamment à Riemann qui a soutenu sa thèse de doctorat sur ce sujet, en 1851 à Göttingen. Les travaux de Cauchy sont très calculatoires et riches en formules mais les concepts généraux qui rendent la théorie plus claire ont été introduits plus tard ; ainsi Riemann a formulé les résultats en termes d'intégrales curvilignes ; la notion d'homologie des lacets a été fournie par H. Poincaré.

La *formule de Cauchy* donnée ci-dessous en termes d'intégrales sur un lacet et son application au développement en série sont des résultats publiés par Cauchy en 1831.

⁽²⁾ A.-L. CAUCHY, *Leçons sur les applications du calcul infinitésimal à la géométrie*, 1822

⁽³⁾ A.-L. CAUCHY, *De l'influence que peut avoir, sur la valeur d'une intégrale double, l'ordre dans lequel on effectue les intégrations*, 1826

⁽⁴⁾ Cauchy avait épousé la fille d'un éditeur (Guillaume de Bure) qui lui imprimait tout et il publiait ainsi une sorte de périodique, les *Exercices de mathématiques*, dont cet article fait partie.

Théorème III.5 Soit $f(z)$ une fonction *analytique* de la variable complexe z dans un domaine Ω (avec ou sans trous). Pour tout lacet γ homologiquement équivalent dans $\Omega - \{z\}$ à un cercle de centre z et de rayon assez petit pour que le disque correspondant soit contenu dans Ω , on a :

$$f(z) = \frac{1}{2i\pi} \int_{\gamma} \frac{f(w)}{w - z} dw \quad (\text{III.41})$$

L'égalité (III.41) s'appelle *formule de Cauchy*. On remarquera que la fonction sous le signe $\int$, $w \mapsto f(w)/(w - z)$, est analytique dans $\Omega - \{z\}$; elle devient singulière en $w = z$ et c'est pourquoi son intégration le long de γ ne donne pas zéro.

Preuve Puisque la fonction $w \mapsto f(w)/(w - z)$ sous le signe d'intégration est analytique dans $\Omega - \{z\}$, il suffit d'établir la relation (III.41) sur un cercle assez petit de centre z . Paramétrons ce cercle en coordonnées polaires $w = z + re^{i\theta}$. On obtient (par définition des intégrales curvilignes) :

$$\frac{1}{2i\pi} \int_{\gamma} \frac{f(w)}{w - z} dw = \frac{1}{2i\pi} \int_0^{2\pi} \frac{f(z + re^{i\theta})}{re^{i\theta}} ire^{i\theta} d\theta = \frac{1}{2\pi} \int_0^{2\pi} f(z + re^{i\theta}) d\theta \quad (\text{III.42})$$

Les cercles de centre z et de rayon assez petit étant forcément homologiquement équivalents, la valeur de l'intégrale ne peut pas dépendre de r et est donc égale à sa limite quand r tend vers zéro. Or f est analytique, donc continue, ce qui entraîne que :

$$\lim_{r \rightarrow 0} \frac{1}{2\pi} \int_0^{2\pi} f(z + re^{i\theta}) d\theta = \frac{1}{2\pi} \int_0^{2\pi} f(z) d\theta = f(z) \cdot \frac{1}{2\pi} \int_0^{2\pi} d\theta = f(z) \quad (\text{III.43})$$

Cette égalité est bien la formule de Cauchy annoncée. •

Cette formule de Cauchy permet de montrer qu'une fonction analytique, c'est-à-dire une fonction qui satisfait aux relations de Cauchy et Riemann, est nécessairement égale (dans un certain disque) à la somme d'une série entière. De façon précise :

Théorème III.6 Soit $f(z)$ une fonction analytique dans un domaine Ω (avec ou sans trous), soit z_0 un point de Ω et R_0 le rayon du plus grand disque ouvert de centre z_0 qui soit contenu dans Ω (R_0 est la distance de z_0 à la frontière de Ω). Alors la série :

$$\sum_{n \geq 0} a_n (z - z_0)^n \quad \text{avec} \quad a_n = \frac{1}{2i\pi} \int_{\gamma} \frac{f(w)}{(w - z_0)^{n+1}} dw \quad (\text{III.44})$$

est convergente dans le disque de centre z_0 et de rayon R_0 (c'est-à-dire que son rayon de convergence est $\geq R_0$) et sa somme est égale à $f(z)$ dans ce disque.

Preuve On n'avait pas besoin de la théorie des fonctions analytiques pour savoir que la fonction $1/(w - z)$ est la somme d'une série entière :

$$\frac{1}{w - z} = \frac{1}{(w - z_0) - (z - z_0)} = \frac{1}{w - z_0} \frac{1}{1 - \frac{z - z_0}{w - z_0}} = \frac{1}{w - z_0} \sum_{n \geq 0} \left[\frac{z - z_0}{w - z_0} \right]^n = \sum_{n \geq 0} \frac{(z - z_0)^n}{(w - z_0)^{n+1}} \quad (\text{III.45})$$

Dans la formule de Cauchy (III.41), remplaçons sous le signe $\int$ le facteur $1/(w - z)$ par cette série; cela donne après l'intervention de $\sum$ et de $\int$ (qui est correcte le long d'un lacet γ égal au cercle de centre z_0 et de rayon $r > |z - z_0|$, puisque la série géométrique (III.45) converge pour $|w - z_0| > |z - z_0|$, donc en particulier le long du cercle $|w - z_0| = r$) :

$$f(z) = \sum_{n \geq 0} \frac{1}{2i\pi} \int_{\gamma} \frac{f(w)}{(w - z_0)^{n+1}} dw \cdot (z - z_0)^n \quad (\text{III.46})$$

ce qui permet de conclure. •

- Ⓡ La condition $r > |z - z_0|$ sur le lacet γ signifie que le point z doit être à l'intérieur du cercle γ ; mais ensuite, les intégrales (III.44) prennent évidemment la même valeur sur n'importe quel autre lacet homologiquement équivalent (dans $\Omega - \{z\}$) à ce cercle. Cette condition qui permet l'échange des signes intégral et somme dans la série géométrique qui développe $f(w)/(z - w)$, nous dit aussi pour quels z cet échange est possible et par conséquent pour quels z l'égalité entre $f(z)$ et la somme de la série a lieu : il faut en effet que le cercle γ soit contenu dans Ω et entoure le point z et un tel cercle ne peut être trouvé que si $|z - z_0| < R_0$: son rayon r sera tel que $|z - z_0| < r < R_0$.

On a ainsi montré que $f(z)$ est bien la somme d'une série entière convergente dans le disque ouvert $|z - z_0| < R_0$.

On déduit du théorème III.6 un grand nombre de corollaires qui expriment les propriétés remarquables des fonctions analytiques.

Corollaire III.6.1 Les coefficients a_n de (III.44) vérifient les inégalités :

$$|a_n| \leq \frac{M_r}{r^n} \quad (\text{III.47})$$

où M_r est la valeur maximum prise par la fonction $|f(z)|$ sur le cercle $|z - z_0| = r$.

Il suffit d'utiliser l'inégalité de la moyenne pour les intégrales (III.44) : en effet, en les paramétrant par $w = z_0 + re^{i\theta}$:

$$a_n = \frac{1}{2\pi} \int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{r^n e^{in\theta}} d\theta \quad (\text{III.48})$$

L'inégalité de la moyenne donne alors :

$$|a_n| \leq \frac{1}{2\pi} \int_0^{2\pi} \frac{M_r}{r^n} d\theta = \frac{M_r}{r^n} \quad (\text{III.49})$$

Ces inégalités sur les coefficients sont appelées *inégalités de Cauchy*.

On déduit de ce corollaire une propriété importante des fonctions analytiques, qui généralise une propriété connue des polynômes : une fonction analytique dans tout le plan ne peut pas être bornée (à moins d'être constante). En effet, si elle était bornée dans tout le plan par une constante M , celle-ci serait supérieure à tous les M_r ; on aurait donc d'après les inégalités de Cauchy $|a_n| \leq M/r^n$ pour tout r et donc (en faisant tendre r vers l'infini) on aurait $a_n = 0$ pour tout $n > 0$. Ce résultat est connu sous le nom de *théorème de Liouville*.

On peut aussi en déduire que $|a_n|^{1/n} \leq M_r^{1/n}/r$; cela montre que $\limsup_{n \rightarrow \infty} |a_n|^{1/n} \leq 1/r$, puisque $\lim_{n \rightarrow \infty} M_r^{1/n} = 1$. Mais cette inégalité signifie simplement que le rayon de convergence de la série $\sum a_n (z - z_0)^n$ est $\geq R_0$, ce qui était déjà contenu dans le théorème III.6.

Corollaire III.6.2 — Principe du maximum Si $f(z)$ est une fonction analytique, $|f(z)|$ ne peut pas avoir de maximum local ; il en est de même de $\Re(f)$ et $\Im(f)$ (parties réelle et imaginaire).

- Ⓡ On entend ici par *maximum* une valeur *strictement* supérieure aux valeurs prises tout autour. Toutefois, le même argument que dans la démonstration ci-dessous montre que si la fonction $f(z)$ est analytique et que $|f(z)|$ reste constant dans un voisinage du maximum, alors $f(z)$ est elle-même constante.

Preuve S'il existait un point z_0 dans le domaine Ω où f est analytique, tel que $|f(z)|$ y devient maximum, cela voudrait dire qu'il existe un cercle de centre z_0 le long duquel $|f(z)|$ serait partout strictement inférieur à $|f(z_0)|$. D'après la formule de Cauchy, on aurait :

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{i\theta}) d\theta \quad (\text{III.50})$$

et l'inégalité de la moyenne donnerait :

$$|f(z_0)| \leq \frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + re^{i\theta})| d\theta \quad (\text{III.51})$$

c'est-à-dire que $|f(z_0)|$ serait inférieur à la moyenne de $|f(z)|$ sur ledit cercle ; cela contredit évidemment que $|f(z_0)|$ soit un maximum.

Pour la partie réelle et la partie imaginaire, on se ramène au cas précédent en considérant les fonctions $\exp(f(z))$ et $\exp(-if(z))$; celles-ci sont évidemment analytiques si f l'est et $|\exp(f(z))| = \exp(\Re f)$, $|\exp(-if(z))| = \exp(\Im f)$. •

Corollaire III.6.3 Soit f une fonction sur un domaine Ω supposée *a priori* une fois continûment ⁽⁵⁾ dérivable (c'est-à-dire que seules les dérivées partielles au premier ordre existent et sont continues). Si ces dérivées partielles vérifient les relations de Cauchy et Riemann dans le domaine, alors elles sont à leur tour dérivables et la fonction f est en fait infiniment dérivable.

Preuve Puisque la fonction f vérifie les relations de Cauchy et Riemann, elle est analytique, donc égale à une série entière dans un disque ; or une série entière peut être dérivée autant de fois qu'on veut (théorème III.4), donc la fonction est infiniment dérivable dans ce disque. En outre, pour tout point z_0 de Ω on a une série entière qui converge dans un disque de centre z_0 et contenu dans Ω . On recouvre ainsi la totalité du domaine. •

La fin de cette démonstration fait appel à un recouvrement de Ω par des disques. Ce point est essentiel et sous-tend de nombreux aspects de la théorie des fonctions analytiques, comme on le verra. Donnons-en une idée plus concrète sur l'exemple de la fonction $f(z) = 1/z$. Celle-ci est analytique sur le domaine $\Omega = \mathbb{C} - \{0\}$. Soit z_0 un point de Ω . En écrivant $1/z$ sous la forme :

$$\frac{1}{z_0 + (z - z_0)} = \frac{1}{z_0} \frac{1}{1 + \frac{(z - z_0)}{z_0}} \quad (\text{III.52})$$

on voit immédiatement que cette fonction se développe en série géométrique ; ainsi pour tout $z_0 \neq 0$:

$$\frac{1}{z} = \sum_{n=0}^{\infty} \frac{(-1)^n}{z_0^{n+1}} (z - z_0)^n \quad (\text{III.53})$$

Le rayon de convergence de la série est $|z_0|$: c'est le rayon du plus grand disque possible de centre z_0 qui soit contenu dans Ω . On ne peut espérer prolonger la fonction $1/z$ en dehors de Ω , puisqu'il ne reste plus que le point zéro qui est singulier. Chacune des séries (III.53) converge dans un disque qui ne couvre qu'une partie de Ω , mais en faisant varier z_0 on peut recouvrir peu à peu la totalité du domaine : il faut pour cela une infinité de disques, voir la figure III.1.

(5) Conformément à un choix d'exposition expliqué au chapitre II, nous faisons l'hypothèse que les dérivées partielles sont continues car cela simplifie les démonstrations sans nuire aux applications utiles ; mais en fait cette hypothèse pourrait se déduire des autres.

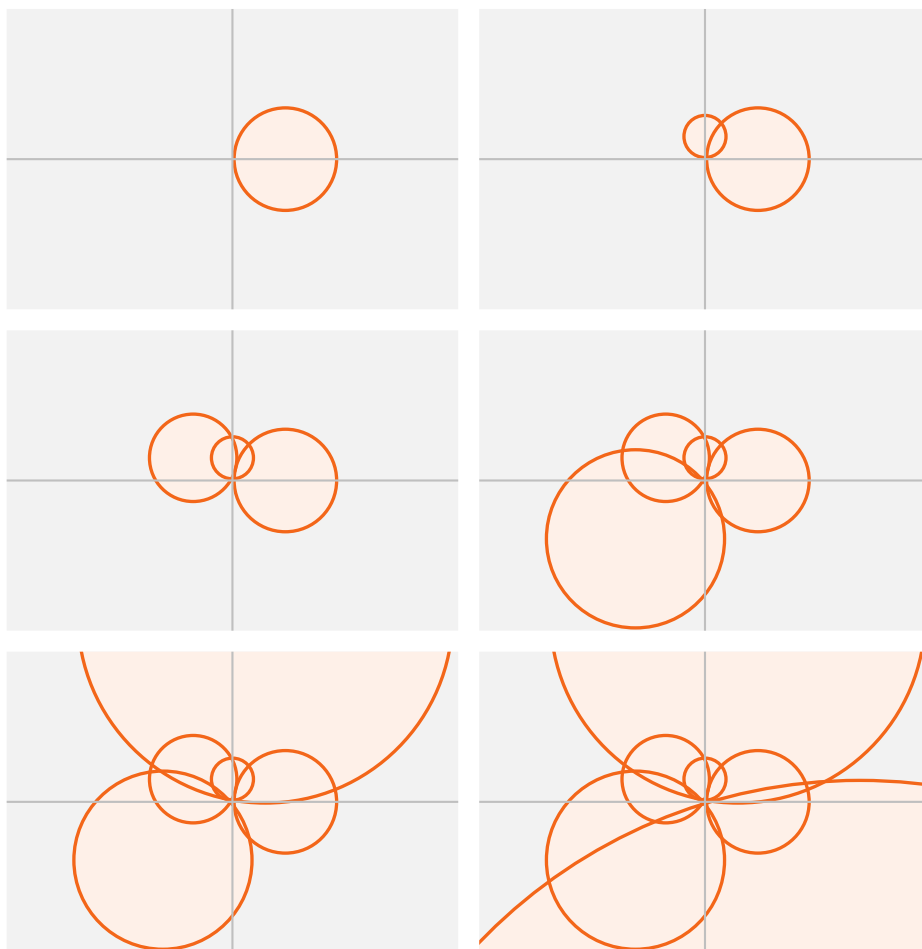


Figure III.1 – Recouvrement progressif d'un domaine d'holomorphie par des disques de convergence d'une série entière

Cet exemple est là pour montrer que, si le domaine maximal dans lequel une série entière converge est nécessairement un disque, le domaine maximal dans lequel une fonction est analytique n'est généralement pas un disque ; le développement en série entière est possible autour d'un centre z_0 quelconque, mais celle-ci ne convergera que dans une partie du domaine.

Une fonction analytique f étant donnée, on appelle *domaine d'holomorphie* de f le plus grand domaine sur lequel on peut avoir un prolongement de f qui soit analytique. Par exemple la série entière $\sum_{n \geq 0} (z - 1)^n$ converge dans le disque de centre 1 et de rayon 1 et y définit donc une fonction analytique ; mais celle-ci se prolonge au-delà et son domaine d'holomorphie est $\mathbb{C} - \{0\}$. Pourtant *le plus grand* domaine n'existe pas toujours (voir le cas de la fonction logarithme dans la section suivante)

Voici encore un corollaire *extrêmement important* du théorème III.6.

Corollaire III.6.4 — Théorème des zéros isolés Une fonction analytique, à moins d'être la fonction identiquement nulle, ne peut s'annuler qu'en des points isolés.

En pratique, cela signifie qu'une fonction analytique qui serait nulle sur un continuum de points serait forcément nulle partout. On va commencer par donner la démonstration de

ce corollaire, puis revenir ensuite à des explications complémentaires.

Preuve (Corollaire III.6.4) Soit $f(z)$ une fonction analytique au voisinage d'un point z_0 où elle s'annule. D'après le théorème III.6, $f(z)$ est égale, dans un disque D non vide de centre z_0 , à la somme d'une série entière qui converge dans ce disque :

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad (\text{III.54})$$

Si tous les a_n sont nuls, la fonction $f(z)$ est la constante zéro ; si donc on suppose que $f(z)$ n'est pas identiquement nulle, les a_n ne sont pas tous nuls, il existe au moins un n tel que $a_n \neq 0$. Soit alors n_0 le plus petit entier pour lequel $a_{n_0} \neq 0$. Comme on a supposé $f(z_0) = 0$, n_0 est forcément ≥ 1 . On a alors :

$$\begin{aligned} f(z) &= a_{n_0} (z - z_0)^{n_0} + a_{n_0+1} (z - z_0)^{n_0+1} + a_{n_0+2} (z - z_0)^{n_0+2} + \dots \\ &= (z - z_0)^{n_0} \cdot \left\{ a_{n_0} + a_{n_0+1} (z - z_0) + a_{n_0+2} (z - z_0)^2 + \dots \right\} \end{aligned} \quad (\text{III.55})$$

Appelons $h(z)$ l'expression entre accolades :

$$h(z) = a_{n_0} + a_{n_0+1} (z - z_0) + a_{n_0+2} (z - z_0)^2 + a_{n_0+3} (z - z_0)^3 + \dots \quad (\text{III.56})$$

Ceci est une série entière qui a le même rayon de convergence que (III.54) puisque ses coefficients sont simplement décalés de n_0 ; elle converge donc dans le même disque D que (III.54), ce qui signifie que la fonction $h(z)$ est analytique dans ce disque. Mais on a aussi $h(z_0) = a_{n_0} \neq 0$, qui signifie que $h(z)$ est non nulle en z_0 . $h(z)$ étant analytique, elle est en particulier continue et par conséquent va rester non nulle dans tout un voisinage de z_0 : on peut dire qu'il existe ε tel que pour $|z - z_0| < \varepsilon$, on aura $|h(z)| > \frac{1}{2}|a_{n_0}|$. Ainsi on a un disque D_ε , de rayon ε , dans lequel $h(z)$ ne s'annule pas. D'autre part le facteur $(z - z_0)^{n_0}$ est non nul partout, sauf en $z = z_0$ (car $n_0 \geq 1$).

On a ainsi montré que dans le disque D_ε , le produit $f(z) = (z - z_0)^{n_0} \cdot h(z)$ est non nul partout, sauf en $z = z_0$.

Autrement dit, si une fonction analytique $f(z)$ non identiquement nulle s'annule en un point z_0 , ce point est entouré d'un disque dans lequel il est le seul où f s'annule, ce qu'on exprime en disant que z_0 est un *zéro isolé*. •

Ce *théorème des zéros isolés* généralise aux fonctions analytiques une propriété des polynômes : on sait qu'un polynôme de degré N ne peut s'annuler qu'en au plus N points, à moins d'être le polynôme nul. Donc les zéros d'un polynôme sont en nombre fini et par conséquent forcément isolés. Si on interprète les fonctions analytiques comme des polynômes de degré infini, on comprend qu'il puisse y avoir une infinité de zéros, mais ils sont toujours isolés.

Il est par exemple impossible que les zéros d'une fonction analytique forment une suite convergente de points en incluant leur limite : ainsi la fonction $\sin(1/z)$ s'annule aux points $1/n\pi$ qui forment une suite convergente, mais elle n'est pas analytique au point limite $z = 0$. à plus forte raison, il est impossible qu'une fonction analytique s'annule sur une droite, un segment de droite, ou un arc de courbe, à moins d'être nulle partout.

Ce fait a une conséquence pratique pour les calculs, dont nous ferons fréquemment usage dans les chapitres ultérieurs. Supposons que nous ayons deux expressions, par exemple une intégrale et une série, dont nous voudrions prouver l'égalité. Si les deux expressions dépendent analytiquement d'un certain paramètre complexe z , il nous suffira de prouver l'égalité pour z réel, ou pour $0 < z < 1$, ou pour $|z| = 1$ (par exemple). En effet les points de ces ensembles ne sont pas isolés les uns des autres ; la différence entre les deux expressions étant nulle sur ces ensembles de points, est alors forcément nulle partout.

Cette méthode ⁽⁶⁾ est souvent désignée sous le nom de *principe du prolongement analytique*. Par exemple, on démontre la formule dite des compléments (théorème V.2) :

$$\Gamma(x) \cdot \Gamma(1-x) = \frac{\pi}{\sin \pi x}, \quad 0 < x < 1 \quad (\text{III.57})$$

Sachant que les deux membres de cette égalité sont des fonctions analytiques de x dans $\{\mathbb{C} - \pi\mathbb{Z}\}$, on en déduit que l'égalité s'étend à tout ce domaine complexe.

Il sera fait de ce principe un usage fréquent.

Voici le cinquième et dernier corollaire du théorème III.6. Il concerne les fonctions qui sont le quotient de deux fonctions analytiques et qu'on appelle *méromorphes*; ce sont donc les fonctions de la forme $h(z) = f(z)/g(z)$, où $f(z)$ et $g(z)$ sont toutes deux analytiques sur l'ensemble du plan complexe. Cela implique, d'après le corollaire III.6.4, que h est analytique en dehors d'un ensemble de points isolés. Le domaine d'holomorphie de h est donc $\Omega(h) = \mathbb{C} - Z(g)$, $Z(g)$ étant l'ensemble discret des zéros de g .

Corollaire III.6.5 Soit une fonction méromorphe $h(z) = f(z)/g(z)$, où $f(z)$ et $g(z)$ sont toutes deux analytiques sur l'ensemble du plan complexe. Soit z_0 un point de $\Omega(h) = \mathbb{C} - Z(g)$. Alors le rayon de convergence de la série entière de h de centre z_0 est la distance de z_0 au point singulier le plus proche.

Preuve Le point singulier le plus proche existe forcément puisque ces points singuliers forment un ensemble discret : dans tout disque borné de centre z_0 il ne peut, en effet, qu'y avoir un nombre fini de ces points (dans le cas contraire la propriété de Bolzano-Weierstrass entraînerait que ces points s'accumulent, ce qui contredit le corollaire III.6.4). Si g possède au moins un zéro (sinon il n'y a pas de point singulier du tout et le corollaire III.6.5 devient un truisme), un disque de rayon assez grand le contiendra forcément. Parmi tous les zéros de g qui seront dans un tel disque, il y en a nécessairement un, appelons-le $z_{\min}$, qui est le plus proche possible de z_0 , éventuellement ex-aequo. Ceux qui seraient à l'extérieur de ce disque étant encore plus éloignés, $z_{\min}$ est donc le point singulier le plus proche de z_0 .

Considérons alors le disque ouvert de rayon $R = |z_0 - z_{\min}|$. Ce disque ne contient aucun point singulier puisqu'il n'y en a pas de plus proche que $z_{\min}$, donc h est analytique sur ce disque. D'après le théorème III.6, le rayon de convergence de la série entière de h de centre z_0 est supérieur ou égal à R . Mais si ce rayon était $> R$, cela voudrait dire que la fonction h est analytique dans un disque ouvert qui contient $z_{\min}$, ce qui est absurde. Donc le rayon de convergence est égal à $R = |z_0 - z_{\min}|$. •

Ⓜ Cette démonstration ne marcherait plus si les points singuliers ne formaient pas un ensemble discret. Par exemple la fonction $\ln_1(z)$ (l'un des logarithmes de z , que nous introduisons dans la section suivante) est analytique sur le domaine Ω_1 égal au plan $\mathbb{C}$ privé de l'intervalle réel $]-\infty; 0]$ et est discontinue sur cet intervalle (donc non analytique); ainsi les points réels négatifs avec 0 inclus, sont des points singuliers où la fonction $\ln_1(z)$ cesse d'être analytique. Prenons par exemple $z_0 = -1 + i$. La distance de ce point z_0 à la frontière de Ω_1 est 1; mais le rayon de convergence de la série entière, centrée en $z_0 = -1 + i$, de $\ln_1(z)$ n'est pas 1, mais $\sqrt{2}$. Cela peut sembler paradoxal en vertu du raisonnement (faux) que voici : la somme d'une série entière de centre $z_0 = -1 + i$ et de rayon $\sqrt{2}$ est forcément analytique dans son disque, donc sur le segment $]-2; 0[$, alors que $\ln_1(z)$ ne l'est pas. Ce qui est faux est l'idée (implicitement considérée comme évidente) que la somme de la série est forcément égale à $\ln_1(z)$. Le paradoxe apparent provient du fait que dans le domaine complexe il y a une infinité de fonctions logarithme qui se valent toutes; lorsqu'on traverse l'intervalle $]-2; 0[$ (tout en restant dans le disque de convergence de la série), la série cesse d'être égale à $\ln_1(z)$ et devient brusquement égale à un autre logarithme, $\ln_2(z)$. Voir la section III.5 pour comprendre le phénomène.

Il est donc capital pour le corollaire III.6.5 que les singularités de la fonction soient des points isolés. Le corollaire resterait vrai pour des fonctions ayant des points singuliers *essentiels* (voir chapitre suivant), qui ne sont pas des quotients f/g , pourvu que ces points soient

(6) Elle consiste à démontrer une formule pour certaines valeurs particulières mais non isolées de z , puis en déduire qu'elle est vraie pour tout z complexe.

isolés. La différence décisive entre des singularités isolées et des singularités en intervalle comme pour $\ln_1(z)$, c'est que des singularités isolées ne pourront jamais séparer un disque en deux parties déconnectées : l'intersection d'un disque avec un domaine du type $\mathbb{C} - Z(g)$ sera toujours connexe ; tandis que l'intersection d'un disque avec Ω_1 peut être non connexe. Or, le corollaire III.6.4 nous dit que si une fonction coïncide avec sa série entière sur un petit disque contenu dans un domaine connexe, elle coïncidera avec la même série dans tout le domaine. Par contre, si la fonction coïncide avec sa série entière sur un petit disque contenu dans un domaine *non* connexe, on pourra seulement conclure qu'elle doit coïncider avec la même série dans celle des composantes connexes qui contient le petit disque.

En tenant compte de ces remarques, il serait donc possible de proposer une version beaucoup plus générale du corollaire III.6.5, qui ne prendrait en compte que la propriété topologique de l'ensemble des points singuliers de ne pouvoir diviser aucun disque en deux parties déconnectées. Ceci est laissé en exercice pour les passionnés ; un corollaire aussi général ne présente aucun intérêt dans le cadre de ce cours orienté vers les applications.

La théorie de Cauchy est l'un des sujets les plus classiques des mathématiques, les livres qui l'exposent sont extrêmement nombreux et il est évidemment impossible de le présenter sous une forme originale ; je ne ferai pas mieux ici que les meilleurs auteurs, donc je me contente de ce résumé des propriétés essentielles.

Pour un exposé plus complet de tout ce qui peut se déduire de la formule de Cauchy :

- M. Lavrentiev, B. Chabat, *Méthodes de la théorie des fonctions d'une variable complexe*
- V. Smirnov, *Cours de mathématiques supérieures, tome III*
- E. T. Whittaker et G. N. Watson, *A Course of Modern Analysis*

Dans la suite, j'insisterai sur les aspects de la théorie qui sont souvent négligés dans les livres parce qu'ils ne sont pas assez généraux mais qui sont essentiels lorsqu'on veut *effectivement* calculer.

III.5 Fonctions multiformes

Reprenons l'exemple de la fonction $1/z$, analytique dans $\mathbb{C} - \{0\}$. Il n'y a aucun problème pour trouver sa dérivée, qui est la fonction analytique $-1/z^2$, ni les dérivées suivantes qui sont $2/z^3$, $-3!/z^4$, $4!/z^5$...

Lorsqu'on considère une intégrale curviligne $\int_{\gamma} f(z) dz$ sur un lacet γ , mais que f devient non analytique en certains points intérieurs au domaine circonscrit par le lacet γ , alors l'intégrale n'est pas forcément nulle : on ne peut en effet garantir que l'intégrale est nulle (en invoquant la formule de Green) que si les conditions de Cauchy et Riemann sont satisfaites *partout* à l'intérieur de γ . Toutefois si la fonction $f(z)$ est déjà la dérivée d'une fonction analytique F ($f = F'$), alors l'intégrale sera nulle *même si* F est non analytique sur un trou entouré par γ , car on peut l'intégrer par quadrature :

$$\int_{\gamma} f(z) dz = \int_0^T F'(z(t)) z'(t) dt = F(z(T)) - F(z(0)) = 0 \quad (\text{III.58})$$

Cela étant dit, calculons à la main les intégrales :

$$I_n = \int_{\gamma} \frac{dz}{z^n} \quad \text{pour } n \geq 1 \quad (\text{III.59})$$

Comme l'intégrale ne dépend pas du lacet, mais seulement de sa classe d'homologie, on va prendre le lacet particulier qu'est le cercle de centre 0 et de rayon r . cela se paramètre par $z(\theta) = re^{i\theta}$, donc :

$$I_n = \int_0^{2\pi} \frac{1}{r^n e^{in\theta}} ire^{i\theta} d\theta = \frac{i}{r^{n-1}} \int_0^{2\pi} e^{-i(n-1)\theta} d\theta = \begin{cases} 2i\pi & \text{si } n = 1 \\ 0 & \text{si } n > 1 \end{cases} \quad (\text{III.60})$$

On observe que pour $n > 1$, I_n est nulle bien que la fonction $1/z^n$ cesse d'être analytique en zéro ; cela n'a rien d'étonnant après ce qui a été dit ci-dessus puisque pour $n > 1$ la fonction $1/z^n$ est la dérivée de $-1/(n-1)z^{n-1}$, qui est analytique dans $\Omega = \mathbb{C} - \{0\}$. Mais I_1 n'est pas nulle, donc on peut en déduire qu'il n'existe pas de fonction analytique dans Ω dont la dérivée serait $1/z$.

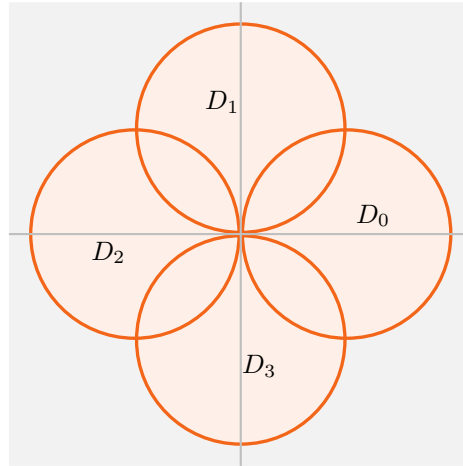


Figure III.2 – Recouvrement d'une région autour de l'origine par quatre disques ouverts D_0, D_1, D_2, D_3 de rayon 1, dont aucun ne contient l'origine elle-même. Il est possible de contourner l'origine en restant dans ces disques

Toutefois, comme cela a été vu en section III.4, la fonction $1/z$ est la somme de séries entières sur des cercles qui recouvrent Ω et les séries entières ont des primitives. En partant de la relation (III.53), il advient que $1/z$ est forcément la dérivée de :

$$\sum_{n \geq 0} \frac{(-1)^n}{(n+1)z_0^{n+1}} (z - z_0)^{n+1} = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n z_0^n} (z - z_0)^n \quad (\text{III.61})$$

puisque'en dérivant terme par terme cette dernière série, on retrouve bien celle de $1/z$. Or, on peut recouvrir le domaine Ω par des disques où ces séries convergent.

La clé du paradoxe est que les primitives sont définies à une constante près : si on considère deux de ces séries primitives, chacune sur leur disque de convergence, mais avec des centres différents, z_0 et z_1 , alors, sur l'intersection de leurs disques de convergence respectifs, on peut seulement affirmer qu'elles diffèrent d'une constante (puisqu'elles ont la même dérivée), mais on ne peut affirmer qu'elles y sont égales. Prenons quatre centres, par

exemple 1, i , -1 et $-i$, afin que les disques de convergence correspondants D_0 , D_1 , D_2 et D_3 recouvrent tout le tour de zéro — figure III.2. Si on veut définir une primitive unique sur $D_0 \cup D_1$, on pourra la définir par :

$$\sum_{n \geq 1} \frac{(-1)^{n-1}}{n} (z-1)^n \quad (\text{III.62})$$

sur D_0 , puis par :

$$\sum_{n \geq 1} \frac{(-1)^{n-1}}{ni^n} (z-i)^n + A_1 \quad (\text{III.63})$$

sur D_1 , en ajustant la constante A_1 pour que les deux définitions coïncident dans l'intersection $D_0 \cap D_1$. On peut prolonger cette primitive à D_2 en la définissant par :

$$- \sum_{n \geq 1} \frac{1}{n} (z+1)^n + A_2 \quad (\text{III.64})$$

et en ajustant A_2 pour faire coïncider sa somme avec la fonction déjà définie sur $D_0 \cup D_1$ dans l'intersection $D_1 \cap D_2$. Si on prolonge encore à D_3 , par :

$$- \sum_{n \geq 1} \frac{1}{ni^n} (z+i)^n + A_3 \quad (\text{III.65})$$

on pourra ajuster la constante A_3 pour que ce soit compatible avec la valeur précédente dans $D_2 \cap D_3$, mais il faut aussi faire coïncider avec la définition prise initialement sur D_0 , puisque D_3 a aussi une intersection avec D_0 : rien ne prouve qu'on peut avoir *à la fois* la compatibilité dans $D_2 \cap D_3$ et dans $D_3 \cap D_0$. Au contraire nous avons la preuve que cela n'est pas compatible, puisque nous avons déjà démontré qu'il n'existe pas de primitive de $1/z$ dans tout Ω .

Pour résumer, appelons $F(z)$, la somme de la série (III.62) dans D_0 , prolongée par (III.63) dans D_1 et par (III.64) dans D_2 , avec les constantes A_1 et A_2 ajustées. Si on prolonge encore dans D_3 avec la constante A_3 ajustée pour $D_2 \cap D_3$, la fonction prolongée différera sur $D_3 \cap D_0$ de la fonction initiale $F(z)$ par une constante *non nulle*.

On peut calculer cette constante non nulle ; ce serait extrêmement compliqué en sommant les séries (III.62), (III.63), (III.64) et (III.65) pour déterminer A_1 , A_2 , puis A_3 ⁽⁷⁾ ; on peut procéder autrement en se souvenir que la primitive $F(z)$ est :

$$F(z) = F(1) + \int_{\gamma} \frac{1}{z} dz \quad (\text{III.66})$$

où γ est un chemin d'origine 1 et d'extrémité z ⁽⁸⁾.

Lorsqu'on prolonge de D_2 à D_3 , la valeur de $F(z)$ est donnée par un chemin γ_1 d'origine 1 qui passe par D_1 , puis par D_2 , puis par D_3 , c'est-à-dire au-dessus de l'origine. Dans $D_3 \cap D_0$, la fonction initiale définie par la série (III.62) est donnée par $\int_{\gamma_2} (1/z) dz$, où γ_2 est un chemin qui va de 1 à z en restant dans D_0 — figure III.3. La différence :

$$\int_{\gamma_1} \frac{dz}{z} - \int_{\gamma_2} \frac{dz}{z} \quad (\text{III.67})$$

(7) La valeur précise des trois constantes A_1 , A_2 et A_3 est sans intérêt pour la présente argumentation ; pour les curieux, signalons quand même qu'elles valent respectivement $i\pi/2$, $i\pi$ et $3i\pi/2$.

(8) $F(1) = 0$ puisque $F(z)$ est donné dans D_0 par la série (III.62) de centre 1, qui est nulle pour $z = 1$.

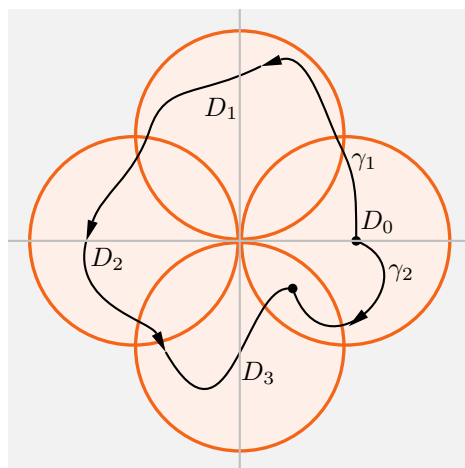


Figure III.3 – γ_1 est un chemin dont l'origine est le point 1 et l'extrémité un point z situé dans $D_3 \cap D_0$; le chemin γ_1 passe par les disques D_0 , D_1 , D_2 et D_3 , en passant au-dessus de l'origine. Le chemin γ_2 va de 1 à z en restant dans D_0 . La concaténation $\gamma_1 - \gamma_2$ est le lacet consistant à parcourir γ_1 de 1 à z , puis γ_2 en sens inverse de z à 1. Ce lacet fait le tour de l'origine dans le sens direct et est homotopiquement équivalent au lacet $z(\theta) = e^{i\theta}$ ($0 \leq \theta < 2\pi$)

est donc une intégrale sur un chemin qui est la concaténation de γ_1 et $-\gamma_2$, qui va de 1 à 1 en faisant le tour de l'origine et qui est donc un lacet entourant l'origine ; cette intégrale vaut $2i\pi$ d'après ce qui a été calculé précédemment.

Ainsi, si on poursuivait l'ajustement de constantes pour passer de D_3 à D_0 , la condition de compatibilité avec la fonction déjà définie sur D_3 imposerait de prendre sur D_0 :

$$F(z) = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} (z-1)^n + A_4 \quad (\text{III.68})$$

avec $A_4 = 2i\pi$, c'est-à-dire la fonction qu'on avait au départ, augmentée de $2i\pi$.

Il est essentiel de bien comprendre ce phénomène, si on veut maîtriser les techniques de calcul que nous verrons dans les chapitres suivants.

Le phénomène consiste en ce qu'une fonction, définie au départ dans un domaine (ici le disque D_0) et prolongée de proche en proche, ne reprend pas les mêmes valeurs lorsqu'on revient au point de départ. Dans la littérature ancienne d'avant 1945, on appelait *multiformes* de telles fonctions, avec l'idée que la fonction $F(z)$ n'aurait pas (pour un z donné) une valeur unique : ainsi si on fait $z = 1$ dans la série (III.62), on trouve $F(1) = 0$, mais après un tour complet autour de l'origine on obtient $F(1) = 2i\pi$; après deux tours, on aurait $F(1) = 4i\pi$ etc. Au cours du dernier demi-siècle, les mathématiciens ont codifié la notion de fonction. On enseigne aujourd'hui qu'une fonction a, pour une valeur donnée de la variable, une valeur unique. On a donc rejeté cette ancienne terminologie afin de ne pas perturber chez les jeunes élèves l'assimilation déjà difficile des notions élémentaires de mathématiques. C'est pourquoi, en dehors du titre de la section, nous ne parlerons jamais de fonction multiforme. Dans le cas de l'exemple précédent, on prendra un domaine dans lequel on ne peut pas faire le tour de l'origine, par exemple $\Omega_1 = \mathbb{C} -]-\infty; 0]$ et on dira que la fonction $F(z)$ est une vraie fonction définie sur Ω_1 , par la valeur unique qu'elle prend sur Ω_1 .

En conclusion : il n'existe pas de fonction analytique dans $\mathbb{C} - \{0\}$, dont la dérivée serait $1/z$ mais il en existe une sur $\Omega_1 = \mathbb{C} -]-\infty; 0]$. On peut en donner une expression simple :

Théorème III.7 Soit $z = re^{i\theta}$ un nombre complexe dans Ω_1 ($r > 0$, $-\pi < \theta < \pi$) et de même $z_0 = r_0 e^{i\theta_0}$. La fonction $F(z) = \ln r + i\theta$ est analytique dans Ω_1 , a pour dérivée $1/z$. Dans tout disque $|z - z_0| < |z_0|$, elle est égale, soit à la série entière :

$$\ln(r_0) + i\theta_0 + \sum_{n \geq 1} \frac{(-1)^{n-1}}{n z_0^n} (z - z_0)^n \quad (\text{III.69})$$

soit à cette série augmentée de $2i\pi$.

Preuve La fonction $F(z) = \ln r + i\theta$ satisfait dans Ω_1 aux conditions de Cauchy et Riemann ; on le vérifie très facilement avec l'expression (III.21) de ces conditions en coordonnées polaires r, θ ; en effet, ici la partie réelle est $u = \ln r$ et la partie imaginaire $v = \theta$ et on a bien :

$$\frac{\partial \ln r}{\partial r} = \frac{1}{r} = \frac{\partial \theta}{r \partial \theta} \quad \text{et} \quad \frac{\partial \theta}{\partial r} = 0 = -\frac{\partial \ln r}{r \partial \theta} \quad (\text{III.70})$$

car ici $u = \ln r$ et $v = \theta$. De plus, la dérivée de la fonction $F(z)$ est donnée par la relation (III.24), d'où :

$$F'(z) = e^{-i\theta} \frac{\partial F}{\partial r} = e^{-i\theta} \frac{\partial \ln r}{\partial r} = \frac{e^{-i\theta}}{r} = \frac{1}{z} \quad (\text{III.71})$$

Pour ce qui est de la série, il suffit de remarquer que sa dérivée est la série de $1/z$ et que la constante d'intégration ne peut être que 0 ou $2i\pi$; cela ayant déjà été fait dans l'exemple étudié plus haut, il est inutile d'y revenir. •

La possibilité de définir une primitive de $1/z$ sur un domaine impose qu'il soit impossible de faire le tour de l'origine dans ce domaine et c'est pourquoi Ω_1 a été construit en enlevant au plan $\mathbb{C}$ toute une demi-droite, qui va du point 0 jusqu'à l'infini. On appelle cela une *coupure* du plan. Rien n'impose que cette coupure soit la demi-droite $]-\infty; 0]$; n'importe quelle autre demi-droite allant de l'origine à l'infini conviendrait tout aussi bien pour empêcher qu'on puisse faire le tour de l'origine, ou même n'importe quelle courbe. Voici quatre cas particuliers :

1. Ω_1 est le domaine $\mathbb{C} -]-\infty; 0]$, c'est-à-dire l'ensemble des nombres complexes $z = re^{i\theta}$ tels que $r > 0$, $-\pi < \theta < \pi$.
2. Ω_2 est le domaine $\mathbb{C} - [0; \infty[$, c'est-à-dire l'ensemble des nombres complexes $z = re^{i\theta}$ tels que $r > 0$, $0 < \theta < 2\pi$.
3. Ω_3 est le domaine formé de $\mathbb{C}$ privé de la demi-droite verticale positive. Ω_3 est l'ensemble des nombres complexes $z = re^{i\theta}$ tels que $r > 0$, $-3\pi/2 < \theta < \pi/2$.
4. Ω_4 est le domaine formé de $\mathbb{C}$ privé de la demi-droite verticale négative. Ω_4 est l'ensemble des nombres complexes $z = re^{i\theta}$ tels que $r > 0$, $-\pi/2 < \theta < 3\pi/2$.
5. Ω_5 est le domaine formé de $\mathbb{C}$ privé de la spirale d'Archimède, d'équation $r = \theta$. Ω_5 est l'ensemble des nombres complexes $z = re^{i\theta}$ où θ peut prendre n'importe quelle valeur réelle positive non nécessairement inférieure à 2π , telle que $\theta \neq r$.

On a représenté ces différents domaines sur les figures III.4(a), III.4(b), III.4(c) et III.4(d). Le disque D_0 des figures III.2 et III.3 est visiblement inclus dans les trois domaines Ω_1 , Ω_3 et Ω_4 ; il n'est pas inclus dans Ω_2 et n'est pas non plus inclus dans Ω_5 . Appelons $\ln_1$ la

fonction que nous avons définie précédemment dans Ω_1 ; si $z = re^{i\theta}$, $\ln_1(z) = \ln r + i\theta$. La partie imaginaire de $\ln_1(z)$ est toujours comprise entre $-\pi$ et $+\pi$, puisque sur Ω_1 l'angle θ est compris entre ces bornes. Posons alors :

- $\ln_2(z) = \ln r + i\theta$ pour $0 < \theta < 2\pi$;
- $\ln_3(z) = \ln r + i\theta$ pour $-\pi/2 < \theta < \pi/2$;
- $\ln_4(z) = \ln r + i\theta$ pour $-\pi/2 < \theta < 3\pi/2$.

Ces fonctions ont toutes pour dérivée $1/z$ et diffèrent donc d'une constante. Le point $z = 1$ est contenu dans les trois domaines Ω_1 , Ω_3 et Ω_4 , mais pas dans Ω_2 . Les trois fonctions $\ln_1$, $\ln_3$ et $\ln_4$ s'annulent en ce point, mais la fonction $\ln_2(z)$ n'y est pas définie. La série :

$$\sum_{n \geq 1} \frac{(-1)^{n-1}}{n} (z-1)^n \quad (\text{III.72})$$

s'annule aussi en $z = 1$ et a pour dérivée $1/z$; par conséquent sur D_0 , on a :

$$\ln_1(z) = \ln_3(z) = \ln_4(z) = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} (z-1)^n \quad (\text{III.73})$$

Pour ce qui est de $\ln_2$, on a vu plus haut que la série ci-dessus n'est égale à $\ln_2(z)$ que dans la moitié supérieure du disque D_0 (celle qui est formée des nombres de partie imaginaire positive) et qu'elle est égale à $\ln_2(z) - 2i\pi$ dans la moitié inférieure de D_0 .

Il est aisé de voir cela directement sur les expressions $\ln r + i\theta$; le plan étant divisé en quatre quadrants $Q_1 : 0 < \theta < \pi/2$, $Q_2 : \pi/2 < \theta < \pi$, $Q_3 : \pi < \theta < 3\pi/2$ et enfin $Q_4 : 3\pi/2 < \theta < 2\pi$, on voit que sur le quadrant Q_1 , les quatre fonctions $\ln_1$, $\ln_2$, $\ln_3$, $\ln_4$, sont égales puisque les nombres complexes du premier quadrant ont un argument compris entre 0 et $\pi/2$, donc compatible avec chacun des quatre intervalles qui caractérisent les quatre fonctions. Par contre dans le quadrant Q_4 par exemple, θ doit être pris entre $-\pi/2$ et 0 pour $\ln_1$, $\ln_3$ et $\ln_4$, mais entre $3\pi/2$ et 2π pour $\ln_2$. Le même nombre complexe z aura un argument supérieur de $2i\pi$ si on le considère dans Ω_2 au lieu de le considérer dans Ω_1 , Ω_3 , ou Ω_4 . Donc pour z dans Q_4 , on aura $\ln_1(z) = \ln_3(z) = \ln_4(z) = \ln_2(z) - 2i\pi$. De la même façon, on vérifie (par la prise en compte directe des angles θ) que sur Q_2 on a $\ln_1(z) = \ln_2(z) = \ln_3(z) + 2i\pi = \ln_4(z)$ et sur Q_3 on a $\ln_1(z) = \ln_2(z) - 2i\pi = \ln_3(z) = \ln_4(z) - 2i\pi$.

Le cas de $\ln_5$ est plus difficile à décrire par écrit, mais pas par l'image ; c'est pourquoi nous avons reporté sa comparaison avec $\ln_2$ sur la figure III.4(d).

Quelqu'un qui partirait du point $z = 1$ en restant dans le domaine Ω_5 (ce qui signifie qu'il ne pourrait pas franchir la ligne spirale) serait obligé, pour s'éloigner de l'origine, d'augmenter continuellement sa coordonnée θ ; donc la partie imaginaire de $\ln_5(z)$ augmenterait sans limite. La k^e fois qu'il franchirait la demi-droite $[0; \infty[$, θ vaudrait $2k\pi$, il aurait aussi fait k fois un tour complet de l'origine. Supposons qu'il aboutisse au point $z_1 = re^{i\theta} = re^{i(\theta-2k\pi)}$ et soit γ_1 le chemin qu'il aura suivi — figure III.5(a). Appelons γ_2 un chemin qui va de 1 à z_1 en franchissant les bras de spirale, mais en restant dans Ω_2 (à l'exception du point $z = 1$ lui-même) — figure III.5(b). Il est clair qu'on a :

$$\ln_2(z_1) = \int_{\gamma_2} \frac{1}{z} dz \quad \text{et} \quad \ln_5(z_1) = \int_{\gamma_1} \frac{1}{z} dz \quad (\text{III.74})$$

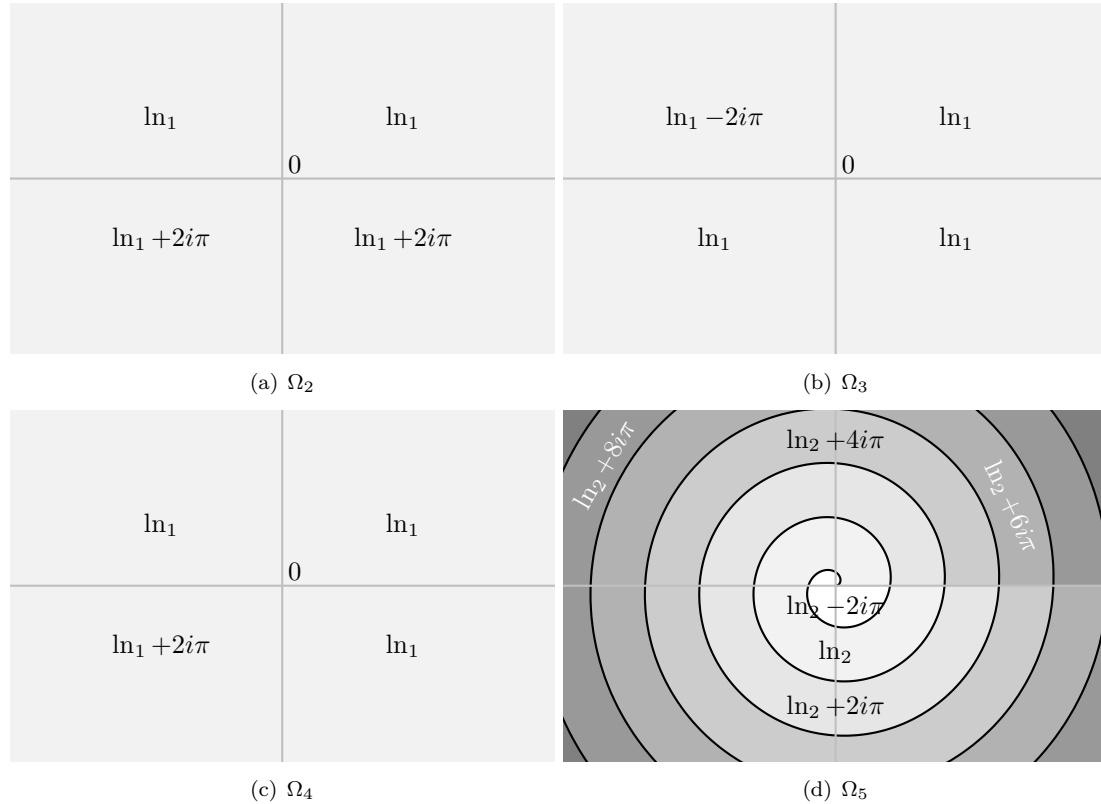


Figure III.4 – Déterminations $\ln_1$, $\ln_3$, $\ln_4$ et $\ln_5$ du logarithme complexe, comparées à la détermination $\ln_2$

puisque $\ln_5(1) = 0$ et $\ln_2$, quoique non défini en $z = 1$, y a pour limite 0 si z tend vers 1 par le demi-plan supérieur. Donc

$$\ln_5(z_1) - \ln_2(z_1) = \int_{\gamma} \frac{1}{z} dz \quad (\text{III.75})$$

où γ est la concaténation de γ_1 et de $-\gamma_2$ (γ_2 parcouru en sens inverse) et est donc un lacet (il part de 1 et aboutit à 1). On peut le décomposer en une autre concaténation de lacets dont chacun est homologiquement équivalent à un cercle de centre 0 — figures III.5(c), III.5(d) et III.5(e) : il suffit de considérer les portions du chemin γ_1 entre deux franchissements consécutifs du chemin γ_2 et les portions de γ_2 séparées par les spires de γ_1 : chaque paire de ces portions forme par concaténation un lacet simple qui fait une fois le tour de l'origine. On en déduit alors que $\ln_5(z_1) - \ln_2(z_1) = 2ik\pi$ comme détaillé sur la figure III.5.

Une autre manière de calculer la différence $\ln_5(z_1) - \ln_2(z_1)$ est de considérer la discontinuité de $\ln_2$ le long de l'axe réel positif. En effet, la fonction $\ln_5(z) - \ln_2(z)$ a une dérivée nulle puisque $\ln_5(z)$ et $\ln_2(z)$ ont toutes deux la même dérivée $1/z$. Par conséquent $\ln_5(z) - \ln_2(z)$ sera constante dans son domaine de définition, $\Omega_2 \cup \Omega_5$. Ce dernier n'est pas connexe puisqu'il n'est pas d'un seul tenant et la constante peut donc changer lorsqu'on passe d'un morceau à l'autre. Cela se produit précisément lorsqu'on traverse l'axe réel positif. Le changement de constante est une discontinuité qu'on peut calculer en remarquant que $\ln_5(z)$ reste continue lorsqu'on traverse l'axe réel positif (c'est à travers la spirale qu'elle a une discontinuité). Donc la discontinuité de $\ln_5(z) - \ln_2(z)$ à travers l'axe réel positif est

due uniquement à la discontinuité de $\ln_2(z)$, dont on sait qu'elle est de $+2i\pi$ (du bas vers le haut). Lorsqu'on parcourt Ω_5 en avançant entre les spires, $\ln_5(z) - \ln_2(z)$ augmentera donc *brusquement* de $2i\pi$ chaque fois qu'on traversera l'axe réel positif, ce qui permet de calculer aisément $\ln_5(z)$ à partir de $\ln_2(z)$ (voir figure III.4(d)).

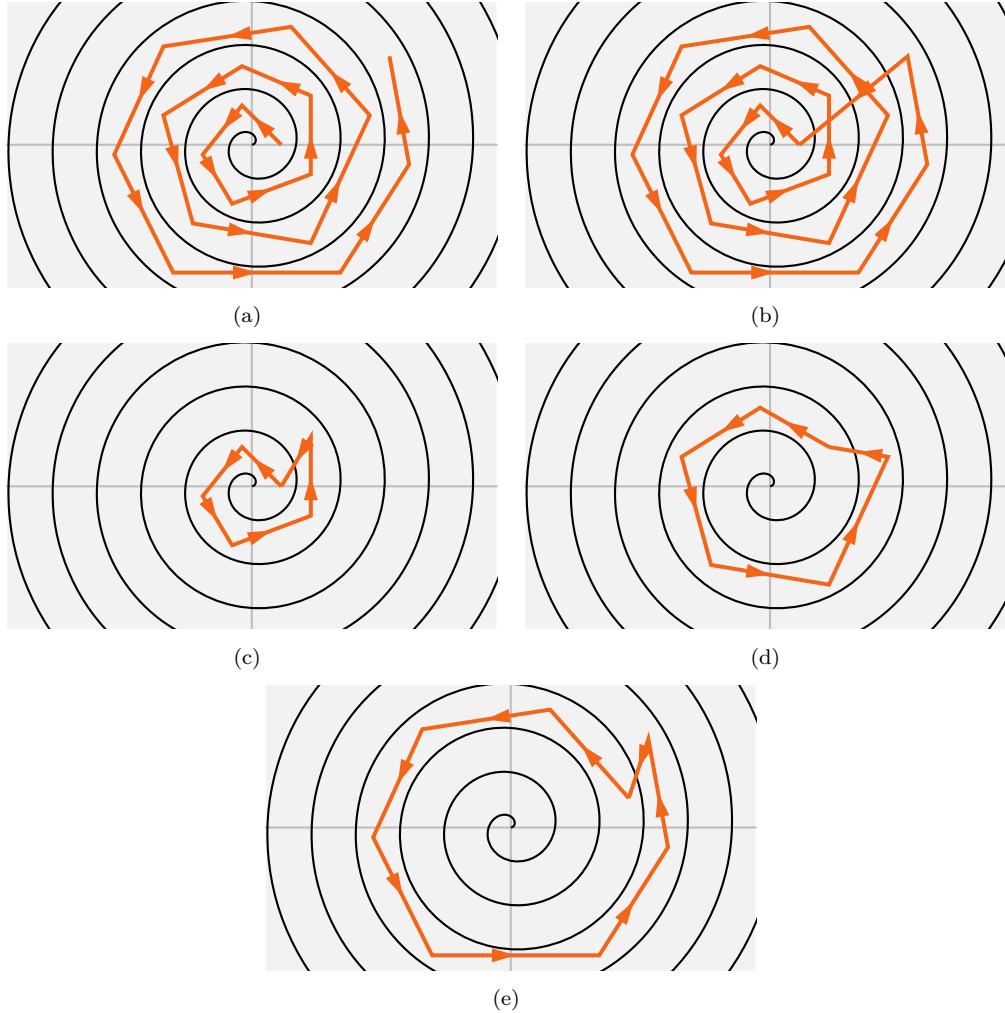


Figure III.5 – Le domaine Ω_5 est ce qui reste du plan $\mathbb{C}$ lorsqu'on a enlevé la spirale. Pour aller du point $z = 1$ au point $z = z_1$ tout en restant dans Ω_5 , on doit suivre un chemin tel que γ_1 (a). Par contre dans Ω_2 , on peut aller du point $z = z_1$ au point $z = 1$ par un chemin beaucoup plus direct γ_2 (b). La fonction $F(z)$, primitive complexe de $1/z$ qui s'annule en $z = 1$, peut donc être définie dans Ω_2 par $\ln_2(z_1) = -\int_{\gamma_2} \frac{1}{z} dz$ et dans Ω_5 par $\ln_5(z_1) = \int_{\gamma_1} \frac{1}{z} dz$. Par conséquent $\ln_5(z_1) - \ln_2(z_1) = \int_{\gamma} \frac{1}{z} dz$, où γ est la concaténation $\gamma_1 + \gamma_2$ (b) et est donc fermé. Or γ est par ailleurs aussi la concaténation de plusieurs lacets Γ_j tous homologiquement équivalents à un cercle qui entoure l'origine : $\gamma = \gamma_1 + \gamma_2 = \Gamma_1 + \Gamma_2 + \Gamma_3$ (c), (d) et (e). On en déduit que $\ln_5(z_1) - \ln_2(z_1) = \sum_j \int_{\Gamma_j} \frac{1}{z} dz = 6i\pi$.

Bien entendu, si z_1 avait été dans la millièmes spire au lieu d'être dans la troisième, on aurait dû concaténer mille lacets Γ_j pour avoir γ et $\ln_5(z_1) - \ln_2(z_1)$ aurait été égal à $2000i\pi$.

IV — Calcul des résidus

IV.1 Séries de Laurent

Le théorème III.6 dit qu'une fonction analytique dans un disque y est développable en série entière. Nous allons voir maintenant qu'on peut généraliser ce résultat lorsque la fonction $f(z)$ est analytique dans une couronne comme celle de la figure IV.1. Commençons par établir une version étendue de la formule de Cauchy. Si z_0 est un point de la couronne, on peut dire, d'après la formule de Cauchy (III.41), que :

$$f(z_0) = \frac{1}{2i\pi} \int_{\gamma} \frac{f(z)}{z - z_0} dz \quad (\text{IV.1})$$

où γ est le bord d'un domaine entièrement contenu dans la couronne et parcouru en sens

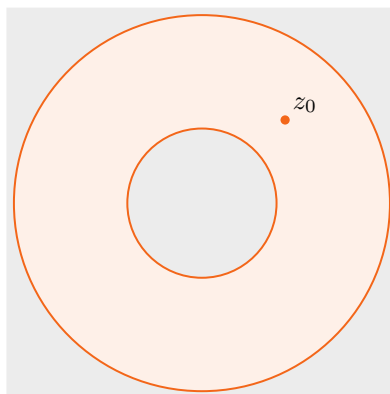


Figure IV.1 – Couronne dans le plan complexe

direct. Les chemins représentés sur les figures IV.2(a), IV.2(b) et IV.2(c), sont des exemples

de tels chemins. On comprend aisément en comparant les figures IV.2(c) et IV.2(d), que le lacet γ est homologiquement équivalent (dans la couronne) à la concaténation des deux lacets γ_1 et γ_2 ; on en déduit que :

$$f(z_0) = \frac{1}{2i\pi} \int_{\gamma_1} \frac{f(z)}{z - z_0} dz + \frac{1}{2i\pi} \int_{\gamma_2} \frac{f(z)}{z - z_0} dz \quad (\text{IV.2})$$

Il faut faire attention à l'orientation : le lacet γ_1 est orienté dans le sens positif, tandis que le lacet γ_2 est orienté dans le sens négatif (voir les figures IV.2(c) et IV.2(d)) ; si on oriente γ_2 dans le sens positif, il faudra mettre le signe '−' devant l'intégrale correspondante dans (IV.2).

Pour utiliser cette formule de Cauchy généralisée, il ne faut pas oublier que le point z_0 doit impérativement se trouver entre les chemins γ_1 et γ_2 , de même que dans la formule de Cauchy simple (III.41), il est impératif que le point z_0 se trouve à l'intérieur du domaine délimité par le chemin.

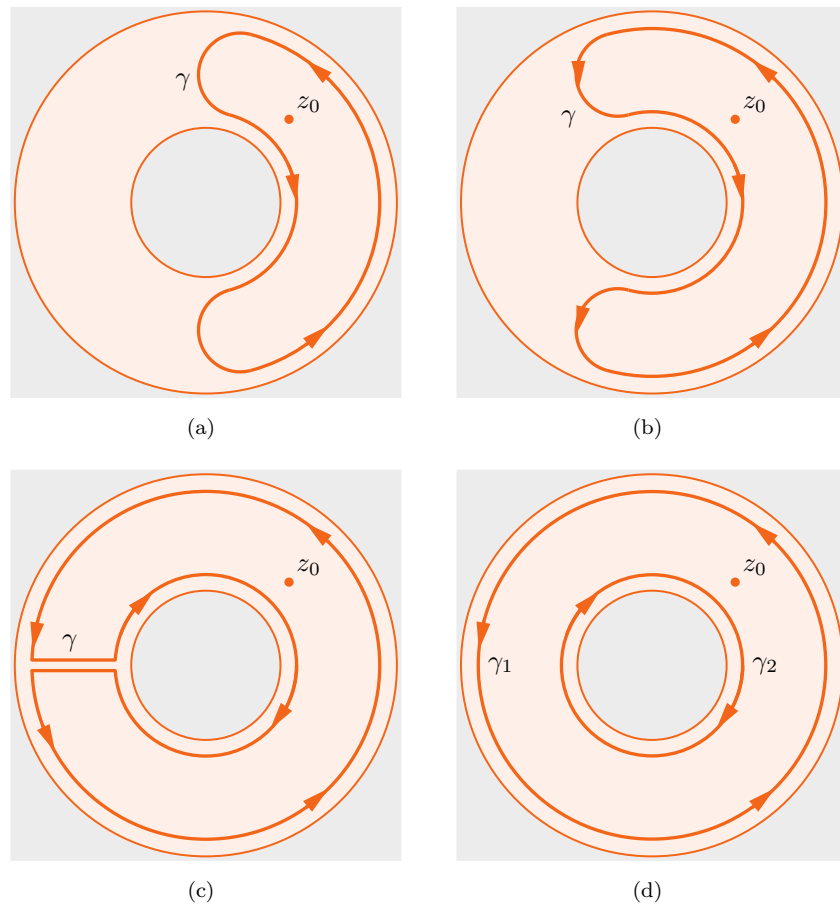


Figure IV.2 – La concaténation des deux lacets séparés γ_1 et γ_2 de la figure IV.2(d) (noter que γ_1 est orienté positivement tandis que γ_2 est orienté négativement) est homologiquement équivalente au lacet γ des figures IV.2(a), IV.2(b) et IV.2(c)

De même qu'en section III.4, nous avons utilisé la formule de Cauchy simple pour obtenir un développement en série, nous allons maintenant utiliser la formule généralisée. Pour fixer les idées, supposons que la couronne soit de centre 0, de petit rayon r et de grand rayon R (si le centre était un point z_1 autre que 0, il suffirait de remplacer dans ce qui suit z par

$z - z_1$ et z_0 par $z_0 - z_1$). Substituons, dans l'intégrale sur le chemin γ_1 de l'expression (IV.2) :

$$\frac{1}{z - z_0} = \frac{1}{z} \cdot \frac{1}{1 - z_0/z} = \frac{1}{z} \cdot \sum_{n=0}^{\infty} \frac{z_0^n}{z^n} = \sum_{n=0}^{\infty} \frac{z_0^n}{z^{n+1}} \quad (\text{IV.3})$$

(ceci ne diffère pas de ce qui avait été fait en III.4), puis dans l'intégrale sur le chemin γ_2 :

$$\frac{1}{z - z_0} = -\frac{1}{z_0} \cdot \frac{1}{1 - z/z_0} = -\frac{1}{z_0} \cdot \sum_{n=0}^{\infty} \frac{z^n}{z_0^n} = -\sum_{n=0}^{\infty} \frac{z^n}{z_0^{n+1}} \quad (\text{IV.4})$$

Ceci par contre diffère de ce qui avait été fait en III.4, car on obtient une série en puissances de $1/z_0$. Pour pouvoir intervertir la sommation de série avec l'intégration, il faut que chaque série soit normalement convergente sur le chemin d'intégration concerné ; or la série (IV.3) converge pour $|z| > |z_0|$, il suffit donc de choisir le chemin γ_1 de telle sorte qu'il reste à l'extérieur du disque $|z| > |z_0|$ (cette condition est satisfaite si γ_1 reste assez proche du bord extérieur de la couronne) et le chemin γ_2 de telle sorte qu'il reste à l'intérieur du disque $|z| < |z_0|$ (cette condition est satisfaite si γ_2 reste assez proche du bord intérieur de la couronne).

On obtient alors :

$$f(z_0) = \sum_{n=0}^{\infty} a_n z_0^n + \sum_{n=1}^{\infty} \frac{b_n}{z_0^n} \quad (\text{IV.5})$$

avec :

$$a_n = \frac{1}{2i\pi} \int_{\gamma_1} \frac{f(z)}{z^{n+1}} dz \quad \text{et} \quad b_n = \frac{1}{2i\pi} \int_{-\gamma_2} f(z) \cdot z^{n-1} dz \quad (\text{IV.6})$$

Si on prend pour γ_1 un cercle de rayon ρ_1 , pour γ_2 un cercle de rayon ρ_2 , de sorte que $r < \rho_2 < |z_0| < \rho_1 < R$ et qu'on appelle M_ρ le maximum de $|f(z)|$ sur le cercle $|z| = \rho$, on obtient des inégalités analogues à (III.49) :

$$|a_n| \leq \frac{M_{\rho_1}}{\rho_1^n} \quad \text{et} \quad |b_n| \leq M_{\rho_2} \cdot \rho_2^n \quad (\text{IV.7})$$

On constate dans (IV.5) que $f(z_0)$ est donné par la somme de *deux* séries, l'une entière et l'autre entière en $1/z_0$. Ce résultat est connu sous le nom du théorème de P. Laurent :

Théorème IV.1 Une fonction $f(z)$ analytique dans une couronne $r < |z| < R$ est dans cette couronne égale à la somme d'une série entière en z et d'une série entière en $1/z$, de sorte que la série entière en z converge dans le disque $|z| < R$ et la série entière en $1/z$ dans le disque $|1/z| < 1/r$ (c'est-à-dire à l'extérieur du disque $|z| > r$).

Définition IV.1 On appelle *série de Laurent*, une série de puissances de z , $\sum c_n z^n$, mais qui (à la différence des séries entières), comporte aussi bien des puissances négatives que positives de z . La somme des termes correspondant à $n \geq 0$ est appelée la *partie entière* ou *régulière* et la somme des termes correspondant à $n \leq -1$ est appelée la *partie singulière*.

L'égalité (IV.5) est appelée le développement de $f(z_0)$ en série de Laurent. Les a_n sont les coefficients de la partie régulière, les b_n ceux de la partie singulière. On voit sans difficulté

qu'il ne peut pas y avoir plusieurs développements de Laurent différents : si une série de la forme (IV.5) est identiquement nulle (dans une couronne de centre 0), alors ses coefficients c_n sont tous nuls ; par exemple en utilisant les inégalités (IV.7).

Pour compléter la démonstration du théorème, il reste seulement à vérifier les affirmations concernant la convergence. Cela résulte des inégalités (IV.7). En effet, on peut prendre dans ces inégalités ρ_1 aussi proche qu'on veut de R et ρ_2 aussi proche qu'on veut de r . La formule de Cauchy-Hadamard du théorème III.2, combinée avec les inégalités (IV.7) donne :

$$\limsup_{n \rightarrow \infty} |a_n|^{1/n} \leq \limsup_{n \rightarrow \infty} M_{\rho_1}^{1/n} / \rho_1 = \frac{1}{\rho_1} \quad (\text{IV.8})$$

qui prouve que le rayon de convergence de la série en z est supérieur ou égal à ρ_1 ; ce dernier pouvant être pris aussi proche qu'on veut de R , cela entraîne évidemment que le rayon de convergence est au moins égal à R , autrement dit que la série est convergente pour $|z| < R$. On voit de la même façon :

$$\limsup_{n \rightarrow \infty} |b_n|^{1/n} \leq \limsup_{n \rightarrow \infty} M_{\rho_2}^{1/n} \cdot \rho_2 = \rho_2 \quad (\text{IV.9})$$

qui prouve que le rayon de convergence de la série en $1/z$ est supérieur ou égal à $1/\rho_2$ etc.

Ⓡ Les déductions précédentes ne montrent pas que les rayons de convergence sont *égaux* à (respectivement) $1/r$ et R ; il se peut évidemment, si $f(z)$ est analytique pour $r < |z| < R$, qu'elle soit prolongeable sur un domaine plus grand. Nous y reviendrons quand nous discuterons des singularités des fonctions analytiques.

IV.2 Théorème des résidus

Nous avons observé en section III.5 et plus particulièrement l'équation (III.60), que l'intégrale $1/z^n$ sur un lacet entourant l'origine est nulle sauf pour $n = 1$. Ceci a une conséquence remarquable si on le combine avec le théorème de Laurent IV.1. Soit en effet $f(z)$ une fonction analytique dans la couronne $r < |z| < R$ et soit γ un lacet simple contenu dans cette couronne et entourant le petit disque $|z| \leq r$, qu'on supposera orienté comme toujours dans le sens direct (par exemple le lacet γ_1 de la figure IV.2(d), ou le lacet $-\gamma_2$ qui lui est homologiquement équivalent).

L'intégrale de $f(z) dz$ le long d'un tel lacet γ n'a évidemment aucune raison d'être nulle, puisque la fonction $f(z)$ n'est pas supposée analytique dans le disque $|z| < r$. Par contre, si on développe $f(z)$ en série de Laurent conformément à (IV.5), on obtient :

$$\int_{\gamma} f(z) dz = \sum_{n=0}^{\infty} a_n \int_{\gamma} z^n dz + \sum_{n=1}^{\infty} b_n \int_{\gamma} \frac{1}{z^n} dz = 2i\pi b_1 \quad (\text{IV.10})$$

Ainsi, l'intégrale sur γ de $f(z) dz$ ne dépend que du coefficient b_1 . Cauchy a appelé ce coefficient le *résidu* de la fonction $f(z)$ sur le disque $|z| \leq r$. Les premières approches de Cauchy (y compris le tout premier mémoire de 1814, où ce mot ne figure pas encore) étaient conçues comme des méthodes pour calculer des intégrales définies.

L'idée est la suivante. Dans les cas simples, il est très facile de calculer le coefficient b_1 ; l'unicité du développement de Laurent garantit que n'importe quelle manière de le calculer donnera le même résultat que les formules (IV.6).

La série de Laurent permet d'étudier aisément les points singuliers isolés d'une fonction analytique. Dire que z_0 est un point singulier isolé de la fonction $f(z)$, équivaut à dire que $f(z)$ est analytique dans une couronne de la forme $0 < |z - z_0| < R$. Le fait que le point singulier soit isolé se traduit par le fait que la partie singulière de la série de Laurent $\sum_{n \geq 0} b_n / (z - z_0)^n$ a un rayon de convergence infini, ou encore $\limsup |b_n|^{1/n} = 0$. On distingue les points singuliers isolés en deux classes :

- si les b_n sont tous nuls pour $n > n_0$, on dit que z_0 est un pôle d'ordre n_0 .
- si une infinité des b_n sont non nuls, on dit que z_0 est un point singulier essentiel.

Nous verrons cela plus concrètement à l'occasion d'exemples.

Bien entendu, les points singuliers ne sont pas forcément isolés. Lorsque l'on dit "soit $f(z)$ une fonction analytique dans la couronne $r < |z| < R$ ", les points singuliers de $f(z)$ peuvent former un ensemble absolument arbitraire dans le disque $|z| \leq r$. Pour donner un exemple, considérons les fonctions $1/(1 - z^n)$; ces fonctions sont toutes analytiques dans le disque $|z| < 1$ mais aussi à l'extérieur de ce disque. Leurs points singuliers sont sur le cercle $|z| = 1$, ce sont les racines de l'unité. Posons :

$$\Phi_n(z) = \prod_{j=1}^n \frac{1}{1 - z^j} \quad \text{et} \quad \Psi_n(z) = \prod_{j=1}^n \frac{1}{1 - z^{-j}} \quad (\text{IV.11})$$

Ces fonctions Φ_n et Ψ_n sont elles aussi analytiques à l'intérieur et à l'extérieur du cercle $|z| = 1$. Leurs points singuliers sont tous les nombres complexes de la forme $e^{2i\pi(j/k)}$ avec $1 \leq j \leq k \leq n$. On peut montrer que lorsque n tend vers l'infini, $\Phi_n(z)$ tend dans $|z| < 1$ vers une limite analytique et $\Psi_n(z)$ tend dans $|z| > 1$ vers une limite également analytique. La limite $\Psi(z) = \lim \Psi_n(z)$ est donc un exemple de fonction analytique dans l'extérieur du cercle $|z| = 1$ mais pour laquelle tous les points du cercle $|z| = 1$ sont singuliers.

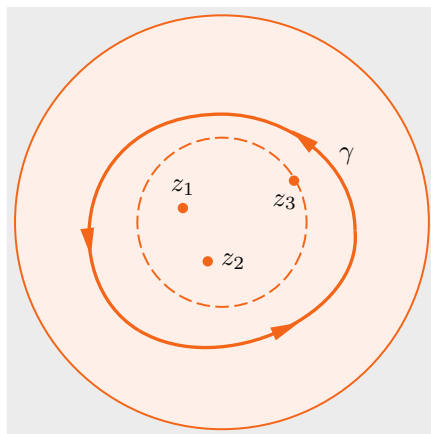


Figure IV.3

Toutefois les fonctions telles que $\Psi(z)$ sont peu courantes : en mathématique, on se sert plutôt de fonctions assez simples, de sorte qu'en pratique, les fonctions qu'on risque de rencontrer effectivement ont généralement un nombre fini de points singuliers isolés. Supposons que la fonction $f(z)$, qui est analytique dans la couronne $r < |z| < R$, n'ait dans le disque $|z| \leq r$ qu'un nombre fini de points singuliers $z_1, z_2, \dots$ comme sur la figure IV.4.

D'après le théorème IV.1, la fonction $f(z)$ sera développable en une série de Laurent $\sum a_n z^n + \sum b_n/z^n$ sur la couronne $r < |z| < R$. La fonction $f(z)$ se prolonge en une fonction analytique dans le disque $|z| \leq r$, dont seuls les points isolés $z_1, z_2, \dots$ sont exclus. La série $\sum b_n/z^n$ ne convergera pas pour $|z| \leq r$ (plus exactement, elle deviendra divergente pour $|z|$ inférieur au module du point singulier le plus éloigné de 0), mais autour de chacun des z_j il y a bien une couronne $0 < |z - z_j| < \varepsilon$ dans laquelle $f(z)$ est analytique et où on peut donc aussi développer $f(z)$ en série de Laurent. Autour de chaque z_j , on a donc :

$$f(z) = \sum_{n=0}^{\infty} a_n^{(j)} (z - z_j)^n + \sum_{n=1}^{\infty} \frac{b_n^{(j)}}{(z - z_j)^n} \quad (\text{IV.12})$$

Bien entendu, les coefficients $a_n^{(j)}$ et $b_n^{(j)}$ sont différents pour chaque j et différents aussi de a_n et b_n , mais il existe une relation entre eux, dont la plus simple et de très loin la plus importante est :

Théorème IV.2 Si $f(z)$ est analytique dans le disque $|z| < R$, à l'exception d'un nombre fini de points singuliers $z_1, z_2, \dots$ isolés et tous contenus dans le disque $|z| \leq r$ (avec $r < R$), alors le résidu du développement de Laurent de $f(z)$ sur la couronne $r < |z| < R$ est égal à la somme des résidus de $f(z)$ en chacun des points singuliers z_j .

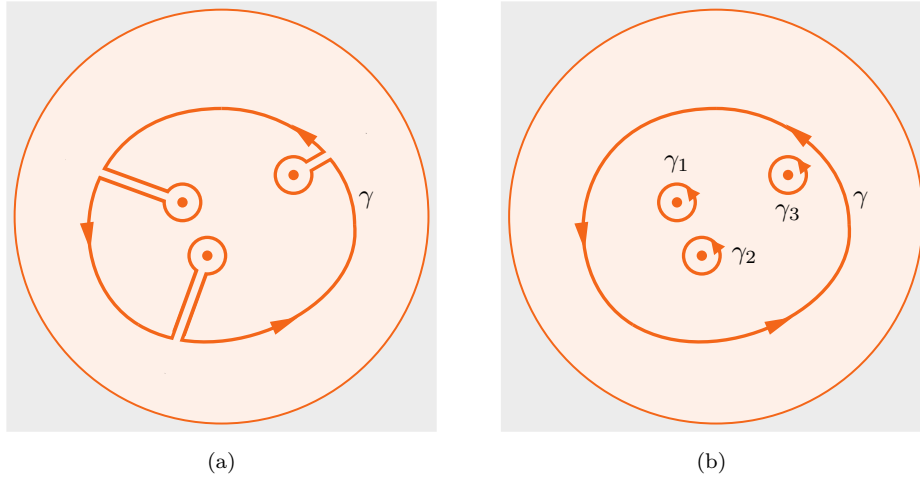


Figure IV.4

Preuve Soit γ un chemin simple contenu dans la couronne $r < |z| < R$ parcouru en sens direct et γ_j des chemins simples, également parcourus en sens direct, entourant chacun le seul point singulier z_j — figure IV.4(b). Il est immédiat que la concaténation $\gamma_1 + \gamma_2 + \gamma_3 + \dots$ est homologiquement équivalente à γ et il suffit de regarder la figure IV.4(a) pour comprendre pourquoi.

Par conséquent :

$$\frac{1}{2i\pi} \int_{\gamma} f(z) dz = \frac{1}{2i\pi} \int_{\sum_j \gamma_j} f(z) dz = \sum_j \frac{1}{2i\pi} \int_{\gamma_j} f(z) dz \quad (\text{IV.13})$$

d'où le résultat. •

Il existe bien sûr pour chaque n une relation analogue entre les coefficients a_n , b_n et les coefficients $a_n^{(j)}$, $b_n^{(j)}$. Pour l'obtenir, remarquons que $f(z)/z^{n+1}$ a, outre les points $z_1, z_2, \dots$,

aussi 0 pour point singulier. Supposons que 0 est distinct de tous les z_j (c'est-à-dire que 0 n'est pas un point singulier de $f(z)$). Comme pour les γ_j , on va considérer un lacet γ_0 entourant seulement 0 (et aucun des z_j). Alors :

$$a_n = \frac{1}{2i\pi} \int_{\gamma} \frac{f(z)}{z^{n+1}} dz = \frac{1}{2i\pi} \int_{\sum_j \gamma_j} \frac{f(z)}{z^{n+1}} dz = \frac{1}{2i\pi} \sum_j \int_{\gamma_j} \frac{f(z)}{z^{n+1}} dz \quad (\text{IV.14})$$

où cette fois la sommation inclut $j = 0$. Le terme correspondant à $j = 0$ est simplement $f^{(n)}(0)/n!$. Pour les autres termes, puisqu'on a supposé les z_j tous différents de 0, on peut développer $1/z^{n+1}$ au voisinage de chacun des z_j en série entière :

$$\frac{1}{z^{n+1}} = \frac{1}{[z_j + (z - z_j)]^{n+1}} = \frac{1}{z_j^{n+1}} \cdot \frac{1}{[1 + \frac{z-z_j}{z_j}]^{n+1}} = \sum_{\ell=0}^{\infty} c_{\ell}^{(j)} (z - z_j)^{\ell} \quad (\text{IV.15})$$

avec :

$$c_{\ell}^{(j)} = \frac{(-1)^{\ell} \binom{n+\ell}{\ell}}{z_j^{n+\ell+1}} \quad (\text{IV.16})$$

La série de Laurent de $f(z)/z^{n+1}$ au voisinage de z_j s'obtient alors en faisant le produit de la série de Laurent de $f(z)$ par celle (entière) de $1/z^{n+1}$ obtenue en (IV.15). L'intégrale :

$$\frac{1}{2i\pi} \int_{\gamma_j} \frac{f(z)}{z^{n+1}} dz \quad (\text{IV.17})$$

sera le coefficient de $1/(z - z_j)$ dans le produit de ces deux séries, soit :

$$\int_{\gamma_j} \frac{f(z)}{z^{n+1}} dz = \sum_{k=1}^{\infty} b_k^{(j)} c_{k-1}^{(j)} \quad (\text{IV.18})$$

où les $c_{k-1}^{(j)}$ sont les coefficients mentionnés en (IV.16). Il ne reste plus qu'à faire la somme de ces résultats pour $j = 0, 1, 2, \dots$. Si l'un des z_j est confondu avec 0, le raisonnement précédent doit être un peu modifié, mais reste valable dans son principe.

On obtient une relation analogue pour les b_n .

Nous mentionnons ces relations pour rappeler une fois de plus la propriété fondamentale des fonctions analytiques, que leur comportement dans une région est lié au comportement dans les autres. Sauf problème très spécial, seule la relation énoncée par le théorème IV.2 est intéressante.

On peut aussi énoncer le théorème IV.2 sous la forme du très célèbre :

Théorème IV.3 — Théorème des résidus Soit Ω un domaine du plan et $f(z)$ une fonction analytique dans Ω moins un nombre fini de points. Alors l'intégrale de $f(z) dz$ sur un lacet entourant ces points est égale à $2i\pi$ fois la somme des résidus de $f(z)$ en ces points.

IV.3 Calculs d'intégrales définies

La première application du théorème des résidus IV.3 est le calcul d'intégrales définies. Les premières publications de Cauchy sur le sujet se présentent en effet ainsi ⁽¹⁾ : *Mémoire sur*

(1) A.-L. CAUCHY, *Mémoire sur les intégrales définies*, Œuvres complètes de Cauchy, 1814, série I, tome 1, déjà cité.

les intégrales définies où l'on donne une formule générale de laquelle se déduisent les valeurs de la plupart des intégrales définies déjà connues et celles d'un grand nombre d'autres⁽²⁾ etc. Nous consacrons la présente section à cette première application, mais nous verrons par la suite des applications bien plus intéressantes.

En gros, le principe est le suivant : pour des points singuliers isolés de fonctions qui s'expriment par des formules algébriques simples, il est généralement aisé d'obtenir un développement en série de Laurent, surtout lorsqu'on se contente de calculer le terme en $1/(z - z_0)$, alors que les intégrales sont beaucoup plus difficiles.

Voici un exemple : soit l'intégrale :

$$\int_{-\infty}^{+\infty} \frac{dx}{x^4 + 1} \quad (\text{IV.19})$$

On peut calculer cette intégrale par quadratures ; pour cela, on doit d'abord, sachant que $x^4 + 1 = (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1)$, décomposer la fonction en :

$$\begin{aligned} \frac{1}{x^4 + 1} &= \frac{-\frac{1}{2\sqrt{2}}x + \frac{1}{2}}{x^2 - \sqrt{2}x + 1} + \frac{\frac{1}{2\sqrt{2}}x + \frac{1}{2}}{x^2 + \sqrt{2}x + 1} \\ &= \frac{1}{2} \frac{(-x\sqrt{2} + 1) + 1}{(-x\sqrt{2} + 1)^2 + 1} + \frac{1}{2} \frac{(x\sqrt{2} + 1) + 1}{(x\sqrt{2} + 1)^2 + 1} \end{aligned} \quad (\text{IV.20})$$

que l'on intègre par quadratures avec les fonctions $\arctan(-x\sqrt{2} + 1)$, $\ln(-x\sqrt{2} + 1)$, $\arctan(x\sqrt{2} + 1)$, $\ln(x\sqrt{2} + 1)$. Par la méthode des résidus, il suffit de dire que l'intégrale sur un lacet Γ en demi-cercle (figure IV.5(a)) est égale à $2i\pi$ fois la somme des résidus des points singuliers situés à l'intérieur du lacet et qui sont $z_1 = e^{i\pi/4}$ et $z_2 = e^{i3\pi/4}$ — ce sont les deux nombres complexes de partie imaginaire positive, tels que $z^4 + 1 = 0$.

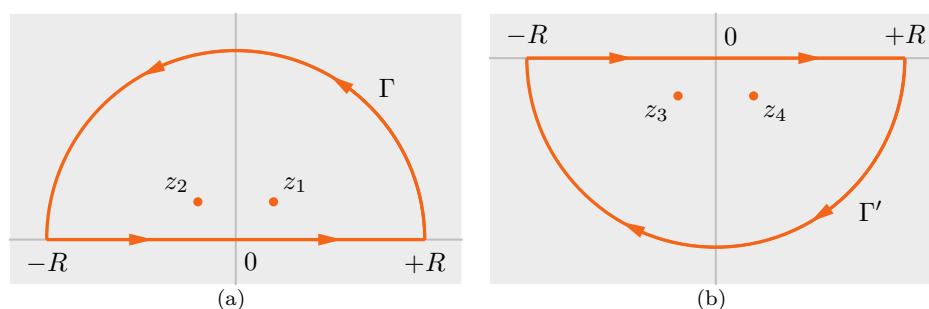


Figure IV.5 – Lacet en demi-cercle

Contrairement au calcul précédent qui exigeait des dénominateurs du second degré, on obtient les résidus de la fonction $1/(z^4 + 1)$ par la décomposition en éléments simples complexes du premier degré :

$$\frac{1}{z^4 + 1} = \frac{A_1}{z - z_1} + \frac{A_2}{z - z_2} + \frac{A_3}{z - z_3} + \frac{A_4}{z - z_4} \quad (\text{IV.21})$$

où on sait que $A_j = 1/4 z_j^3$ (le coefficient de la décomposition en éléments simples d'une fraction rationnelle $P(z)/Q(z)$ est donné par la formule $A_j = P(z_j)/Q'(z_j)$). Le coefficient

(2) A.-L. CAUCHY, *Œuvres complètes de Cauchy*, 1825, série II, tome 2, p. 343–387. La formule générale dont il est question dans le titre est le théorème des résidus, non encore nommé ainsi.

A_j est le résidu, puisque le terme $A_j/(z - z_j)$ de la décomposition est le terme en $1/(z - z_j)$ du développement en série de Laurent. Ainsi, les résidus sont obtenus presque sans calculs.

On voit sans difficultés que l'intégrale sur la partie semi-circulaire du chemin Γ tend vers zéro quand le rayon R tend vers l'infini, donc :

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{dx}{x^4 + 1} &= \lim_{R \rightarrow \infty} \int_{\Gamma} \frac{dz}{z^4 + 1} = 2i\pi \cdot \left[\frac{1}{4e^{i3\pi/4}} + \frac{1}{4e^{i9\pi/4}} \right] \\ &= \frac{\pi}{2} \left[\sin \frac{3\pi}{4} + \sin \frac{9\pi}{4} + i \cos \frac{3\pi}{4} + i \cos \frac{9\pi}{4} \right] = \frac{\pi}{\sqrt{2}} \end{aligned} \quad (\text{IV.22})$$

On aurait tout aussi bien pu passer par l'intermédiaire du lacet Γ' de la figure IV.5(b). Dans ce cas, il aurait fallu prendre en compte les résidus aux points $z_3 = e^{i5\pi/4}$ et $z_4 = e^{i7\pi/4}$.

Dans l'exemple que nous venons d'étudier, on pouvait aussi calculer l'intégrale par quadratures. En voici un autre, où on ne le peut pas :

$$\int_{-\infty}^{+\infty} \frac{e^{itx}}{x^4 + 1} dx \quad (\text{IV.23})$$

En tant que fonction de t , cette intégrale est la transformée de Fourier de la fonction $1/(x^4 + 1)$. Nous verrons plus tard que beaucoup de transformées de Fourier se calculent par la méthode des résidus.

La première chose à faire est de chercher les résidus ; les points singuliers sont les mêmes, puisque e^{itz} n'introduit aucun nouveau point singulier. En utilisant à nouveau la décomposition en éléments simples de $1/(z^4 + 1)$, on peut écrire :

$$\frac{e^{itz}}{z^4 + 1} = \left[\frac{A_1}{z - z_1} + \frac{A_2}{z - z_2} + \frac{A_3}{z - z_3} + \frac{A_4}{z - z_4} \right] e^{itz} \quad (\text{IV.24})$$

Au voisinage de z_1 , e^{itz} se développe en série entière de $z - z_1$:

$$e^{itz} = \sum_{n=0}^{\infty} \frac{(it)^n e^{itz_1}}{n!} (z - z_1)^n \quad (\text{IV.25})$$

On voit bien que si on multiplie cela par $A_1/(z - z_1)$, le terme en $1/(z - z_1)$ sera $A_1 e^{itz_1} = e^{itz_1}/4z_1^3$. Les autres termes $A_2/(z - z_2), \dots$ sont analytiques au point z_1 , donc ne contribuent qu'à la partie régulière de la série de Laurent. Ainsi le résidu de $e^{itz}/(z^4 + 1)$ au point z_1 est $e^{itz_1}/4z_1^3$. On trouve de la même façon que le résidu aux autres points est $e^{itz_j}/4z_j^3$.

Si nous voulons utiliser le même procédé que dans l'exemple précédent, il faut que l'intégrale sur la partie semi-circulaire du lacet tende vers zéro quand R tend vers l'infini et pour cela il faut que la fonction sous le signe $\int$ tende vers zéro (assez rapidement). On voit sans difficulté que cela ne peut être le cas que pour l'un à la fois des deux chemins Γ ou Γ' . En effet, si $t \geq 0$, $|e^{itz}|$ reste borné dans le demi-plan $\Im(z) > 0$ (mais croît exponentiellement vers l'infini dans le demi-plan $\Im(z) < 0$), tandis que si $t \leq 0$, $|e^{itz}|$ reste borné dans le demi-plan $\Im(z) < 0$ (mais croît exponentiellement vers l'infini dans le demi-plan $\Im(z) > 0$). Comme une croissance exponentielle ne peut pas être compensée par $1/(z^4 + 1)$, nous n'avons pas le choix (sauf pour $t = 0$) et nous devons prendre Γ lorsque $t > 0$ et Γ' lorsque $t < 0$. Ceci implique que notre intégrale (divisée par $2i\pi$) sera égale à la somme des résidus du demi-plan $\Im(z) > 0$ pour $t > 0$ et à la somme des résidus du demi-plan $\Im(z) < 0$

pour $t < 0$. Toutefois on peut aussi remarquer que si on remplace t par $-t$, la valeur de l'intégrale (IV.23) sera transformée en sa conjuguée, donc il suffit de faire le calcul pour $t > 0$. Sachant que $z_1 = e^{i\pi/4} = (1+i)/\sqrt{2}$ et $z_2 = e^{3i\pi/4} = (-1+i)/\sqrt{2}$, cela donne :

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{e^{itx}}{x^4+1} dx &= 2i\pi \left[\frac{e^{itz_1}}{4z_1^3} + \frac{e^{itz_2}}{4z_2^3} \right] = \frac{i\pi}{2} \left[e^{itz_1-i3\pi/4} + e^{itz_2-i9\pi/4} \right] \\ &= \frac{i\pi}{2} e^{-t/\sqrt{2}} \left[e^{it/\sqrt{2}-i3\pi/4} + e^{-it/\sqrt{2}-i9\pi/4} \right] \\ &= \pi \cos\left(\frac{t}{\sqrt{2}} - \frac{\pi}{4}\right) e^{-t/\sqrt{2}} \end{aligned} \quad (\text{IV.26})$$

L'intégrale est réelle, donc égale à sa conjuguée ; par conséquent, pour $t < 0$ elle vaudra $\pi \cos(-t/\sqrt{2} - \pi/4) e^{t/\sqrt{2}}$, soit pour t quelconque :

$$\int_{-\infty}^{+\infty} \frac{e^{itx}}{x^4+1} dx = \pi \cos\left(\frac{|t|}{\sqrt{2}} - \frac{\pi}{4}\right) e^{-|t|/\sqrt{2}} \quad (\text{IV.27})$$

On peut généraliser ces exemples sous la forme suivante :

Théorème IV.4 Soit la fraction rationnelle $P(x)/Q(x)$ telle que :

1. le dénominateur $Q(x)$ n'a que des racines simples z_j ;
2. le degré de $Q(x)$ dépasse celui de $P(x)$ d'au moins deux unités ;
3. aucune racine de $Q(x)$ n'est réelle.

alors :

$$\int_{-\infty}^{+\infty} \frac{P(x)}{Q(x)} dx = 2i\pi \sum_{+} \frac{P(z_j)}{Q'(z_j)} = -2i\pi \sum_{-} \frac{P(z_j)}{Q'(z_j)} \quad (\text{IV.28})$$

où $\sum_{+}$ désigne la sommation sur toutes les racines z_j de partie imaginaire positive et $\sum_{-}$ la sommation sur toutes les racines z_j de partie imaginaire négative.

Preuve La décomposition en éléments simples de la fraction rationnelle $P(x)/Q(x)$ montre que le résidu au point singulier z_j est $P(z_j)/Q'(z_j)$; le théorème IV.3 dit que l'intégrale sur le lacet Γ est égale à $2i\pi$ fois la somme des résidus aux points singuliers qu'il entoure, qui sont (pour R assez grand) ceux du demi-plan $\Im(z) > 0$; il dit aussi que l'intégrale sur le lacet Γ' est égale à $-2i\pi$ fois la somme des résidus aux points singuliers qu'il entoure, qui sont (pour R assez grand) ceux du demi-plan $\Im(z) < 0$. Il ne reste plus qu'à vérifier que l'intégrale sur la partie semi-circulaire de l'un ou l'autre de ces deux lacets tend vers zéro quand R tend vers l'infini ; or cela résulte de l'hypothèse 2, puisqu'en paramétrant la partie semi-circulaire de Γ , l'intégrale devient :

$$\int_0^\pi \frac{P(Re^{i\theta})}{Q(Re^{i\theta})} iRe^{i\theta} d\theta \quad (\text{IV.29})$$

L'hypothèse 2 entraîne que pour z grand on a $|P(z)/Q(z)| \leq M/|z|^2$, donc l'intégrale (en module) est majorée par :

$$\int_0^\pi \frac{M}{R^2} R d\theta = \pi \frac{M}{R} \quad (\text{IV.30})$$

ce qui permet de conclure. •

Théorème IV.5 $P(x)/Q(x)$ étant une fraction rationnelle satisfaisant aux mêmes conditions que dans le théorème IV.4, on a :

$$\int_{-\infty}^{+\infty} \frac{P(x)}{Q(x)} e^{itx} dx = \begin{cases} 2i\pi \sum_{+} \frac{P(z_j) e^{itz_j}}{Q'(z_j)} & \text{si } t > 0 \\ -2i\pi \sum_{-} \frac{P(z_j) e^{itz_j}}{Q'(z_j)} & \text{si } t < 0 \end{cases} \quad (\text{IV.31})$$

où $\sum_{+}$ désigne comme avant la sommation sur toutes les racines z_j de partie imaginaire positive et $\sum_{-}$ la sommation sur toutes les racines z_j de partie imaginaire négative.

Preuve décomposons $P(z)/Q(z)$ en éléments simples ; on obtient :

$$\frac{P(z)}{Q(z)} e^{itz} = \sum_{j=1}^N \frac{A_j e^{itz}}{z - z_j} \quad (\text{IV.32})$$

où N est le degré de Q et $A_j = P(z_j)/Q'(z_j)$. Soit z_{j_0} l'un des points singuliers ; on peut développer e^{itz} en série entière autour de ce point :

$$e^{itz} = \sum_{n=0}^{\infty} \frac{(it)^n e^{itz_{j_0}}}{n!} (z - z_{j_0})^n \quad (\text{IV.33})$$

donc la série de Laurent de la fonction $\frac{A_{j_0} e^{itz}}{z - z_{j_0}}$ est :

$$\frac{A_{j_0} e^{itz_{j_0}}}{z - z_{j_0}} + \sum_{n=1}^{\infty} \frac{(it)^n e^{itz_{j_0}}}{n!} (z - z_{j_0})^{n-1} \quad (\text{IV.34})$$

Les autres termes $\frac{A_j e^{itz}}{z - z_j}$ pour $j \neq j_0$ sont analytiques au voisinage de z_{j_0} et y ont donc un développement en série entière, sans partie singulière. On en conclut que la partie singulière du développement en série de Laurent de $[P(z)/Q(z)]e^{itz}$ se réduit au seul terme $A_{j_0} e^{itz_{j_0}}/(z - z_{j_0})$, ce qui prouve que le résidu est $A_{j_0} e^{itz_{j_0}}$.

Il faut encore prouver que l'intégrale sur la partie semi-circulaire de Γ ou Γ' tend vers zéro. Or pour $t > 0$, $|e^{itz}| = e^{-ty}$ (y étant la partie imaginaire de z), ce qui est partout ≤ 1 dans le demi-plan $y > 0$; de même pour $t < 0$, $|e^{itz}|$ est partout ≤ 1 dans le demi-plan $y < 0$. Ainsi la fonction à intégrer est dans chaque cas majorée par M/R^2 et tout se passe comme dans le théorème IV.4 (excepté que cette fois on n'a pas le choix entre Γ et Γ'). •

Lorsque les racines du dénominateur $Q(x)$ ne sont pas simples, on peut procéder de façon analogue : on utilisera aussi la décomposition en éléments simples, mais le résidu au point z_j ne sera pas donné par une formule aussi commode. En effet, la partie de la décomposition en éléments simples qui concerne le point z_j sera de la forme :

$$\sum_{k=1}^{\alpha_j} \frac{A_{j,k}}{(z - z_j)^k} \quad (\text{IV.35})$$

où α_j est la multiplicité de la racine z_j . En faisant le produit de cette expression avec la série :

$$e^{itz} = \sum_{n=0}^{\infty} \frac{(it)^n e^{itz_j}}{n!} (z - z_j)^n \quad (\text{IV.36})$$

et en regroupant les termes en $1/(z - z_j)$, on aura pour le résidu :

$$\sum_{k=1}^{\alpha_j} \frac{(it)^{k-1} e^{itz_j}}{(k-1)!} A_{j,k} \quad (\text{IV.37})$$

Bien entendu, il n'est pas possible de donner un aperçu exhaustif de toutes les intégrales définies calculables par la formule des résidus. Dans chaque cas particulier, il faut trouver l'astuce spécifique la mieux adaptée (généralement un choix astucieux du lacet). La section IV.4 traite du calcul des résidus ; les sections IV.5 et IV.6 développent quelques cas particuliers qui pourront donner des idées.

IV.4 Comment calculer pratiquement les résidus

Les théorèmes IV.4 et IV.5 concernaient les cas où le dénominateur de la fonction à intégrer est un polynôme qui n'a que des racines simples ; les théorèmes fournissent alors des formules simples. On peut aussi avoir affaire à un dénominateur qui, au lieu d'être un *polynôme*, est une fonction analytique non polynomiale : par exemple si le dénominateur est $e^z + 1$, celui-ci a bien des racines simples $z_n = i\pi(2n + 1)$ pour $n \in \mathbb{Z}$ mais leur nombre est infini car $e^z + 1$ n'est pas un polynôme. Il peut aussi arriver que le dénominateur, polynôme ou non, ait des racines doubles, ou triples etc. Le numérateur peut aussi être autre chose qu'un polynôme comme dans le théorème IV.4, ou que le produit d'un polynôme par une exponentielle comme dans le théorème IV.5. On ne peut évidemment pas énoncer un théorème général qui fournit une formule de résolution pour n'importe quelle intégrale ; déjà le fait de pouvoir appliquer le théorème des résidus IV.3 en prenant un lacet en demi-cercle est particulier aux fonctions du type envisagé par les théorèmes IV.4 et IV.5 : si le dénominateur est $e^z + 1$, l'intégrale sur le demi-cercle est divergente et cet artifice, qui marchait très bien pour les polynômes, ne marche plus du tout. En général on devra trouver des lacets astucieux et adaptés au cas qu'on veut traiter.

Si on pose le problème général d'une intégrale sur un intervalle réel, on ne peut espérer se ramener au théorème des résidus IV.3 qu'en construisant un lacet astucieux qui, ou bien contient l'intervalle d'intégration et donne zéro sur la partie ajoutée, ou bien s'y ramène d'une façon ou d'une autre par paramétrage. La résolution d'un tel problème est une question d'imagination et ne peut pas être enfermée dans une formule magique.

On va donc laisser de côté la question du choix astucieux d'un lacet et expliquer seulement comment calculer les résidus d'une fonction. On va même se restreindre ici au cas où les points singuliers sont isolés, plus précisément au cas où la fonction dont on cherche les résidus est de la forme $f(z)/g(z)$ (quotient de deux fonctions analytiques). Les points singuliers sont alors les zéros du dénominateur g et sont par conséquent forcément isolés. On dit que ce sont des pôles. On verra par la suite quelques cas de singularités plus compliquées (qui ne sont pas les zéros d'un dénominateur analytique) et qui ne peuvent pas non plus être traités par des formules prévues pour des pôles. Par exemple $z = 0$ est un point singulier de la fonction $\exp(1/z)$; ce point est certes isolé, mais on ne peut pas écrire $\exp(1/z)$ sous la forme $f(z)/g(z)$ avec f et g analytiques et $g(0) = 0$. La fonction d'Euler $\text{Eu}(z)$ de la section suivante possède aussi un point singulier de ce type, dit *essentiel*. En section IV.5, on verra aussi le cas de la fonction $F(z) = [z^\alpha]_1 \cdot [(z-1)^\beta]_1$, dont la singularité n'est pas un point isolé, mais tout un segment (cela n'empêche pas d'avoir un résidu).

Pour une fonction $h(z) = f(z)/g(z)$, c'est-à-dire le quotient de deux fonctions analy-

tiques, il existe une formule générale qu'on trouvera dans la plupart des manuels :

$$\text{rés}(h, z_0) = \lim_{z \rightarrow z_0} \frac{1}{(n-1)!} \frac{d^{n-1}}{dz^{n-1}} \left((z - z_0)^n h(z) \right) \quad (\text{IV.38})$$

où n est la multiplicité du pôle z_0 . Dans le cas particulier d'un pôle simple, cela devient :

$$(z - z_0)h(z) = \frac{f(z)}{g(z)/(z - z_0)} \quad (\text{IV.39})$$

et comme la limite de $g(z)/(z - z_0)$ quand z tend vers z_0 est $g'(z_0)$, on voit que (IV.38) se réduit à :

$$\text{rés}(h, z_0) = \frac{f(z_0)}{g'(z_0)} \quad (\text{IV.40})$$

La formule (IV.40) est très facile à utiliser et généralise la formule $P(z_0)/Q(z_0)$ qu'on a vue au théorème IV.4 pour des polynômes. Par contre, pour $n > 1$, la formule (IV.38) est rarement commode et sa commodité diminue factoriellement en n : non seulement il faut dériver l'expression assez complexe $(z - z_0)^n h(z)$, mais ensuite celle-ci est une forme indéterminée dont il faudra trouver la limite — règle de l'Hospital, développements limités etc.

Partons du principe que le résidu de $h = f/g$ au point z_0 est le coefficient du terme en $1/(z - z_0)$ dans le développement de h en série de Laurent. Si z_0 est un pôle d'ordre n , c'est qu'il est un zéro d'ordre n du dénominateur $g(z)$. Autrement dit, le premier terme non nul de la série entière de g autour de z_0 est le terme d'ordre n :

$$g(z) = (z - z_0)^n \times \left(a_n + a_{n+1}(z - z_0) + a_{n+2}(z - z_0)^2 + a_{n+3}(z - z_0)^3 + \dots \right) \quad (\text{IV.41})$$

avec $a_n \neq 0$. Par conséquent la fonction réduite :

$$\hat{g}(z) = a_n + a_{n+1}(z - z_0) + a_{n+2}(z - z_0)^2 + a_{n+3}(z - z_0)^3 + \dots \quad (\text{IV.42})$$

telle que $g(z) = (z - z_0)^n \cdot \hat{g}(z)$ est analytique dans un disque autour de z_0 et ne s'y annule pas, en sorte que le quotient $\hat{h}(z) = f(z)/\hat{g}(z)$ est analytique dans ce disque.

On peut alors obtenir la série de Laurent de h en calculant la série entière de $\hat{h}(z)$:

Le coefficient de $1/(z - z_0)$ de la série de Laurent de $h(z)$ est le même que le coefficient de $(z - z_0)^{n-1}$ dans la série entière de $\hat{h}(z)$. □

Ainsi, pour obtenir le résidu de $h(z)$ au point z_0 , il suffit de calculer le développement en série entière de $\hat{h}(z)$ de centre z_0 jusqu'à l'ordre $n - 1$.

Ceci permet déjà de démontrer la formule (IV.38). En effet, la fonction $(z - z_0)^n h(z)$ n'est autre que la fonction $\hat{h}(z)$. Le $(n - 1)^{\text{e}}$ coefficient de Taylor de la série entière de $\hat{h}(z)$ est donc bien $1/(n - 1)!$ fois la $(n - 1)^{\text{e}}$ dérivée de $\hat{h}(z)$ au point z_0 , qu'on peut *en théorie* obtenir comme la limite indiquée dans (IV.38). *En pratique*, il est presque toujours bien plus efficace de calculer la série entière de $\hat{h}(z)$ en effectuant la division par puissances croissantes de la série entière de $f(z)$ par celle de $\hat{g}(z)$. Ainsi, si on a :

$$\begin{aligned} f(z) &= b_0 + b_1(z - z_0) + b_2(z - z_0)^2 + b_3(z - z_0)^3 + \dots \\ g(z) &= a_n(z - z_0)^n + a_{n+1}(z - z_0)^{n+1} + a_{n+2}(z - z_0)^{n+2} + \dots \end{aligned} \quad (\text{IV.43})$$

on posera $\zeta = z - z_0$ et on effectuera la division :

$$\begin{array}{r|l} b_0 + b_1 \zeta + b_2 \zeta^2 + b_3 \zeta^3 + \dots & a_n + a_{n+1} \zeta + a_{n+2} \zeta^2 + a_{n+3} \zeta^3 + \dots \\ \hline & \frac{b_0}{a_n} + \dots \end{array} \quad (\text{IV.44})$$

Dans certains cas, on pourra avoir directement la série entière de la fonction $1/\hat{g}(z)$, par exemple si c'est une série géométrique ou binomiale. Dans ce cas, on peut effectuer le produit des séries de $f(z)$ et de $1/\hat{g}(z)$, si cela semble moins lourd que la division. Un peu de pratique montrera que ces opérations sont — sauf dans quelques rares cas très particuliers — bien moins pénibles que les calculs de dérivées et de limites exigés par la formule (IV.38).

IV.5 Fonction d'Euler

Afin d'illustrer l'ensemble des propriétés des fonctions analytiques que nous avons rencontrées jusqu'ici, voyons un exemple très intéressant et instructif. On appelle *intégrale d'Euler* l'intégrale suivante :

$$\text{Eu}(z) = \int_0^\infty \frac{e^{-t}}{1 + zt} dt \quad (\text{IV.45})$$

où z est un nombre complexe. L'intégrale devient divergente lorsque le dénominateur peut

Ⓡ Il y a beaucoup d'intégrales d'Euler. Il ne faut pas confondre celle-ci avec celle de la fonction $\Gamma(x)$ qui sera étudiée au chapitre V. L. Euler, mathématicien suisse (1707–1783) a trouvé d'innombrables formules de l'analyse, qu'il a étendues aux nombres complexes sans posséder la notion de fonction analytique ; mais toutes les fonctions qu'il envisageait étaient des expressions algébriques : polynômes, fractions rationnelles, exponentielles, intégrales dépendant d'un paramètre etc. Il a donc pratiqué le *prolongement analytique* sans le savoir.

s'annuler, c'est-à-dire lorsque z est un nombre réel négatif. Pour toute autre valeur complexe de z , le dénominateur ne peut s'annuler et la fonction $\text{Eu}(z)$ est alors parfaitement définie. Cette fonction est analytique.

Le domaine où tout se passe bien est le plan complexe privé de l'intervalle $]-\infty; 0]$, soit le domaine Ω_1 que nous avons déjà introduit à la fin du chapitre III. Commençons par minorer le dénominateur $1 + zt$ en cherchant son minimum lorsque t parcourt l'intervalle d'intégration $[0; \infty[$. Si on sépare partie réelle et partie imaginaire en posant $z = x + iy$, on aura $|1 + zt|^2 = 1 + 2xt + (x^2 + y^2)t^2$. Le minimum de cette expression lorsque t parcourt $[0; \infty[$ est :

$$d(z)^2 = \begin{cases} 1 & \text{si } x \geq 0 \\ \frac{y^2}{x^2 + y^2} & \text{si } x < 0 \end{cases} \quad (\text{IV.46})$$

Par conséquent le minimum du module du dénominateur sera :

$$d(z) = \begin{cases} 1 & \text{si } x \geq 0 \\ |\Im(z)|/|z| & \text{si } x < 0 \end{cases} \quad (\text{IV.47})$$

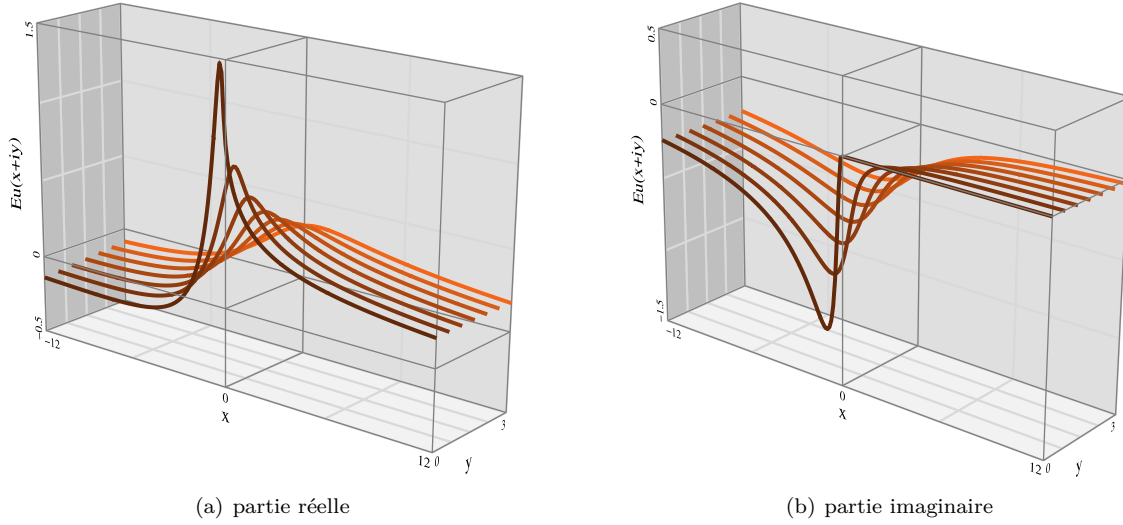


Figure IV.6 – Différentes sections horizontales de la fonction $\text{Eu}(z)$: chaque graphique représente la fonction $x \mapsto \text{Eu}(x + i\varepsilon)$, pour différentes valeurs de ε (à gauche la partie réelle, à droite la partie imaginaire). Pour $\varepsilon < 0$, les figures seraient symétriques car $\text{Eu}(\bar{z}) = \overline{\text{Eu}(z)}$.

On voit aisément que $d(z)$ est > 0 sur Ω_1 , mais devient nul sur $]-\infty; 0[$. Le cas où z tend vers zéro est spécial : l'expression $|\Im(z)|/|z|$ devient indéterminée et il n'y a pas de limite. Pour z dans Ω_1 on peut alors avoir une majoration de $|\text{Eu}(z)|$:

$$|\text{Eu}(z)| \leq \int_0^\infty \frac{e^{-t}}{d(z)} dt = \frac{1}{d(z)} \quad (\text{IV.48})$$

Voyons maintenant pourquoi $\text{Eu}(z)$ est analytique dans Ω_1 . On va vérifier que $[\text{Eu}(z+h) - \text{Eu}(z)]/h$ a une limite quand h tend vers zéro et que cette limite est :

$$\text{Eu}'(z) = \int_0^\infty \frac{-te^{-t}}{(1+zt)^2} dt \quad (\text{IV.49})$$

Il faut donc montrer que :

$$\frac{1}{h} \int_0^\infty \left[\frac{1}{1+(z+h)t} - \frac{1}{1+zt} + \frac{th}{(1+zt)^2} \right] e^{-t} dt \quad (\text{IV.50})$$

tend vers zéro quand h tend vers zéro. Or, l'expression entre crochets dans (IV.50) devient, si on la réduit au même dénominateur :

$$\frac{h^2 t^2}{[1+(z+h)t][1+zt]^2} \quad (\text{IV.51})$$

Si z est dans Ω_1 , $d(z)$ est > 0 ; puisque h doit tendre vers zéro, on peut le choisir tel que $d(z+h) > \frac{1}{2}d(z)$ ⁽³⁾. Par conséquent, le module de l'expression (IV.50) est majoré par :

$$\int_0^\infty \frac{|h|t^2 e^{-t}}{\frac{1}{2}d(z)^3} dt = \frac{4|h|}{d(z)^3} \quad (\text{IV.52})$$

qui tend bien vers zéro quand $|h|$ tend vers zéro.

(3) C'est un argument de continuité classique : l'expression (IV.47) montre que $z \mapsto d(z)$ est continue sur Ω_1 , donc si $d(z) > 0$, il existe un voisinage de z dans lequel $d > d(z)/2$ et il suffit de prendre h assez petit pour que $z+h$ soit dans ce voisinage.

On a ainsi montré que $\text{Eu}(z)$ est analytique dans Ω_1 et que sa dérivée y est donnée par (IV.49).

A priori, il n'y a aucune raison que $\text{Eu}(z)$ puisse se prolonger par continuité sur la demi-droite $]-\infty; 0]$; on va voir, d'ailleurs, qu'on ne le peut effectivement pas et que la fonction $\text{Eu}(z)$ a un saut de discontinuité à travers cette demi-droite, que nous allons calculer. Pour cela on va considérer un point $z = -a$ (avec a réel > 0) sur cette demi-droite et chercher les deux limites suivantes :

$$\text{Eu}^+(-a) = \lim_{h \rightarrow 0+} \text{Eu}(-a + h) \quad \text{et} \quad \text{Eu}^-(-a) = \lim_{h \rightarrow 0-} \text{Eu}(-a + h) \quad (\text{IV.53})$$

où la notation $h \rightarrow 0+$ signifie que h tend vers 0 dans le demi-plan $\Im(z) > 0$ (et inversement pour $h \rightarrow 0-$).

On peut donc dire que :

$$\text{Eu}^\pm(-a) = \lim_{h \rightarrow 0^\pm} \int_0^\infty \frac{e^{-t}}{1 + (-a + h)t} dt \quad (\text{IV.54})$$

Malheureusement, on ne peut pas passer à la limite dans l'intégrale car l'intégrale devient divergente pour $h = 0$: lorsque le point $z = -a + h$ tend vers $-a$, le point complexe $t = -1/z = 1/(a - h)$ où le dénominateur s'annule tend vers $1/a$ et devient donc réel positif.

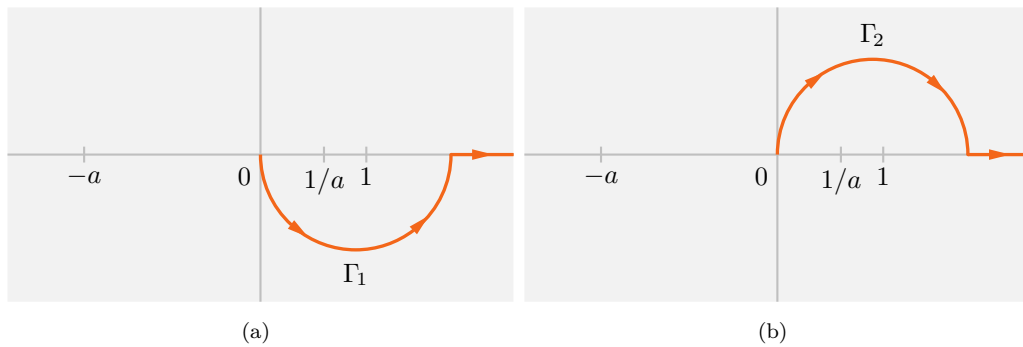


Figure IV.7 – Chemins homologiquement équivalents à $[0; \infty[$

Toutefois, la fonction à intégrer, $t \mapsto e^{-t}/(1 + zt)$, est analytique dans tout le plan excepté le point $-1/z$. L'intégrale qui apparaît dans (IV.54) est l'intégrale de cette fonction sur le chemin $[0; \infty[$ et la valeur de cette intégrale ne change pas si on la remplace par un chemin homologiquement équivalent. Or, ce qui empêche de passer à la limite sous le signe d'intégration est le fait que le point $t = -1/z = 1/(a - h)$ devient réel lorsque h tend vers zéro, c'est-à-dire que le dénominateur $1 + zt$ devient nul sur le chemin d'intégration $[0; \infty[$; si donc on remplace ce chemin par un chemin homologiquement équivalent, mais qui évite le point $1/a$, on n'aura plus cet inconvénient et on pourra passer à la limite dans l'intégrale sur un tel chemin. On voit sur les figures IV.7(a) et IV.7(b) que les chemins désignés respectivement par Γ_1 et Γ_2 remplissent ces conditions. Lorsque h est de partie imaginaire > 0 , $-1/z$ est aussi de partie imaginaire positive, donc au-dessus du point limite $1/a$. Dans ce cas, le chemin Γ_1 sur la figure IV.7(a), qui contourne la singularité

par en-dessous, reste homologiquement équivalent au chemin $[0; \infty[$; de même, Γ_2 sur la figure IV.7(b), lorsque h est de partie imaginaire négative.

On peut écrire cela sous une forme mathématiquement plus précise en minorant le dénominateur $1 + zt$ le long des chemins Γ_1 et Γ_2 . Il suffit d'écrire $|1 + zt| = |z| \cdot |t + \frac{1}{z}|$. En effet, $|t + \frac{1}{z}|$ est la distance (euclidienne) du point t sur le chemin au point $-1/z$; il suffit donc que la distance de Γ_1 (ou Γ_2) au point $-1/z$ reste constamment supérieure à un minimum α non nul, pour qu'ait lieu l'inégalité :

$$\left| \int_{\Gamma_{1,2}} \frac{e^{-t}}{1 + zt} dt \right| \leq \frac{1}{|z|\alpha} \int_0^\infty e^{-t(s)} |t'(s)| ds \quad (\text{IV.55})$$

où $t(s)$ est un paramétrage de Γ_1 (ou Γ_2). En vertu des théorèmes généraux sur le passage à la limite sous le signe d'intégration, cette inégalité montre clairement ce que nous voulions. Ainsi, on peut réécrire (IV.53) sous la forme :

$$\text{Eu}^+(-a) = \int_{\Gamma_1} \frac{e^{-t}}{1 - at} dt; \quad \text{Eu}^-(-a) = \int_{\Gamma_2} \frac{e^{-t}}{1 - at} dt \quad (\text{IV.56})$$

Le saut de discontinuité de la fonction $\text{Eu}(z)$ à travers la coupure $] -\infty; 0[$ au point $-a$ est alors égal à la différence des deux limites :

$$\text{saut} = \text{Eu}^+(-a) - \text{Eu}^-(-a) = \int_{\Gamma_1} \frac{e^{-t}}{1 - at} dt - \int_{\Gamma_2} \frac{e^{-t}}{1 - at} dt = \int_{\Gamma_1 - \Gamma_2} \frac{e^{-t}}{1 - at} dt \quad (\text{IV.57})$$

Or, le chemin $\Gamma_1 - \Gamma_2$ est formé d'une portion circulaire qui entoure le point $1/a$ et de deux portions rectilignes qui s'annulent mutuellement, puisqu'il s'agit d'un intervalle commun parcouru en des sens opposés. On en déduit que le saut au point $-a$ est égal au résidu de la fonction $e^{-t}/(1 - at)$ au point $t = 1/a$, multiplié par $2i\pi$.

Pour calculer ce résidu, il suffit de développer en série au voisinage de $t = 1/a$ et de chercher le coefficient du terme $1/(t - 1/a)$:

$$e^{-t} = e^{-1/a} \cdot e^{-(t-1/a)} = e^{-1/a} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left(t - \frac{1}{a}\right)^n \quad (\text{IV.58})$$

d'où :

$$\frac{e^{-t}}{1 - at} = -\frac{1}{a} e^{-1/a} \cdot \frac{e^{-(t-1/a)}}{t - \frac{1}{a}} = -\frac{\frac{1}{a} e^{-1/a}}{t - \frac{1}{a}} - \frac{1}{a} e^{-1/a} \sum_{n \geq 1} \frac{(-1)^n}{n!} \left(t - \frac{1}{a}\right)^{n-1} \quad (\text{IV.59})$$

La partie $\sum_{n \geq 1}$ est la partie régulière qui ne nous intéresse pas et on voit ainsi que le résidu est $-\frac{1}{a} e^{-1/a}$.

On a donc démontré que le saut de discontinuité de la fonction $\text{Eu}(z)$ au point $-a$ est égal à $-2i\pi \frac{1}{a} e^{-1/a}$. Ce saut est à comparer à celui du logarithme : d'après ce qui a été vu à la fin du chapitre III, la détermination $\ln_1$ du logarithme est elle aussi définie sur Ω_1 et :

$$\lim_{h \rightarrow 0+} \ln_1(-a + h) - \lim_{h \rightarrow 0-} \ln_1(-a + h) = 2i\pi \quad (\text{IV.60})$$

Par conséquent la fonction $\frac{1}{z} e^{1/z} \ln_1(z)$ aura au point $z = -a$ un saut de discontinuité égal à $-2i\pi \frac{1}{a} e^{-1/a}$, ce qui est exactement la même chose que pour la fonction $\text{Eu}(z)$.

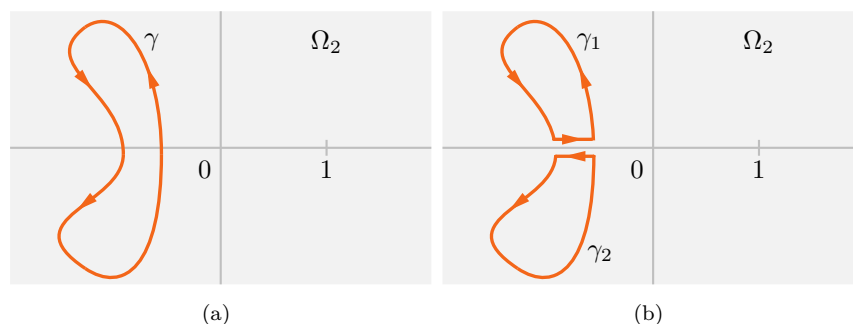


Figure IV.8

Cela entraîne que la différence $G(z) = \text{Eu}(z) - \frac{1}{z}e^{1/z}\ln_1(z)$, qui est aussi une fonction analytique sur Ω_1 , a une discontinuité nulle à travers la coupure $]-\infty; 0]$.

Il est facile de se convaincre que $G(z)$ est en fait analytique sur $\mathbb{C} - \{0\}$. Une preuve que $G(z)$ est analytique à travers la coupure serait par exemple que l'intégrale de $G(z) dz$ sur tout lacet γ dont l'intérieur est contenu dans $\mathbb{C} - \{0\}$ soit nulle. Si un lacet ne traverse pas la coupure $]-\infty; 0]$ c'est évident, donc seuls sont à considérer les lacets qui traversent la coupure comme sur la figure IV.8(a). Or, un tel lacet peut être décomposé en deux lacets γ_1 et γ_2 dont chacun entoure une région entièrement contenue dans Ω_1 (figure IV.8(b)), mais dont un segment est contenu dans la coupure. Sur ce segment on donne à $G(z)$ la valeur limite obtenue en faisant tendre z d'un seul côté (le dessus pour γ_1 et le dessous pour γ_2) vers le point du segment. Comme les deux limites coïncident, les intégrales sur ces deux portions rectilignes s'annulent mutuellement. L'intégrale de $G(z) dz$ sur γ est alors égale à la somme des intégrales sur γ_1 et γ_2 — si les deux limites n'étaient pas égales, la somme $\int_{\gamma_1} + \int_{\gamma_2}$ serait égale à $\int_{\gamma}$ plus l'intégrale du saut de discontinuité sur le segment. Comme les domaines délimités par γ_1 et γ_2 sont entièrement contenus dans Ω_1 , les intégrales correspondantes sont nulles et par conséquent aussi celle sur γ .

On a ainsi prouvé que $G(z)$ est analytique dans $\mathbb{C} - \{0\}$, c'est-à-dire dans la couronne $0 < |z| < \infty$. Par conséquent, $G(z)$ peut être développée en série de Laurent ; par contre $\text{Eu}(z)$ ne pouvait pas l'être.

Cherchons une expression simple de $G(z)$ qui permettra de trouver aisément ce développement. Puisque :

$$G(z) = \text{Eu}(z) - \frac{1}{z}e^{1/z}\ln_1(z) \quad (\text{IV.61})$$

remplaçons $\text{Eu}(z)$ par l'intégrale (IV.45) et supposons z réel positif, de sorte que :

$$\ln_1(z) = \ln(z) = \int_1^z \frac{1}{t} dt \quad (\text{IV.62})$$

L'hypothèse que z est réel positif permet aussi d'effectuer dans l'intégrale (IV.45) le changement de variable $s = \frac{1}{z} + t$, de sorte que :

$$\text{Eu}(z) = \int_0^\infty \frac{e^{-t}}{1+zt} dt = \frac{1}{z}e^{1/z} \int_{1/z}^\infty \frac{e^{-s}}{s} ds \quad (\text{IV.63})$$

On a aussi :

$$\ln_1(z) = \ln(z) = \int_1^z \frac{1}{s} ds = \int_{1/z}^1 \frac{1}{s} ds \quad (\text{IV.64})$$

En combinant (IV.61), (IV.63) et (IV.64), on obtient :

$$G(z) = \frac{1}{z} e^{1/z} \left[\int_{1/z}^{\infty} \frac{e^{-s}}{s} ds - \int_{1/z}^1 \frac{1}{s} ds \right] = \frac{1}{z} e^{1/z} \left[\int_{1/z}^1 \frac{e^{-s} - 1}{s} ds + \int_1^{\infty} \frac{e^{-s}}{s} ds \right] \quad (\text{IV.65})$$

Introduisons la constante :

$$\gamma = \int_0^1 \frac{1 - e^{-s}}{s} ds - \int_1^{\infty} \frac{e^{-s}}{s} ds \quad (\text{IV.66})$$

On remarque que les deux intégrales convergent (mais on ne sait pas les calculer par quadratures à partir des fonctions élémentaires). Si on les calcule numériquement, on obtient la valeur approchée $\gamma \simeq 0,577215$. Avec cette constante, on peut écrire (IV.65) sous la forme équivalente :

$$G(z) = \frac{1}{z} e^{1/z} \left[-\gamma - \int_0^{1/z} \frac{e^{-s} - 1}{s} ds \right] \quad (\text{IV.67})$$

On obtient aisément le développement en série de Laurent de l'expression entre crochets, en intégrant terme par terme la série entière de la fonction $[e^{-s} - 1]/s$ qui est $\sum_{n \geq 1} \frac{(-1)^n}{n!} s^{n-1}$; ce qui donne :

$$G(z) = \frac{1}{z} e^{1/z} \left[-\gamma - \sum_{n \geq 1} \frac{(-1)^n}{n n!} \frac{1}{z^n} \right] \quad (\text{IV.68})$$

On obtient alors la série de Laurent de $G(z)$ en faisant le produit de la série entre crochets et du développement de $\frac{1}{z} e^{1/z}$ en puissances de $1/z$. On voit immédiatement que le terme en $1/z$ dans ce produit a pour coefficient la constante $-\gamma$, qui est donc le résidu de $G(z)$ au point $z = 0$. On constate aussi que la série en puissances de $1/z$ aura une infinité de termes, c'est-à-dire que $z = 0$ est un point singulier essentiel.

La constante γ introduite ci-dessus mérite une mention spéciale. Elle est connue sous le nom de constante d'Euler et nous la rencontrerons à nouveau quand nous étudierons la fonction $\Gamma(x)$. On en obtient une autre expression en intégrant par parties les deux intégrales qui figurent au second membre de (IV.66) :

$$\begin{aligned} \int_0^1 \frac{e^{-s} - 1}{s} ds &= \int_0^1 [e^{-s} - 1] \frac{d}{ds} \ln(s) ds = [e^{-s} - 1] \ln(s) \Big|_0^1 + \int_0^1 e^{-s} \ln(s) ds \\ \int_1^{\infty} \frac{e^{-s}}{s} ds &= \int_1^{\infty} e^{-s} \frac{d}{ds} \ln(s) ds = e^{-s} \ln(s) \Big|_1^{\infty} + \int_1^{\infty} e^{-s} \ln(s) ds \end{aligned} \quad (\text{IV.69})$$

Ce qui conduit à :

$$-\gamma = \int_0^{\infty} e^{-s} \ln(s) ds \quad (\text{IV.70})$$

On montrera aussi en section V.4 que :

$$\gamma = \lim_{n \rightarrow \infty} 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n} - \ln(n) \quad (\text{IV.71})$$

Cette étude a montré que la fonction $\text{Eu}(z)$ a une singularité assez complexe en $z = 0$, puisqu'il y a à la fois la discontinuité logarithmique le long de $]-\infty; 0[$ et une singularité essentielle. Toutefois si on se restreint aux valeurs réelles et > 0 de z , la fonction est très régulière car sa limite quand z tend vers 0 (en restant > 0) est 1 ; cela se voit facilement à partir de (IV.45) : la fonction sous l'intégrale se majore uniformément en $z \geq 0$ par :

$$\frac{e^{-t}}{1+zt} \leq e^{-t} \quad (\text{IV.72})$$

donc en appliquant le théorème I.2, on voit que :

$$\lim_{z \rightarrow 0^+} \text{Eu}(z) = \int_0^\infty e^{-t} dt = 1 \quad (\text{IV.73})$$

On peut obtenir un développement limité (à n'importe quel ordre fini) en intégrant terme par terme l'identité :

$$\frac{e^{-t}}{1+zt} = \sum_{k=0}^{n-1} (-zt)^k e^{-t} + \frac{(-zt)^n e^{-t}}{1+zt} \quad (\text{IV.74})$$

ce qui donne :

$$\text{Eu}(z) = \sum_{k=0}^{n-1} (-z)^k \int_0^\infty t^k e^{-t} dt + (-z)^n \int_0^\infty \frac{t^n e^{-t}}{1+zt} dt \quad (\text{IV.75})$$

Or on sait que :

$$\int_0^\infty t^k e^{-t} dt = k! \quad (\text{IV.76})$$

d'où :

$$\text{Eu}(x) = \sum_{k=0}^{n-1} k! (-z)^k + R_n(z) \quad (\text{IV.77})$$

avec pour le reste la majoration (on suppose ici $z > 0$) :

$$|R_n(z)| = \left| (-z)^n \int_0^\infty \frac{t^n e^{-t}}{1+zt} dt \right| \leq z^n \int_0^\infty \frac{t^n e^{-t}}{1+zt} dt \leq z^n \int_0^\infty t^n e^{-t} dt = n! z^n \quad (\text{IV.78})$$

Notez bien que cette majoration du reste n'est valable que pour $z \geq 0$; pour $z < 0$ ou pour z complexe elle serait fautive.

D'autre part, on ne peut pas faire tendre n vers l'infini pour transformer (IV.77) en une série entière car cette série serait divergente pour tout $z > 0$: son rayon de convergence serait nul (ce qui est logique puisque la fonction $\text{Eu}(z)$ n'est pas analytique en $z = 0$).

Cependant, quoique la série *infinie* soit divergente, le développement *limité* (IV.77) permet de calculer des valeurs approchées de $\text{Eu}(z)$ pour $z > 0$. En effet pour $0 < z < 1$, la quantité $n! z^n$ qui majore le reste dans (IV.78) est minimum lorsque n est égal à la partie entière de $1/z$ et ce minimum est alors de l'ordre de $e^{-n} \sqrt{2\pi n}$.

Exercice IV.1 Le vérifier à l'aide de la formule de Stirling. ■

Ainsi pour $z = 0,1$, le développement (IV.77) jusqu'à l'ordre $n = 10$ fournit une valeur approchée de $\text{Eu}(0,1)$ à quatre décimales et pour $z = 0,05$, le développement à l'ordre $n = 20$ fournit une valeur approchée à huit décimales.

IV.6 Fonctions puissance non entière

Une catégorie de fonctions très importante pour les applications est celle des fonctions z^α , α étant un nombre réel ou complexe, non nécessairement entier. Ces fonctions sont multiformes, c'est-à-dire admettent plusieurs déterminations que la notation z^α ne précise pas ; c'est pourquoi cette notation est à proscrire, excepté si on la place dans un contexte qui précise la détermination. Ici, nous ferons comme pour le logarithme complexe, nous distinguerons les diverses déterminations par un indice, en écrivant $[z^\alpha]_j$.

On peut définir ces fonctions en les ramenant à un logarithme ; il suffit de poser :

$$[z^\alpha]_j = e^{\alpha \ln_j(z)} \quad (\text{IV.79})$$

Ainsi, la détermination du logarithme étant fixée, celle de $[z^\alpha]_j$ l'est aussi. Comme les diverses déterminations du logarithme diffèrent toujours d'un multiple entier de $2i\pi$, on voit que si α est entier, les différentes déterminations de $[z^\alpha]_j$ sont toutes identiques et on retrouve alors la fonction $z^\alpha = z \times z \times z \times \cdots \times z$. Si α est rationnel, de la forme irréductible p/q , il arrivera que deux déterminations différentes du logarithme conduisent à la même détermination de $[z^\alpha]_j$: il suffit pour cela que les deux déterminations du logarithme diffèrent d'un multiple de $2i\pi q$.

Les fonctions $[z^\alpha]_j$ sont, avec $\exp$, $\sin$..., les plus fréquemment utilisées en analyse complexe. Il est donc indispensable de savoir les manipuler.

Se ramener à un logarithme par (IV.79) est généralement peu commode. La meilleure approche est d'utiliser la représentation trigonométrique des nombres complexes. En effet, si $z = re^{i\theta}$, on aura toujours $[z^\alpha]_j = r^\alpha e^{i\alpha\theta}$. Cette expression sera toujours valable et les diverses déterminations ne différeront que par les valeurs respectives autorisées pour θ .

Les deux déterminations les plus courantes sont :

- $[z^\alpha]_1$, définie sur $\Omega_1 = \mathbb{C} -]-\infty; 0[$, où l'on représente z par $re^{i\theta}$ avec $-\pi < \theta < +\pi$;
- $[z^\alpha]_2$, définie sur $\Omega_2 = \mathbb{C} - [0; \infty[$, où l'on représente z par $re^{i\theta}$ avec $0 < \theta < 2\pi$.

Il n'y a pas de détermination *officielle* : pour un problème donné, il faut trouver celle qui convient le mieux et qui n'est pas forcément l'une des deux précédentes.

Notez bien qu'on pourrait aussi représenter les éléments de Ω_2 par $z = re^{i\theta}$ avec $2\pi < \theta < 4\pi$, ou par $z = re^{i\theta}$ avec $14\pi < \theta < 16\pi$. Cela conduirait à d'autres déterminations, disons $[z^\alpha]_{101}$ et $[z^\alpha]_{107}$; en prenant $2\pi < \theta < 4\pi$ on aurait :

$$[z^\alpha]_{101} = r^\alpha e^{i\alpha\theta} = r^\alpha e^{i\alpha(\theta-2\pi)} e^{+i\alpha \cdot 2\pi} = [z^\alpha]_1 e^{+i\alpha \cdot 2\pi} \quad (\text{IV.80})$$

et en prenant $14\pi < \theta < 16\pi$:

$$[z^\alpha]_{107} = r^\alpha e^{i\alpha\theta} = r^\alpha e^{i\alpha(\theta-14\pi)} e^{+i\alpha \cdot 14\pi} = [z^\alpha]_1 e^{+i\alpha \cdot 14\pi} \quad (\text{IV.81})$$

Ainsi, bien qu'elles soient elles aussi définies sur le domaine Ω_2 , $[z^\alpha]_{101}$ et $[z^\alpha]_{107}$ diffèrent de $[z^\alpha]_2$ par un facteur multiplicatif, qui n'est égal à 1 que si α est entier.

On devine les difficultés qu'on rencontrerait en ne respectant pas scrupuleusement les spécificités de chaque détermination.

La détermination $[z^\alpha]_1$ est définie pour z , réel strictement positif, où elle se réduit à la fonction usuelle z^α . La détermination $[z^\alpha]_2$ n'est pas définie pour z réel strictement positif.

Si on prend z réel strictement supérieur à 1, $[z^\alpha]_1 = z^\alpha$ et $[(z-1)^\alpha]_1 = (z-1)^\alpha$, de sorte que $z^\alpha \cdot (z-1)^\alpha = [z(z-1)]^\alpha = [[z(z-1)]^\alpha]_1$. Cela résulte du fait connu que le produit de puissances est la puissance du produit : $a^\alpha \cdot b^\alpha = (a \cdot b)^\alpha$, qui est valable pour a et b réels > 0 . La relation $[z^\alpha]_1 \cdot [(z-1)^\alpha]_1 = [[z(z-1)]^\alpha]_1$ est donc vraie pour z réel > 1 , mais elle n'est pas vraie pour n'importe quel z complexe. Prenons par exemple $\alpha = \frac{1}{2}$, la fonction $[z^\alpha]_1$ devient alors $\sqrt{z}^1$. Soit $z = re^{i(\pi-\varepsilon)}$, avec ε petit et > 0 . Cela signifie que z est tout près de l'axe réel (juste au-dessus), mais du côté négatif : $z \simeq -r$. Alors $z-1$ aura un module $\rho \simeq r+1$ et un argument ξ également très proche de π ($\xi = \pi - \eta$). On constate que :

$$\begin{aligned}\sqrt{z}^1 &= \sqrt{r} e^{i(\pi-\varepsilon)/2} = i\sqrt{r} e^{-i\varepsilon/2} \\ \sqrt{z-1}^1 &= \sqrt{\rho} e^{i(\pi-\eta)/2} = i\sqrt{\rho} e^{-i\eta/2} \\ \sqrt{z}^1 \cdot \sqrt{z-1}^1 &= -\sqrt{r\rho} e^{-i(\varepsilon+\eta)/2}\end{aligned}\tag{IV.82}$$

ce qui est proche de $-\sqrt{r\rho}$ puisque ε et η sont petits. Mais :

$$z(z-1) = r\rho e^{i2\pi-i\varepsilon-i\eta} = r\rho e^{-i\varepsilon-i\eta}\tag{IV.83}$$

et comme $-\pi < -\varepsilon - \eta < +\pi$ (mais non $-\pi < 2\pi - \varepsilon - \eta < +\pi$), on a :

$$\sqrt{z(z-1)}^1 = \sqrt{r\rho} e^{-i(\varepsilon+\eta)/2}\tag{IV.84}$$

ce qui est proche de $+\sqrt{r\rho}$.

Ce qui a causé la différence entre les deux expressions est que $\sqrt{z(z-1)}^1$ est égal à $\sqrt{r\rho} e^{-i(\varepsilon+\eta)/2}$ et non à $\sqrt{r\rho} e^{i(2\pi-\varepsilon-\eta)/2}$. Si on avait eu $-\pi < 2\pi - \varepsilon - \eta < +\pi$ (c'est-à-dire $\pi < \varepsilon + \eta < 3\pi$, c'est-à-dire encore $-\pi < \theta + \xi < +\pi$), les deux résultats auraient été égaux. On peut donc conclure que l'égalité :

$$\sqrt{z(z-1)}^1 = \sqrt{z}^1 \cdot \sqrt{z-1}^1\tag{IV.85}$$

n'a lieu que si $-\pi < \theta + \xi < +\pi$. Un raisonnement géométrique élémentaire permet de voir que cette condition équivaut à $\Re(z) > \frac{1}{2}$. En conclusion :

- si $\Re(z) > \frac{1}{2}$, $\sqrt{z(z-1)}^1 = \sqrt{z}^1 \cdot \sqrt{z-1}^1$
- si $\Re(z) < \frac{1}{2}$, $\sqrt{z(z-1)}^1 = -\sqrt{z}^1 \cdot \sqrt{z-1}^1$
- si $\Re(z) = \frac{1}{2}$, $z(z-1)$ est réel < 0 et $\sqrt{z(z-1)}^1$ n'est pas définie.

Il ne faut donc jamais appliquer la règle $[(ab)^\alpha]_j = [a^\alpha]_j \cdot [b^\alpha]_j$ dans le domaine complexe.

Afin de montrer les méthodes qu'il convient d'employer avec les fonctions puissance, voyons l'exemple suivant :

$$F(z) = [z^\alpha]_1 \cdot [(z-1)^\beta]_1\tag{IV.86}$$

C'est donc le produit de deux fonctions puissance. Pour ne pas se tromper dans les valeurs, la meilleure méthode est de représenter z et $z-1$ sous forme trigonométrique, soit $z = re^{i\theta}$ et $z-1 = \rho e^{i\xi}$. $[z^\alpha]_1$ n'est pas définie pour z réel < 0 et $[(z-1)^\alpha]_1$ n'est pas définie pour z réel < 1 ; donc le produit $F(z)$ des deux est non défini pour z réel < 1 .

Cherchons la discontinuité de $F(z)$ à travers cette coupure. Le plus commode pour comprendre est de procéder en représentation géométrique : le vecteur-image de z est le

vecteur d'origine 0 et d'extrémité z , le vecteur-image de $z - 1$ est le vecteur d'origine 1 et d'extrémité z , comme illustré sur la figure IV.9(a). Lorsque z tend vers un point de l'intervalle $]-\infty; 1]$, il y a deux cas de figure : lorsque ce point est sur $]-\infty; 0[$, sur la figure IV.9(b) et lorsqu'il est sur $]0; 1[$ sur la figure IV.9(c).

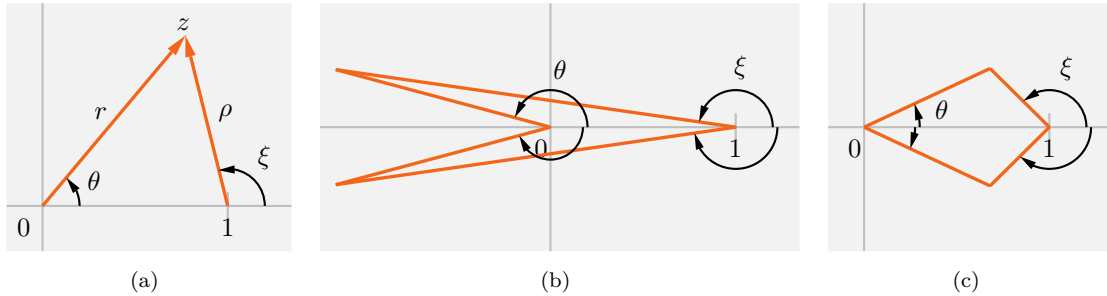


Figure IV.9 – Passage à la limite dans les expressions $[z^\alpha]_1 \cdot [(z-1)^\beta]_1$

La figure IV.9(b) montre vers quelle valeur limite tendent les angles θ et ξ (ainsi que les modules r et ρ) lorsque z tend vers un point $-a$ ($a > 0$) de l'intervalle $]-\infty; 0[$:

- si $z \rightarrow -a$ par au-dessus, $\theta \rightarrow +\pi$ et $\xi \rightarrow +\pi$;
- si $z \rightarrow -a$ par en-dessous, $\theta \rightarrow -\pi$ et $\xi \rightarrow -\pi$;

dans les deux cas, $r \rightarrow a^\alpha$ et $\rho \rightarrow (1+a)^\beta$. Par conséquent :

- si $z \rightarrow -a$ par au-dessus, $F(z) \rightarrow a^\alpha(1+a)^\beta e^{i(\alpha+\beta)\pi}$;
- si $z \rightarrow -a$ par en-dessous, $F(z) \rightarrow a^\alpha(1+a)^\beta e^{-i(\alpha+\beta)\pi}$.

On constate que si $\alpha + \beta$ est un entier algébrique, ces deux limites sont égales, c'est-à-dire que dans ce cas il n'y a pas de discontinuité et la fonction $F(z)$ se prolonge en une fonction analytique dans $\mathbb{C} - [0; 1]$, comme nous l'avons vu à la section IV.4 pour la fonction $G(z) = \text{Eu}(z) - \frac{1}{z} e^{1/z} \ln_1(z)$.

De même, la figure IV.9(c) montre vers quelles valeurs limite tendent les angles θ et ξ (ainsi que les modules r et ρ) lorsque z tend vers un point a du segment $]0; 1[$:

- si $z \rightarrow a$ par au-dessus, $\theta \rightarrow 0^+$ et $\xi \rightarrow +\pi$;
- si $z \rightarrow a$ par en-dessous, $\theta \rightarrow 0^-$ et $\xi \rightarrow -\pi$;

dans les deux cas, $r \rightarrow a^\alpha$ et $\rho \rightarrow (1-a)^\beta$. Par conséquent :

- si $z \rightarrow a$ par au-dessus, $F(z) \rightarrow a^\alpha(1-a)^\beta e^{i\beta\pi}$;
- si $z \rightarrow a$ par en-dessous, $F(z) \rightarrow a^\alpha(1-a)^\beta e^{-i\beta\pi}$.

Ces deux limites sont distinctes, sauf évidemment si β est entier.

Lorsque $\alpha + \beta = k$ est entier, la fonction $F(z)$ est analytique dans $\mathbb{C} - [0; 1]$; elle est donc à plus forte raison analytique dans la couronne $|z| > 1$. Par conséquent, d'après le théorème IV.1, elle peut être développée en série de Laurent. Nous nous proposons de trouver cette série et en particulier, le résidu.

Lorsque z est réel > 1 , $[z^\alpha]_1 = z^\alpha$ et $[(z-1)^\beta]_1 = (z-1)^\beta$, donc :

$$F(z) = z^\alpha \cdot (z-1)^\beta = z^{(\alpha+\beta)} \left[1 - \frac{1}{z}\right]^\beta = z^k \left[1 - \frac{1}{z}\right]^\beta \quad (\text{IV.87})$$

On a utilisé ici la relation $(ab)^\beta = a^\beta b^\beta$ (valable pour $a > 0$ et $b > 0$ et non pour a et b complexes) : on a en effet remplacé $(z-1)^\beta$ par $z^\beta \cdot (1 - 1/z)^\beta$ et comme z a été supposé

réel > 1 , z et $1 - 1/z$ sont tous deux réels > 0 .

La fonction $X \mapsto (1 - X)^\beta$ est développable en série de Taylor pour $X \in]-1; 1[$:

$$(1 - X)^\beta = \sum_{n=0}^{\infty} a_n X^n \quad \text{avec} \quad a_n = (-1)^n \frac{\beta(\beta-1)(\beta-2)\cdots(\beta-n+1)}{n!} \quad (\text{IV.88})$$

On suppose ce résultat connu (formule du binôme de Newton étendue au cas d'un exposant non entier).

Le rayon de convergence de la série entière $\sum a_n X^n$ est égal à 1, donc la somme de cette série définit pour X complexe une fonction $H(X)$ analytique dans le disque $\{X \in \mathbb{C} \mid |X| < 1\}$. Par conséquent la fonction $h(z) = H(1/z)$ est analytique dans la couronne $D = \{z \in \mathbb{C} \mid |z| > 1\}$: c'est la composée de deux fonctions analytiques, $z \mapsto 1/z$ et $X \mapsto H(X)$. On a évidemment :

$$h(z) = \sum_{n=0}^{\infty} a_n z^{-n} \quad (\text{IV.89})$$

D'après (IV.87), on conclut que :

$$F(z) = z^k h(z) = \sum_{n=0}^{\infty} a_n z^{k-n} \quad (\text{IV.90})$$

pour $z \in]1; \infty[$.

Et voici le point le plus important de l'argumentation : d'après (IV.87), notre fonction $F(z)$ est égale à $z^k h(z)$ pour z réel > 1 ; or l'intervalle $]1; \infty[$ est un ensemble non discret. Donc $F(z)$ et $z^k h(z)$ sont égales dans tout le domaine complexe $|z| > 1$ (théorème des zéros isolés ou principe du prolongement analytique). En conclusion, la série (IV.90) est la série de Laurent de $F(z)$. En particulier, le résidu de $F(z)$ est :

$$a_{k+1} = (-1)^{k+1} \frac{\beta(\beta-1)(\beta-2)\cdots(\beta-k)}{(k+1)!} \quad (\text{IV.91})$$

On en déduit que, si γ est n'importe quel lacet (simple et orienté dans le sens direct) entourant le segment $[0; 1]$, l'intégrale de $F(z) dz$ sur ce lacet vaut $2i\pi a_{k+1}$. Lorsqu'un tel lacet est complètement aplati sur le segment, l'intégrale devient celle de 0 à 1 des valeurs limite de $F(z)$ par dessous, plus celle de 1 à 0 des valeurs limite de $F(z)$ par dessus, soit :

$$\begin{aligned} 2i\pi a_{k+1} &= \int_0^1 a^\alpha (1-a)^\beta e^{-i\beta\pi} da + \int_1^0 a^\alpha (1-a)^\beta e^{+i\beta\pi} da \\ &= \int_0^1 a^\alpha (1-a)^\beta [e^{-i\beta\pi} - e^{+i\beta\pi}] da \\ &= -2i \sin(\beta\pi) \int_0^1 a^\alpha (1-a)^\beta da \end{aligned} \quad (\text{IV.92})$$

On obtient ainsi la formule suivante, valable pour $\alpha + \beta = k$ entier :

$$\int_0^1 a^\alpha (1-a)^\beta da = -\frac{\pi}{\sin(\beta\pi)} a_{k+1} \quad (\text{IV.93})$$

Un cas particulièrement simple, et donc très intéressant, est celui où $k = -1$; alors $a_{k+1} = a_0 = 1$:

$$\int_0^1 a^\alpha (1-a)^{-1-\alpha} da = -\frac{\pi}{\sin(\alpha\pi)} \quad (\text{IV.94})$$

Bien entendu, il est impossible de calculer ces intégrales par des primitives connues.



V — Fonctions eulériennes

V.1 Présentation

Définition V.1 On appelle *intégrale eulérienne de première espèce*, ou *fonction Beta*, l'intégrale (dépendant des deux paramètres x et y) :

$$\beta(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt \quad (\text{V.1})$$

Cette intégrale diverge pour $x \leq 0$ (en $t = 0$) et pour $y \leq 0$ (en $t = 1$). La fonction $\beta(x, y)$ est donc définie *a priori* pour x et y positifs. Toutefois, elle possède un prolongement analytique au-delà comme nous verrons dans la section V.2.

Le changement de variable $t \mapsto 1-t$ dans l'intégrale montre que $\beta(x, y) = \beta(y, x)$. Le changement de variable $t = \sin^2 \theta$ conduit à une expression équivalente :

$$\beta(x, y) = 2 \int_0^{\frac{\pi}{2}} \sin^{2x-1} \theta \cos^{2y-1} \theta d\theta \quad (\text{V.2})$$

Définition V.2 On appelle *intégrale eulérienne de deuxième espèce*, ou *fonction Gamma*, l'intégrale (dépendant du paramètre x) :

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \quad (\text{V.3})$$

Cette intégrale diverge pour $x \leq 0$ (en $t = 0$). Comme pour $\beta(x, y)$, $\Gamma(x)$ n'est définie *a priori* que pour $x > 0$, mais se prolonge analytiquement.

L'intégration par parties :

$$\int_0^\infty t^{x-1} e^{-t} dt = \frac{t^x}{x} e^{-t} \Big|_0^\infty + \int_0^\infty \frac{t^x}{x} e^{-t} dt \quad (\text{V.4})$$

montre que :

$$\Gamma(x) = \frac{1}{x} \Gamma(x+1) \quad \Longleftrightarrow \quad \Gamma(x+1) = x\Gamma(x) \quad (\text{V.5})$$

On en déduit immédiatement que pour $x = n$ entier, $\Gamma(n+1) = n!$. Le changement de variable $t = u^2$ conduit à une autre expression équivalente de $\Gamma(x)$:

$$\Gamma(x) = 2 \int_0^\infty u^{2x-1} e^{-u^2} du \quad (\text{V.6})$$

Théorème V.1 Entre les fonctions eulériennes de première et de deuxième espèce, on a la relation suivante :

$$\beta(x, y) \cdot \Gamma(x+y) = \Gamma(x) \cdot \Gamma(y) \quad (\text{V.7})$$

Preuve L'intégrale double :

$$\int_0^\infty \int_0^\infty u^{2x-1} v^{2y-1} e^{-u^2-v^2} du dv \quad (\text{V.8})$$

d'une part, se factorise en $\Gamma(x) \cdot \Gamma(y)$, d'après (V.6), et d'autre part s'écrit :

$$\int_0^\infty \int_0^{\pi/2} r^{2(x+y)-1} \sin^{2x-1} \theta \cos^{2y-1} \theta e^{-r^2} d\theta dr \quad (\text{V.9})$$

en coordonnées polaires $u = r \sin \theta$, $v = r \cos \theta$, ce qui se factorise en $\Gamma(x+y) \cdot \beta(x, y)$ d'après (V.2) et (V.6). •

Théorème V.2 — Formule des compléments pour $0 < x < 1$ on a :

$$\Gamma(x) \cdot \Gamma(1-x) = \frac{\pi}{\sin \pi x} \quad (\text{V.10})$$

Preuve On utilise la formule du résidu appliquée à la fonction $[z^{-x}]_1 \cdot [z - 1^{x-1}]_1$. Le détail du calcul ayant déjà été vu, il suffit de reprendre la formule (IV.94) avec $x = -\alpha$. •

V.2 Prolongements analytiques

Théorème V.3 La fonction $\Gamma(x)$ se prolonge en une fonction analytique sur le domaine D , égal au plan complexe $\mathbb{C}$ privé des entiers négatifs et nul. Les points $z = -n$, où n est un entier positif ou nul, sont des pôles simples, et le résidu de la fonction au point $z = -n$ est $(-1)^n/n!$.

Preuve pour x réel > 0 , on peut écrire :

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt = \int_0^1 t^{x-1} e^{-t} dt + \int_1^\infty t^{x-1} e^{-t} dt \quad (\text{V.11})$$

Posons :

$$\Phi(x) = \int_0^1 t^{x-1} e^{-t} dt \quad \text{et} \quad \Psi(x) = \int_1^\infty t^{x-1} e^{-t} dt \quad (\text{V.12})$$

Dans la première intégrale, on peut remplacer e^{-t} par le développement en série $\sum (-1)^n \frac{1}{n!} t^n$ et intégrer terme par terme, puisque la série est uniformément convergente sur l'intervalle $[0; 1]$, d'où :

$$\Phi(x) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{1}{x+n} \quad (\text{V.13})$$

Considérons maintenant les fonctions *de la variable complexe* z :

$$\Phi(z) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{1}{z+n} \quad \text{et} \quad \Psi(z) = \int_1^\infty t^{z-1} e^{-t} dt = \int_1^\infty e^{-t+(z-1)\ln(t)} dt \quad (\text{V.14})$$

Il est clair que pour tout z autre qu'un entier négatif ou nul, la série qui définit $\Phi(z)$ converge (grâce au coefficient $1/n!$ qui décroît très rapidement) ; de même la série :

$$\Upsilon(z) = \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{1}{(z+n)^2} \quad (\text{V.15})$$

On va montrer que les fonctions $\Phi(z)$ et $\Psi(z)$ sont toutes deux analytiques, la première en dehors des entiers ≤ 0 , la seconde partout. Pour cela, écrivons :

$$\begin{aligned} \Phi(z+h) - \Phi(z) + h\Upsilon(z) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \left[\frac{1}{z+h+n} - \frac{1}{z+n} + \frac{h}{(z+n)^2} \right] \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{h^2}{(z+n+h)(z+n)^2} \end{aligned} \quad (\text{V.16})$$

Si $z \in D$, soit ε la distance de z au point singulier le plus proche ($\varepsilon = \inf_n \{|z+n|\}$) ; si on fait tendre h vers zéro, on peut supposer $|h| < \frac{1}{2}\varepsilon$ d'où $|z+n+h| > \frac{1}{2}\varepsilon$; par conséquent il résulte de (V.16) que :

$$\left| \frac{\Phi(z+h) - \Phi(z)}{h} + \Upsilon(z) \right| \leq |h| \sum_{n=0}^{\infty} \frac{1}{n!} \frac{2}{\varepsilon^3} \quad (\text{V.17})$$

ce qui prouve que $[\Phi(z+h) - \Phi(z)]/h$ a une limite lorsque h tend vers zéro, qui est $-\Upsilon(z)$. Ainsi, $\Phi(z)$ est analytique dans D . Du coup, cela implique que $\Upsilon(z)$, qui est sa dérivée, est aussi analytique dans D .

Montrons maintenant que $\Psi(z)$ est analytique dans $\mathbb{C}$ tout entier. De l'identité connue :

$$e^w - 1 - w = w^2 e^w \int_0^1 t e^{-tw} dt \quad (\text{V.18})$$

que l'on obtient en intégrant par parties, valable pour tout nombre complexe w , on déduit, par l'inégalité de la moyenne, que :

$$\begin{aligned} |e^w - 1 - w| &\leq |w|^2 e^{\Re(w)} & \text{si } \Re(w) \geq 0 \\ |e^w - 1 - w| &\leq |w|^2 & \text{si } \Re(w) \leq 0 \end{aligned} \quad (\text{V.19})$$

Par conséquent, d'après (V.14) :

$$\begin{aligned} \left| \frac{\Psi(z+h) - \Psi(z)}{h} - \int_1^\infty \ln(t) e^{-t+(z-1)\ln(t)} dt \right| &= \left| \int_1^\infty e^{-t+(z-1)\ln(t)} \left[\frac{e^{h\ln(t)} - 1 - h\ln(t)}{h} \right] dt \right| \\ &\leq \int_1^\infty \left| \frac{e^{h\ln(t)} - 1}{h} - \ln(t) \right| e^{-t+(\Re(z)-1)\ln(t)} dt \\ &\leq \int_1^\infty |h| [\ln(t)]^2 e^{\Re(h)\ln(t)} e^{-t+(\Re(z)-1)\ln(t)} dt \\ &= |h| \int_1^\infty [\ln(t)]^2 t^{\Re(z+h)-1} e^{-t} dt \end{aligned} \quad (\text{V.20})$$

On a ici pris le cas où $\Re(h) \geq 0$, mais le cas où $\Re(h) \leq 0$ est analogue. L'intégrale qui figure au dernier membre est convergente et reste bornée quand h reste borné, donc $(\Psi(z+h) - \Psi(z))/h$ tend bien vers une limite lorsque h tend vers zéro.

En ce qui concerne les points singuliers, on remarque qu'au voisinage point $z = -n_0$ (n_0 entier ≥ 0), la fonction

$$G(z) = \sum_{\substack{n \neq n_0 \\ n \geq 0}} \frac{(-1)^n}{n!} \frac{1}{z+n} + \Psi(z) = \Phi(z) - \frac{(-1)^{n_0}}{n_0!} \frac{1}{z+n_0} + \Psi(z) \quad (\text{V.21})$$

est analytique; elle est en fait analytique dans tout le disque $|z+n_0| < 1$, puisque la fonction $\Phi(z)$ n'avait pas d'autre point singulier que $-n_0$ dans ce disque et par conséquent sa série entière en puissance de $z+n_0$ a pour rayon de convergence 1. On peut donc écrire :

$$\Gamma(z) = G(z) + \frac{(-1)^{n_0}}{n_0!} \frac{1}{z+n_0} \quad (\text{V.22})$$

ce qui (en imaginant que $G(z)$ est remplacée par sa série entière en puissances de $z+n_0$) est le développement de Laurent de $\Gamma(z)$ autour du point $-n_0$. On voit que le résidu est bien celui annoncé. •

Pour prouver que $\Psi(z)$ est analytique dans $\mathbb{C}$ tout entier, on aurait aussi pu montrer que ses dérivées partielles sont continues et que son intégrale sur n'importe quel lacet γ est nulle; or cela est aisé, car l'intégrale double $\int_{\gamma} \int_1^{\infty} e^{-t+(z-1)\ln(t)} dt dz$ est toujours absolument convergente (il faut considérer l'intégrale sur γ comme paramétrée par $z(s)$, $0 \leq s \leq 2\pi$). En effet, si z parcourt un domaine $\mathcal{D}_A$ de la forme $\mathcal{D}_A = \{z \in \mathbb{C} \mid \Re(z) - 1 \leq A\}$, on peut majorer la fonction à intégrer :

$$|e^{-t+[z(s)-1]\ln(t)} z'(s)| \leq t^{A-1} |z'(s)| e^{-t} \quad (\text{V.23})$$

et la fonction du membre de droite est évidemment intégrable sur $[0; 2\pi] \times [1; \infty[$.

On en déduit donc que :

$$\int_{\gamma} \Psi(z) dz = \int_1^{\infty} \left[\int_{\gamma} e^{-t+(z-1)\ln(t)} dz \right] dt \quad (\text{V.24})$$

ce qui est nul, puisque pour tout t , $z \mapsto e^{-t+(z-1)\ln(t)}$ est analytique. Pour vérifier la continuité des dérivées partielles, on utilise les théorèmes de passage à la limite sous le signe $\int$ du chapitre I.

On peut aussi utiliser le théorème suivant, qui est une variante du théorème I.8 :

Théorème V.4 Soit $f(z, t)$ une fonction telle que :

- pour tout A , il existe une fonction $F_A(t) \geq 0$ telle que l'intégrale de $F_A(t) dt$ converge et telle que l'inégalité $|f(z, t)| \leq F_A(t)$ ait lieu uniformément pour $|z| \leq A$ et t dans l'intervalle d'intégration ;
- pour tout t dans l'intervalle d'intégration, $z \mapsto f(z, t)$ est analytique dans le domaine D (indépendant de t).

Alors $z \mapsto \int f(z, t) dt$ est analytique dans D .

Ce théorème, appliqué aux intégrales (V.1) et (V.3) montre d'emblée que les fonctions Beta et Gamma sont analytiques pour $\Re(x) > 0$, $\Re(y) > 0$. De toute façon, comme le montre le théorème V.3, elles sont analytiques dans un domaine plus étendu encore.

Ainsi la fonction $\Gamma(z)$ est analytique partout excepté aux points entiers positifs ou nul. L'intégrale (V.3) définissait une fonction analytique *a priori* pour $\Re(z) > 0$ seulement, mais on voit que la fonction se prolonge à un domaine où l'intégrale (V.3) ne converge pas.

La relation $\Gamma(z+1) = z\Gamma(z)$ que nous avons démontrée pour $z > 0$, ainsi que la relation $\Gamma(z)\Gamma(1-z) = \pi/\sin \pi z$ que nous avons démontrée pour $0 < z < 1$, sont donc vraies dans la totalité des domaines d'analyticité respectifs ($\mathbb{C} - \{0, -1, -2, -3, \dots\}$ pour la première et $\mathbb{C} - \mathbb{Z}$ pour la seconde) en vertu du théorème des zéros isolés III.6.4. C'est le principe du prolongement analytique.

On a aussi un prolongement analytique pour $\beta(w, z)$: *a priori*, l'intégrale (V.1) fournit une fonction analytique en w ou en z pour $\Re(w) > 0$ et $\Re(z) > 0$; mais en posant $\beta(w, z) = \Gamma(z)\Gamma(w)/\Gamma(w+z)$, on obtient une fonction analytique dans $D \times D$.

V.3 Formule d'Euler

Revenons encore au cas où la variable est réelle :

Théorème V.5 Pour x strictement positif, on a :

$$\Gamma(x) = \lim_{n \rightarrow +\infty} \int_0^n t^{x-1} \left(1 - \frac{t}{n}\right)^n dt = \lim_{n \rightarrow +\infty} n^x \beta(x, n+1) \quad (\text{V.25})$$

Preuve On considère sur l'intervalle $[0; n]$, la fonction croissante :

$$\varphi_n(t) = 1 - e^t \left(1 - \frac{t}{n}\right)^n \Rightarrow \varphi'_n(t) = \frac{t}{n} e^t \left(1 - \frac{t}{n}\right)^{n-1} \geq 0 \quad (\text{V.26})$$

Par ailleurs, il est bien connu que pour tout t fixé, $(1 - t/n)^n$ tend vers e^{-t} quand n tend vers l'infini, donc pour tout t , $\varphi_n(t)$ tend vers zéro. Décomposons :

$$\begin{aligned} \Gamma(x) - \int_0^n t^{x-1} \left(1 - \frac{t}{n}\right)^n dt &= \int_0^\infty t^{x-1} e^{-t} dt - \int_0^n t^{x-1} \left(1 - \frac{t}{n}\right)^n dt \\ &= \int_0^A t^{x-1} e^{-t} \varphi_n(t) dt + \int_A^n t^{x-1} e^{-t} \varphi_n(t) dt + \int_n^\infty t^{x-1} e^{-t} dt \end{aligned} \quad (\text{V.27})$$

où A est une constante arbitraire (mais $n > A$). Il faut montrer que pour tout ε , il existe un $N(\varepsilon)$ tel que pour $n \geq N(\varepsilon)$, cette expression soit plus petite que ε .

Puisque la fonction $\varphi_n(t)$ est croissante, on peut dire que sur $[0; A]$ elle est majorée par $\varphi_n(A)$, sur $[A; n]$ elle est majorée par $\varphi_n(n) = 1$. Donc d'après l'inégalité de la moyenne, la première intégrale de (V.36) se majore comme suit :

$$\left| \int_0^A t^{x-1} e^{-t} \varphi_n(t) dt \right| \leq \varphi_n(A) \int_0^A t^{x-1} e^{-t} dt \leq \varphi_n(A) \int_0^\infty t^{x-1} e^{-t} dt = \varphi_n(A) \Gamma(x) \quad (\text{V.28})$$

La seconde se majore de façon similaire :

$$\left| \int_A^n t^{x-1} e^{-t} \varphi_n(t) dt \right| \leq \varphi_n(n) \int_A^n t^{x-1} e^{-t} dt = \int_A^n t^{x-1} e^{-t} dt \quad (\text{V.29})$$

On a donc :

$$\begin{aligned} \left| \Gamma(x) - \int_0^n t^{x-1} \left(1 - \frac{t}{n}\right)^n dt \right| &\leq \varphi_n(A) \Gamma(x) + \int_A^n t^{x-1} e^{-t} dt + \int_n^\infty t^{x-1} e^{-t} dt \\ &= \varphi_n(A) \Gamma(x) + \int_A^\infty t^{x-1} e^{-t} dt \end{aligned} \quad (\text{V.30})$$

Comme l'intégrale $\int_0^\infty t^{x-1} e^{-t} dt$ converge, on peut choisir la constante A telle que $\int_A^\infty t^{x-1} e^{-t} dt \leq \frac{\varepsilon}{2}$. Cette constante *étant ainsi choisie* il existe un n_0 tel que, pour $n \geq n_0$, on ait $\varphi_n(A) \leq \varepsilon/2\Gamma(x)$ (puisque $\lim \varphi_n(A) = 0$). Si n est supérieur à ce n_0 , on aura bien :

$$\left| \Gamma(x) - \int_0^n t^{x-1} \left(1 - \frac{1}{n}\right)^n dt \right| \leq \varepsilon \quad (\text{V.31})$$

Pour obtenir la seconde relation de l'énoncé, il suffit d'effectuer dans l'intégrale $\int_0^\infty t^{x-1} \left[1 - \frac{t}{n}\right] dt$ le changement de variable $t = ns$. •

Théorème V.6 — formule d'Euler

$$\lim_{n \rightarrow \infty} \frac{x(x+1)(x+2) \cdots (x+n)}{n! n^x} = \frac{1}{\Gamma(x)} \quad (\text{V.32})$$

Preuve En itérant (V.5) on obtient $\Gamma(x+n+1) = (x+n)(x+n-1) \cdots (x+1)x\Gamma(x)$ d'où :

$$\frac{\Gamma(x)}{\Gamma(x+n+1)} = \frac{1}{x(x+1)(x+2) \cdots (x+n)} \quad (\text{V.33})$$

En combinant le théorème V.5 avec le théorème V.1, on obtient :

$$\Gamma(x) = \lim_{n \rightarrow \infty} n^x \beta(x, n+1) = \lim_{n \rightarrow \infty} n^x \frac{\Gamma(x)\Gamma(n+1)}{\Gamma(x+n+1)} = \lim_{n \rightarrow \infty} \frac{n^x n!}{x(x+1)(x+2) \cdots (x+n)} \quad (\text{V.34})$$

ce qui permet de conclure. •

Théorème V.7 La fonction $1/\Gamma(z)$ est analytique dans $\mathbb{C}$ tout entier et ne s'annule qu'aux points $z = -n$ avec n , entier positif ou nul.

Preuve La fonction $\sin \pi z$ est analytique dans $\mathbb{C}$ tout entier, donc son inverse ne peut pas s'annuler (si $f(z) = 1/\sin \pi z$ s'annulait en un point, alors $\sin \pi z = 1/f(z)$ serait singulière en ce point). Or d'après le théorème 2 (formule des compléments) on a $\Gamma(x)\Gamma(1-x) = \pi/\sin \pi x$ pour $0 < x < 1$; par prolongement analytique, cette égalité est donc vraie aussi dans tout le domaine où $\Gamma(z)\Gamma(1-z)$ et $\pi/\sin \pi z$ sont analytiques, c'est-à-dire dans $\mathbb{C} - \mathbb{Z}$. S'il existait un point z_0 de $\mathbb{C} - \mathbb{Z}$ tel que $\Gamma(z_0) = 0$, on aurait donc aussi $\pi/\sin \pi z_0 = 0$. D'autre part, pour $z = n$ entier > 0 , $\Gamma(z) = (n-1)! \neq 0$. •

Cela prouve que $1/\Gamma(z)$ est définie et analytique dans D . Elle ne peut s'y annuler, pour la même raison que $f(z) = 1/\sin \pi z$. Reste à voir les points $z = -n$. Or, $\Gamma(1-z)$ est analytique en ces points (car $z = -n \Rightarrow 1-z = n+1$) donc la fonction $\Gamma(1-z) \cdot \frac{1}{\pi} \sin \pi z$ l'est aussi. Il suffit alors de poser $1/\Gamma(z) = \Gamma(1-z) \cdot \frac{1}{\pi} \sin \pi z$ pour avoir le prolongement analytique de $1/\Gamma(z)$ à $\mathbb{C}$ tout entier. Cette dernière égalité montre aussi que ce prolongement est nul pour $z = -n$; autrement dit les zéros de $1/\Gamma(z)$ sont les points $z = -n$. On pouvait aussi le déduire plus directement du fait du théorème V.3 qui montre que $\Gamma(z)$ y devient infinie puisque ce sont des pôles.

V.4 Dérivée de $\Gamma(z)$

Théorème V.8 Soit $\Gamma'(z)$, la dérivée de $\Gamma(z)$. Alors la fonction $\Gamma'(z)/\Gamma(z)$ est analytique dans D et :

$$\begin{aligned} -\frac{\Gamma'(z)}{\Gamma(z)} &= \lim_{n \rightarrow \infty} \left\{ \frac{1}{z} + \frac{1}{z+1} + \frac{1}{z+2} + \cdots + \frac{1}{z+n} - \ln(n+1) \right\} \\ &= \frac{1}{z} + \sum_{k=1}^{\infty} \left[\frac{1}{z+k} - \ln \left(1 + \frac{1}{k} \right) \right] \end{aligned} \quad (\text{V.35})$$

la série dans le dernier membre étant convergente (uniformément sur tout domaine borné).

Preuve Commençons par le démontrer dans un petit voisinage du point $z = 1$, après quoi cela s'étendra par prolongement analytique. Soit donc $\varepsilon < 1$ et supposons $|z - 1| < \varepsilon$. Reprenons la relation (V.36), mais cette fois pour z complexe :

$$\begin{aligned}\Gamma(z) - \int_0^n t^{z-1} \left(1 - \frac{1}{n}\right)^n dt &= \int_0^\infty t^{z-1} e^{-t} dt - \int_0^n t^{z-1} \left(1 - \frac{1}{n}\right)^n dt \\ &= \int_0^A t^{z-1} e^{-t} \varphi_n(t) dt + \int_A^n t^{z-1} e^{-t} \varphi_n(t) dt + \int_n^\infty t^{z-1} e^{-t} dt\end{aligned}\quad (\text{V.36})$$

Dans la première des trois intégrales de la deuxième ligne, on peut majorer $|t^{z-1}|$ par $t^{+\varepsilon}$ quand $t \geq 1$ et par $t^{-\varepsilon}$ quand $0 < t \leq 1$; en effet, $|z - 1| < \varepsilon \Rightarrow 1 - \varepsilon < \Re(z) < 1 + \varepsilon$ et $|t^{z-1}| = t^{\Re(z)-1}$. Cette première intégrale peut donc se majorer comme suit :

$$\left| \int_0^A t^{z-1} e^{-t} \varphi_n(t) dt \right| \leq \varphi_n(1) \int_0^1 t^{-\varepsilon} e^{-t} dt + \varphi_n(A) \int_1^A t^{\varepsilon} e^{-t} dt \quad (\text{V.37})$$

Cela prouve, puisque $\varphi_n(1)$ comme $\varphi_n(A)$ tendent vers zéro lorsque n tend vers l'infini, que cette intégrale tend *uniformément* vers zéro quand z est dans le disque $|z - 1| < \varepsilon$. Dans la seconde et la troisième intégrale, t est toujours ≥ 1 et on peut aussi les majorer uniformément par les intégrales obtenues en remplaçant t^{z-1} par $t^{+\varepsilon}$. Ainsi, lorsque $|z - 1| < \varepsilon$, la suite des fonctions :

$$H_n(z) = n^z \beta(z)(n+1) = \frac{n!n^z}{z(z+1)(z+2)\cdots(z+n)} \quad (\text{V.38})$$

converge *uniformément* vers $\Gamma(z)$. •

Cela implique que la suite des dérivées des H_n va converger vers $\Gamma'(z)$. En effet, la convergence uniforme implique que l'intégrale de la limite est la limite de l'intégrale, de sorte que

$$\frac{1}{2i\pi} \int_\gamma \frac{\Gamma(w)}{w - z} dw = \lim_{n \rightarrow \infty} \frac{1}{2i\pi} \int_\gamma \frac{H_n(w)}{w - z} dw \quad (\text{V.39})$$

Bien entendu, comme la convergence uniforme n'a été prouvée que dans le disque $|z - 1| < \varepsilon < 1$, il faut prendre le lacet γ dans ce disque; si ce lacet entoure le point z , l'égalité ci-dessus est équivalente à $\Gamma'(z) = \lim H'_n(z)$. Autrement dit : le fait que pour les fonctions analytiques, la dérivée est donnée par une intégrale, permet de dériver les suites uniformément convergentes.

L'argument utilisé ci-dessus, qu'on peut intégrer terme par terme une série uniformément convergente et que par conséquent, si chaque terme est analytique, la somme est elle aussi analytique et on peut passer à la limite pour les dérivées, a une validité générale et constitue le *théorème de Weierstrass* :

Théorème V.9 Soit $\sum f_n(z)$ une série de fonctions analytiques dans un domaine U , uniformément convergente sur tout disque borné $|z| \leq R$. Alors sa somme $F(z)$ est analytique sur U et les dérivées $F^{(k)}(z)$ sont égales aux séries $\sum f_n^{(k)}(z)$ des dérivées.

Preuve Il suffit de remarquer que pour tout lacet γ dont l'intérieur est entièrement contenu dans U :

$$\frac{1}{2i\pi} \int_\gamma F(z) dz = \sum_{n=0}^\infty \frac{1}{2i\pi} \int_\gamma f_n(z) dz = 0 \quad (\text{V.40})$$

et que pour tout lacet simple γ qui entoure le point z_0 :

$$\frac{1}{2i\pi} \int_\gamma \frac{F(z)}{(z - z_0)^{k+1}} dz = \sum_{n=0}^\infty \frac{1}{2i\pi} \int_\gamma \frac{f_n(z)}{(z - z_0)^{k+1}} dz \quad (\text{V.41})$$

En toute rigueur, il faut signaler l'argument de la continuité des dérivées partielles (voir le chapitre II) mais cela résulte aussi de la convergence uniforme : on considère la série $\sum f'_n(z)$, puis on intègre terme par terme sur un chemin de z_0 à z , etc. •

Pour dériver $H_n(z)$, réécrivons (V.38) comme le produit des facteurs :

$$H_n(z) = n! e^{z \ln(n)} \cdot \frac{1}{z} \cdot \frac{1}{(z+1)} \cdot \frac{1}{(z+2)} \cdot \frac{1}{(z+3)} \cdots \frac{1}{(z+n)} \quad (\text{V.42})$$

Ce qui donne :

$$\begin{aligned} H'_n(z) &= \frac{d}{dz} \left(n! e^{z \ln(n)} \cdot \frac{1}{z} \cdot \frac{1}{(z+1)} \cdot \frac{1}{(z+2)} \cdot \frac{1}{(z+3)} \cdots \frac{1}{(z+n)} \right) \\ &= \ln(n) H_n(z) - \frac{1}{z} H_n(z) - \frac{1}{(z+1)} H_n(z) - \\ &\quad - \frac{1}{(z+2)} H_n(z) - \cdots - \frac{1}{(z+n)} H_n(z) \end{aligned} \quad (\text{V.43})$$

Ce qu'on peut encore écrire sous la forme :

$$\frac{H'_n(z)}{H_n(z)} = \ln(n) - \frac{1}{z} - \frac{1}{z+1} - \frac{1}{z+2} - \cdots - \frac{1}{z+n} \quad (\text{V.44})$$

Le théorème de Weierstrass dit que le membre de gauche a pour limite $\Gamma'(z)/\Gamma(z)$, ce qui montre la première assertion du théorème — le fait d'avoir $\ln(n+1)$ au lieu de $\ln(n)$ ne change rien à la limite puisque $\ln(n+1) - \ln(n)$ tend vers zéro. Pour avoir la deuxième, il suffit de remarquer que :

$$\begin{aligned} \ln(n+1) &= \ln\left(\frac{2}{1}\right) + \ln\left(\frac{3}{2}\right) + \ln\left(\frac{4}{3}\right) + \cdots + \ln\left(\frac{n+1}{n}\right) \\ &= \ln\left(1 + \frac{1}{1}\right) + \ln\left(1 + \frac{1}{2}\right) + \ln\left(1 + \frac{1}{3}\right) + \cdots + \ln\left(1 + \frac{1}{n}\right) \end{aligned} \quad (\text{V.45})$$

On a alors :

$$\frac{1}{z} + \frac{1}{z+1} + \frac{1}{z+2} + \cdots + \frac{1}{z+n} - \ln(n+1) = \frac{1}{z} + \sum_{k=1}^n \left[\frac{1}{z+k} - \ln\left(1 + \frac{1}{k}\right) \right] \quad (\text{V.46})$$

La série converge uniformément pour $|z| \leq R$, quel que soit R . En effet, on peut écrire :

$$\left| \frac{1}{z+k} - \ln\left(1 + \frac{1}{k}\right) \right| \leq \left| \frac{1}{z+k} - \frac{1}{k} \right| + \left| \frac{1}{k} - \ln\left(1 + \frac{1}{k}\right) \right| \quad (\text{V.47})$$

Pour $k > R$, $|z+k| \geq k-R$ donc le premier de ces deux termes, égal à $|z|/k|z+k|$, est inférieur ou égal à $R/k(k-R)$; le second se majore par l'inégalité connue $0 \leq x - \ln(1+x) \leq x^2/2$, de sorte que le terme général de la série, pour les rangs $k > R$, est majoré par :

$$\frac{R}{k(k-R)} + \frac{1}{2k^2} \quad (\text{V.48})$$

ce qui entraîne bien la convergence. Mais le fait que cette série soit uniformément convergente pour $|z| \leq R$, et cela quel que soit R , entraîne que la limite est analytique (théorème de Weierstrass), c'est-à-dire que la fonction :

$$f(z) = \frac{1}{z} + \sum_{k=1}^n \left[\frac{1}{z+k} - \ln\left(1 + \frac{1}{k}\right) \right] \quad (\text{V.49})$$

est analytique dans D . La fonction $-\Gamma'(z)/\Gamma(z)$ l'est également, et les deux coïncident sur le disque $|z - 1| < \varepsilon$, donc elles coïncident partout.

On déduit du théorème V.8 une expression de la constante d'Euler $\gamma \simeq 0,577\,215$ rencontrée en IV.5. Rappelons que l'on avait alors obtenu $\int_0^\infty \ln(t) e^{-t} dt = -\gamma$. Or cette dernière relation signifie que $-\gamma = \Gamma'(1)$, qu'on obtient en dérivant sous le signe d'intégration. Pour justifier cette dérivation sans recourir à des théorèmes généraux, on peut partir de l'inégalité (valable pour h complexe et α réel) :

$$\left| \frac{e^{h\alpha} - 1}{h} - \alpha \right| \leq \frac{1}{2} |h| \alpha^2 e^{|\alpha h|} \quad (\text{V.50})$$

On a alors :

$$\frac{\Gamma(z+h) - \Gamma(z)}{h} - \int_0^\infty t^{z-1} e^{-t} dt = \int_0^\infty t^{z-1} e^{-t} \left[\frac{e^{h \ln(t)} - 1}{h} - \ln(t) \right] dt \quad (\text{V.51})$$

En combinant l'inégalité précédente (en prenant $\ln(t) = \alpha$), l'inégalité de la moyenne, et le fait que $|t^{z-1}| = t^{\Re(z)-1}$, on obtient :

$$\left| \frac{\Gamma(z+h) - \Gamma(z)}{h} - \int_0^\infty \ln(t) t^{z-1} e^{-t} dt \right| \leq \frac{1}{2} |h| \int_0^\infty t^{|h|+\Re(z)-1} e^{-t} [\ln(t)]^2 dt \quad (\text{V.52})$$

L'intégrale du second membre étant convergente, le facteur $|h|/2$ garantit que le premier membre tend vers zéro, c'est-à-dire qu'on a bien :

$$\Gamma'(z) = \int_0^\infty \ln(t) t^{z-1} e^{-t} dt \quad (\text{V.53})$$

Bien entendu, on obtiendrait de la même façon ⁽¹⁾ :

$$\frac{d^n \Gamma(z)}{dz^n} = \int_0^\infty [\ln(t)]^n t^{z-1} e^{-t} dt \quad (\text{V.54})$$

On voit qu'en prenant $z = 1$ dans (V.53), on trouve $\Gamma'(1) = -\gamma$. Par ailleurs, la formule (V.35) du théorème V.8 donne, pour $z = 1$:

$$\gamma = -\frac{\Gamma'(1)}{\Gamma(1)} = \lim_{n \rightarrow \infty} \left\{ 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln(n+1) \right\} \quad (\text{V.55})$$

En utilisant l'équation (V.45), (V.55) s'écrit sous la forme d'une série :

$$\gamma = \sum_{n=1}^{\infty} \left[\frac{1}{n} - \ln \left(1 + \frac{1}{n} \right) \right] \quad (\text{V.56})$$

En introduisant maintenant γ , on peut donner une forme équivalente du théorème V.8 :

Théorème V.10

$$\frac{\Gamma'(z)}{\Gamma(z)} = -\gamma + (z-1) \sum_{n=0}^{\infty} \frac{1}{(n+1)(z+n)} \quad (\text{V.57})$$

(1) Il suffit de remarquer que les intégrales telles que :

$$\int_0^\infty [\ln(t)]^n t^{x-1} e^{-t} dt$$

sont convergentes pour tout $n \geq 0$ et tout $x > 0$.

Ceci montre que les points singuliers $z = -n$ (n entier positif ou nul) de la fonction $\Gamma'(z)/\Gamma(z)$ sont des pôles simples, et que le résidu en chacun de ces points est -1 .

Preuve D'après l'équation (V.35) du théorème V.8, $-\Gamma'(z)/\Gamma(z)$ est la somme de la série de terme général :

$$u_n = \frac{1}{z+n} - \ln\left(1 + \frac{1}{n+1}\right) = \left[\frac{1}{z+n} - \frac{1}{n+1}\right] + \left[\frac{1}{n+1} - \ln\left(1 + \frac{1}{n+1}\right)\right] \quad (\text{V.58})$$

Or la série, de $n = 0$ à l'infini, des seconds termes entre crochets de (V.58) pour somme γ d'après (V.56) ; en réduisant au même dénominateur, le premier terme devient $(1-z)/(n+1)(z+n)$, de sorte que :

$$-\frac{\Gamma'(z)}{\Gamma(z)} = \gamma + \sum_{n=0}^{\infty} \frac{1-z}{(n+1)(z+n)} \quad (\text{V.59})$$

ce qui conduit bien à (V.57). Au point $z = -n_0$, la fonction :

$$G(z) = \sum_{n \neq n_0} \frac{1-z}{(n+1)(z+n)} \quad (\text{V.60})$$

est analytique (donc se développe en série entière de $z + n_0$), et :

$$\frac{\Gamma'(z)}{\Gamma(z)} = -\gamma - \frac{1-z}{(n_0+1)(z+n_0)} - G(z) = -\gamma - G(z) + \frac{1}{1+n_0} - \frac{1}{z+n_0} \quad (\text{V.61})$$

La dernière expression est bien un développement de Laurent, dont la partie singulière est le terme $-1/(z+n_0)$. Cela montre que le résidu est -1 , et cela pour tout entier $n_0 \geq 0$. •

V.5 Développements eulériens

Nous avons vu au cours de la démonstration du théorème V.3 que :

$$\Gamma(z) = \Psi(z) + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \frac{1}{z+n} \quad (\text{V.62})$$

où $\Psi(z) = \int_1^{\infty} t^{z-1} e^{-t} dt$ est une fonction entière de z . Nous avons vu aussi au théorème V.10, que :

$$\frac{\Gamma'(z)}{\Gamma(z)} = -\gamma + (z-1) \sum_{n=0}^{\infty} \frac{1}{(n+1)(z+n)} \quad (\text{V.63})$$

Ces deux formules sont analogues au développement en éléments simples d'une fraction rationnelle : on écrit en effet une fonction analytique, qui a des pôles simples en $z = a_n$, en une somme de fonctions (les éléments simples) de la forme $A_n/(z - a_n)$, plus une fonction entière. La différence avec les fractions rationnelles est qu'il y a une infinité d'éléments simples, et leur somme doit donc être une série convergente. C'est cette contrainte qui dans le second cas empêchait de sommer les éléments simples $-1/(z+n)$ sans leur compagnon $1/(n+1)$.

On appelle *développements eulériens* de tels développements en une série d'éléments simples associés à chaque pôle. Ces développements font apparaître les résidus en ces pôles, mais sont rarement utilisables pour le calcul numérique de la fonction, car ils convergent en général très lentement.

Un procédé découvert par Cauchy et utilisant le théorème des résidus permet, pour des fonctions assez régulières à l'infini, de trouver de tels développements eulériens, à condition toutefois de connaître déjà les résidus. Voici comment.

Soit une fonction $f(z)$ analytique en dehors d'une infinité de points singuliers isolés, $a_0, a_1, a_2, \dots$. La suite a_n tend forcément vers l'infini, sinon il y aurait des points d'accumulation qui seraient nécessairement des points singuliers essentiels. On suppose que la suite des a_n est ordonnée de sorte que $|a_n| \geq |a_m|$ si $n > m$ (le module de a_n croît avec n). Nous supposons aussi pour simplifier que ce sont tous des pôles d'ordre 1 (mais, à condition de calculer un peu plus, le procédé fonctionne pour des pôles d'ordre quelconque). Il est évident que la fonction $f(z)$ ne peut pas tendre vers zéro à l'infini puisqu'elle devient infinie aux points a_n et que ceux-ci sont arbitrairement proches de l'infini. Toutefois, elle peut tendre vers zéro quand z tend vers l'infini sans sortir d'un domaine qui ne contient pas les pôles. Laissons d'abord de côté la question du comportement de $f(z)$ à l'infini, et considérons la fonction $g(z) = f(z)/(z - w)$, pour w distinct de tous les pôles de f .

Il est clair que $g(z)$ n'a que des pôles simples : ceux de $f(z)$, plus w . Le résidu de $g(z)$ en w est $f(w)$: en effet, au voisinage de w , $f(z)$ est analytique, donc $f(z) = \sum_{n \geq 0} c_n(z - w)^n$ avec $c_0 = f(w)$ et alors $g(z) = c_0/(z - w) +$ série entière. De façon analogue on trouve le résidu de $g(z)$ au point $z = a_n$: il est égal au résidu A_n de $f(z)$, multiplié par $1/(a_n - w)$.

D'après le théorème des résidus, l'intégrale de $g(z) dz$ sur un cercle de centre 0 et de rayon R (R étant choisi de telle manière que le cercle ne rencontre aucun pôle) est égale à la somme des résidus de $g(z)$ aux pôles contenus dans le disque de rayon R . Supposons maintenant qu'on puisse trouver une suite de rayons R_k qui tende vers l'infini et tels que les cercles correspondants ne rencontrent pas les pôles (que ces cercles passent entre les pôles). Si l'intégrale sur ces cercles tend vers zéro lorsque k tend vers l'infini, on aura nécessairement :

$$\lim_{k \rightarrow \infty} \sum_{|a_n| \leq R_k} \frac{A_n}{w - a_n} = f(w) \quad (\text{V.64})$$

Ce qui signifie que la série :

$$\sum_{n=0}^{\infty} \frac{A_n}{w - a_n} \quad (\text{V.65})$$

converge et a pour somme $f(w)$. Compte tenu de la forme de la série (V.65), elle fournit donc un développement eulérien de la fonction $f(w)$. Bien entendu n'importe quel lacet, de forme quelconque, convient ; le cas d'un cercle n'est envisagé ici que pour fixer les idées.

Ce développement eulérien ne comporte pas de partie entière, comme cela avait été le cas pour le développement (V.62) ; cela est évidemment dû à l'hypothèse que nous avons faite pour l'intégrale sur le cercle de rayon R_k , lorsque R_k tend vers l'infini. Cette hypothèse ne saurait être satisfaite pour $f(z) = \Gamma(z)$, mais le serait pour $f(z) = \Gamma(z) - \Psi(z) = \Phi(z)$. On peut donc utiliser le procédé de Cauchy lorsque l'hypothèse mentionnée n'est pas satisfaite, à condition de trouver une fonction entière $G(z)$ telle que $f(z) - G(z)$ satisfasse l'hypothèse.

À titre d'exemple, appliquons le procédé à la fonction $f(z) = 1/\sin \pi z$. Pour contrôler la décroissance à l'infini, on va calculer le module $|\sin \pi z|$. En posant $z = x + iy$, on a :

$$\sin \pi z = \sin \pi x \cosh \pi y + i \cos \pi x \sinh \pi y \quad \Rightarrow \quad |\sin \pi z|^2 = \sin^2 \pi x + \sinh^2 \pi y \quad (\text{V.66})$$

Le chemin le plus commode pour vérifier notre hypothèse n'est pas un cercle, mais un carré de demi-côté $R_k = k + \frac{1}{2}$. Sur les côtés verticaux, on aura $x = \pm(k + \frac{1}{2})$, soit $\sin^2 \pi x = 1$,

d'où $|\sin \pi z|^2 = 1 + \sinh^2 \pi y = \cosh^2 \pi y$; sur les côtés horizontaux, $\sin \pi x$ sera variable, mais $\sin^2 \pi x$ toujours ≥ 0 donc $|\sin \pi z|^2 \geq \sinh^2 \pi y = \sinh^2 \pi R_k$. Il est donc évident que l'intégrale de $f(z)/(z-w)$ sur les côtés horizontaux sera majorée en module par :

$$\frac{1}{d_k(w) \sinh \pi R_k} \int_{-R_k}^{+R_k} dy = \frac{2R_k}{d_k(w) \sinh \pi R_k} \quad (\text{V.67})$$

où $d_k(w)$ est la distance du point w au périmètre du carré. L'expression ci-dessus tend visiblement vers zéro quand R_k tend vers l'infini; sur les côtés verticaux l'intégrale sera majorée en module par :

$$\frac{1}{d_k(w)} \int_{-R_k}^{+R_k} \frac{1}{\cosh \pi y} dy \leq \frac{1}{d_k(w)} \int_{-\infty}^{+\infty} \frac{1}{\cosh \pi y} dy = \frac{1}{d_k(w)} \quad (\text{V.68})$$

ce qui tend aussi vers zéro puisque $d_k(w)$ tend vers l'infini : le périmètre du carré s'écarte indéfiniment du point (fixe) w . On peut donc appliquer le procédé de Cauchy et on obtient le développement eulérien :

$$\frac{1}{\sin \pi y} = \sum_{n \in \mathbb{Z}} \frac{A_n}{w - n} \quad (\text{V.69})$$

Le résidu A_n de la fonction $1/\sin \pi z$ au point $z = n$ est immédiat en remarquant que $\sin \pi z = (-1)^n \sin \pi(z - n) = (-1)^n \pi(z - n) \times$ une fonction analytique égale à 1 pour $z = n$: $A_n = (-1)^n/\pi$. Pour faire apparaître la convergence de la série il faut aussi grouper deux par deux les termes correspondant à $+n$ et $-n$. Ainsi on obtient le développement eulérien :

$$\frac{\pi}{\sin \pi w} = \frac{1}{z} + \sum_{n \geq 1} (-1)^n \frac{2z}{z^2 - n^2} \quad (\text{V.70})$$

L'inconvénient du procédé est qu'il faut déjà connaître les résidus de la fonction à développer (dans cet exemple, c'était facile) et on ne peut donc pas l'utiliser pour calculer ces résidus s'ils sont inconnus (rappelons que dans les théorèmes V.3 et V.8, nous avons utilisé le développement eulérien pour obtenir les résidus).

V.6 Intégrale de Hankel

Soit H un chemin infini contenu dans Ω_1 et encadrant la coupure $]-\infty; 0]$ dans le sens positif — figure V.1(a). Considérons l'intégrale :

$$\mathcal{H}(z) = \frac{1}{2i\pi} \int_H [w^{-z}]_1 e^w dw = \frac{1}{2i\pi} \int_H e^{w-z \ln_2(w)} dw \quad (\text{V.71})$$

Contrairement aux intégrales sur des chemins bornés, il se pose ici la question de la convergence. Pour y répondre, voyons ce que l'intégrale devient après paramétrage $t \mapsto w(t)$ avec $-\infty < t < +\infty$:

$$\mathcal{H}(z) = \frac{1}{2i\pi} \int_{-\infty}^{+\infty} e^{w(t)-z \ln_2(w(t))} w'(t) dt \quad (\text{V.72})$$

Le nombre complexe $w(t)$ étant assujéti à se promener dans Ω_1 , on peut l'écrire sous forme trigonométrique $w(t) = r(t) e^{i\theta(t)}$, avec $\forall t, -\pi < \theta(t) < \pi$ et $r(t) > 0$. La fonction que l'on

intégrale est donc :

$$e^{r(t)} e^{i\theta(t) - z \ln r(t) - iz\theta(t)} w'(t) \quad (V.73)$$

dont le module est :

$$e^{r(t) \cos \theta(t) - x \ln r(t) + y\theta(t)} |w'(t)| \quad (V.74)$$

avec $x = \Re(z)$ et $y = \Im(z)$. Le facteur $|w'(t)|$ ne pose pas de problème car on peut toujours prendre un paramétrage pour lequel $|w'(t)|$ reste borné ; cela signifie simplement que la vitesse du point courant sur H — t étant le temps — reste bornée. Reste à voir ce qui

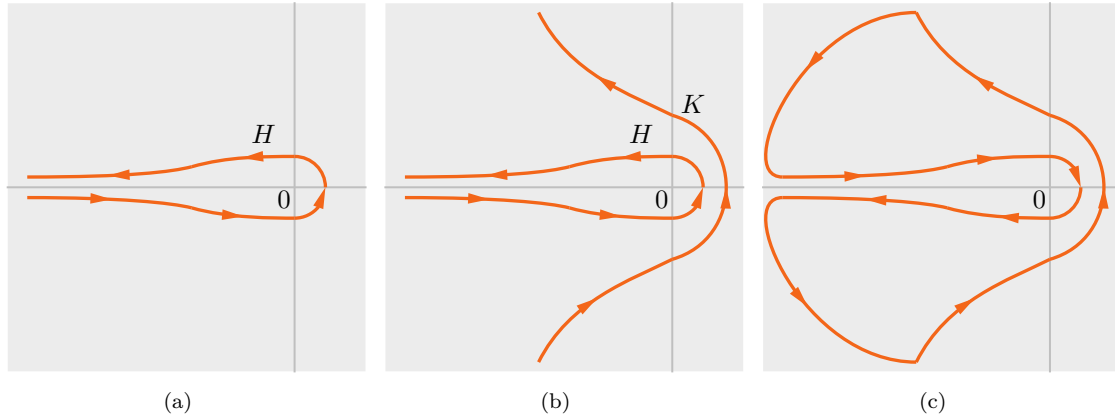


Figure V.1 – Chemin d'intégration

est en exposant. Le terme $y\theta(t)$ reste toujours compris entre $-y\pi$ et $+y\pi$ et n'intervient donc pas dans la caractérisation de la convergence. Comme le chemin H va à l'infini (aussi bien pour $t \rightarrow -\infty$ que pour $t \rightarrow \infty$), il est clair que $r(t) \rightarrow \infty$; par conséquent le terme $\ln r(t)$ peut être négligé devant $r(t)$, quelle que soit la fonction $r(t)$, pourvu qu'elle tende vers l'infini quand $t \rightarrow \pm\infty$, on aura toujours $\ln r(t)/r(t) \rightarrow 0$. Cela n'entraîne pas qu'il puisse être négligé devant $r(t) \cos \theta(t)$, car $\cos \theta(t)$ pourrait tendre vers zéro assez vite pour compenser $r(t)$. Par contre, il est évident que l'intégrale va diverger si $\cos \theta(t)$ est positif ; une condition qui garantit la convergence est donc que :

$$\limsup_{t \rightarrow \pm\infty} \cos \theta(t) < 0 \quad (V.75)$$

cela veut dire qu'il existe $\varepsilon > 0$ tel que pour $|t|$ grand, $\cos \theta(t) \leq -\varepsilon$, ou encore qu'il existe $\eta > 0$ tel que, pour t grand négatif, $-\pi < \theta(t) < -\frac{\pi}{2} - \eta$ et pour t grand positif, $\frac{\pi}{2} + \eta < \theta(t) < +\pi$. Géométriquement, cela signifie que le chemin H doit, le long de ses deux branches infinies, *pencher* vers la gauche du plan — l'angle θ doit rester $> \frac{\pi}{2} + \eta$ ou $< -\frac{\pi}{2} - \eta$. Peu importe son comportement à distance finie, pourvu qu'il ne traverse pas la coupure du plan.

On a ainsi exprimé une condition suffisante de convergence. Un chemin qui les vérifie sera appelé un *chemin de Hankel*.

- Ⓡ On pourrait étendre la condition de convergence à des chemins qui deviennent verticaux ($\theta \rightarrow \pm \frac{\pi}{2}$) à l'infini, en demandant que $\cos \theta(t)$ ne tende pas trop vite vers zéro ; cela est peu intéressant pour l'étude générale qui nous occupe, mais on pourrait y être amené pour traiter une application où cela serait nécessaire.

Lorsque deux chemins H et K vérifient tous deux la condition de convergence (V.75) — figure V.1(b), ils donnent la même valeur pour l'intégrale :

$$\frac{1}{2i\pi} \int_H [w^{-z}]_1 e^w dw = \frac{1}{2i\pi} \int_K [w^{-z}]_1 e^w dw \quad (\text{V.76})$$

Cela se montre aisément en considérant un lacet comme celui de la figure V.1(c), et sur lequel l'intégrale est nulle, puisque la fonction $w \mapsto [w^{-z}]_1 e^w$ est analytique dans Ω_1 ; la condition de convergence entraîne en effet que la partie de l'intégrale prise sur les ponts en quart de cercle — figure V.1(c) — tend vers zéro lorsque ceux-ci tendent vers l'infini.

D'autre part, il est clair que si H satisfait lesdites conditions de convergence, l'intégrale converge *quel que soit* z . En utilisant l'argument déjà fréquemment utilisé, que si γ est n'importe quel lacet dans $\mathbb{C}$, l'intégrale double $\int_\gamma \int_H dw dz$ peut être intégrée dans n'importe quel ordre, on voit que la fonction $\mathcal{H}(z)$ est analytique dans $\mathbb{C}$ tout entier (vérifier que les conditions pour qu'il en soit ainsi sont ici satisfaites est une affaire de routine).

Voyons maintenant comment calculer la fonction $\mathcal{H}(z)$. Puisque l'intégrale ne dépend pas du chemin choisi, on va choisir un chemin particulièrement commode. Un tel chemin est celui qui consiste à suivre l'axe $]-\infty; 0]$ par en-dessous en donnant à la fonction $w \mapsto [w^{-z}]_1 e^w$ ses valeurs-limite par dessous, puis par dessus (mais de 0 vers $-\infty$) en donnant à la fonction ses valeurs-limite par dessus. L'intégrale sera alors la limite des intégrales sur des chemins H qui contournent l'axe en le suivant de plus en plus près, et cette limite sera évidemment $\mathcal{H}(z)$. Toutefois, ce passage à la limite ne sera possible que pour une plage de valeurs de z : $\Re(z) < 1$; en effet, en dehors de cette plage, l'intégrale-limite sera divergente en 0, car la fonction $[w^{-z}]_1$ devient infinie en zéro pour $\Re(z) > 0$ et cette singularité n'est intégrable que pour $\Re(z) < 1$. Voyons cela en détail.

À cet effet, supposons à partir d'ici que z est réel (pour la commodité) et que $z < 1$. Pour les autres valeurs, on utilisera ensuite le principe du prolongement analytique.

Le chemin mentionné à l'instant, consistant à parcourir deux fois l'intervalle $]-\infty; 0]$, est un chemin-limite, mais n'est pas un chemin du type H , puisqu'il n'est pas inclus dans Ω_1 ; pour avoir un chemin du type H , il faudrait prendre par exemple :

- pour $\varepsilon < r < +\infty$: $w(r) = re^{i(-\pi+\alpha)}$;
- pour $-\pi + \alpha < \theta < +\pi - \alpha$: $w(\theta) = \varepsilon e^{i\theta}$;
- pour $\varepsilon < r < +\infty$: $w(r) = re^{i(\pi-\alpha)}$.

comme sur la figure V.2(a), ou bien :

- pour $-\infty < t < 0$: $w(t) = t - i\varepsilon$;
- pour $-\frac{\pi}{2} < \theta < +\frac{\pi}{2}$: $w(\theta) = \varepsilon e^{i\theta}$;
- pour $0 < t < +\infty$: $w(t) = t + i\varepsilon$.

comme sur la figure V.2(b).

Le premier chemin de la figure V.2(a) est formé d'une demi-droite d'argument angulaire constant égal à $-\pi + \alpha$ (avec α petit), allant de l'infini au cercle $|w| = \varepsilon$, puis de ce cercle

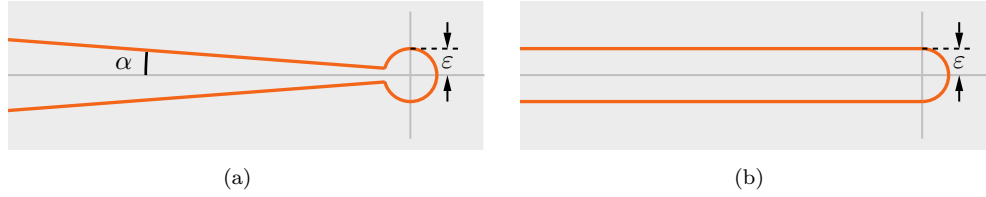


Figure V.2 – Chemins de Hankel

parcouru depuis l'angle $-\pi + \alpha$ jusqu'à l'angle $+\pi - \alpha$, et enfin d'une demi-droite d'argument angulaire constant égal à $+\pi - \alpha$ allant du cercle $|w| = \varepsilon$ à l'infini.

Le second chemin, sur la figure V.2(b), est formé d'une demi-droite située juste en-dessous (à la distance ε) de la coupure, puis d'un demi-cercle de rayon ε contournant 0, puis d'une demi-droite située juste au-dessus (à la distance ε) de la coupure.

On a vu à la section IV.6 que les fonctions puissance s'exprimaient le plus aisément en représentation trigonométrique, de sorte que le chemin du premier type sera le plus commode ; on posera en effet $w = re^{\pm i(\pi - \alpha)}$ sur les parties rectilignes, et $w = \varepsilon e^{i\theta}$ sur la partie circulaire, de sorte que la fonction à intégrer se paramétrera comme suit :

$$[w^{-z}]_1 e^w = \begin{cases} r^{-z} e^{iz(\pi - \alpha)} e^{-r e^{i\alpha}} & \text{sur la première demi-droite} \\ \varepsilon^{-z} e^{-iz\theta} e^{\varepsilon e^{i\theta}} & \text{sur la partie circulaire} \\ r^{-z} e^{-iz(\pi - \alpha)} e^{-r e^{-i\alpha}} & \text{sur la seconde demi-droite} \end{cases} \quad (\text{V.77})$$

(on a remplacé partout ci-dessus $e^{i\pi}$ par -1). Avec ce paramétrage, l'intégrale devient :

$$\begin{aligned} \mathcal{H}(z) &= \frac{1}{2i\pi} \int_{\varepsilon}^{\infty} r^{-z} [e^{i\alpha + iz(\pi - \alpha) - r e^{i\alpha}} - e^{-i\alpha - iz(\pi - \alpha) - r e^{-i\alpha}}] dr \\ &\quad + \frac{1}{2\pi} \varepsilon^{1-z} \int_{-\pi + \alpha}^{+\pi - \alpha} e^{i(1-z)\theta + \varepsilon e^{i\theta}} d\theta \\ &= \frac{1}{\pi} \int_{\varepsilon}^{\infty} r^{-z} e^{-r \cos \alpha} \sin(\alpha + z(\pi - \alpha) - r \sin \alpha) dr \\ &\quad + \frac{1}{2\pi} \varepsilon^{1-z} \int_{-\pi + \alpha}^{+\pi - \alpha} e^{i(1-z)\theta + \varepsilon e^{i\theta}} d\theta \end{aligned} \quad (\text{V.78})$$

Pour $\Re(z) < 1$, la partie de l'intégrale provenant du demi-cercle tend vers zéro, à cause du facteur ε^{1-z} , et les valeurs de la fonction à intégrer sur les deux parties rectilignes tendent, uniformément sur tout intervalle $]-\infty; -\eta]$, vers les valeurs-limites sur la coupure lorsque α tend vers zéro (de sorte qu'on peut passer à la limite dans les intégrales). Par contre pour $\Re(z) \geq 1$, la partie de l'intégrale provenant du demi-cercle tend vers l'infini et compense les divergences des intégrales sur les parties rectilignes. C'est pourquoi le cas où $\Re(z) < 1$ est bien plus commode pour calculer ⁽²⁾.

Si (pour $z < 1$) on effectue ce passage à la limite, avec $\alpha \rightarrow 0$ puis $\varepsilon \rightarrow 0$, on obtient :

$$\mathcal{H}(z) = \frac{1}{\pi} \int_0^{\infty} r^{-z} e^{-r} dr \times \sin \pi z \quad (\text{V.79})$$

Dans le membre de gauche ci-dessus on reconnaît l'intégrale eulérienne de deuxième espèce,

(2) Il dispense de calculer la contribution semi-circulaire, ce qui serait très difficile.

de sorte que :

$$\mathcal{H}(z) = \frac{1}{\pi} \Gamma(1-z) \sin \pi z \quad (\text{V.80})$$

En utilisant la formule des compléments du théorème V.2, cette expression devient :

$$\mathcal{H}(z) = \frac{1}{\Gamma(z)} \quad (\text{V.81})$$

Cette égalité s'appelle *formule de Hankel*. Elle a été obtenue en supposant $z < 1$, mais nous savons que $1/\Gamma(z)$ est analytique dans $\mathbb{C}$ tout entier (théorème V.7) ; nous avons vu aussi plus haut que $\mathcal{H}(z)$ est analytique dans $\mathbb{C}$ tout entier. Par conséquent, d'après le théorème des zéros isolés (principe du prolongement analytique), $\mathcal{H}(z)$ est *partout* égale à $1/\Gamma(z)$.

La formule de Hankel présente l'avantage d'être une représentation intégrale de $1/\Gamma(z)$, alors que les autres représentations intégrales de ce chapitre donnent $\Gamma(z)$. Nous verrons l'aspect pratique de cette représentation lorsque nous calculerons les transformées de Fourier des fonctions puissance. Il s'agit en effet d'intégrales telles que :

$$\int_{-\infty}^{+\infty} \frac{e^{ix\xi}}{[(ix)^\alpha]_1} dx \quad (\text{V.82})$$

pour α réel (ou complexe). Ces intégrales sont toujours divergentes : à l'infini si $\alpha \leq 1$ (ou $\Re(\alpha) < 1$), en zéro si $\alpha \geq 1$. Pour $\alpha \geq 1$, on peut la rendre convergente en remplaçant l'intégration sur l'axe réel $]-\infty; \infty[$ par une intégration sur un chemin qui coïncide avec l'axe réel en dehors d'un voisinage de zéro, mais contourne la singularité zéro en quittant l'axe réel. Un tel procédé pour donner un sens à une intégrale divergente est courant, surtout en physique (Cauchy l'avait également employé). Si pour $\xi > 0$ on remplace le paramétrage de x le long d'un tel chemin par celui de $w = ix\xi$ (ne pas confondre cette opération avec un changement de variable), on voit que (V.82) est la *même* chose que :

$$\xi^{\alpha-1} \int_H e^w [(w)^{-\alpha}]_1 dw \quad (\text{V.83})$$

où H est un chemin de Hankel. On reconnaît ci-dessus l'intégrale de Hankel.

Nous reviendrons de manière plus approfondie sur les détails quand nous calculerons ces transformées de Fourier.



Figure V.3 – Fonction $\mathcal{H}(x) = 1/\Gamma(x)$ pour $x \in]-4,5; 4,5[$. Les extrema entre les entiers négatifs augmentent factoriellement et il devient difficile de montrer un graphique pour $x < -5$.

VI — Transformations conformes

VI.1 Transformations géométriques du plan

Les transformations classiques du plan euclidien sont la translation, la rotation et l'homothétie. En composant une rotation et une homothétie, on obtient une similitude.

Si on représente chaque point du plan par un nombre complexe z (le point est alors l'*affiche* de z), une translation correspond à la transformation $z \mapsto z + a$ où a , également un nombre complexe, représente le vecteur de translation. Une rotation de centre 0 et d'angle θ correspond à la transformation $z \mapsto e^{i\theta} z$. Une homothétie de centre 0 et de rapport r correspond à la transformation $z \mapsto r z$.

Ces transformations correspondent toutes à une fonction *linéaire* de z (polynôme du premier degré). Il est clair qu'en les composant de toutes les manières possibles, on obtiendra toujours une fonction linéaire de z . Ainsi, en composant successivement :

- une rotation de centre 0 et d'angle θ ;
- une translation de vecteur a ;
- une homothétie de centre 0 et de rapport r ;
- une translation de vecteur $-a$;
- une rotation de centre 0 et d'angle $-\theta$;
- une homothétie de centre 0 et de rapport $1/r$.

on obtient la transformation :

$$z \mapsto \left[(e^{i\theta} z + a)r - a \right] e^{-i\theta} \frac{1}{r} \quad (\text{VI.1})$$

En développant l'expression, on lui trouve la forme équivalente :

$$z \mapsto z - a \left(\frac{1}{r} - 1 \right) e^{-i\theta} \quad (\text{VI.2})$$

autrement dit, une translation de vecteur $-a(\frac{1}{r} - 1)e^{-i\theta}$.

En composant ces transformations, on reste toujours dans domaine des transformations linéaires parce que l'ensemble des transformations linéaires forme un groupe.

Parmi les transformations géométriques classiques, il en existe aussi de non linéaires ; la plus connue est l'inversion. Elle se définit ainsi : soit O un point du plan (le *centre* de l'inversion) et m un nombre réel strictement positif (le *coefficient* de l'inversion) ; le transformé d'un point A du plan est le point B situé sur la demi-droite OA , tel que $OA \cdot OB = m$.

En prenant le centre de l'inversion comme origine du plan, l'inversion correspond aussi à une fonction simple de la variable complexe z . En effet, écrivons z en représentation trigonométrique $z = re^{i\theta}$. Si le point A est l'afixe de z , il est clair que le point B , situé sur la demi-droite OA , aura le même argument θ . D'autre part, sa distance à l'origine, OB , sera égale à $m/OA = m/r$. Par conséquent le point B est l'afixe du nombre complexe $\frac{m}{r}e^{i\theta} = m/\bar{z}$. Ainsi, l'inversion correspond à la fonction $z \mapsto m/\bar{z}$.

Il est connu, en géométrie élémentaire, que les inversions transforment généralement les droites en cercles ; plus précisément, elles transforment en cercles les droites qui ne passent pas par le centre de l'inversion (une droite qui passe par le centre de l'inversion se transforme en elle-même). Elles transforment aussi les cercles qui ne passent pas par le centre de l'inversion en d'autres cercles et les cercles qui passent par le centre de l'inversion en droites. Une autre propriété connue des inversions est de conserver les angles : par exemple, si deux droites sont perpendiculaires, les cercles transformés de ces droites se couperont à angle droit.

La composition de deux inversions de même centre donne une homothétie : en effet, $m' / \overline{(m/\bar{z})} = \frac{m'}{m} z$. La composition de deux inversions de centres différents donne une transformation plus générale :

$$\frac{\frac{m'}{z-a} - b}{\frac{m}{z-a} - b} = \frac{m'}{\frac{m}{z-a} - b} = \frac{m'z - m'a}{-bz + m + ab} \quad (\text{VI.3})$$

C'est donc une transformation *homographique* mais z n'y apparaît plus par l'intermédiaire de son conjugué.

Les transformations homographiques sont les transformations de la forme :

$$z \mapsto \frac{Az + B}{Cz + D} \quad (\text{VI.4})$$

La matrice :

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \quad (\text{VI.5})$$

est appelée matrice de la transformation. Lorsqu'on compose deux transformations homographiques, $z \mapsto (Az + B)/(Cz + D)$ puis $w \mapsto (A'w + B')/(C'w + D')$, on obtient :

$$\begin{aligned} z \mapsto \frac{A' \frac{Az+B}{Cz+D} + B'}{C' \frac{Az+B}{Cz+D} + D'} &= \frac{A'(Az+B) + B'(Cz+D)}{C'(Az+B) + D'(Cz+D)} \\ &= \frac{(A'A + B'C)z + (A'B + B'D)}{(C'A + D'C)z + (C'B + D'D)} \end{aligned} \quad (\text{VI.6})$$

c'est-à-dire que *la matrice de la composée est le produit des matrices*.

Cette propriété montre que les transformations homographiques forment un groupe. À une même transformation peuvent correspondre des matrices différentes : si on multiplie simultanément tous les éléments de la matrice par un même nombre complexe, on ne change pas la transformation. Autrement dit, la matrice d'une transformation homographique n'est définie qu'à un coefficient multiplicatif près.

Pour avoir une relation bi-univoque avec les transformations homographiques, il faudrait considérer non les matrices mais leurs classes d'équivalence par la relation :

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \sim \begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix} \Leftrightarrow \exists \lambda \in \mathbb{C} - \{0\}, \quad \begin{bmatrix} A' = \lambda A & B' = \lambda B \\ C' = \lambda C & D' = \lambda D \end{bmatrix} \quad (\text{VI.7})$$

Le groupe quotient est l'espace projectif des matrices.

Les translations, les rotations, les homothéties, sont des cas particuliers de transformations homographiques : les matrices associées sont respectivement (à un coefficient multiplicatif près) :

$$\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} e^{i\theta} & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix} \quad (\text{VI.8})$$

On aurait tout aussi bien pu prendre les matrices :

$$\begin{bmatrix} \frac{1}{a} & 1 \\ 0 & \frac{1}{a} \end{bmatrix} \quad \begin{bmatrix} e^{\frac{i\theta}{2}} & 0 \\ 0 & e^{-\frac{i\theta}{2}} \end{bmatrix} \quad \begin{bmatrix} \sqrt{r} & 0 \\ 0 & \frac{1}{\sqrt{r}} \end{bmatrix} \quad (\text{VI.9})$$

L'inversion n'est pas à strictement parler une homographie, puisque les homographies sont des fonctions de z et non de $\bar{z}$; mais les transformations du type $z \mapsto m/z$ (au lieu de $z \mapsto m/\bar{z}$) sont des homographies particulières, de matrice :

$$\begin{bmatrix} 0 & m \\ 1 & 0 \end{bmatrix} \quad (\text{VI.10})$$

On peut appeler *anti-inversion* ce type de transformation ; géométriquement, c'est la composée d'une inversion suivie d'une symétrie par rapport à l'axe réel.

De même que la géométrie euclidienne considère les propriétés qui sont invariantes par rotation et translation, on peut considérer les propriétés qui sont invariantes par les transformations homographiques : on appelle *anallagmatique* ⁽¹⁾ la géométrie (non-euclidienne) correspondante. La géométrie anallagmatique ne distingue pas les cercles des droites : une droite est simplement un cercle qui passe par le point à l'infini. La notion de longueur n'existe pas dans cette géométrie (alors qu'elle est essentielle en géométrie euclidienne) ; par contre la notion d'angle subsiste, car les transformations homographiques conservent les angles. Ainsi, la propriété fondamentale de la géométrie anallagmatique est la suivante : *si trois cercles se coupent en un même point P , ils se coupent deux à deux en trois autres points et la somme des angles de ces intersections est égale à π* (figure VI.1).

(1) En grec, ce mot signifie *invariante*.

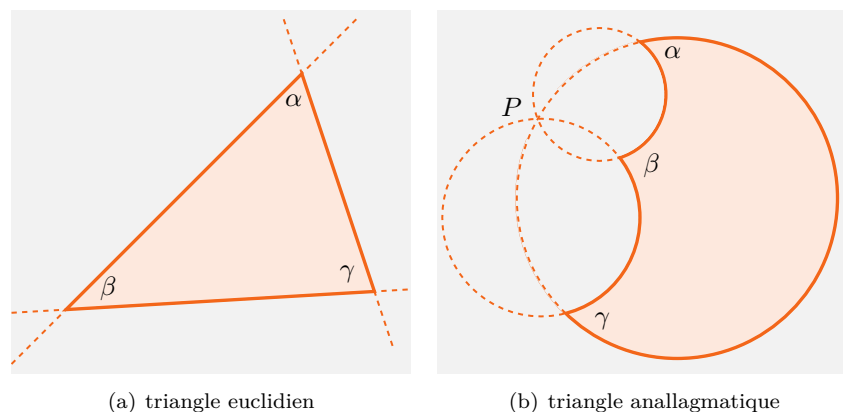


Figure VI.1 – Le triangle anallagmatique est le transformé du triangle euclidien par une homographie, qui conserve les angles. Le point P où les trois cercles se rencontrent est l'image du point à l'infini, où les trois droites du triangle euclidien se rencontrent.

VI.2 Propriétés géométriques des fonctions analytiques

Les transformations homographiques sont analytiques. On peut avoir une transformation des points du plan en considérant n'importe quelle fonction analytique. Toutefois, les transformations homographiques transforment n'importe quel point du plan ⁽²⁾ et sont, en outre, inversibles. On ne pourra pas en général en demander autant à une fonction analytique.

On peut montrer ⁽³⁾ que les seules transformations analytiques inversibles (bijectives) du plan sur lui-même (sans extension au point à l'infini) sont les transformations linéaires, c'est-à-dire euclidiennes ; et aussi que les seules transformations inversibles du plan étendu au point à l'infini sont les homographies.

Autrement dit, les transformations analytiques autres que les linéaires et les homographies, ne seront définies que dans une partie du plan, ou ne seront bijectives que dans une partie du plan. Par exemple, $z \mapsto z^2$ est analytique pour tout z mais n'est bijective que sur une partie du plan : ainsi, elle transforme bijectivement le demi-plan $\Im(z) > 0$ dans le domaine Ω_1 — son inverse est alors $[z^{1/2}]_1$, ou le demi-plan $\Re(z) > 0$ dans le domaine Ω_2 — son inverse est alors $[z^{1/2}]_2$. La fonction $z \mapsto \exp z$ est elle aussi analytique pour tout z mais n'est bijective que dans des domaines restreints tels que $0 < \Im(z) < 2\pi$ — l'inverse est alors la fonction $\ln_1 z$ — ou $-\pi < \Im(z) < \pi$ — l'inverse est alors la fonction $\ln_2 z$.

Si U et V sont deux domaines du plan et f une fonction analytique, définie sur U , telle que $V = f(U)$ et bijective en tant que fonction de U dans V , on dit que f est une transformation conforme de U sur V .

Théorème VI.1 Si $f(w)$ est une fonction analytique sur U , bijective de U dans V , sa dérivée $f'(w)$ ne peut s'annuler dans U .

Prouvons d'abord le lemme suivant :

⁽²⁾ Même celui qui annule le dénominateur si on lui attribue le point à l'infini comme image.

⁽³⁾ H. CARTAN, *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*, éditions Hermann, Paris, 1961, p. 182–183. Ces démonstrations ne sont pas très difficiles

Lemme VI.1 Soit $h(w)$ une fonction analytique dans le disque $|w| < \rho$. Si γ est un lacet simple contenu dans ce disque et ne passant par aucun de ces zéros, alors l'intégrale :

$$\frac{1}{2i\pi} \int_{\gamma} \frac{h'(w)}{h(w)} dw \quad (\text{VI.11})$$

est égale au nombre de zéros situés à l'intérieur du lacet γ , les zéros multiples étant comptés selon leur multiplicité.

Ⓡ La fonction h'/h a des pôles là où h a ses zéros ; comme par hypothèse les zéros ne sont pas situés sur le lacet, l'intégrale est bien définie.

Preuve (Lemme VI.1) On va montrer que le résidu de h'/h en chacun de ses pôles est égal à la multiplicité du zéro correspondant. Soit w_j , l'un des zéros de h dans le disque $|w| < \rho$ et $N \geq 1$ sa multiplicité ; la série entière de h au voisinage de w_j est :

$$h(w) = a_N(w - w_j)^N + a_{N+1}(w - w_j)^{N+1} + a_{N+2}(w - w_j)^{N+2} + \dots \quad (\text{VI.12})$$

avec $a_N \neq 0$ (puisque w_j est par hypothèse un zéro d'ordre N), de sorte que $h(w) = a_N(w - w_j)^N h_0(w)$, où $h_0(w)$ est une fonction analytique égale à 1 en w_j et par conséquent (par continuité) non nulle dans tout un voisinage de w_j . La dérivée est :

$$h'(w) = N a_N(w - w_j)^{N-1} h_0(w) + a_N(w - w_j)^N h'_0(w) \quad (\text{VI.13})$$

De sorte que le quotient h'/h s'écrit :

$$\frac{h'(w)}{h(w)} = \frac{N}{w - w_j} + \frac{h'_0(w)}{h_0(w)} \quad (\text{VI.14})$$

étant donné que $h_0(w) \neq 0$ (pour $|w - w_j|$ assez petit), la fonction $h'_0(w)/h_0(w)$ est analytique dans un voisinage de w_j , donc égale à une série entière. En conséquence :

$$\frac{h'(w)}{h(w)} = \frac{N}{w - w_j} + \text{série entière} \quad (\text{VI.15})$$

ce qui montre que le résidu de h'/h au point w_j est bien égal à N .

Le théorème des résidus dit alors que l'intégrale (VI.11) est égale à la somme des résidus aux points w_j intérieurs à γ , c'est-à-dire à la somme des multiplicités. •

Preuve (Théorème VI.1) S'il existe un point w_0 de U tel que $f'(w_0) = 0$, l'équation $f(w) = f(w_0)$ a pour racine w_0 avec une multiplicité au moins égale à deux puisque si w_0 était une racine simple, $f'(w_0)$ ne serait pas nul ; soit N cette multiplicité ($N \geq 2$). Le théorème des zéros isolés entraîne que $f'(w_0) \neq 0$ dans une couronne $\rho > |w - w_0| > 0$, avec ρ suffisamment petit ; par conséquent les racines éventuelles de l'équation $f(w) = z$ dans cette couronne sont forcément toutes simples parce qu'une racine multiple annule la dérivée.

Le nombre de ces racines simples est alors donné par l'intégrale :

$$\frac{1}{2i\pi} \int_{\gamma} \frac{f'(w)}{f(w) - z} dw \quad (\text{VI.16})$$

où γ est par exemple le cercle $|w - w_0| = \rho$. Par continuité de la fonction $f(w)$, tant que $z - a_0$ reste assez petit, les racines w de $f(w) = z$ sont toutes dans le disque $|w - w_0| < \frac{1}{2}\rho$, donc on peut minorer uniformément le long de γ le dénominateur $|f(w) - z|$ dans cette intégrale ; ce qui implique que l'intégrale (VI.16) est une fonction continue de z (pour $z - a_0$ assez petit). Or, lorsque $z = a_0$, la racine w_0 est une racine de multiplicité $N \geq 2$, de sorte que l'intégrale (VI.16) vaut N . Une conséquence de la continuité est alors que la valeur de l'intégrale, qui est un entier, reste constante (l'entier ne fera un saut que lorsque, $|z - a_0|$ augmentant, une racine traversera le lacet γ). Puisque ces racines deviennent simples dès que $z \neq a_0$, il y aura N racines simples (et donc distinctes) de $f(w) = z$ au voisinage de w_0 , c'est-à-dire au moins deux.

Autrement dit, si $f'(w)$ s'annule en un point w_0 de U , il y aura au voisinage de w_0 au moins deux racines distinctes de l'équation $f(w) = z$, ce qui signifie que f n'est pas injective. •

Le théorème VI.1 dit seulement que *si* f est bijective (de U dans V), *alors* sa dérivée ne peut s'annuler dans U ; mais :

1. la réciproque est fausse; le fait que la dérivée f' ne s'annule pas dans U ne garantit absolument pas que f soit injective (par exemple si $f(z) = z^3$ sur $U = \{z \mid \Im(z) > 0\}$, on voit bien que $e^{i\frac{\pi}{6}} \neq e^{i\frac{5\pi}{6}}$, alors que $f(e^{i\frac{\pi}{6}}) = f(e^{i\frac{5\pi}{6}}) = e^{i\frac{\pi}{2}}$ et pourtant $f'(z) = 3z^2$ ne s'annule en aucun point du demi-plan U);
2. il existe cependant une réciproque partielle : si $f'(w_0) \neq 0$, il existe un voisinage de w_0 dans lequel f est injective (théorème *d'inversion locale*); cela implique par exemple dans le cas de $f(z) = z^3$ que f est injective si on la restreint à des domaines assez petits (plus petits que $U = \{z \mid \Im(z) > 0\}$);
3. si f est bijective de U dans V , la dérivée f' peut s'annuler *sur la frontière* de U ; l'hypothèse que U est un domaine *ouvert* est essentielle dans ce contexte et la démonstration du théorème l'illustre : on utilise l'intégrale (VI.11) sur le chemin γ qui *entoure* le point w_0 où la dérivée s'annule, ce qui serait impossible si w_0 était sur la frontière.

Le mot conforme provient de ce que les transformations analytiques conservent les angles, avec leur orientation. Si tel est le cas, il est aisé d'en déduire que les transformations anti-analytiques, qui sont des fonctions analytiques de $\bar{z}$, conservent les valeurs absolues des angles mais inversent leur orientation : en effet, la conjugaison complexe correspond géométriquement à une symétrie par rapport à l'axe réel, qui inverse l'orientation des angles.

Afin que l'expression « conserver les angles » ait un sens précis, nous donnons l'énoncé suivant :

Théorème VI.2 Soient U et V deux domaines tels que la fonction analytique f réalise une transformation conforme de U sur V . Soient $t \mapsto w_1(t)$ et $t \mapsto w_2(t)$ deux chemins dans U qui se coupent au point $w = w_1(0) = w_2(0)$, où ils ont des vecteurs-vitesse $w'_1(0)$ et $w'_2(0)$ non nuls. Alors $z_1(t) = f(w_1(t))$ et $z_2(t) = f(w_2(t))$ sont deux chemins dans V qui se coupent au point $z = f(w)$, ayant en z des vecteurs-vitesse $z'_1(0)$ et $z'_2(0)$ non nuls et l'angle orienté entre les nombres complexes $z'_1(0)$ et $z'_2(0)$ est égal à l'angle orienté entre les nombres complexes $w'_1(0)$ et $w'_2(0)$.

Ⓡ Cet énoncé précise la notion d'angle entre deux courbes *au point où elles se coupent* comme étant l'angle de leurs vecteurs tangents respectifs en ce point.

Preuve f étant différentiable, il est clair que si $t \mapsto w(t)$ est un chemin différentiable, il en sera de même de $t \mapsto z(t) = f(w(t))$. Le vecteur-vitesse du chemin $t \mapsto f(w(t))$ est alors la dérivée, qui s'obtient par la règle de composition : $z'(t) = f'(w(t)) w'(t)$. Le vecteur-vitesse $z'(t)$ sera non nul (lorsque $w'(t)$ l'est) si et seulement si, $f'(w(t)) \neq 0$, ce qui résulte du théorème VI.1, puisque nous avons l'hypothèse que f est une bijection de U dans V .

Appliquons ces résultats aux deux chemins w_1 et w_2 , au point w correspondant à $t = 0$; on obtient :

$$\begin{aligned} z'_1(0) &= f'(w_1(0)) \cdot w'_1(0) = f'(w) \cdot w'_1(0) \\ z'_2(0) &= f'(w_2(0)) \cdot w'_2(0) = f'(w) \cdot w'_2(0) \end{aligned} \quad (\text{VI.17})$$

Ces deux égalités montrent que les nombres complexes $z'_1(0)$ et $z'_2(0)$ sont les images de $w'_1(0)$ et $w'_2(0)$ par la transformation linéaire $Z \mapsto f'(w) \cdot Z$. Si on représente trigonométriquement le nombre complexe $f'(w)$ par $re^{i\theta}$, cette transformation (la multiplication par $f'(w) = re^{i\theta}$) est une similitude d'angle θ et de rapport r , qui évidemment conserve les angles entre les vecteurs. •

L'hypothèse que f est injective (bijective de U dans V) intervient ici par sa conséquence que $f'(w)$ ne s'annule pas dans U . Toutefois la bijectivité interviendra plus loin par elle-même, car sans elle on ne peut pas parler de transformation. Le théorème VI.2 n'exclut pas que $f'(w)$ puisse s'annuler sur le *bord* du domaine U . Non seulement cela *peut* se produire mais c'est même, comme nous allons voir, une possibilité qui est systématiquement exploitée dans les applications des transformations conformes. En effet, le théorème VI.1 ne garantit pas la conservation des angles lorsque la dérivée f' s'annule. Une fonction dont la dérivée est nulle se comporte localement comme la fonction z^2 (si f'' ne s'annule pas en même temps) qui *double* les angles. Si cela se produit sur le bord de U , l'effet sera de replier ce bord, comme la fonction z^2 replie l'axe $\Re(z) = 0$ en deux fois la demi-droite $]-\infty; 0]$ (figure VI.2). C'est ce phénomène qui explique comment la transformation $z \mapsto \cosh z$ transforme les deux droites $\Im(z) = 0$ et $\Im(z) = \pi$ en les demi-droites dédoublées $]-\infty; -1]$ et $[1; \infty[$ (figure VI.4), ou bien la formation des points de rebroussement dans les rosettes des figures VI.6, VI.7 et VI.9.

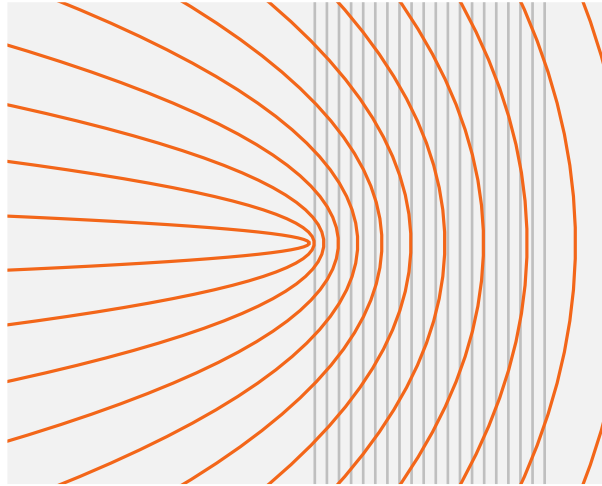


Figure VI.2 – Courbes images des droites $\Re(z) = C^{\text{te}}$ par la transformation conforme $z \mapsto z^2$

VI.3 Fonctions harmoniques

La propriété des transformations conformes la plus importante pour les applications est la *conservation de l'harmonicité*.

Une fonction *harmonique* $U(x_1, x_2, x_3, \dots, x_n)$ (à valeurs réelles) est une fonction vérifiant l'équation de Poisson :

$$\frac{\partial^2 U}{\partial x_1^2} + \frac{\partial^2 U}{\partial x_2^2} + \frac{\partial^2 U}{\partial x_3^2} + \dots + \frac{\partial^2 U}{\partial x_n^2} = 0 \quad (\text{VI.18})$$

Pour $n = 3$, c'est l'équation du potentiel électrostatique en dehors de tout milieu continu et électriquement chargé.

Nous considérons uniquement le cas du plan. Soient $Q(R, S)$ et $P(x, y)$ deux fonctions de deux variables, à valeurs réelles, dont les domaines de définition sont respectivement V et U . Si $R(x, y)$ et $S(x, y)$ sont elles-mêmes des fonctions de x, y , alors P est la fonction

composée de $f : (x, y) \mapsto (R, S)$ et de $(R, S) \mapsto Q$. On suppose donc que l'application $f : (x, y) \mapsto (R, S)$ transforme le domaine U en le domaine V :

$$\left[\begin{array}{ccccc} & f & & Q & \\ U & \rightarrow & V & \rightarrow & \mathbb{R} \\ (x, y) & \mapsto & (R, S) & \mapsto & P(x, y) = Q(R, S) \end{array} \right] \quad (\text{VI.19})$$

Ainsi $P(x, y) = Q(R(x, y), S(x, y))$. La règle de dérivation des fonctions composées donne :

$$\frac{\partial P}{\partial x} = \frac{\partial P}{\partial R} \frac{\partial R}{\partial x} + \frac{\partial P}{\partial S} \frac{\partial S}{\partial x} \quad (\text{VI.20})$$

Pour avoir la dérivée seconde par rapport à x , on applique à nouveau la règle de dérivation des fonctions composées :

$$\begin{aligned} \frac{\partial^2 P}{\partial x^2} &= \frac{\partial}{\partial x} \frac{\partial P}{\partial R} \frac{\partial S}{\partial x} + \frac{\partial P}{\partial R} \frac{\partial^2 R}{\partial x^2} + \frac{\partial}{\partial x} \frac{\partial P}{\partial S} \frac{\partial S}{\partial x} + \frac{\partial P}{\partial S} \frac{\partial^2 S}{\partial x^2} \\ &= \left[\frac{\partial^2 P}{\partial R^2} \frac{\partial R}{\partial x} + \frac{\partial^2 P}{\partial R \partial S} \frac{\partial S}{\partial x} \right] \frac{\partial R}{\partial x} + \left[\frac{\partial^2 P}{\partial R \partial S} \frac{\partial R}{\partial x} + \frac{\partial^2 P}{\partial S^2} \frac{\partial S}{\partial x} \right] \frac{\partial S}{\partial x} + \frac{\partial P}{\partial R} \frac{\partial^2 R}{\partial x^2} + \frac{\partial P}{\partial S} \frac{\partial^2 S}{\partial x^2} \end{aligned} \quad (\text{VI.21})$$

Il est clair que la dérivation par rapport à y aurait donné la même expression, sauf que x y serait remplacé par y . Par conséquent le laplacien sera :

$$\begin{aligned} \Delta P &= \frac{\partial^2 P}{\partial R^2} \left[\left(\frac{\partial R}{\partial x} \right)^2 + \left(\frac{\partial R}{\partial y} \right)^2 \right] + 2 \frac{\partial^2 P}{\partial R \partial S} \left[\frac{\partial R}{\partial x} \frac{\partial S}{\partial x} + \frac{\partial R}{\partial y} \frac{\partial S}{\partial y} \right] + \\ &\quad + \frac{\partial^2 P}{\partial S^2} \left[\left(\frac{\partial S}{\partial x} \right)^2 + \left(\frac{\partial S}{\partial y} \right)^2 \right] + \frac{\partial P}{\partial R} \Delta R + \frac{\partial P}{\partial S} \Delta S \end{aligned} \quad (\text{VI.22})$$

Ceci est l'expression *générale* du laplacien pour une fonction composée (elle ne se simplifie pas davantage). Supposons maintenant que la fonction $f : (x, y) \mapsto (R, S)$, ou en termes équivalents $f : x + iy \mapsto R + iS$, est analytique dans U . Cela implique que les relations de Cauchy-Riemann sont satisfaites, à savoir :

$$\frac{\partial R}{\partial x} = \frac{\partial S}{\partial y} \quad \text{et} \quad \frac{\partial R}{\partial y} = -\frac{\partial S}{\partial x} \quad (\text{VI.23})$$

Ces relations entraînent :

$$\frac{\partial R}{\partial x} \frac{\partial S}{\partial x} + \frac{\partial R}{\partial y} \frac{\partial S}{\partial y} = 0 \quad (\text{VI.24})$$

et :

$$\left(\frac{\partial R}{\partial x} \right)^2 + \left(\frac{\partial R}{\partial y} \right)^2 = \left(\frac{\partial S}{\partial x} \right)^2 + \left(\frac{\partial S}{\partial y} \right)^2 \quad (\text{VI.25})$$

Les deux expressions égales de (VI.25) ne sont autres que $|f'(x + ix)|^2$; en effet, la dérivée $f'(z)$ d'une fonction analytique $f(z) = R + iS$ est donnée par n'importe laquelle des expressions suivantes :

$$f'(z) = \frac{\partial f}{\partial x} = -i \frac{\partial f}{\partial y} = \frac{\partial R}{\partial x} - i \frac{\partial R}{\partial y} = \frac{\partial S}{\partial y} + i \frac{\partial S}{\partial x} = \text{etc.} \quad (\text{VI.26})$$

Enfin, les relations de Cauchy et Riemann entraînent aussi que $\Delta R = \Delta S = 0$. Par conséquent, si f est analytique, l'expression (VI.22) se simplifie énormément :

$$\Delta P = |f'|^2 \Delta Q \quad (\text{VI.27})$$

En particulier, si P est harmonique, Q l'est aussi (et vice-versa). On utilise cette propriété pour résoudre l'équation de Poisson $\Delta P = 0$ avec des conditions aux limites : si cette équation se résout facilement sur le domaine U et qu'on trouve une transformation conforme de U sur V , il suffira de transporter la solution de U à V par f , pour avoir une solution dans V .

On utilise cette propriété des transformations conformes pour résoudre bon nombre de problèmes d'électrostatique (P étant le *potentiel électrostatique*) ou d'hydrodynamique (P étant alors le *potentiel des vitesses* d'un fluide, dont le gradient est le *champ des vitesses*).

Il existe deux sortes de configurations dans le plan où l'équation de Poisson se résout trivialement :

- lorsqu'il y a invariance par translation : le domaine U est une bande du type $a < \Re(z) < b$ (ou $a < \Im(z) < b$) avec des conditions aux limites également invariantes sur les bords du type $P(a, y) = p_0$ et $P(b, y) = p_1$ (ou $P(x, a) = p_0$ et $P(x, b) = p_1$) ;
- lorsqu'il y a invariance par rotation : le domaine U est une couronne du type $a < |z| < b$ avec des conditions aux limites du type $P(z) = p_0$ pour $|z| = a$ et $P(z) = p_1$ pour $|z| = b$.

En effet, dans le premier cas le potentiel $P(x, y)$ ne dépend que de x (ou que de y), donc l'équation de Poisson se réduit à $P'' = 0$, dont les solutions sont $P = \lambda x + \mu$ (λ et μ sont alors déterminés par les conditions aux limites). Dans le second cas, le potentiel $P(x, y)$ ne dépend que de $r = \sqrt{x^2 + y^2}$ et l'équation de Poisson se réduit à :

$$P'' + \frac{1}{r}P' = 0 \quad (\text{laplacien en coordonnées polaires}) \quad (\text{VI.28})$$

dont les solutions sont $P = \lambda \ln r + \mu$ (λ et μ étant pareillement déterminés par les conditions aux limites).

Ⓡ Les deux configurations envisagées se transforment l'une en l'autre par la transformation conforme $f(z) = \exp z$.

L'exemple le plus simple de transformation conforme non triviale (autrement dit, dans le présent contexte, ni linéaire, ni exponentielle) est la transformation $f(z) = z^2$. Soit donné le problème suivant :

Problème VI.1 Résoudre l'équation de Poisson $\Delta P = 0$ sur le domaine Ω_2 , avec la condition aux limites $P = 0$ sur $]-\infty; 0]$.

Ce problème fournit un modèle mathématique du champ électrique autour du bord d'une tôle. Pour le traiter, on remarque que la fonction $z \mapsto z^2$ réalise une transformation conforme du demi-plan $\Re(z) > 0$ sur le domaine Ω_2 , dont l'inverse est $w \mapsto \sqrt{w}^2$.

Ⓡ La fonction $z \mapsto z^2$ réalise aussi une transformation conforme du demi-plan $\Im(z) > 0$ sur le domaine Ω_1 , dont l'inverse est alors $w \mapsto \sqrt{w}^1$.

Dans le demi-plan $\Re(z) > 0$, l'équation de Poisson a pour solutions $\lambda x + \mu$; la condition aux limites $P = 0$ pour $x = 0$ entraîne que $\mu = 0$ mais λ reste indéterminé⁽⁴⁾. Par conséquent, le potentiel $Q(R, S)$ s'obtient par $Q(x^2 - y^2, 2xy) = \lambda x$ ou, ce qui est équivalent :

(4) Il n'y a pas unicité de la solution car il n'y a pas suffisamment de conditions aux limites.

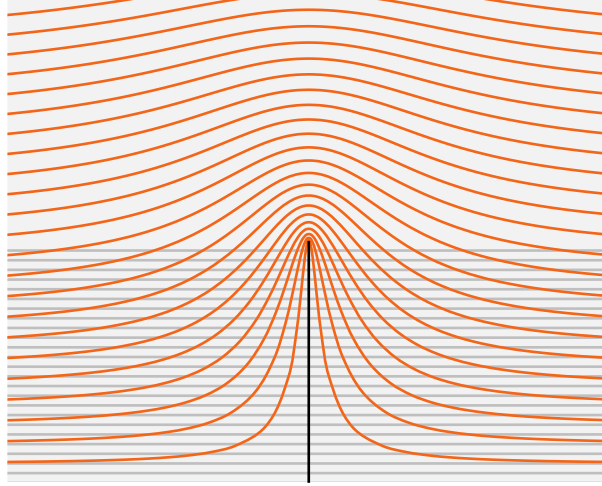


Figure VI.3 – Image du demi-plan $\Im(z) > 0$ par la transformation $z \mapsto \sqrt{z-1}^2 \cdot \sqrt{z+1}^2$. Les courbes sont les transformées des droites $\Im(z) = k/25$, pour $k = 1, 2, 3, \dots$

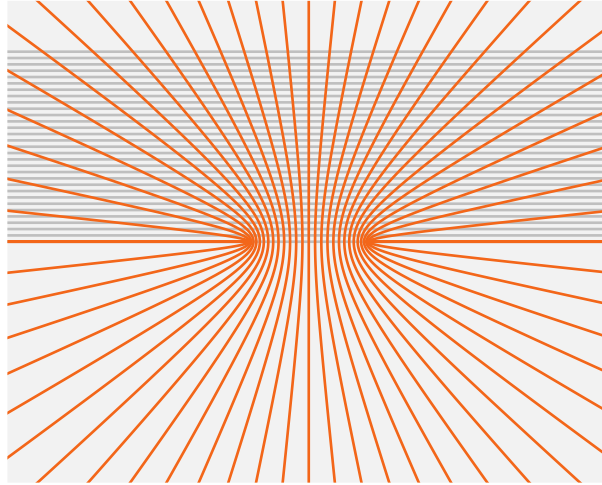


Figure VI.4 – Image de la bande $\pi > \Im(z) > 0$ par la transformation $z \mapsto \cosh z$. Les courbes sont les transformées des droites $\Im(z) = k\pi/30$, pour $k = 1, 2, 3, \dots 29$

$$Q(R, S) = \lambda \sqrt{\frac{\sqrt{R^2 + S^2} + R}{2}} \quad (\text{VI.29})$$

On a utilisé les expressions de $\sqrt{R + iS}^1$ et $\sqrt{R + iS}^2$ en coordonnées cartésiennes qui sont :

$$\begin{aligned} \sqrt{R + iS}^1 &= \sigma \sqrt{\frac{\sqrt{R^2 + S^2} + R}{2}} + i \sqrt{\frac{\sqrt{R^2 + S^2} - R}{2}} \\ \sqrt{R + iS}^2 &= \sqrt{\frac{\sqrt{R^2 + S^2} + R}{2}} + i\sigma \sqrt{\frac{\sqrt{R^2 + S^2} - R}{2}} \end{aligned} \quad (\text{VI.30})$$

où σ est le signe de S ; on remarquera que ce signe est indéterminé, soit pour $S = 0$, quand le facteur de σ est nul ou bien quand $R + iS$ est sur la coupure, de sorte que l'indétermination du signe reste toujours sans effet.

Ces expressions de $\sqrt{R + iS}^1$ et $\sqrt{R + iS}^2$ en coordonnées cartésiennes s'obtiennent sim-

plement en résolvant le système d'équations du second degré :

$$x^2 - y^2 = R; \quad 2xy = S \quad (\text{VI.31})$$

et en choisissant correctement celle des deux solutions qui correspond à la détermination retenue.

Les courbes $Q = C^{\text{te}}$ (courbes équipotentielles dans Ω_2) sont les images par f des droites $x = C^{\text{te}}$ (courbes équipotentielles dans $\Re(z) > 0$). Si on paramètre ces droites par $y \mapsto x + iy$, on obtient un paramétrage de leurs images $R = x^2 - y^2$ et $S = 2xy$. En éliminant le paramètre y , on obtient l'équation des courbes $R = x^2 - S^2/4x^2$, paraboles dont l'axe est l'axe réel et le sommet, le point d'abscisse x^2 (figure VI.2).

Les figures VI.3, VI.4 et VI.5 montrent les exemples les plus simples.

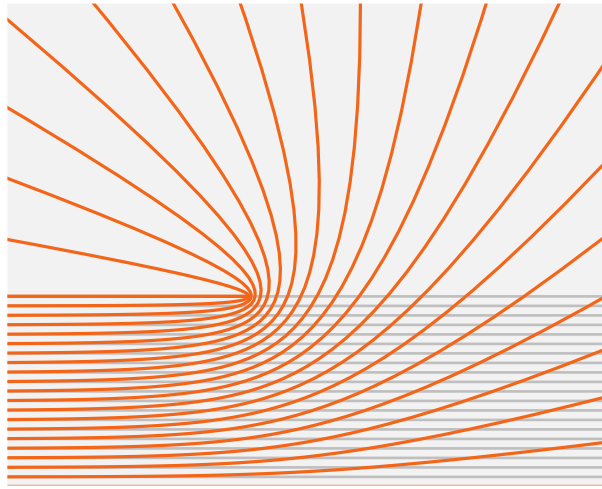


Figure VI.5 – Image de la bande $\pi > \Im(z) > 0$ par la transformation $z \mapsto z + \exp z$. Les courbes sont les transformées des droites $\Im(z) = k\pi/20$, pour $k = 1, 2, 3, \dots, 19$

VI.4 Autres exemples

VI.4.1 Polynômes

Il s'agit des figures VI.6 et VI.7.

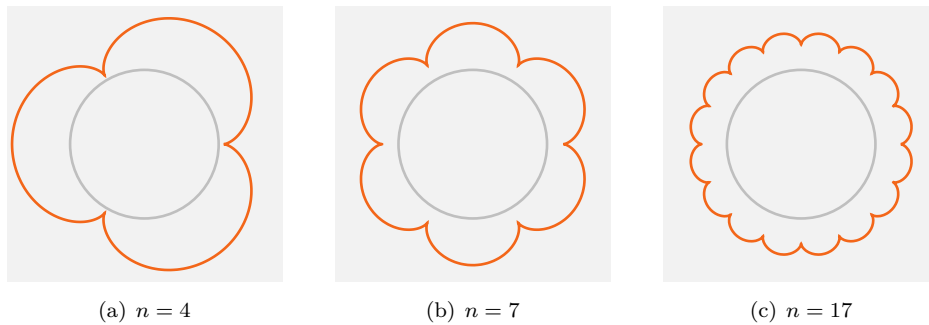


Figure VI.6 – Rosaces transformées du cercle $|z| = 1$ par les polynômes $z - \frac{z^n}{n}$ dont la dérivée, $1 - z^{n-1}$, s'annule sur les $n - 1$ racines de l'unité

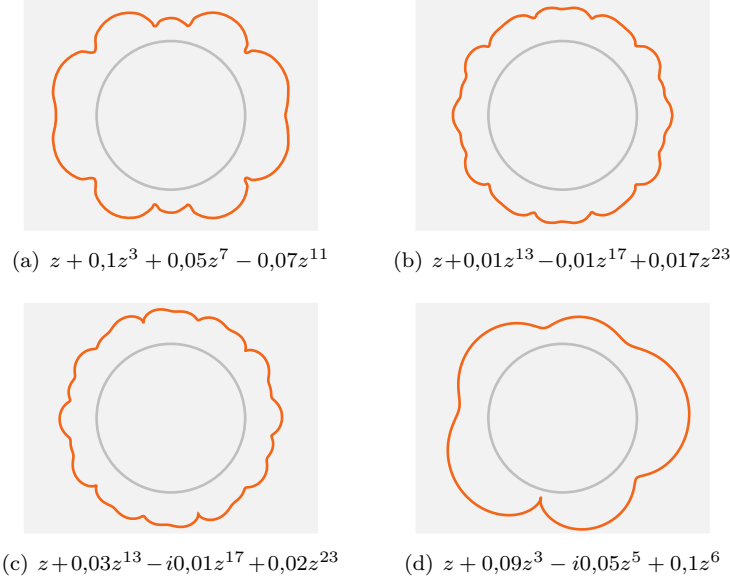


Figure VI.7 – Transformées du cercle $|z| = 1$ par les polynômes indiqués. Ces polynômes réalisent ainsi une transformation conforme du disque $|z| < 1$ sur la région délimitée par les courbes.

VI.4.2 Rosettes

Il s'agit des figures VI.6 et VI.9. Pour un entier $n \geq 1$, soient :

$$\omega_n^j = e^{\frac{2i\pi}{n}j} \quad (\text{VI.32})$$

les racines n^{e} de l'unité, c'est-à-dire les solutions complexes de l'équation $z^n = 1$. Considérons les fonctions suivantes :

$$f_n(z) = z - \frac{z^{n+1}}{n+1} \quad \text{et} \quad g_n(z) = z - \frac{a^n + 1}{na^{n-1}} \sum_{j=0}^{n-1} \omega_n^j \ln_2 \left(1 - \frac{z}{a\omega_n^j} \right) \quad (\text{VI.33})$$

Leurs dérivées sont respectivement :

$$f'_n(z) = 1 - z^n$$

$$g'_n(z) = 1 - \frac{a^n + 1}{na^{n-1}} \sum_{j=0}^{n-1} \frac{\omega_n^j}{z - a\omega_n^j} = 1 + \frac{a^n + 1}{z^n - a^n} = \frac{z^n + 1}{z^n - a^n} \quad (\text{VI.34})$$

Ces transformations ont pour $a \simeq 1$ une valeur (complexe) grande, de l'ordre de $\ln(1 - \frac{1}{a})$, au voisinage des points $z = e^{\frac{2i\pi}{n}(j+\frac{1}{2})}$ et une dérivée nulle aux points $z = e^{\frac{2i\pi}{n}j}$; elles vont donc étirer considérablement le disque au voisinage des premiers et le replier sans étirement au voisinage des seconds. Ceci explique la forme des rosettes de la figure VI.9.

On constate sans difficulté que la dérivée $f'_n(z)$ s'annule pour $z = \omega_n^j$ et qu'il s'agit de racines simples du fait de la factorisation bien connue :

$$z^n - 1 = (z - 1)(z - \omega_n^1)(z - \omega_n^2) \cdots (z - \omega_n^{n-1}) \quad (\text{VI.35})$$

À propos de $g'_n(z)$, son dénominateur s'annule pour $z = a\omega_n^j$ et son numérateur, pour $z = \omega_n^{j+\frac{1}{2}}$

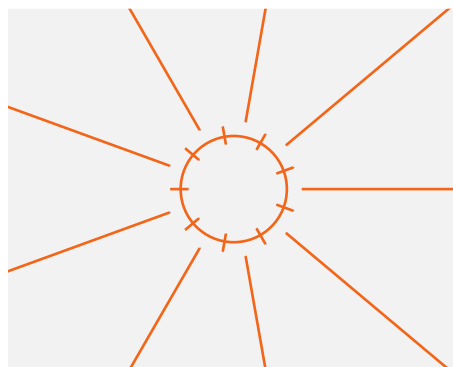
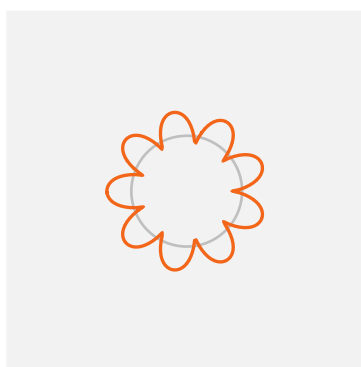


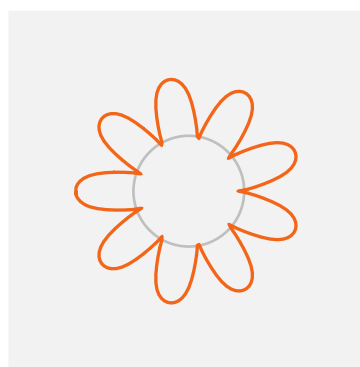
Figure VI.8 – Domaine d'holomorphie de la fonction :

$$f_n(z) = z + q \sum_{j=0}^{n-1} \omega_n^j \ln_2 \left(1 - \frac{z}{a\omega_n^j} \right)$$

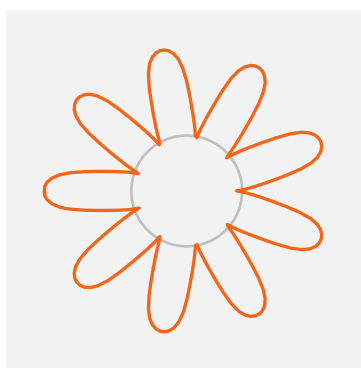
La fonction $f_n(z)$ est une somme de logarithmes de détermination $\ln_2$, de sorte que chaque terme exclut une coupure du plan qui est une demi-droite de la forme $\omega_n^j \times [a; \infty[$ correspondant aux valeurs réelles négatives de $1 - z/a\omega_n^j$. La dérivée $f'_n(z)$ s'annule aux points $z = \omega_n^{j+\frac{1}{2}}$ marqués d'un trait court.



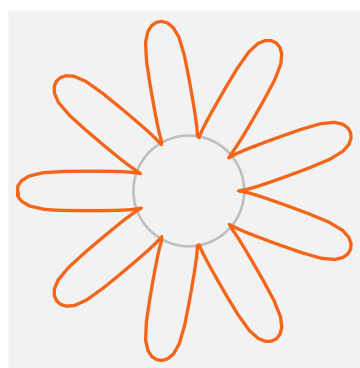
(a) $a = 1,01$



(b) $a = 1,001$



(c) $a = 1,0001$



(d) $a = 1,00001$

Figure VI.9 – Transformées du cercle $z = |1|$ par les fonctions :

$$z + q \sum_{j=0}^{n-1} \omega_n^j \ln_2 \left(1 - \frac{z}{a\omega_n^j} \right)$$

où ω_n^j ($j = 0, 1, \dots, n-1$) désignent les racines n^{es} de l'unité (ici $n = 9$) et $q = (a^n + 1)/na^{n-1}$.

VII — Transformation de Fourier

VII.1 Équation de la chaleur

En 1807, J. Fourier a proposé, en s'inspirant d'une méthode déjà évoquée par D. Bernoulli pour résoudre l'équation des cordes vibrantes, une méthode pour résoudre le problème suivant :

Problème VII.1 Trouver les solutions $P(t, x)$ de l'équation :

$$\frac{\partial P}{\partial t} = \frac{\partial^2 P}{\partial x^2} \quad (\text{VII.1})$$

sur le domaine $t \geq 0$, $0 \leq x \leq a$ avec les conditions aux limites $P(0, x) = p_0(x)$ où $p_0(x)$ est une fonction donnée *a priori*.

Il s'agit de chercher les solutions $P(t, x)$ sous la forme de séries trigonométriques du type :

$$P(t, x) = \sum_{n=0}^{\infty} a_n(t) \cos \frac{2\pi n}{a} x + b_n(t) \sin \frac{2\pi n}{a} x \quad (\text{VII.2})$$

En substituant (VII.2) dans (VII.1), on obtient :

$$\begin{aligned} \sum_{n=0}^{\infty} a'_n(t) \cos \frac{2\pi n}{a} x + b'_n(t) \sin \frac{2\pi n}{a} x &= \\ &= \sum_{n=0}^{\infty} -\left(\frac{2\pi n}{a}\right)^2 a_n(t) \cos \frac{2\pi n}{a} x - \left(\frac{2\pi n}{a}\right)^2 b_n(t) \sin \frac{2\pi n}{a} x \end{aligned} \quad (\text{VII.3})$$

En identifiant les coefficients des deux séries, on voit que ces coefficients doivent vérifier

pour tout indice n les équations :

$$a'_n(t) = -\left(\frac{2\pi n}{a}\right)^2 a_n(t); \quad b'_n(t) = -\left(\frac{2\pi n}{a}\right)^2 b_n(t) \quad (\text{VII.4})$$

dont la résolution est immédiate :

$$a_n(t) = a_n(0) \exp\left(-\frac{4\pi^2}{a^2}t\right); \quad b_n(t) = b_n(0) \exp\left(-\frac{4\pi^2}{a^2}t\right) \quad (\text{VII.5})$$

Le problème posé se résout donc comme suit ⁽¹⁾ :

- écrire la fonction donnée $p_0(x)$ sous forme de série trigonométrique :

$$p_0(x) = \sum a_n(0) \cos nx + b_n(0) \sin nx \quad (\text{VII.6})$$

- substituer dans (VII.5) les coefficients $a_n(0)$ et $b_n(0)$ ainsi obtenus.

Ce qui donne :

$$P(t, x) = \sum_{n=0}^{\infty} [a_n(0) \cos nx + b_n(0) \sin nx] e^{-n^2 t} \quad (\text{VII.7})$$

La difficulté du problème n'était pas dans les calculs exposés ci-dessus, mais dans la découverte des deux *évidences* suivantes :

- que la fonction p_0 est effectivement développable en série trigonométrique ;
- que, p_0 étant donnée, il existe un moyen mathématique simple d'exprimer les coefficients $a_n(0)$ et $b_n(0)$.

Autrement dit, la véritable découverte de Fourier est qu'une fonction $p(x)$ étant donnée, alors :

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} p(x) dx; \quad a_n = \frac{1}{\pi} \int_0^{2\pi} p(x) \cos nx dx; \quad b_n = \frac{1}{\pi} \int_0^{2\pi} p(x) \sin nx dx \quad (\text{VII.8})$$

et $p(x)$ est la somme de la série :

$$\sum_{n=0}^{\infty} a_n \cos nx + b_n \sin nx \quad (\text{VII.9})$$

On appelle *série de Fourier* de p la série (VII.9) avec les coefficients a_n et b_n de (VII.8).

Toutefois cet énoncé soulève beaucoup de difficultés : quel est le sens précis de l'affirmation " $p(x)$ est la somme de la série" ? C'est pourquoi, après les travaux de Fourier, beaucoup de mathématiciens ont étudié les conditions de validité de cette affirmation. Le bilan de cette lente maturation qui s'étend jusqu'au milieu du XX^e siècle est qu'on peut lui donner d'innombrables sens différents, dont voici quelques exemples :

Convergence simple — pour tout $x \in [0; 2\pi]$, la série (VII.9) converge dans $\mathbb{R}$ (c'est-à-dire au sens habituel des séries numériques) vers le nombre $p(x)$. Dirichlet a montré qu'il en est bien ainsi si $p(x)$ est continue sur $[0; 2\pi]$, si $p(0) = p(2\pi)$ et si p est monotone par morceaux sur $[0; 2\pi]$ (conditions de Dirichlet).

(1) Afin d'alléger l'écriture on prend à partir de maintenant, $a = 2\pi$.

Convergence au sens de Cesàro — pour tout $x \in [0; 2\pi]$, la *suite* numérique :

$$Q_N = \frac{1}{N} \sum_{n=0}^N a_n \cos nx + b_n \sin nx \quad (\text{VII.10})$$

converge dans $\mathbb{R}$ vers le nombre $p(x)$. On peut montrer que si p est continue sur $[0; 2\pi]$ et si $p(0) = p(2\pi)$ (sans autre condition) il en est bien ainsi.

Convergence en moyenne quadratique — l'intégrale :

$$\int_0^{2\pi} \left| p(x) - \sum_{n=0}^N a_n \cos nx + b_n \sin nx \right|^2 dx \quad (\text{VII.11})$$

tend vers zéro lorsque N tend vers l'infini. Cette forme de convergence est moins contraignante que les précédentes, car elle n'exige pas que la série converge pour *chaque* valeur de x . Elle est aussi la plus naturelle pour les séries de Fourier, celle qui exige le moins d'hypothèses compliquées. Pour que la série converge, il suffit que les séries $\sum |a_n|^2$ et $\sum |b_n|^2$ soient convergentes ; pour que la somme de (VII.9) soit égale à p , il suffit, même si $p(x)$ devient infini en certains points, que l'intégrale $\int_0^{2\pi} |p(x)|^2 dx$ soit convergente.

Convergence au sens des distributions — Pour que (VII.9) converge au sens des distributions, il suffit que les suites $|a_n|$ et $|b_n|$ soient à croissance polynomiale. Par exemple la série :

$$\sum_{n=0}^{\infty} \cos nx \quad (\text{VII.12})$$

est convergente au sens des distributions et sa somme est alors $\delta(x)$ (distribution de Dirac). Elle est aussi convergente au sens de Cesàro et sa somme est alors la fonction égale à 0 pour $0 < x < 2\pi$ et à 1 pour $x = 0$ ou $x = 2\pi$, mais *n'est pas* la série de Fourier de cette fonction. Autre exemple, la série :

$$\sum_{n=0}^{\infty} n \sin nx \quad (\text{VII.13})$$

est aussi convergente au sens des distributions (sa somme est alors $-\delta'(x)$, dérivée de la distribution de Dirac) mais n'est pas convergente au sens de Cesàro.

Selon le type de convergence qu'on attribue aux séries trigonométriques, leurs sommes seront des fonctions continues, des fonctions de carré intégrable, des distributions etc. La méthode de Fourier et Bernoulli s'applique donc essentiellement à un *champ de solutions*. La manière moderne de poser le problème VII.1 est la suivante :

Problème VII.2 Trouver les solutions continues (resp : de carré intégrable, distributions, infiniment dérivables etc.) de l'équation donnée.

Les réponses dépendent de la notion de convergence retenue. La clé de la méthode de Fourier–Bernoulli est que la dérivation des fonctions $\cos nx$ et $\sin nx$ équivaut à les multiplier par une constante. On peut procéder ainsi avec *n'importe* quel opérateur différentiel :

$$D = \sum_{k=0}^m a_m \frac{d^k}{dx^k} \quad (\text{VII.14})$$

sauf qu'il est alors plus commode d'utiliser les fonctions e^{inx} au lieu de $\cos nx$ et $\sin nx$. On aura en effet :

$$De^{inx} = \sum_{k=0}^m a_m (in)^k e^{inx} = \lambda_n e^{inx} \quad (\text{VII.15})$$

et par conséquent l'équation :

$$\frac{\partial P}{\partial t} = D P \quad (\text{VII.16})$$

avec la condition initiale :

$$P(0, x) = p_0(x) = \sum_{n \geq 0} a_n \cos nx + b_n \sin nx = \sum_{n \in \mathbb{Z}} c_n e^{inx} \quad (\text{VII.17})$$

(avec $c_{-n} = \overline{c_n} = \frac{1}{2}[a_n + ib_n]$) se résoudra par :

$$P(t, x) = \sum c_n e^{\lambda_n t + inx} \quad (\text{VII.18})$$

La méthode repose sur la réduction d'une équation aux dérivées partielles (qui comporte à la fois des dérivées par rapport à x et par rapport à t), à une équation différentielle ordinaire (qui ne comporte plus que des dérivées par rapport à t) que l'on sait résoudre par quadrature. La disparition de la dérivation par rapport à x provient de la propriété des fonctions trigonométriques, d'être égales à leurs dérivées multipliées par une constante.

L'idée originale de Fourier et Bernoulli a été généralisée ; supposons qu'étant donné un opérateur D , pas forcément différentiel, on puisse trouver une famille de fonctions $\varphi_n(x)$ telles que $D\varphi_n = \lambda_n \varphi_n$ (on dit que les φ_n sont des fonctions propres de D), alors l'équation :

$$\frac{\partial P}{\partial t} = D P \quad (\text{VII.19})$$

se résoudra de façon analogue, dans le champ des fonctions qui sont la somme d'une série de la forme $\sum c_n \varphi_n$. Voici l'exemple de l'hamiltonien de l'oscillateur harmonique quantique :

$$\mathcal{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} kx^2 \quad (\text{VII.20})$$

Les fonctions :

$$\varphi_n(x) = e^{-\frac{\sqrt{mk}}{2\hbar} x^2} H_n \left(\frac{1}{\sqrt{\hbar}} (mk)^{1/4} x \right) \quad (\text{VII.21})$$

où les H_n sont les polynômes d'Hermite, sont des fonctions propres de l'opérateur $\mathcal{H}$ ⁽²⁾ de sorte que l'équation aux dérivées partielles :

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + \frac{1}{2} kx^2 \psi \quad \text{— équation de Schrödinger de l'oscillateur} \quad (\text{VII.22})$$

se résout d'une manière analogue à la méthode de Fourier-Bernoulli. Il suffit de faire jouer aux fonctions φ_n le rôle des fonctions trigonométriques.

(2) Le calcul montre que :

$$\mathcal{H}\varphi_n = \left[\left(n + \frac{1}{2} \right) \frac{\sqrt{k/m}}{\hbar} \right] \varphi_n$$

VII.2 Transformation intégrale

La méthode consistant à chercher les solutions possibles sous la forme de séries de Fourier fournit, comme on a vu en section VII.1, des solutions sur un intervalle borné — voir (VII.5). Si l'intervalle devient infini, c'est-à-dire que a devient infini, les expressions (VII.5) cessent d'être utilisables.

Au lieu de considérer des séries, on considère alors des intégrales. C'est-à-dire qu'au lieu de chercher des solutions de la forme :

$$P(t, x) = \sum_{n \in \mathbb{Z}} c_n(t) \exp\left(-\frac{2i\pi}{a} n x\right) \quad (\text{VII.23})$$

on va les chercher sous la forme :

$$P(t, x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} C(\xi, t) e^{-ix\xi} d\xi \quad (\text{VII.24})$$

le coefficient $\frac{1}{2\pi}$ devant l'intégrale n'étant qu'une convention usuelle.

Ⓡ Si on discrétise l'intégrale (VII.24) en posant $\xi = n\varepsilon$, $d\xi = \varepsilon$ et $c_n = C(\xi) = C(\varepsilon n)$, on obtient sa somme de Riemann de pas ε . Si on prend $\varepsilon = 2\pi/a$, cette somme de Riemann est égale à la série (VII.23). Autrement dit, l'intégrale (VII.24) (sans son coefficient) est la limite de la série (VII.23) lorsque a tend vers l'infini. Il est donc naturel, lorsqu'on cherche des solutions de l'équation (VII.1) sur un intervalle de longueur infinie, de les prendre sous la forme de l'intégrale (VII.24).

La fonction $C(\xi, t)$ de (VII.24), définie sur $]-\infty; \infty[$, est appelée *transformée de Fourier en x* de la fonction $P(x, t)$. Toutefois les conventions telles que le coefficient $\frac{1}{2\pi}$ devant l'intégrale sont variables selon les auteurs, ou plutôt selon les spécialités. Il y a donc plusieurs *transformations de Fourier* possibles, dont les plus courantes sont les suivantes. Celle qui est utilisée en calcul des probabilités, à savoir la *fonction caractéristique* d'une densité de probabilité, est $f \mapsto \hat{f}$, ainsi que son inverse $f \mapsto \tilde{f}$:

$$\hat{f}(\xi) = \int_{-\infty}^{+\infty} f(x) e^{ix\xi} dx; \quad \tilde{f}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\xi) e^{-ix\xi} d\xi \quad (\text{VII.25})$$

La transformation $f \mapsto \tilde{f}$ est en effet la transformation inverse de $f \mapsto \hat{f}$, comme on le montrera plus loin (théorème VII.4). Dans d'autres domaines on préfère une version plus symétrique, où la transformation inverse ne diffère de la transformation directe que par le signe dans l'exposant et non par un coefficient $\frac{1}{2\pi}$ qui apparaît devant l'intégrale. Ainsi, la variante utilisée en traitement du signal est :

$$\Phi f(\xi) = \int_{-\infty}^{+\infty} f(x) e^{-2\pi i x \xi} dx \quad (\text{VII.26})$$

dont l'inverse est :

$$\Phi^{-1} f(\xi) = \int_{-\infty}^{+\infty} f(x) e^{+2\pi i x \xi} dx \quad (\text{VII.27})$$

On constate en effet une meilleure symétrie entre Φ et Φ^{-1} . Une autre variante qu'on peut rencontrer parfois est :

$$\mathcal{F}_1 f(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{ix\xi} dx \quad (\text{VII.28})$$

dont l'inverse est :

$$\mathcal{F}_1^{-1}f(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x) e^{-ix\xi} dx \quad (\text{VII.29})$$

Signalons encore le cas très important de la mécanique quantique, où la transformation de Fourier est :

$$\mathcal{F}_h f(\xi) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} f(x) e^{-\frac{ix\xi}{\hbar}} dx \quad (\text{VII.30})$$

Son inverse est :

$$\mathcal{F}_h^{-1}f(\xi) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} f(x) e^{-\frac{i}{\hbar}x\xi} dx \quad (\text{VII.31})$$

Elle est symétrique et la relation de Parseval correspondante est :

$$\|\mathcal{F}_h f\|_2 = \|f\|_2 \quad (\text{VII.32})$$

Comme on sait, cette transformation fait passer de la représentation d'espace à la représentation d'impulsion.

La fonction $C(\xi, t)$ de (VII.24), définie sur $]-\infty; \infty[$, est donc l'une (parmi d'autres) des transformées de Fourier en x de la fonction $P(x, t)$ — t est considéré comme un paramètre fixe. La transformée de Fourier en t de la fonction $P(x, t)$, où cette fois x serait le paramètre fixe, serait la fonction $D(x, \tau)$ telle que :

$$P(t, x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} D(x, \tau) e^{-it\tau} d\tau \quad (\text{VII.33})$$

Si les conditions pour pouvoir dériver sous le signe intégral sont satisfaites, on peut substituer l'intégrale (VII.24) dans l'équation (VII.1) et on obtient :

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\partial C}{\partial t}(\xi, t) e^{-ix\xi} d\xi = \frac{1}{2\pi} \int_{-\infty}^{+\infty} -\xi^2 C(\xi, t) e^{-ix\xi} d\xi \quad (\text{VII.34})$$

On arrive donc à la conclusion que *pour tout* x :

$$\int_{-\infty}^{+\infty} \left[\frac{\partial C}{\partial t}(\xi, t) + \xi^2 C(\xi, t) \right] e^{-ix\xi} d\xi = 0 \quad (\text{VII.35})$$

Il n'est pas évident (quoique vrai sous des conditions que nous préciserons) que si (VII.35) est vrai *pour tout* x , alors :

$$\frac{\partial C}{\partial t}(\xi, t) + \xi^2 C(\xi, t) = 0 \quad (\text{VII.36})$$

est vrai pour tout ξ . La réciproque est évidente : si (VII.36) est vrai pour tout ξ , alors (VII.35) sera vrai pour tout x . Si en outre les deux intégrales de (VII.34) sont convergentes, on peut en conclure aussi que (VII.34) est vrai.

Or, l'équation (VII.36) est facile à résoudre pour toute valeur fixée de ξ ; la solution en est :

$$C(\xi, t) = C(\xi, 0) e^{-\xi^2 t} \quad (\text{VII.37})$$

On retrouve le même principe qui faisait la puissance de la méthode de Bernoulli–Fourier : transformer une équation aux dérivées partielles en une équation différentielle ordinaire, intégrable par quadratures.

Toutefois ce principe ne peut être *effectivement* appliqué que si les conditions que nous avons supposées sont satisfaites, à savoir :

- les intégrales qui interviennent dans (VII.24) et (VII.34) sont convergentes, ou du moins, si elles ne le sont pas, on arrive à leur donner un sens ;
- la dérivation sous le signe d'intégration est justifiée. Une question qu'on peut se poser est la suivante : *la méthode décrite, consistant à chercher les solutions sous la forme (VII.24), permet-elle de trouver toutes les solutions, ou existe-t-il d'autres solutions de l'équation (VII.1) qui ne peuvent être mises sous la forme (VII.24) ?*

On pouvait déjà se poser cette question pour les intervalles bornés : *ne pourrait-il pas y avoir des solutions de l'équation (VII.1) qui ne peuvent pas être développées en séries trigonométriques ?*

Il y a deux manières de comprendre cette question. La première est la version dogmatique : on décrète que seules sont dignes du nom de *solution* des fonctions de $\mathbb{R} \times \mathbb{R}$ dans $\mathbb{R}$ (ou dans $\mathbb{C}$) qui sont différentiables (au moins une fois par rapport à t et deux fois par rapport à x) et satisfont (VII.1). On peut alors démontrer qu'on obtient bien toutes les solutions sous la forme (VII.24), avec $C(\xi, t)$ donnée par (VII.37). La deuxième est la version ouverte : on part du principe que si par exemple l'intégrale (VII.24) est divergente, on peut lui donner un sens cohérent et ensuite chercher des solutions correspondant à ce sens élargi. La version ouverte conduit (entre autres) à la théorie des distributions.

Pour mieux comprendre la seconde voie, il vaut mieux partir de l'équation :

$$\frac{\partial P}{\partial t} = i \frac{\partial^2 P}{\partial x^2} \quad (\text{VII.38})$$

car l'équation (VII.1) n'a que des solutions très régulières, quel que soit le sens, même élargi, qu'on peut donner au mot solution.

Si on cherche les solutions de (VII.38) par le même procédé, on obtient des intégrales du type (VII.24), mais avec :

$$C(\xi, t) = C(\xi, 0) e^{-i\xi^2 t} \quad (\text{VII.39})$$

La différence est importante : l'intégrale (VII.24) avec $C(\xi, t)$ de la forme (VII.37) est convergente ; par contre, avec $C(\xi, t)$ de la forme (VII.39), elle ne l'est pas. Selon qu'on considère la question envisagée ci-dessus dans sa version dogmatique ou dans sa version ouverte, on rejettera ou on acceptera les solutions (VII.39). Dans le second cas, on devra donner un sens cohérent aux intégrales :

$$\int_{-\infty}^{+\infty} e^{-i(\xi^2 t + x\xi)} d\xi \quad (\text{VII.40})$$

qui de surcroît justifie les dérivations sous le signe intégral.

C'est la seconde voie qui a été choisie historiquement. Nous verrons dans la section suivante comment interpréter des intégrales telles que (VII.40).

Pour le moment, nous nous intéresserons uniquement au cas où les intégrales convergent au sens usuel. L'intégrale (VII.24) dépend du paramètre x et est donc une fonction de x . Cette fonction dépend évidemment de la fonction $C(\xi)$, le paramètre t étant fixé. On peut donc considérer qu'il s'agit d'une *application* de l'ensemble des fonctions de ξ dans l'ensemble des fonctions de x , dont le domaine de définition est l'ensemble des fonctions $C(\xi)$ qui rendent l'intégrale absolument convergente.

De façon plus précise, soit $\mathbb{L}_1(]-\infty; \infty[)$ l'ensemble ⁽³⁾ des fonctions $C(\xi)$ définies sur $]-\infty; \infty[$ et telles que l'intégrale $\int |C(\xi)| d\xi$ soit convergente. Par exemple, la fonction :

$$C(\xi) = \begin{cases} 0 & \text{si } \xi \leq 0 \\ \frac{1}{\sqrt{\xi}} e^{-\xi} & \text{si } \xi > 0 \end{cases} \quad (\text{VII.41})$$

est dans $\mathbb{L}_1(]-\infty; \infty[)$.

On appellera *intégrables* sur $]-\infty; \infty[$ les fonctions appartenant à cet espace. Cet espace peut désigner les fonctions à valeurs réelles ou complexes : lorsqu'on dit que l'intégrale $\int |C(\xi)| d\xi$ converge, le symbole $| |$ peut désigner aussi bien la valeur absolue d'un nombre réel que le module d'un nombre complexe. Toutefois, l'intégrale qui définit la transformation de Fourier contient le facteur $e^{-ix\xi}$, de sorte que si une fonction est réelle, sa transformée de Fourier ne l'est pas en général. En effectuant le changement de variable $\xi \mapsto -\xi$ dans l'intégrale, on voit que si C est réelle *et paire*, alors sa transformée de Fourier est réelle aussi.

Il est clair que si $C(\xi)$ est intégrable sur $]-\infty; \infty[$, la fonction :

$$x \mapsto f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} C(\xi) e^{-ix\xi} d\xi \quad (\text{VII.42})$$

est définie pour tout $x \in]-\infty; \infty[$. Rien ne prouve que la nouvelle fonction f est intégrable sur $]-\infty; \infty[$ (c'est d'ailleurs en général faux), mais on peut montrer facilement qu'elle est continue. En effet, d'après l'inégalité de la moyenne :

$$|f(x) - f(y)| \leq \frac{1}{2\pi} \int_{-\infty}^{+\infty} |C(\xi)| \cdot |e^{ix\xi} - e^{iy\xi}| d\xi \quad (\text{VII.43})$$

On a aussi l'inégalité bien connue :

$$|e^{-ix\xi} - e^{-iy\xi}| \leq 2 \left| \sin \frac{(x-y)\xi}{2} \right| \quad (\text{VII.44})$$

On peut majorer $2 \left| \sin \frac{(x-y)\xi}{2} \right|$ à la fois par 2 et par $|x-y||\xi|$; divisons alors l'intervalle $]-\infty; \infty[$ en deux parties, l'une étant l'intervalle borné $[-A; A]$, l'autre le reste et majorons $2 \left| \sin \frac{(x-y)\xi}{2} \right|$ par $|x-y||\xi|$ sur $[-A; A]$ et par 2 sur le reste. On obtient :

$$\begin{aligned} \int_{-\infty}^{+\infty} |C(\xi)| \cdot |e^{-ix\xi} - e^{-iy\xi}| d\xi &\leq \int_{-A}^{+A} |C(\xi)| \cdot |x-y||\xi| d\xi + 2 \int_{|\xi| \geq A} |C(\xi)| d\xi \\ &\leq |x-y| A \int_{-\infty}^{+\infty} |C(\xi)| d\xi + 2 \int_{|\xi| \geq A} |C(\xi)| d\xi \end{aligned} \quad (\text{VII.45})$$

(3) En toute rigueur, il faut la théorie de l'intégrale de H. LEBESGUE pour définir l'espace $\mathbb{L}_1(]-\infty; \infty[)$; mais faute de temps, la nature exacte de cet espace sera laissée dans l'ombre. Il suffira d'admettre que toutes les fonctions qu'on peut rencontrer en font partie, à la seule condition que leur intégrale sur $]-\infty; \infty[$ soit absolument convergente.

Ceci étant vrai quel que soit A , on peut prendre par exemple $A = 1/\sqrt{|x-y|}$; lorsque $|x-y|$ tend vers zéro, le deuxième terme de (VII.45) tend vers zéro (puisque A tend vers l'infini et que l'intégrale converge) et le premier terme aussi puisque $|x-y|A$ tend vers zéro. On a ainsi montré que la fonction $f(x)$ est uniformément continue.

On voit de la même façon que la nouvelle fonction $f(x)$ est bornée uniformément sur $]-\infty; \infty[$; en effet, l'inégalité de la moyenne donne :

$$|f(x)| \leq \int_{-\infty}^{+\infty} |C(\xi)| d\xi \quad (\text{VII.46})$$

et le membre de droite ne dépend pas de x . On peut énoncer :

Lemme VII.1 La fonction (VII.37) est uniformément continue sur $]-\infty; \infty[$ et uniformément majorée par l'intégrale de $|C(\xi)|$.

Appelons $\mathbb{C}_0(]-\infty; \infty[)$ l'ensemble des fonctions continues et bornées sur $]-\infty; \infty[$. Ici aussi il peut s'agir aussi bien des fonctions à valeurs réelles que des fonctions à valeurs complexes. On peut résumer les résultats précédents en disant que (VII.42) définit une application de $\mathbb{L}_1(]-\infty; \infty[)$ dans $\mathbb{C}_0(]-\infty; \infty[)$.

Voyons deux exemples qui resserviront plus tard.

Exemple VII.1 Soit $C(\xi) = e^{-\alpha\xi^2}$. On se propose de calculer la fonction :

$$f(x) = \frac{1}{2\pi} \int e^{-\alpha\xi^2 - ix\xi} d\xi \quad (\text{VII.47})$$

On sait déjà que $\int e^{-\alpha\xi^2} d\xi = \sqrt{\pi/\alpha}$; remarquons que :

$$\alpha\xi^2 + ix\xi = \alpha \left(\xi + i\frac{x}{2\alpha} \right)^2 + \frac{x^2}{4\alpha} \quad (\text{VII.48})$$

et que $\xi \mapsto \xi + i\frac{x}{2\alpha}$ est un paramétrage de la droite $\Im(z) = \frac{x}{2\alpha}$. Cela suggère de calculer l'intégrale de $e^{-\alpha z^2} dz$ sur cette droite. Comme $e^{-\alpha z^2}$ est analytique dans tout le plan, on peut dire que $\int e^{-\alpha z^2} dz = 0$ si on prend un chemin fermé. Afin de retrouver à la fois l'intégrale connue et l'intégrale sur la droite $\Im(z) = \frac{x}{2\alpha}$, prenons un rectangle de sommets $\pm A$ et $\pm A + i\frac{x}{2\alpha}$, après quoi on fera tendre A vers l'infini. On a donc :

$$\int_{\text{rect}} e^{-\alpha z^2} dz = I_1 + I_2 + I_3 + I_4 = 0 \quad (\text{VII.49})$$

avec :

$$\begin{aligned} I_1 &= \int_{-A}^A e^{-\alpha\xi^2} d\xi & I_2 &= \int_0^{\frac{x}{2\alpha}} e^{-\alpha(A+it)^2} i dt = e^{-\alpha A^2} \int_0^{\frac{x}{2\alpha}} e^{-2i\alpha A t + \alpha t^2} i dt \\ I_3 &= e^{\frac{x^2}{4\alpha}} \int_A^{-A} e^{-\alpha\xi^2 - ix\xi} d\xi & I_4 &= \int_{\frac{x}{2\alpha}}^0 e^{-\alpha(-A+it)^2} i dt = -e^{-\alpha A^2} \int_0^{\frac{x}{2\alpha}} e^{-2i\alpha A t + \alpha t^2} i dt \end{aligned}$$

Il est évident que lorsque A tend vers l'infini, les intégrales I_2 et I_4 tendent vers zéro à cause du facteur $e^{-\alpha A^2}$. Puisque $I_1 + I_2 + I_3 + I_4$ est constamment nulle, cela implique que $I_1 + I_3$ tend vers zéro, autrement dit :

$$\int_{-\infty}^{+\infty} e^{-\alpha\xi^2} d\xi = e^{\frac{x^2}{4\alpha}} \int_{+\infty}^{-\infty} e^{-\alpha\xi^2 - ix\xi} d\xi \quad (\text{VII.50})$$

L'intégrale du premier membre est déjà connue et vaut $\sqrt{\pi/\alpha}$; celle du second membre est celle que nous cherchons (à l'inversion des bornes près). Donc :

$$\int_{-\infty}^{+\infty} e^{-\alpha\xi^2 - ix\xi} d\xi = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{x^2}{4\alpha}} \quad (\text{VII.51})$$

d'où $f(x)$, qui est égale à $\frac{1}{2\pi}$ fois cette intégrale :

$$f(x) = \frac{1}{\sqrt{4\pi\alpha}} e^{-\frac{x^2}{4\alpha}} \quad (\text{VII.52})$$

□

Exemple VII.2 Soit $C(\xi) = 1/(\xi^2 + a^2)$. On doit donc calculer l'intégrale :

$$\int_{-\infty}^{+\infty} \frac{e^{-ix\xi}}{\xi^2 + a^2} d\xi \quad (\text{VII.53})$$

qui est du type envisagé au théorème VII.5 du chapitre IV en section IV.3. Ce théorème fournit la réponse : le dénominateur s'annule pour $x = \pm ia$, donc l'intégrale vaut $\frac{\pi}{a}e^{-ax}$ si $x \geq 0$ et $\frac{\pi}{a}e^{+ax}$, autrement. En définitive :

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{-ix\xi}}{\xi^2 + a^2} d\xi = \frac{1}{2a} e^{-a|x|} \quad (\text{VII.54})$$

□

Les fonctions de ces deux exemples sont bien des fonctions intégrables sur $]-\infty; \infty[$, c'est-à-dire des fonctions appartenant à l'espace $\mathbb{L}_1(]-\infty; \infty[)$ puisque les intégrales $\int e^{-\alpha\xi^2} d\xi$ et $\int 1/(\xi^2 + a^2) d\xi$ sont absolument convergentes. Leurs images par la transformation intégrale sont respectivement $f(x) = (1/\sqrt{4\pi\alpha})e^{-x^2/4\alpha}$ et $f(x) = \frac{\pi}{a}e^{-a|x|}$. Ces fonctions sont elles aussi dans l'espace $\mathbb{L}_1(]-\infty; \infty[)$, mais cela ne correspond à aucune vérité générale.

VII.3 Principales propriétés de la transformation intégrale

Nous avons vu en section VII.2, qu'en admettant la dérivation sous le signe $\int$, la dérivation de $f(x)$ se traduisait, pour la fonction $C(\xi)$, par la multiplication par $i\xi$. Nous allons étudier dans cette section les conditions de validité de ces opérations, ainsi que d'autres propriétés de la transformation, en prouvant une série de théorèmes.

Définition VII.1 Une fonction $C(\xi)$ sur $]-\infty; \infty[$ est dite à décroissance rapide si pour tout entier $n > 0$, il existe une constante M_n telle que $\forall \xi \ C(\xi) \leq M_n/(1 + |\xi|^n)$.

Concrètement cela signifie qu'à l'infini, la fonction tend vers zéro plus vite que n'importe quelle puissance $1/\xi^n$. Ainsi les fonctions $e^{-\alpha\xi^2}$ ou $e^{-a|\xi|}$ sont à décroissance rapide; mais $1/(\xi^2 + a^2)$, $1/(\xi^4 + 1)$ ou encore $1/(\xi^{12} + 1)$ ne le sont pas.

Théorème VII.1 Si $C(\xi)$ est à décroissance rapide, alors la fonction :

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} C(\xi) e^{-ix\xi} d\xi \quad (\text{VII.55})$$

est infiniment dérivable en tout point. Les dérivées sont données par :

$$f^{(n)}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} (-i\xi)^n C(\xi) e^{-ix\xi} d\xi \quad (\text{VII.56})$$

Preuve On montre que $(f(x+h) - f(x))/h$ tend vers l'intégrale $\int -i\xi C(\xi) e^{-ix\xi} d\xi$ pour tout x . Pour cela, considérons la différence :

$$\frac{f(x+h) - f(x)}{h} - \int_{-\infty}^{+\infty} -i\xi C(\xi) e^{-ix\xi} d\xi = \int_{-\infty}^{+\infty} \left[\frac{e^{ih\xi} - 1}{h} + i\xi \right] C(\xi) e^{-ix\xi} d\xi \quad (\text{VII.57})$$

On connaît l'identité :

$$\int_0^t s e^{-i(t-s)} ds = 1 - e^{-it} - it \quad (\text{VII.58})$$

qu'on obtient en intégrant par parties et dont on déduit par l'inégalité de la moyenne que :

$$|e^{-it} - 1 + it| \leq \frac{1}{2} t^2 \quad (\text{VII.59})$$

Si on reporte cette inégalité dans (VII.57) en utilisant à nouveau l'inégalité de la moyenne, on obtient :

$$\begin{aligned} \left| \frac{f(x+h) - f(x)}{h} - \int_{-\infty}^{+\infty} -i\xi C(\xi) e^{-ix\xi} d\xi \right| &\leq \int_{-\infty}^{+\infty} \left| \frac{e^{ih\xi} - 1}{h} + i\xi \right| |C(\xi)| d\xi \leq \\ &\leq \frac{h}{2} \int_{-\infty}^{+\infty} \xi^2 |C(\xi)| d\xi \end{aligned} \quad (\text{VII.60})$$

Par hypothèse, $C(\xi)$ est à décroissance rapide, donc l'intégrale $\int \xi^2 |C(\xi)| d\xi$ est convergente. Le facteur $\frac{1}{2}h$ devant cette intégrale dans (VII.60) montre que l'expression (VII.57) tend vers zéro.

On a ainsi prouvé que f est dérivable. Il suffit de reconduire la même argumentation autant de fois qu'on veut, en remplaçant successivement $C(\xi)$ par $-i\xi C(\xi)$, puis par $-\xi^2 C(\xi)$ etc : f' sera à son tour dérivable si l'intégrale $\int |\xi^3 C(\xi)| d\xi$ est convergente, f'' sera dérivable si l'intégrale $\int |\xi^4 C(\xi)| d\xi$ est convergente etc. Or l'hypothèse que C est à décroissance rapide garantit qu'on peut poursuivre indéfiniment, car les fonctions $|\xi|^k / (1 + |\xi|^n)$ sont toutes intégrables si $n \geq k + 2$. •

Ⓡ La démonstration précédente montre plus précisément que si l'intégrale :

$$\int |\xi^n C(\xi)| d\xi \quad (\text{VII.61})$$

est convergente, alors $f(x)$ est $n-1$ fois dérivable. Si (VII.61) cesse d'être convergente pour $n+1$, on ne peut plus répéter *indéfiniment* le même argument et rien ne prouve alors que f est n fois dérivable. On peut constater cela directement sur l'exemple VII.2 : si $C(\xi)$ est la fonction $1/(\xi^2 + a^2)$, on voit bien que la fonction $C(\xi)$ est intégrable, mais pas la fonction $\xi^2 C(\xi)$; et en effet, sa transformée $f(x) = \frac{1}{2a} e^{-a|x|}$ est continue mais non dérivable en tout point. Troisième exemple plus sophistiqué : l'intégrale (IV.23) dont la valeur est donnée en (IV.27) :

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{-ix\xi}}{\xi^4 + 1} d\xi = \frac{1}{2} \cos\left(\frac{|x|}{\sqrt{2}} + \frac{\pi}{4}\right) e^{-(|x|/\sqrt{2})} \quad (\text{VII.62})$$

La fonction $C(\xi) = 1/(\xi^4 + 1)$ est intégrable, ainsi que $\xi C(\xi)$ et $\xi^2 C(\xi)$, mais non $\xi^3 C(\xi)$. La vérification directe de la dérivabilité en $x = 0$ montre que $f(x)$ est dérivable deux fois (mais non trois) ; on notera que d'après la remarque qui suit le théorème VII.1, seule la première dérivation était garantie.

Théorème VII.2 Si C est infiniment dérivable et que toutes ses dérivées (ordre zéro inclus) sont intégrables et nulles à l'infini, alors f est à décroissance rapide.

Preuve On procède en intégrant par parties :

$$\int_{-A}^{+A} C(\xi) e^{-ix\xi} d\xi = \frac{1}{ix} \left\{ -C(\xi) e^{-ix\xi} \Big|_{-A}^{+A} + \int_{-A}^{+A} C'(\xi) e^{-ix\xi} d\xi \right\} \quad (\text{VII.63})$$

L'expression entre accolades du membre de droite reste bornée uniformément en A puisque $C(\xi)$ est nulle à l'infini (donc le premier terme disparaît quand A tend vers l'infini) et $C'(\xi)$ intégrable (donc le second terme reste borné). Il existe donc une constante M_1 telle que $|f(x)| \leq M_1/|x|$. En intégrant à nouveau par parties, on voit qu'il existe une constante M_2 telle que $|f(x)| \leq M_2/|x|^2$ et ainsi de suite. Comme par

ailleurs $f(x)$ est bornée ($\leq M_0 = \int |C(\xi)|$) d'après ce qui a été vu en section VII.2, cela prouve que pour tout n ; on a :

$$|f(x)| \leq \inf \{M_0, M_n/|x|^n\} \leq \frac{M_0 + M_n}{1 + |x|^n} \quad (\text{VII.64})$$

donc que f est à décroissance rapide. •

On peut constater aussi, d'après cette démonstration, que si l'hypothèse que C et ses dérivées C' , $C'' \dots$ sont intégrables et nulles à l'infini n'était vérifiée que jusqu'à l'ordre k , alors on pourrait seulement conclure que f décroît plus vite à l'infini que $1/x^k$.

Les théorèmes VII.1 et VII.2 présentés ici sont des versions fortement réduites de théorèmes plus généraux : il existe d'innombrables généralisations de ces théorèmes, mais bien entendu les démonstrations sont alors beaucoup plus compliquées et sans intérêt pour une formation d'ingénieur.

En réunissant les théorèmes VII.1 et VII.2, on voit que les fonctions qui ont la propriété d'être infiniment dérivables et d'avoir toutes leurs dérivées à décroissance rapide, auront pour transformées des fonctions jouissant de la même propriété. Ces propriétés sont vraies également pour les transformations Φ , $\mathcal{F}_1$ et $\mathcal{F}_h$. Les fonctions infiniment dérivables ayant toutes leurs dérivées à décroissance rapide jouent un rôle important en analyse fonctionnelle et c'est pourquoi on introduit un espace spécial pour ces fonctions :

Définition VII.2 On appelle *espace de Schwartz*, l'espace de toutes les fonctions infiniment dérivables dont toutes les dérivées (ordre zéro compris) sont à décroissance rapide. On note $\mathcal{S}(\mathbb{R})$ ou $\mathcal{S}(-\infty; \infty)$ cet espace. Les éléments de cet espace seront appelés les *bonnes fonctions* ou *fonctions régulières*.

Ⓡ On dit *espace* plutôt qu'*ensemble* parce que c'est un espace vectoriel; en outre il sera muni d'une notion spéciale de convergence.

Nous avons vu, à la section VII.2, que la transformation intégrale de Fourier était une application de $\mathbb{L}_1(\mathbb{R})$ dans $\mathbb{C}_0(\mathbb{R})$; nous pouvons maintenant ajouter que l'image du sous-espace $\mathcal{S}(\mathbb{R})$ est $\mathcal{S}(\mathbb{R})$.

Théorème VII.3 Si f et g sont deux fonctions de l'espace $\mathbb{L}_1(\mathbb{R})$, on a toujours :

$$\int_{-\infty}^{+\infty} \widehat{f}(\xi) g(\xi) e^{-iy\xi} d\xi = \int_{-\infty}^{+\infty} f(x+y) \widehat{g}(x) dx \quad (\text{VII.65})$$

où l'on a posé :

$$\widehat{f}(\xi) = \int_{-\infty}^{+\infty} f(x) e^{ix\xi} dx \quad \text{et} \quad \widehat{g}(x) = \int_{-\infty}^{+\infty} g(\xi) e^{ix\xi} d\xi \quad (\text{VII.66})$$

Preuve On considère l'intégrale double :

$$\iint_{\mathbb{R}^2} f(x) g(\xi) e^{i(x-y)\xi} dx d\xi \quad (\text{VII.67})$$

Il est clair que cette intégrale double est absolument convergente, puisque f et g sont dans $\mathbb{L}_1(\mathbb{R})$ et que $|e^{i(x-y)\xi}| = 1$. On obtient donc le même résultat en intégrant d'abord par rapport à ξ puis par rapport à

x , qu'en intégrant d'abord par rapport à x puis par rapport à ξ . Or :

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x) \left[\int_{-\infty}^{+\infty} g(\xi) e^{i(x-y)\xi} d\xi \right] dx &= \int_{-\infty}^{+\infty} f(x) \widehat{g}(x-y) dx = \int_{-\infty}^{+\infty} f(x+y) \widehat{g}(x) dx \\ \int_{-\infty}^{+\infty} g(\xi) e^{-iy\xi} \left[\int_{-\infty}^{+\infty} f(x) e^{ix\xi} dx \right] d\xi &= \int_{-\infty}^{+\infty} \widehat{f}(\xi) g(\xi) e^{-iy\xi} d\xi \end{aligned} \quad (\text{VII.68})$$

ce qui permet de conclure. •

Théorème VII.4 Si $f \in \mathcal{S}(\mathbb{R})$, on a :

$$f(y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \widehat{f}(\xi) e^{-iy\xi} d\xi \quad (\text{VII.69})$$

Cette relation est appelée *formule d'inversion* ; en effet, si $\widehat{f}$ est donnée à partir de f par (VII.66), on retrouve f à partir de $\widehat{f}$ grâce à (VII.69).

Preuve On écrit (VII.65) avec $g(\xi) = e^{-\alpha\xi^2}$, puis on fait tendre α vers zéro, alors, pour la première intégrale de (VII.65), on obtient :

$$\int_{-\infty}^{+\infty} \widehat{f}(\xi) e^{-\alpha\xi^2} e^{-iy\xi} d\xi \xrightarrow{\alpha \rightarrow 0} \int_{-\infty}^{+\infty} \widehat{f}(\xi) e^{-iy\xi} d\xi \quad (\text{VII.70})$$

En effet, puisque la famille des fonctions $|\widehat{f}(\xi) e^{-\alpha\xi^2} e^{-iy\xi}|$ est majorée uniformément en α par une fonction intégrable (par $|\widehat{f}(\xi)|$), on peut passer à la limite sous le signe intégral.

Voyons l'autre intégrale ; la fonction $\widehat{g}(x)$ a été calculée dans l'exemple VII.1 de la section VII.2 :

$$\int_{-\infty}^{+\infty} e^{-\alpha\xi^2 - ix\xi} d\xi = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{x^2}{4\alpha}} \quad (\text{VII.71})$$

ce qui montre, en remplaçant x par $-x$, que :

$$\widehat{g}(x) = \sqrt{\frac{\pi}{\alpha}} e^{-x^2/4\alpha} \quad (\text{VII.72})$$

Par conséquent, pour tout $\alpha > 0$, on a :

$$\int_{-\infty}^{+\infty} \widehat{g}(x) dx = 2\pi \quad (\text{VII.73})$$

D'autre part, lorsque α tend vers zéro, $\widehat{g}(x)$ tend uniformément vers zéro en dehors des intervalles $]-\varepsilon; \varepsilon[$ et cela quel que soit $\varepsilon > 0$ (en effet, $\sup_{|x| \geq \varepsilon} \widehat{g}(x) = \widehat{g}(\varepsilon)$ et cela tend bien vers zéro avec α). On en déduit que pour tout ε :

$$\lim_{\alpha \rightarrow 0} \int_{|x| \geq \varepsilon} \widehat{g}(x) f(x+y) dx = 0 \quad (\text{VII.74})$$

Par ailleurs, puisque $f \in \mathcal{S}(\mathbb{R})$, on peut dire que $|f(x+y) - f(y)| \leq K|x|$ avec $K = \sup_{x \in \mathbb{R}} |f'(x)|$; d'où :

$$\begin{aligned} \left| \int_{-\varepsilon}^{+\varepsilon} f(x+y) \widehat{g}(x) dx - 2\pi f(y) \right| &= \left| \int_{-\varepsilon}^{+\varepsilon} f(x+y) \widehat{g}(x) dx - \int_{-\varepsilon}^{+\varepsilon} f(y) \widehat{g}(x) dx \right| \\ &\leq \int_{-\varepsilon}^{+\varepsilon} |f(x+y) - f(y)| \widehat{g}(x) dx \\ &\leq \int_{-\varepsilon}^{+\varepsilon} K|x| \widehat{g}(x) dx \leq 4K\pi\varepsilon \end{aligned} \quad (\text{VII.75})$$

sachant que $\int_{\mathbb{R}} \widehat{g}(x) dx = 2\pi$. Récapitulons l'ensemble du raisonnement :

– pour tout $\varepsilon > 0$:

$$\lim_{\alpha \rightarrow 0} \int_{|x| \geq \varepsilon} f(x+y) \widehat{g}(x) dx = 0 \quad (\text{VII.76})$$

– pour tout α et tout ε :

$$\left| \int_{|x| \leq \varepsilon} f(x+y) \widehat{g}(x) dx - 2\pi f(y) \right| \leq 4K\pi\varepsilon \quad (\text{VII.77})$$

Comme on peut prendre ε aussi petit qu'on veut, cela entraîne que :

$$\lim_{\alpha \rightarrow 0} \int_{-\infty}^{+\infty} f(x+y) \widehat{g}(x) dx = 2\pi f(y) \quad (\text{VII.78})$$

On a ainsi prouvé que la limite du membre de gauche dans (VII.65) est $\int \widehat{f}(\xi) e^{-iy\xi} d\xi$ et celle du membre de droite $2\pi f(y)$. •

VII.4 Notions de convergence

Dans les sections précédentes, on a introduit des *espaces* de fonctions : $\mathbb{L}_1(\mathbb{R})$, $\mathbb{C}_0(\mathbb{R})$, $\mathcal{S}(\mathbb{R})$. Ces espaces sont des espaces vectoriels : si deux fonctions f et g sont dans $\mathbb{L}_1(\mathbb{R})$, on a aussi :

$$\begin{aligned} \int_{-\infty}^{+\infty} |f(x) + g(x)| dx &\leq \int_{-\infty}^{+\infty} |f(x)| + |g(x)| dx = \\ &= \int_{-\infty}^{+\infty} |f(x)| dx + \int_{-\infty}^{+\infty} |g(x)| dx \end{aligned} \quad (\text{VII.79})$$

et de même :

$$\int_{-\infty}^{+\infty} |\lambda f(x)| dx = |\lambda| \int_{-\infty}^{+\infty} |f(x)| dx \quad (\text{VII.80})$$

On vérifie la même chose pour les deux autres espaces.

Ces espaces vectoriels sont de dimension infinie : il suffit de constater par exemple que les fonctions $\varphi_n(x) = x^n e^{-x^2}$ ou $\chi_a(x) = e^{-x^2 + iax}$ sont linéairement indépendantes : si *pour tout* x $\sum_{k=0}^n \lambda_k \varphi_k(x) = 0$, alors les λ_k sont tous nuls, ou encore, pour toute famille finie $\{a_0, a_1, a_2, \dots, a_n\}$ de nombres réels, on a :

$$\sum_{k=0}^n \lambda_k \chi_{a_k}(x) = 0 \quad \Rightarrow \quad \forall k, \quad \lambda_k = 0 \quad (\text{VII.81})$$

Pour pouvoir parler de la convergence de suites ou de séries de fonctions, on doit préciser ce qu'on entend par *convergent* ou *divergent*. Or, il n'y a pas une seule notion de convergence, mais beaucoup ; et de plus, comme on va le voir, il n'est pas possible de trouver une notion universelle de convergence, qui dans chaque cas serait toujours la meilleure. On rencontrera différents problèmes, pour chacun desquels s'imposera un type différent de convergence :

espace $\mathbb{L}_1(\mathbb{R})$ — la notion naturelle de convergence est la convergence dite *en moyenne* : on dit qu'une suite f_n de fonctions de $\mathbb{L}_1(\mathbb{R})$ converge *en moyenne* vers la fonction f , également dans $\mathbb{L}_1(\mathbb{R})$, si :

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} |f_n(x) - f(x)| dx = 0 \quad (\text{VII.82})$$

espace $\mathbb{C}_0(\mathbb{R})$ — la notion naturelle de convergence est la convergence dite *uniforme* : on dit qu'une suite f_n de fonctions de $\mathbb{C}_0(\mathbb{R})$ converge *uniformément* vers la fonction f , également dans $\mathbb{C}_0(\mathbb{R})$, si :

$$\lim_{n \rightarrow \infty} \left[\sup_{x \in \mathbb{R}} |f_n(x) - f(x)| \right] = 0 \quad (\text{VII.83})$$

espace $\mathcal{S}(\mathbb{R})$ — la notion naturelle de convergence est la suivante (elle n'a pas de nom consacré) : on dit qu'une suite f_n de fonctions de $\mathcal{S}(\mathbb{R})$ converge *dans* $\mathcal{S}(\mathbb{R})$ ou *au sens de* $\mathcal{S}(\mathbb{R})$, ou encore *au sens de Schwartz*, vers la fonction f , également dans $\mathcal{S}(\mathbb{R})$, si pour tout couple d'entiers $j, k \geq 0$:

$$\lim_{n \rightarrow \infty} \left[\sup_{x \in \mathbb{R}} \{ (1 + |x|^k) |f_n^{(j)}(x) - f^{(j)}(x)| \} \right] = 0 \quad (\text{VII.84})$$

La convergence dans $\mathcal{S}(\mathbb{R})$ est une notion très forte de convergence : f_n tend vers f si toutes les dérivées de f_n , multipliées par n'importe quelle puissance de $|x|$, convergent toutes à la fois uniformément.

$\mathcal{S}(\mathbb{R})$ est un sous-espace de $\mathbb{C}_0(\mathbb{R})$ et il est clair que si f_n converge vers f dans $\mathcal{S}(\mathbb{R})$, alors f_n converge vers f dans $\mathbb{C}_0(\mathbb{R})$ (c'est-à-dire uniformément). D'autre part, $\mathcal{S}(\mathbb{R})$ est aussi un sous-espace de $\mathbb{L}_1(\mathbb{R})$ et il est clair aussi que si f_n converge vers f dans $\mathcal{S}(\mathbb{R})$, alors f_n converge vers f dans $\mathbb{L}_1(\mathbb{R})$ (c'est-à-dire en moyenne).

Par contre, si f_n et f sont dans $\mathbb{L}_1(\mathbb{R}) \cap \mathbb{C}_0(\mathbb{R})$, la convergence en moyenne n'entraîne pas la convergence uniforme, ni la convergence uniforme, la convergence en moyenne ⁽⁴⁾.

Appliquons ces nouvelles notions à la transformation intégrale de Fourier.

Théorème VII.5 Si f_n converge vers f dans $\mathbb{L}_1(\mathbb{R})$, alors $\widehat{f}_n$ converge vers $\widehat{f}$ dans $\mathbb{C}_0(\mathbb{R})$. Si f_n converge vers f dans $\mathcal{S}(\mathbb{R})$, alors $\widehat{f}_n$ converge vers $\widehat{f}$ dans $\mathcal{S}(\mathbb{R})$.

Autrement dit, la transformation intégrale de Fourier est une application *continue* de $\mathbb{L}_1(\mathbb{R})$ dans $\mathbb{C}_0(\mathbb{R})$ et de $\mathcal{S}(\mathbb{R})$ dans lui-même.

Preuve En utilisant l'inégalité de la moyenne, on obtient :

$$|\widehat{f}_n(\xi) - \widehat{f}(\xi)| = \left| \int_{-\infty}^{+\infty} [f_n(x) - f(x)] e^{ix\xi} dx \right| \leq \int_{-\infty}^{+\infty} |f_n(x) - f(x)| dx \quad (\text{VII.85})$$

ce qui prouve la première partie du théorème. Ensuite, l'inégalité de la moyenne et les théorèmes VII.3 et VII.4 conduisent aux inégalités suivantes :

$$\begin{aligned} |\xi|^k |\widehat{f}_n(\xi) - \widehat{f}(\xi)| &\leq \int_{-\infty}^{+\infty} |f_n^{(k)}(x) - f^{(k)}(x)| dx \\ |\widehat{f}_n^{(j)}(\xi) - \widehat{f}^{(j)}(\xi)| &\leq \int_{-\infty}^{+\infty} |x|^j |f_n(x) - f(x)| dx \\ |\xi|^k |\widehat{f}_n^{(j)}(\xi) - \widehat{f}^{(j)}(\xi)| &\leq \int_{-\infty}^{+\infty} \left| \frac{d^k}{dx^k} \left\{ x^j [f_n(x) - f(x)] \right\} \right| dx \end{aligned} \quad (\text{VII.86})$$

La dernière expression se majore encore avec la formule de Leibniz :

$$\leq \sum_{\ell=0}^k K_\ell \int_{-\infty}^{+\infty} |x|^{j-\ell} |f_n^{(\ell)}(x) - f^{(\ell)}(x)| dx \quad (\text{VII.87})$$

la somme ne portant (au cas où $j < k$) que sur les $\ell \leq j$; en combinant tout cela :

$$(1 + |\xi|^k) |\widehat{f}_n^{(j)}(\xi) - \widehat{f}^{(j)}(\xi)| \leq \int_{-\infty}^{+\infty} |x|^j |f_n(x) - f(x)| dx + \sum_{\ell=0}^j K_\ell \int_{-\infty}^{+\infty} |x|^{j-\ell} |f_n^{(\ell)}(x) - f^{(\ell)}(x)| dx \quad (\text{VII.88})$$

(4) Toutefois, si on considérait l'espace $\mathbb{L}_1([a; b])$, où l'intervalle borné $[a; b]$ remplace l'intervalle $]-\infty; \infty[$, la convergence uniforme entraînerait la convergence en moyenne.

Si $f_n - f$ tend vers zéro dans $\mathcal{S}(\mathbb{R})$, il existe, pour n'importe quelle paire d'entiers $j, \ell \geq 0$, une suite numérique $M_n^{(j, \ell)}$, qui tend vers zéro et telle que $|f_n^{(\ell)} - f^{(\ell)}| \leq M_n^{(j, \ell)} / (1 + |x|^{j+2})$; la dernière inégalité ci-dessus donne alors :

$$(1 + |\xi|^k) |\widehat{f_n^{(j)}}(\xi) - \widehat{f^{(j)}}(\xi)| \leq M_n^{(j, 0)} \int_{-\infty}^{+\infty} \frac{|x|^j}{1 + |x|^{j+2}} dx + \sum K_\ell M_n^{(j, \ell)} \int_{-\infty}^{+\infty} \frac{|x|^{j-\ell}}{1 + |x|^{j+2}} dx \quad (\text{VII.89})$$

où il est évident que les termes du second membre tendent vers zéro quand n tend vers l'infini. •

On voit que la continuité de la transformation de Fourier $f \mapsto \widehat{f}$ s'obtient par un argument simple parce que les notions de convergence retenues correspondent exactement aux propriétés de l'intégrale. On ne pourrait pas obtenir la continuité de façon aussi simple (on ne l'obtiendrait d'ailleurs pas davantage par des voies compliquées) si par exemple on considérait la transformation comme définie sur $E = \mathbb{L}_1(\mathbb{R}) \cap \mathbb{C}_0(\mathbb{R})$ et à valeurs dans $F = \mathbb{C}_0(\mathbb{R})$ et en considérant dans E et F la convergence uniforme. Tout cela montre que le choix des *bonnes* notions de convergence est essentiel.

Une autre question se pose encore : *la transformation de Fourier est-elle injective, surjective etc ?* On peut déduire facilement du théorème VII.4 que la transformation de Fourier est une bijection de $\mathcal{S}(\mathbb{R})$ sur lui-même. L'existence d'une formule d'inversion le prouve : si $\widehat{f} = 0$, alors $f(x) = \frac{1}{2\pi} \int \widehat{f}(\xi) e^{-ix\xi} d\xi$ est forcément nul pour tout x , donc la transformation est injective; et si $h \in \mathcal{S}(\mathbb{R})$, il est clair qu'en posant $f(x) = \frac{1}{2\pi} \int h(\xi) e^{-ix\xi} d\xi$ on aura automatiquement $h = \widehat{f}$; comme d'après les théorèmes VII.1 et VII.2, f est dans $\mathcal{S}(\mathbb{R})$, la transformation est aussi surjective.

Par contre on ne peut pas déduire de la formule d'inversion que $f \mapsto \widehat{f}$ est une bijection de $\mathbb{L}_1(\mathbb{R})$ dans $\mathbb{C}_0(\mathbb{R})$. D'abord, cette formule n'a pas été prouvée pour $f \in \mathbb{L}_1(\mathbb{R})$ mais seulement si f est une bonne fonction, c'est-à-dire si $f \in \mathcal{S}(\mathbb{R})$. D'autre part, il est évident que les éléments de $\mathbb{C}_0(\mathbb{R})$ ne sont pas tous intégrables et on ne voit pas a priori comment étendre la formule d'inversion à $\mathbb{C}_0(\mathbb{R})$. Nous verrons plus loin qu'on peut cependant étendre la transformation de Fourier à des fonctions non intégrables. Mais il se trouve que même si on introduit ces extensions, il reste des éléments de $\mathbb{C}_0(\mathbb{R})$ qui ne sont l'image d'aucun élément de $\mathbb{L}_1(\mathbb{R})$. C'est-à-dire que la non surjectivité est *intrinsèque* et non due simplement à une insuffisance dans les définitions. Il y a donc un sous-espace (strict) de $\mathbb{C}_0(\mathbb{R})$ qui est l'image de $\mathbb{L}_1(\mathbb{R})$ par la transformation de Fourier.

Ces questions relatives à l'espace $\mathbb{L}_1(\mathbb{R})$ sont *sans intérêt* pour une formation d'ingénieur et ne sont mentionnées ici que pour les curieux. L'espace $\mathcal{S}(\mathbb{R})$ est beaucoup plus simple et suffit pour traiter les problèmes théoriques utiles. Si on veut traiter effectivement les questions relatives à l'espace $\mathbb{L}_1(\mathbb{R})$, on ne peut se contenter de la définition vague qui en a été donnée ici — fonctions *intégrables* — et on doit se placer dans le cadre de la théorie de l'intégrale de Lebesgue, qui donne de l'espace $\mathbb{L}_1(\mathbb{R})$ une définition rigoureuse et opératoire.

VII.5 Espace $\mathbb{L}_2(\mathbb{R})$

Définition VII.3 On désigne par $\mathbb{L}_2(\mathbb{R})$ l'espace vectoriel des fonctions *de carré intégrable* sur $] -\infty ; \infty[$, c'est-à-dire les fonctions $f(x)$ telles que l'intégrale $\int_{-\infty}^{+\infty} |f(x)|^2 dx$ converge.

Cet espace est le plus important de l'analyse fonctionnelle. Son importance est encore renforcée par le rôle essentiel qu'il joue en mécanique quantique. Comme l'espace $\mathbb{L}_1(\mathbb{R})$, il ne peut être défini d'une manière précise et opératoire que dans le cadre de la théorie de l'intégrale de Lebesgue.

Théorème VII.6 — Inégalité de Schwartz Si f et g sont deux éléments de $\mathbb{L}_2(\mathbb{R})$, alors leur produit $f \cdot g$ est élément de $\mathbb{L}_1(\mathbb{R})$ et leur somme $f + g$ est élément de $\mathbb{L}_2(\mathbb{R})$ et par conséquent $\mathbb{L}_2(\mathbb{R})$ est un espace vectoriel. En outre, on a l'inégalité :

$$\left| \int_{-\infty}^{+\infty} f(x) g(x) dx \right| \leq \sqrt{\int_{-\infty}^{+\infty} |f(x)|^2 dx} \cdot \sqrt{\int_{-\infty}^{+\infty} |g(x)|^2 dx} \quad (\text{VII.90})$$

Preuve Pour montrer en toute rigueur que $f \cdot g$ est élément de $\mathbb{L}_1(\mathbb{R})$, il faudrait avoir défini $\mathbb{L}_2(\mathbb{R})$ dans le cadre de la théorie de Lebesgue. C'est pourquoi ce résultat sera admis. On en déduit que ⁽⁵⁾ :

$$\begin{aligned} \int_{-\infty}^{+\infty} |f(x) + g(x)|^2 dx &= \int_{-\infty}^{+\infty} |f(x)|^2 dx + \int_{-\infty}^{+\infty} |g(x)|^2 dx + 2 \int_{-\infty}^{+\infty} f(x) g(x) dx \leq \\ &\leq \int_{-\infty}^{+\infty} |g(x)|^2 dx + 2 \int_{-\infty}^{+\infty} |f(x) g(x)| dx \leq (\|f\|_2 + \|g\|_2)^2 \end{aligned} \quad (\text{VII.91})$$

Les normes de f et g sont finies d'après les hypothèses et le résultat admis, donc l'intégrale du premier membre est finie. Comme il est par ailleurs évident que si $\int |f(x)|^2 dx$ converge, il en est de même de $\int |\lambda f(x)|^2 dx$, on a prouvé par là que $\mathbb{L}_2(\mathbb{R})$ est bien un espace vectoriel.

Enfin, on remarque que l'intégrale convergente $\int |\lambda f(x) + g(x)|^2 dx$ est toujours positive. Or :

$$\int_{-\infty}^{+\infty} |\lambda f(x) + g(x)|^2 dx = \lambda^2 \int_{-\infty}^{+\infty} |f(x)|^2 dx + 2\lambda \int_{-\infty}^{+\infty} f(x) g(x) dx + \int_{-\infty}^{+\infty} |g(x)|^2 dx \quad (\text{VII.92})$$

La condition pour que cela soit positif pour tout λ est que le discriminant de cette expression du second degré en λ soit négatif, ce qui donne exactement l'inégalité (VII.90). •

Sur $\mathbb{L}_2(\mathbb{R})$, on considère, tout comme sur $\mathbb{L}_1(\mathbb{R})$, $\mathbb{C}_0(\mathbb{R})$ et $\mathcal{S}(\mathbb{R})$, une notion de convergence spécifique, appelée *convergence en moyenne quadratique* : une suite f_n de fonctions de carré intégrable converge vers une fonction f également de carré intégrable si :

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} |f_n(x) - f(x)|^2 dx = 0 \quad (\text{VII.93})$$

Pour exprimer commodément ces nouvelles notions de convergence, on introduit pour tout $f \in \mathbb{L}_1(\mathbb{R})$ la notation :

$$\|f\|_1 = \int_{-\infty}^{+\infty} |f(x)| dx \quad (\text{VII.94})$$

et pour tout $f \in \mathbb{L}_2(\mathbb{R})$:

$$\|f\|_2 = \sqrt{\int_{-\infty}^{+\infty} |f(x)|^2 dx} \quad (\text{VII.95})$$

Avec ces nouvelles notations, l'inégalité de Schwartz (VII.90) s'écrit :

$$\|f \cdot h\|_1 \leq \|f\|_2 \cdot \|h\|_2 \quad (\text{VII.96})$$

(5) Ces inégalités correspondent au cas où f et g sont des fonctions à valeurs réelles ; mais si elles sont complexes, il suffit de remplacer $2 \int f(x)g(x) dx$ par $\int \bar{f}(x)g(x) + f(x)\bar{g}(x) dx$ pour que toutes ces inégalités restent valables.

Théorème VII.7 — Inégalité de Minkowski Si f et g sont deux éléments de $\mathbb{L}_2(\mathbb{R})$, alors :

$$\|f + h\|_2 \leq \|f\|_2 + \|h\|_2 \quad (\text{VII.97a})$$

$$\|f - h\|_2 \geq \left| \|f\|_2 - \|h\|_2 \right| \quad (\text{VII.97b})$$

Preuve Cela se déduit facilement de l'inégalité de Schwartz (VII.90). En effet :

$$|f(x) + h(x)|^2 = |f(x)|^2 + |h(x)|^2 + f(x)\overline{h(x)} + \overline{f(x)}h(x) \quad (\text{VII.98})$$

soit, en passant aux intégrales :

$$\int_{-\infty}^{+\infty} |f(x) + h(x)|^2 dx = \int_{-\infty}^{+\infty} |f(x)|^2 dx + \int_{-\infty}^{+\infty} |h(x)|^2 dx + \int_{-\infty}^{+\infty} f(x)\overline{h(x)} dx + \int_{-\infty}^{+\infty} \overline{f(x)}h(x) dx \quad (\text{VII.99})$$

Les deux derniers termes du membre de droite se majorent par l'inégalité de Schwartz, ce qui donne :

$$\int_{-\infty}^{+\infty} |f(x) + h(x)|^2 dx \leq \|f\|_2^2 + \|h\|_2^2 + 2\|f\|_2 \cdot \|h\|_2 \quad (\text{VII.100})$$

et on reconnaît dans le membre de droite ci-dessus le développement de $[\|f\|_2 + \|h\|_2]^2$. On a ainsi obtenu (VII.97a). Pour avoir (VII.97b), il suffit de remplacer dans (VII.97a), f par $f - h$, puis (si $\|f\|_2 < \|h\|_2$) d'échanger f et h . •

Théorème VII.8 — Formule de Parseval Soient f et h deux bonnes fonctions à valeurs réelles ou complexes, alors :

$$\int_{-\infty}^{+\infty} \widehat{f}(\xi) \overline{\widehat{h}(\xi)} d\xi = 2\pi \int_{-\infty}^{+\infty} f(x) \overline{h(x)} dx \quad (\text{VII.101})$$

La barre de $\widehat{h}(\xi)$ ou $\overline{\widehat{h}(\xi)}$ désigne le nombre complexe conjugué et disparaît si les fonctions sont à valeurs réelles.

Preuve Appelons $\widetilde{f}$ la transformation de Fourier inverse :

$$\widehat{f}(\xi) = \int_{-\infty}^{+\infty} f(x) e^{ix\xi} dx; \quad \widetilde{f}(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(\xi) e^{-ix\xi} d\xi \quad (\text{VII.102})$$

On commence par remarquer que :

$$\overline{\widehat{f}(\xi)} = \int_{-\infty}^{+\infty} \overline{f(x)} e^{-ix\xi} dx = 2\pi \widetilde{\widetilde{f}}(\xi) \quad (\text{VII.103})$$

et inversement :

$$\widetilde{\widetilde{f}}(\xi) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \overline{f(x)} e^{ix\xi} dx = \frac{1}{2\pi} \widehat{\widehat{f}}(\xi) \quad (\text{VII.104})$$

Dans l'identité (VII.65), prenons $y = 0$ et $g = \widetilde{\widetilde{h}} = 2\pi \widetilde{\widehat{h}}$. Alors on aura $\widehat{g} = 2\pi \overline{\widehat{h}}$, puisque $\widetilde{\cdot}$ et $\widehat{\cdot}$ se compensent. •

Pour les transformations de Fourier Φ , $\mathcal{F}_1$, ou $\mathcal{F}_h$ (voir section VII.2), la relation de Parseval prend une forme symétrique :

$$\int_{-\infty}^{+\infty} \Phi f(\xi) \overline{\Phi h(\xi)} d\xi = \int_{-\infty}^{+\infty} f(x) \overline{h(x)} dx \quad (\text{VII.105a})$$

$$\int_{-\infty}^{+\infty} \mathcal{F}_1 f(\xi) \overline{\mathcal{F}_1 h(\xi)} d\xi = \int_{-\infty}^{+\infty} f(x) \overline{h(x)} dx \quad (\text{VII.105b})$$

$$\int_{-\infty}^{+\infty} \mathcal{F}_h f(\xi) \overline{\mathcal{F}_h h(\xi)} d\xi = \int_{-\infty}^{+\infty} f(x) \overline{h(x)} dx \quad (\text{VII.105c})$$

En prenant $h = f$, on obtient les relations très importantes :

$$\|\Phi f\|_2 = \|f\|_2, \quad \|\mathcal{F}_1 f\|_2 = \|f\|_2, \quad \text{et} \quad \|\mathcal{F}_h f\|_2 = \|f\|_2 \quad (\text{VII.106})$$

la grandeur $\|f\|_2$ étant appelée la *norme* de f dans $\mathbb{L}_2(\mathbb{R})$; elle est donnée par :

$$\|f\|_2 = \sqrt{\int_{-\infty}^{+\infty} |f(x)|^2 dx} \quad (\text{VII.107})$$

On peut alors énoncer la relation de Parseval en disant que les transformations Φ , $\mathcal{F}_1$ ou $\mathcal{F}_h$ conservent la norme dans $\mathbb{L}_2(\mathbb{R})$: ce sont des *isométries* de l'espace $\mathbb{L}_2(\mathbb{R})$. Cette façon de dire résulte d'une analogie avec les espaces euclidiens où la norme d'un vecteur est sa longueur. Cette analogie conduit à la théorie des *espaces de Hilbert* que nous aborderons plus loin : un espace de Hilbert est analogue aux espaces euclidiens, sauf que sa dimension est infinie. Les isométries de l'espace euclidien sont les rotations et les symétries. Les transformations $\mathcal{F}_1$ et Φ_1 sont donc en quelque sorte des rotations de l'espace $\mathbb{L}_2(\mathbb{R})$ ⁽⁶⁾.

Si on poursuit cette analogie avec la géométrie euclidienne, on peut dire que la transformation $f \mapsto \hat{f}$, pour laquelle la relation de Parseval s'écrivait $\|\hat{f}\|_2 = \sqrt{2\pi} \|f\|_2$, n'est pas exactement une isométrie, mais une sorte de similitude (composée d'une rotation et d'une homothétie — ici de rapport $\sqrt{2\pi}$).

Théorème VII.9 L'espace $\mathcal{S}(\mathbb{R})$ est dense dans $\mathbb{L}_2(\mathbb{R})$: pour toute fonction f de carré intégrable, il existe une suite de bonnes fonctions f_n qui converge en moyenne quadratique vers f .

Preuve La démonstration n'utilise que des techniques élémentaires, mais est fastidieuse. L'idée est la suivante : on construit à partir de f , supposée donnée *a priori* dans $\mathbb{L}_2(\mathbb{R})$, d'abord les fonctions :

$$g_n(x) = \begin{cases} f(x) & \text{si } |x| \leq n \\ 0 & \text{si } |x| > n \end{cases} \quad (\text{VII.108})$$

puis on pose :

$$f_n(x) = \sqrt{\frac{n}{\pi}} \int_{-\infty}^{+\infty} e^{-n(x-y)^2} g_n(y) dy \quad (\text{VII.109})$$

L'opération ci-dessus est ce qu'on appelle un *lissage par convolution*. On montre alors à l'aide d'intégrations par parties successives, accompagnées d'inégalités de la moyenne et de découpages en morceaux, que les fonctions f_n ainsi construites sont de bonnes fonctions, infiniment dérivables et à décroissance rapide à l'infini ainsi que toutes leurs dérivées. •

Ensuite, avec le même type de techniques, on montre que $\int |f_n(x) - f(x)|^2 dx$ tend bien vers zéro quand n tend vers l'infini. La théorie de l'intégrale de Lebesgue est certes implicite par l'évocation de l'espace $\mathbb{L}_2(\mathbb{R})$, mais la démonstration du théorème n'utilise de cette théorie que des propriétés élémentaires de l'intégrale, telles que l'inégalité de la moyenne, l'inégalité de Schwartz, ou l'intégration par parties.

Le même type de démonstration permet d'établir aussi que :

(6) Nous y reviendrons et nous verrons de façon précise qu'elles s'apparentent bien à des rotations d'angle $\pi/2$.

Théorème VII.10 L'espace $\mathcal{S}(\mathbb{R})$ est dense dans $\mathbb{L}_1(\mathbb{R})$: pour toute fonction f intégrable, il existe une suite de bonnes fonctions f_n qui converge en moyenne vers f .

Par contre $\mathcal{S}(\mathbb{R})$ n'est pas dense dans $\mathbb{C}_0(\mathbb{R})$, l'espace des fonctions continues et bornées. Il est d'ailleurs assez facile de comprendre pourquoi : la convergence dans $\mathbb{C}_0(\mathbb{R})$ est la convergence uniforme sur tout $\mathbb{R}$, ce qui veut dire que si f_n tend vers f dans $\mathbb{C}_0(\mathbb{R})$, $\sup_x |f_n(x) - f(x)|$ tend vers zéro ; or les bonnes fonctions sont toutes nulles à l'infini : il est évident que si par exemple $f = 1$ et $f_n \in \mathcal{S}(\mathbb{R})$ on aura toujours $\sup_x |f_n(x) - f(x)| \geq 1$.

Ce qui est vrai est le résultat suivant :

Théorème VII.11 L'espace $\mathcal{S}(\mathbb{R})$ est dense, pour la convergence uniforme, dans l'espace $\mathbb{C}_{00}(\mathbb{R})$ des fonctions continues et nulles à l'infini : pour toute fonction f continue, et telle que $\lim_{x \rightarrow \pm\infty} f(x) = 0$, il existe une suite de bonnes fonctions f_n qui converge uniformément sur tout $\mathbb{R}$ vers la fonction f .

En réunissant les théorèmes VII.8 et VII.9, on obtient le corollaire suivant, qui est l'une des propriétés les plus remarquables de l'espace $\mathbb{L}_2(\mathbb{R})$:

Théorème VII.12 La transformation de Fourier, définie sur $\mathcal{S}(\mathbb{R})$, se prolonge par continuité à $\mathbb{L}_2(\mathbb{R})$; sur l'espace $\mathbb{L}_2(\mathbb{R})$, ce prolongement est inversible et continu ainsi que son inverse : c'est un isomorphisme. La formule de Parseval se prolonge elle aussi à $\mathbb{L}_2(\mathbb{R})$ tout entier, de sorte que (VII.105) est vrai pour f et h quelconques dans $\mathbb{L}_2(\mathbb{R})$.

Preuve D'après la formule de Parseval du théorème VII.8, si $f \in \mathcal{S}(\mathbb{R})$, on a $\|\widehat{f}\|_2 = \sqrt{2\pi} \|f\|_2$. En effet, il suffit de poser $h = f$ dans (VII.105), puis de prendre la racine carrée des deux membres. Si maintenant f est un élément quelconque de $\mathbb{L}_2(\mathbb{R})$, non nécessairement dans le sous-espace $\mathcal{S}(\mathbb{R})$, on peut dire d'après le théorème VII.9, qu'il existe une suite f_n d'éléments de $\mathcal{S}(\mathbb{R})$ qui tend dans $\mathbb{L}_2(\mathbb{R})$ (c'est-à-dire en moyenne quadratique) vers f . Posons alors par définition $\widehat{f} = \lim_n \widehat{f}_n$. La limite $\widehat{f}$ ne dépend pas de la suite f_n choisie, puisque si on en avait pris une autre g_n (ayant également f pour limite), on aurait d'après la formule de Parseval (VII.105), $\|\widehat{f}_n - \widehat{g}_n\|_2 = \sqrt{2\pi} \|f_n - g_n\|_2$, ce qui tend évidemment vers zéro. Cependant, rien ne prouve *a priori* que cette limite existe ; en fait, l'existence de cette limite est garantie pour la raison suivante : la suite f_n étant convergente dans $\mathbb{L}_2(\mathbb{R})$, est automatiquement une suite de Cauchy, c'est-à-dire que $\lim_{n \rightarrow \infty} \sup_{m \geq n} \|f_m - f_n\|_2 = 0$. D'après la relation de Parseval (VII.105), la suite $\widehat{f}_n$ est alors aussi une suite de Cauchy, puisque $\|\widehat{f}_m - \widehat{f}_n\|_2 = \sqrt{2\pi} \|f_m - f_n\|_2$. L'existence d'une limite de la suite $\widehat{f}_n$ dans l'espace $\mathbb{L}_2(\mathbb{R})$ résulte d'une propriété qu'on démontre dans la théorie de Lebesgue, à savoir que l'espace $\mathbb{L}_2(\mathbb{R})$ est complet. •

On a ainsi prolongé par continuité la transformation $f \mapsto \widehat{f}$. Puisque le sous-espace $\mathcal{S}(\mathbb{R})$ est dense dans $\mathbb{L}_2(\mathbb{R})$, ce prolongement par continuité s'étend à $\mathbb{L}_2(\mathbb{R})$ tout entier.

Pour prouver que la formule de Parseval se prolonge aussi, supposons d'abord que les fonctions f et h de (VII.105) sont réelles. On peut alors écrire que :

$$\|f + h\|_2^2 = \|f\|_2^2 + \|h\|_2^2 + 2 \int_{-\infty}^{+\infty} f(x) h(x) dx \quad (\text{VII.110a})$$

$$\|\widehat{f} + \widehat{h}\|_2^2 = \|\widehat{f}\|_2^2 + \|\widehat{h}\|_2^2 + 2 \int_{-\infty}^{+\infty} \widehat{f}(\xi) \widehat{h}(\xi) d\xi \quad (\text{VII.110b})$$

Ces égalités montrent qu'il suffit de prouver que :

$$\|\widehat{f}\|_2 = \sqrt{2\pi} \|f\|_2 \quad (\text{VII.111})$$

En effet, si cela est vrai pour tout f , ce sera vrai aussi pour f , h et $f + h$ et (VII.105) résultera de (VII.110a) et (VII.110b).

Or, d'après l'inégalité de Minkowski (VII.97b), $\lim f_n = f$ entraîne automatiquement que $\lim \|f_n\|_2 = \|f\|_2$ et de même $\lim \widehat{f_n} = \widehat{f}$ entraîne que $\lim \|\widehat{f_n}\|_2 = \|\widehat{f}\|_2$. Si donc la relation (VII.111) est vraie pour les f_n et les $\widehat{f_n}$, qui sont dans $\mathcal{S}(\mathbb{R})$, elle sera vraie aussi pour les limites.

Enfin, si f et h sont complexes, on se ramène au cas qui vient d'être traité en séparant la partie réelle et la partie imaginaire.

La formule d'inversion (VII.69) du théorème VII.4 ne peut pas se prolonger telle quelle, puisque si f est un élément quelconque de $\mathbb{L}_2(\mathbb{R})$, l'intégrale $\int \widehat{f}(\xi) e^{-ix\xi} d\xi$ n'est pas forcément définie. On procède alors ainsi : soit f un élément quelconque de $\mathbb{L}_2(\mathbb{R})$ et g_n une suite d'éléments de $\mathcal{S}(\mathbb{R})$ qui converge (en moyenne quadratique) vers $\widehat{f}$. La formule d'inversion peut s'écrire :

$$f(x) = \lim_{n \rightarrow \infty} \frac{1}{2\pi} \int_{-\infty}^{+\infty} g_n(\xi) e^{-ix\xi} d\xi \quad (\text{VII.112})$$

Il est facile de voir que cela est vrai quelle que soit la suite g_n choisie : d'après la relation de Parseval précédemment étendue à $\mathbb{L}_2(\mathbb{R})$ tout entier, $\|\widetilde{g_n} - f\|_2 = \sqrt{2\pi} \|g_n - \widehat{f}\|_2$ (où $\widetilde{g_n}$ désigne la transformée inverse de g_n). Cela montre bien que si g_n tend vers $\widehat{f}$, alors $\widetilde{g_n}$ tend vers f .

Pour que la démonstration soit intellectuellement honnête, il convient encore d'approfondir un aspect de la question qui est resté dans l'ombre parce que l'espace $\mathbb{L}_2(\mathbb{R})$ n'a pas été construit devant vous, mais renvoyé à la mystérieuse théorie de H. Lebesgue. Revenons à la question de la bijectivité de la transformation de Fourier étendue à $\mathbb{L}_2(\mathbb{R})$. Dans la démonstration ci-dessus, nous avons fait comme si la bijectivité était établie automatiquement par la simple existence de la formule d'inversion étendue (VII.112), ce qui est certes correct, mais cache un point subtil. Pour être bijective, il faut d'abord que la transformation de Fourier étendue à $\mathbb{L}_2(\mathbb{R})$ soit injective ; ce qui signifie que si $\widehat{f} = \widehat{g}$, alors $f = g$. Il est bien clair que $\widehat{f} = \widehat{g}$ entraîne $\|\widehat{f} - \widehat{g}\|_2 = 0$: cela résulte simplement de ce que l'intégrale d'une fonction nulle est nulle. Notre démonstration est entièrement basée sur la formule de Parseval, d'après laquelle, si $\|\widehat{f} - \widehat{g}\|_2 = 0$, on aura aussi $\|f - g\|_2 = 0$. Le point qui est resté dans l'ombre est alors celui-ci : comment pouvons-nous affirmer que $f = g$, alors que la relation de Parseval permet seulement d'obtenir $\|f - g\|_2 = 0$? L'intégrale du carré d'une fonction nulle est nulle, mais la réciproque de cette affirmation est fautive : si par exemple une fonction $h(x)$ est telle que $\int |h(x)|^2 dx = 0$, on ne peut pas conclure que $h(x)$ est nulle en tout point : si $h(x)$ est nulle partout *sauf en un nombre fini ou discret de points*, son intégrale sera nulle et pourtant on ne pourra pas dire que $\forall x, h(x) = 0$. Dans la théorie de Lebesgue, il existe un concept spécial pour cela : un ensemble fini ou discret de points est un *ensemble négligeable* ou *ensemble de mesure nulle* (chapitre II, p. 17). La notion d'intégrale admise dans ce cours est trop vague pour permettre une définition précise, rigoureuse et opératoire des ensembles négligeables. Lorsque $\|f - g\|_2 = 0$, on ne peut donc pas conclure quelque chose de précis. Dans la théorie de Lebesgue, on peut démontrer rigoureusement les deux théorèmes suivants :

1. Si $\int |f(x)| dx = 0$ ou $\int |f(x)|^2 dx = 0$, alors il existe un ensemble négligeable en dehors duquel f est nulle.
2. Entre deux fonctions f et g , la relation : “il existe un ensemble négligeable en dehors duquel $f(x) = g(x)$ ” est une relation d'équivalence.

Les espaces $\mathbb{L}_1(\mathbb{R})$ et $\mathbb{L}_2(\mathbb{R})$ sont en réalité des espaces quotients par cette relation d'équivalence. Ainsi, quand on dit que $f \in \mathbb{L}_2(\mathbb{R})$, on sous-entend que f n'est pas exactement une fonction définie en chaque point de l'intervalle $\mathbb{R}$, mais une fonction définie en *presque* tous les points. Par exemple, la fonction $J(x)$, égale à 1 si x est rationnel et à 0 si x est irrationnel, ne se distingue pas de la fonction égale à zéro partout.

En conclusion, le théorème VII.10 dit très exactement ceci : si deux fonctions f et g , définies en tout point de $\mathbb{R}$ et de carrés intégrables, sont telles que $\hat{f} = \hat{g}$, alors $f - g$ est nulle *presque partout*, i.e. en dehors d'un ensemble négligeable.

La surjectivité de la transformation de Fourier étendue à $\mathbb{L}_2(\mathbb{R})$ exige une interprétation analogue : si $f(x)$ est une fonction définie en tout point x de $\mathbb{R}$ et de carré intégrable, il existe une suite f_n d'éléments de $\mathcal{S}(\mathbb{R})$ qui tend vers f *en moyenne quadratique* ; la suite $\widetilde{f_n}$ est alors une suite d'éléments de $\mathcal{S}(\mathbb{R})$ qui a une limite dans $\mathbb{L}_2(\mathbb{R})$ (car c'est une suite de Cauchy et que $\mathbb{L}_2(\mathbb{R})$ est complet). Mais cela ne signifie pas que pour tout x , la suite numérique $\widetilde{f_n}(x)$ a une limite dans $\mathbb{R}$: on peut seulement affirmer cela en dehors d'un certain ensemble négligeable, dont la théorie de Lebesgue garantit l'existence. De sorte que la limite de $\widetilde{f_n}$ est une fonction définie *presque* partout, au sens de Lebesgue et non absolument partout. Cette limite est alors l'élément de $\mathbb{L}_2(\mathbb{R})$ dont la transformée de Fourier est f , ce qui prouve la surjectivité.

VII.6 Transformation de Laplace

Une autre transformation, très importante pour les applications électroniques par exemple, est la *transformation de Laplace*, qui s'apparente à la transformation de Fourier. étant donnée une fonction $f(t)$ définie sur l'intervalle $[0; \infty[$, on appelle transformée de Laplace de f la fonction :

$$F(z) = \int_0^{+\infty} f(t) e^{-zt} dt \quad (\text{VII.113})$$

Cette définition est correcte si l'intégrale converge ; pour garantir cela, on ne considérera que des fonctions f à croissance au plus exponentielle, c'est-à-dire des fonctions f pour lesquelles il existe des constantes positives M et A telles que $\forall t \geq 0, |f(t)| \leq M e^{At}$. Sous ces conditions, on peut affirmer que la transformée de Laplace $F(z)$ est une fonction analytique de z dans le demi-plan $\Re(z) > A$. Pour $\Re(z) \leq 0$, l'intégrale sera en général divergente, mais cela n'interdit pas que $F(z)$ puisse avoir un prolongement analytique au-delà du demi-plan $\Re(z) > A$. En tous cas, il est clair que la fonction $f(t) = e^{t^2}$ par exemple, n'a pas de transformée de Laplace.

Voyons des exemples :

Exemple VII.3 — $f(t) = t^{\alpha-1}$ Cette fonction est à croissance au plus exponentielle. D'après ce qui a été vu au chapitre V, sa transformée de Laplace est $F(z) = [z^{-\alpha}]_2$. Cette fonction

est analytique dans le domaine Ω_2 ($\mathbb{C}$ privé de la demi-droite $]-\infty; 0]$) mais l'intégrale $\int t^{\alpha-1} e^{-zt} dt$ n'est convergente que pour $\Re(z) > 0$ et par conséquent ne représente une fonction analytique que dans ce demi-plan. La fonction $[z^{-\alpha}]_2$ est donc un prolongement analytique au-delà de ce demi-plan. $\square$

Exemple VII.4 — $f(t) = e^{at}$ a étant un nombre complexe quelconque. En prenant $M = 1$ et $A = \Re a$ on a bien l'inégalité $|f(t)| \leq M e^{At}$ et on s'attend donc à ce que $F(z)$ soit analytique dans le demi-plan $\Re(z) > A$. Le calcul direct donne $F(z) = 1/(z - a)$; cette fonction est en effet analytique dans le demi-plan $\Re(z) > A$, mais se prolonge à $\mathbb{C} - \{a\}$. $\square$

Exemple VII.5 — $f(t) = 1/(1+t)$ Pour z réel > 0 , un changement de variable simple montre que :

$$F(z) = \frac{1}{z} \int_0^{+\infty} \frac{e^{-s}}{1 + \frac{s}{z}} ds = \frac{1}{z} \text{Eu}\left(\frac{1}{z}\right) \quad (\text{VII.114})$$

où $\text{Eu}(w)$ désigne la fonction d'Euler dans la section IV.5 (qui avait été désignée alors par $F(w)$, mais cela n'est plus possible ici). On avait vu en IV.5 que la fonction $\text{Eu}(w)$ se prolonge au domaine Ω_2 ; il en est donc de même pour $F(z) = \frac{1}{z} \text{Eu}\left(\frac{1}{z}\right)$, puisque la transformation $z \mapsto 1/z$ transforme Ω_2 en lui-même. On voit donc une fois de plus que, bien que l'intégrale $\int f(t) e^{-zt} dt$ diverge pour $\Re(z) < 0$, la fonction $F(z)$ se prolonge analytiquement au-delà du demi-plan $\Re(z) > 0$. $\square$

Il y a une parenté entre la transformation de Laplace et la transformation de Fourier. En effet, soit $f(t)$ une fonction définie sur l'intervalle $]0; \infty[$ et $F(z)$ sa transformée de Laplace, définie et analytique au moins dans le demi-plan $\Re(z) > A$. Posons alors :

$$\varphi(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ f(t) & \text{si } t > 0 \end{cases} \quad (\text{VII.115})$$

Il est clair que la transformée de Fourier $\widehat{\varphi}(\xi)$ est égale à $F(-i\xi)$. Autrement dit, les valeurs de la transformée de Laplace $F(z)$ le long de la droite $\Re(z) = 0$ représentent la transformée de Fourier de la fonction $\varphi(x)$; plus généralement, les valeurs de $F(z)$ le long de la droite $\Re(z) = a$ représentent la transformée de Fourier de la fonction $\varphi(x) e^{-ax}$; en effet :

$$F(a + i\xi) = \int_{-\infty}^{+\infty} \varphi(t) e^{-at} e^{-it\xi} dt \quad (\text{VII.116})$$

Pour la commodité, on a pris ici, la transformation avec $e^{-ix\xi}$. D'après la formule d'inversion, on en déduit que :

$$\varphi(t) e^{-at} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(a + i\xi) e^{it\xi} d\xi \quad (\text{VII.117})$$

ou encore :

$$\varphi(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(a + i\xi) e^{t(a+i\xi)} d\xi \quad (\text{VII.118})$$

On peut interpréter le membre de droite de (VII.118) comme l'intégrale obtenue par paramétrage de :

$$\frac{1}{2i\pi} \int_{\Gamma_a} F(z) e^{tz} dz \quad (\text{VII.119})$$

où Γ_a est le chemin (infini) constitué par la droite $\Re(z) = a$ parcourue du bas vers le haut.

Puisque $F(z)$ est analytique et souvent (comme le montrent les exemples) au-delà du demi-plan $\Re(z) > 0$, on peut dans bien des cas déformer le chemin Γ_a sans changer sa classe d'homologie et utiliser le théorème des résidus pour calculer l'intégrale (VII.119) et par conséquent en déduire la fonction $f(t)$. Autrement dit, l'inversion de la transformation de Laplace est souvent possible par la méthode des résidus et fournit ainsi un outil puissant pour les applications, en électronique, notamment.



VIII — Intégrales divergentes

VIII.1 Calcul d'une intégrale semi-convergente

Dans ce chapitre, on propose de calculer des intégrales divergentes par différentes méthodes. Ces procédés conduiront à la *théorie des distributions* qui fait l'objet du chapitre suivant.

Nous avons vu à propos de la transformation de Fourier que celle-ci, définie au départ pour des fonctions intégrables, c'est-à-dire appartenant à l'espace $\mathbb{L}_1(\mathbb{R})$, pouvait s'étendre à l'espace $\mathbb{L}_2(\mathbb{R})$. Par contre, pour une fonction $f(x)$ appartenant à l'espace $\mathbb{L}_2(\mathbb{R})$ et n'appartenant pas à l'espace $\mathbb{L}_1(\mathbb{R})$, on ne peut définir $\hat{f}$ par une intégrale comme dans les expressions (VII.66). Toutefois, on peut justifier les calculs pour des intégrales semi-convergentes, en prenant simplement des limites.

La fonction :

$$f(x) = \frac{1}{1 + |x|} \quad (\text{VIII.1})$$

est dans $\mathbb{L}_2(\mathbb{R})$ mais pas dans $\mathbb{L}_1(\mathbb{R})$. Cependant, on peut *convenir* que :

$$\hat{f}(\xi) = \int_{-\infty}^{+\infty} \frac{e^{ix\xi}}{1 + |x|} dx \quad (\text{VIII.2})$$

représente la limite pour $A \rightarrow \infty$ de :

$$\hat{f}(\xi) = \lim_{A \rightarrow \infty} \int_{-A}^{+A} \frac{e^{ix\xi}}{1 + |x|} dx \quad (\text{VIII.3})$$

car cette limite est bien définie. Remarquons cependant que l'intégrale (VIII.2) n'est semi-convergente que pour $\xi \neq 0$. Pour $\xi = 0$, elle est vraiment divergente. On peut aussi lui donner un sens par un autre passage à la limite. Conformément à ce qui a été vu au

théorème VII.9, on peut définir $\widehat{f}$ en prenant une suite f_n de fonctions de $\mathbb{L}_1(\mathbb{R})$ qui converge en moyenne *quadratique* vers f , puis poser $\widehat{f} = \lim \widehat{f_n}$. On va donc considérer :

$$f_n(x) = \frac{e^{-\frac{1}{n}|x|}}{1+|x|} \quad (\text{VIII.4})$$

Cette suite est bien dans $\mathbb{L}_1(\mathbb{R})$. Sa transformée de Fourier s'écrit :

$$\widehat{f_n}(\xi) = \int_{-\infty}^{+\infty} \frac{e^{-\frac{1}{n}|x|+ix\xi}}{1+|x|} dx \quad (\text{VIII.5})$$

En découpant l'intégrale (VIII.5) en deux morceaux, l'un de $-\infty$ à 0, dans lequel on fait le changement de variable $x \mapsto -x$ et l'autre de 0 à $+\infty$, on voit facilement que :

$$\widehat{f_n}(\xi) = 2\Re \left(\int_0^{+\infty} \frac{e^{-(\frac{1}{n}-i\xi)x}}{1+x} dx \right) \quad (\text{VIII.6})$$

La fonction :

$$z \mapsto \int_0^{+\infty} \frac{e^{-zx}}{1+x} dx \quad (\text{VIII.7})$$

est analytique dans le demi-plan $\Re(z) > 0$ et pour z réel positif, le changement de variable $y = zx$ dans l'intégrale (VIII.7) la ramène à :

$$\frac{1}{z} \int_0^{+\infty} \frac{e^{-y}}{1+\frac{1}{z}y} dy \quad (\text{VIII.8})$$

où l'on reconnaît la fonction d'Euler (IV.45). Or, cette fonction est, comme nous l'avons vu, analytique pour $\zeta \in \Omega_2$, donc (VIII.8) est aussi analytique dans Ω_2 (la transformation $z \mapsto \zeta = 1/z$ transforme Ω_2 en lui-même). D'autre part (VIII.7) est analytique dans $\{\Re(z) > 0\}$ d'après des théorèmes généraux et coïncide avec (VIII.8) pour z réel strictement positif, donc coïncide avec (VIII.8) dans tout le demi-plan $\{\Re(z) > 0\}$ par prolongement analytique. On peut donc exprimer la transformée de Fourier $\widehat{f_n}$ à l'aide de cette fonction Eu :

$$\widehat{f_n}(\xi) = 2\Re \left(\frac{1}{\frac{1}{n}-i\xi} \text{Eu} \left(\frac{1}{\frac{1}{n}-i\xi} \right) \right) \quad (\text{VIII.9})$$

Ⓡ Le raisonnement suivi de (VIII.6) à (VIII.9) est courant : le changement de variable $y = zx$ n'est possible que pour z réel ; si z était complexe, il ne s'agirait plus d'un changement de variable ; la nouvelle variable y parcourrait un chemin du plan complexe et non plus un intervalle réel et une telle opération pourrait modifier la valeur de l'intégrale. Dans des cas analogues, il faut toujours suivre la méthode que nous avons suivie ici : effectuer le changement de variable pour z réel, puis faire jouer la propriété du prolongement analytique. Il faut alors vérifier que les deux membres de l'égalité à prolonger sont bien analytiques.

Lorsque n tend vers l'infini, la limite de (VIII.9) est :

$$\widehat{f}(\xi) = 2\Re \left(\frac{1}{-i\xi} \text{Eu} \left(\frac{1}{-i\xi} \right) \right) \quad (\text{VIII.10})$$

La limite qui est considérée ici est, pour tout $\xi \neq 0$ fixé (et réel), la limite dans $\mathbb{R}$ de la suite numérique $n \mapsto \widehat{f_n}(\xi)$ et non la limite dans $\mathbb{L}_2(\mathbb{R})$ de la suite $\widehat{f_n}$. Toutefois la limite (VIII.10) définit bien, pour $\xi \neq 0$, une fonction $\widehat{f}$ qui est *aussi* la limite dans $\mathbb{L}_2(\mathbb{R})$ de la suite $\widehat{f_n}$.

L'expression (VIII.10) n'est pas clairement définie pour $\xi = 0$, mais on peut avoir une idée plus précise de la singularité au point $\xi = 0$ si on utilise le développement en série de Laurent (IV.68); celui-ci conduit en effet à :

$$\widehat{f}(\xi) = -2 \Re \left(e^{-i\xi} \left[\ln_2(-i\xi) + \gamma + \sum_{n \geq 1} \frac{(-1)^n}{n n!} (-i\xi)^n \right] \right) \quad (\text{VIII.11})$$

Ce développement montre que la singularité en $\xi = 0$ est logarithmique, donc intégrable et de carré intégrable. L'existence du développement (VIII.11), dont la série est entière et de rayon de convergence infini, garantit qu'il n'y a pas d'autre singularité en dehors de $\xi = 0$: la fonction est continue en tout point $\xi \neq 0$. Pour vérifier qu'à l'infini aussi la fonction $\widehat{f}(\xi)$ est de carré intégrable, il faut connaître son comportement pour $\xi \rightarrow \pm\infty$; mais pour cela on peut revenir à la définition initiale de $\text{Eu}(1/-i\xi)$. On a en effet :

$$\widehat{f}(\xi) = 2 \Re \left(\frac{1}{-i\xi} \text{Eu} \left(\frac{1}{-i\xi} \right) \right) = 2 \Re \left(\int_0^\infty \frac{e^{-t}}{t - i\xi} dt \right) \quad (\text{VIII.12})$$

On en déduit, par l'inégalité de la moyenne, que :

$$|\xi \widehat{f}(\xi)| \leq 2 \int_0^\infty \left| \frac{\xi}{t - i\xi} \right| e^{-t} dt \quad (\text{VIII.13})$$

et comme $|t - i\xi| = \sqrt{t^2 + \xi^2} \geq |\xi|$, on peut en déduire que $|\xi \widehat{f}(\xi)| \leq 2$, ce qui prouve bien qu'à l'infini, $\widehat{f}(\xi)$ est de carré intégrable. En calculant l'intégrale de (VIII.13) par parties, on obtient même :

$$\int_0^\infty \frac{e^{-t}}{t - i\xi} dt = \frac{1}{-i\xi} - \int_0^\infty \frac{e^{-t}}{(t - i\xi)^2} dt \quad (\text{VIII.14})$$

En prenant la partie réelle, le terme $1/-i\xi$ disparaît et les mêmes inégalités conduisent à $|\xi|^2 |\widehat{f}(\xi)| \leq 2$, ce qui prouve que $\widehat{f}$ est non seulement dans $\mathbb{L}_2(\mathbb{R})$ mais même dans $\mathbb{L}_1(\mathbb{R})$.

Pour résumer, la transformée de Fourier de la fonction (VIII.1) est une fonction continue sur $\mathbb{R} - \{0\}$, qui devient infinie en $\xi = 0$, où elle se comporte de manière équivalente à $2 \Re \{\ln_2(-i\xi)\} = 2 \ln(|\xi|)$ et qui pour $\xi \rightarrow \pm\infty$, se comporte de manière équivalente à $2/\xi^2$.

Ⓡ Pour terminer on devrait encore vérifier que la fonction $\widehat{f}(\xi)$, qui a été obtenue ici comme la limite dans $\mathbb{R}$ (pour ξ fixé) de la suite numérique $\widehat{f}_n(\xi)$, est bien identique à la limite dans $\mathbb{L}_2(\mathbb{R})$ de la suite $\widehat{f}_n$. Cela consiste simplement à vérifier que :

$$\int_{-\infty}^{+\infty} \left| 2 \Re \left(\frac{1}{-i\xi} \text{Eu} \left(\frac{1}{-i\xi} \right) \right) - \widehat{f}_n(\xi) \right|^2 d\xi \quad (\text{VIII.15})$$

tend vers zéro quand n tend vers l'infini ; cette vérification est laissée en exercice (utiliser les théorèmes de passage à la limite sous le signe $\int$).

On arrive ainsi à la conclusion que l'intégrale (VIII.2), bien que divergente, a une valeur bien définie, excepté pour $\xi = 0$. Une particularité importante des fonctions dans les espaces $\mathbb{L}_1(\mathbb{R})$ ou $\mathbb{L}_2(\mathbb{R})$ est que les fonctions n'ont pas à être définies partout, mais seulement *presque* partout au sens de la théorie de l'intégration de H. Lebesgue. Il importe donc peu que la valeur en $\xi = 0$ manque.

Cet exemple devait montrer que l'on peut donner un sens précis et rigoureux à des intégrales divergentes. Toutefois, ici, l'intégrale était semi-convergente, donc la divergence

n'était pas trop grave. L'idée essentielle était l'extension de la notion de limite, car l'intégrale divergente (VIII.2) est définie comme une limite dans $\mathbb{L}_2(\mathbb{R})$: c'est la limite des intégrales absolument convergentes (VIII.5) lorsque n tend vers l'infini. Cependant, comme l'intégrale était semi-convergente (si on laisse de côté le cas $\xi = 0$), on aurait pu s'en sortir avec la limite au sens usuel : la limite dans $\mathbb{R}$ de la suite numérique $\widehat{f}_n(\xi)$. Elle donne le même résultat qu'avec la limite dans l'espace $\mathbb{L}_2(\mathbb{R})$, ce qui n'est pas une règle générale. Dans le prochain exemple (section VIII.2), nous verrons qu'on ne peut pas du tout se contenter de la notion usuelle de limite.

Ces phénomènes n'ont commencé à être compris que dans les années 1920. Les mathématiciens des XVIII^e et XIX^e siècles ont recouru fréquemment à des ruses diverses pour donner un sens aux intégrales divergentes mais ils ne comprenaient pas pourquoi cela marchait parfois, mais pas toujours. L'explication était liée à l'existence inconnue d'un espace muni d'une notion favorable de limite. L'étude de ces espaces de fonctions et des notions de limites qui leur sont attachées s'appelle l'*analyse fonctionnelle*. Celle-ci serait restée confinée dans une spécialisation étroite si elle n'avait servi qu'à donner un sens aux intégrales divergentes. La mécanique quantique en a fait son principal outil mathématique, d'où l'importance énorme qu'elle a pris aujourd'hui.

VIII.2 Valeur principale de Cauchy

On sait que la fonction $1/x$ n'est pas intégrable en $x = 0$ (ni d'ailleurs à l'infini). Cela veut dire que pour $a > 0$ et $b > 0$ les intégrales :

$$\int_{-a}^{-\eta} \frac{1}{x} dx \quad \text{et} \quad \int_{+\varepsilon}^{+b} \frac{1}{x} dx \quad (\text{VIII.16})$$

n'ont pas de limite quand η et ε tendent vers zéro.

Toutefois si on prend $\eta = \varepsilon$ et qu'on considère leur somme, celle-ci aura une limite car les deux infinis se compensent :

$$\int_{-a}^{-\eta} \frac{1}{x} dx = \int_{+a}^{+\eta} \frac{1}{x} dx = \ln(\eta) - \ln(a) \quad \text{et} \quad \int_{+\varepsilon}^{+b} \frac{1}{x} dx = \ln(b) - \ln(\varepsilon) \quad (\text{VIII.17})$$

Si $\eta = \varepsilon$, les termes $\ln(\eta)$ et $\ln(\varepsilon)$ s'annulent dans la somme et il reste $\ln(b) - \ln(a)$.

Plus généralement, si $\varphi(x)$ est une fonction différentiable en $x = 0$, on peut donner un sens à l'intégrale :

$$\int_{-a}^{+b} \frac{\varphi(x)}{x} dx \quad (\text{VIII.18})$$

en posant qu'elle est la limite, pour $\varepsilon \rightarrow 0$, de :

$$\int_{-a}^{-\varepsilon} \frac{\varphi(x)}{x} dx + \int_{\varepsilon}^{+b} \frac{\varphi(x)}{x} dx \quad (\text{VIII.19})$$

En effet, la fonction $\varphi(x)$ étant différentiable, peut se décomposer sous la forme $\varphi(x) = \varphi(0) + x\chi(x)$, où $\chi(x)$ est régulière. Pour être précis et fixer les idées : admettons que $\varphi(x)$ est continûment différentiable sur $[-a; +b]$; $\chi(x)$ sera continue sur $[-a; +b]$ et (VIII.19)

devient :

$$\varphi(0) \left[\int_{-a}^{-\varepsilon} \frac{1}{x} dx + \int_{\varepsilon}^b \frac{1}{x} dx \right] + \left[\int_{-a}^{-\varepsilon} \chi(x) dx + \int_{\varepsilon}^b \chi(x) dx \right] \quad (\text{VIII.20})$$

Le second terme entre crochets dans (VIII.20) a pour limite $\int \chi(x) dx$, puisque la fonction $\chi(x)$ est partout régulière. Le premier terme entre crochets est déjà calculé et vaut $\ln(b) - \ln(a)$ indépendamment de ε . La limite est donc :

$$\varphi(0) [\ln(b) - \ln(a)] + \int_{-a}^{+b} \chi(x) dx \quad (\text{VIII.21})$$

On appelle cette limite la *valeur principale de Cauchy* de l'intégrale et on utilise souvent la notation :

$$\text{VP} \int_{-a}^{+b} \frac{\varphi(x)}{x} dx \quad (\text{VIII.22})$$

Lorsqu'on trouve le symbole VP devant une intégrale singulière, cela signifie qu'à l'intervalle d'intégration donné on enlève un intervalle *symétrique* $[x_0 - \varepsilon; x_0 + \varepsilon]$ autour de chaque singularité et qu'on fait tendre ensuite ε vers zéro. Il est bien évident que si on prend des intervalles dissymétriques $[x_0 - \eta; x_0 + \varepsilon]$ et qu'on fait tendre η et ε indépendamment l'un de l'autre vers zéro, il n'y aura pas de limite.

Ce procédé de régularisation des intégrales singulières (autre terme pour divergentes) utilise la compensation des infinis. Le procédé de la section VIII.1 consistait à considérer l'intégrale singulière comme une limite d'intégrales convergentes ; pour la valeur principale de Cauchy, on a aussi utilisé cette approche, puisque l'intégrale singulière est bien une limite. On peut interpréter la méthode précédente d'une manière qui la rapprochera de celle de la section VIII.1. Appelons $f(x) = \varphi(x)/x$ la fonction à intégrer ; puis posons :

$$f_n(x) = \begin{cases} f(x) & \text{si } |x| > \frac{1}{n} \\ 0 & \text{si } |x| \leq \frac{1}{n} \end{cases} \quad (\text{VIII.23})$$

Il est clair que pour tout n entier supérieur ou égal à 1, la fonction f_n est dans l'espace $\mathbb{L}_1([-a; +b])$ et (VIII.19) est l'intégrale de f_n (avec $\varepsilon = \frac{1}{n}$). Autrement dit :

$$\text{VP} \int_{-a}^{+b} f(x) dx = \lim_{n \rightarrow \infty} \int_{-a}^{+b} f_n(x) dx \quad (\text{VIII.24})$$

Les intégrales sous le signe lim sont des intégrales régulières et on obtient ainsi l'intégrale divergente comme limite d'intégrales régulières. Toutefois, dans la section VIII.1, la fonction limite f appartenait à l'espace $\mathbb{L}_2$, où la convergence était bien définie et on pouvait dire que *quelle que soit* la suite f_n qui tend vers f , l'intégrale de f_n a toujours la même limite. Ici, il semble qu'on a choisi une suite particulière et rien ne prouve qu'avec une autre suite f_n on aurait eu la même limite. On peut même constater directement que si au lieu de f_n on avait pris :

$$g_n(x) = \begin{cases} f(x) & \text{si } x < -\frac{1}{n} \text{ ou } x > \frac{2}{n} \\ 0 & \text{si } -\frac{1}{n} \leq x \leq \frac{2}{n} \end{cases} \quad (\text{VIII.25})$$

on n'aurait pas trouvé la même limite. En effet :

$$\int_{-a}^{-1/n} \frac{1}{x} dx = \ln\left(\frac{1}{n}\right) - \ln(a) \quad (\text{VIII.26})$$

et :

$$\int_{2/n}^b \frac{1}{x} dx = \ln(b) - \ln\left(\frac{2}{n}\right) = \ln(b) - \ln(2) - \ln\left(\frac{1}{n}\right) \quad (\text{VIII.27})$$

Lorsqu'on fait la somme, les deux infinis $\ln(1/n)$ se compensent effectivement, mais il reste le terme $\ln(2)$ qui provient de la dissymétrie de l'intervalle. On voit bien que la limite dépend fortement du choix de la suite f_n .

Pour avoir une véritable analogie avec l'exemple étudié en section VIII.1, il faudrait avoir un espace de fonctions qui contient f et sur cet espace une notion de limite qui implique la symétrie de compensation des infinis ; au sens de cette limite, f_n tendrait vers f , mais pas g_n . Or tout cela existe (voir plus loin le chapitre sur les distributions). *Tous les procédés de régularisation d'intégrales divergentes se ramènent à une notion de limite adéquate.*

La suite f_n est dans l'espace $\mathbb{L}_1([-a; +b])$, mais est discontinue ; ce détail n'est cependant pas essentiel et on aurait tout aussi bien pu approcher la fonction f avec la suite :

$$h_n(x) = \frac{x}{\frac{1}{n^2} + x^2} \quad (\text{VIII.28})$$

La fonction $h_n(x)$ a un minimum égal à $-n/2$ en $x = -1/n$ et un maximum égal à $n/2$ en $x = 1/n$. Autrement dit, la fonction h_n passe, sur une distance égale à $2/n$, de $-n/2$ à $n/2$, ce qui pour n grand représente une croissance extrêmement rapide ; mais la fonction reste toujours continue et même infiniment dérivable. Lorsque n tend vers l'infini le minimum tend vers $-\infty$ et le maximum vers $+\infty$, en même temps que leurs abscisses tendent vers zéro, de sorte que $h_n(x)$ tend vers $f(x) = 1/x$.

La fonction $h_n(x)$ est la dérivée de $\frac{1}{2} \ln\left(\frac{1}{n^2} + x^2\right)$, donc :

$$\int_{-a}^{+b} h_n(x) dx = \frac{1}{2} \left[\ln\left(\frac{1}{n^2} + b^2\right) - \ln\left(\frac{1}{n^2} + a^2\right) \right] \quad (\text{VIII.29})$$

ce qui tend bien vers $\ln(b) - \ln(a)$ lorsque n tend vers l'infini. On peut approcher la fonction singulière $f(x) = \frac{1}{x}$ de bien d'autres manières, qui toutes donneront la même limite pour l'intégrale ; le tout est de définir la notion adéquate de limite.

VIII.3 Pseudo-fonctions de Hadamard

Le mathématicien français J. Hadamard (1865–1963) est justement un des précurseurs de l'analyse fonctionnelle. Le problème étudié ici (comment donner un sens, puis calculer les intégrales divergentes) était un de ses sujets de recherche favoris. Son élève L. Schwartz, né en 1915, est l'auteur de la *théorie des distributions* dont une version simplifiée sera présentée plus loin ; c'est lui qui a introduit les espaces de Schwartz $\mathcal{S}(\mathbb{R})$.

J. Hadamard, pour donner un sens à certaines intégrales divergentes, a introduit la notion de pseudo-fonction. On ne présentera pas sa théorie sous forme générale et abstraite mais

on la fera apparaître à travers l'exemple suivant, très utile pour les applications, notamment en traitement du signal où elle est tout particulièrement opératoire. Considérons la famille d'intégrales :

$$I_\alpha(x) = \int_{-\infty}^{+\infty} \frac{e^{ix\xi}}{[(i\xi)^\alpha]_2} d\xi \quad (\text{VIII.30})$$

Cette famille est paramétrée par le nombre réel α . On remarque facilement que cette intégrale est toujours divergente : pour $\alpha < 1$ elle diverge à l'infini, pour $\alpha > 1$ elle diverge en zéro et pour $\alpha = 1$ elle diverge à la fois à l'infini et en zéro. Toutefois, pour $0 < \alpha < 1$ elle est semi-convergente.

Pour la calculer, on va mettre en œuvre deux procédés typiques. Si $\alpha < 1$, on interprète l'intégrale comme une intégrale le long de la droite réelle ; pour éviter la divergence en zéro, on calcule l'intégrale le long d'un chemin identique à cette droite, sauf autour de zéro, où il fait un détour par le plan complexe, comme illustré sur la figure VIII.1. La fonction $\xi \mapsto [(i\xi)^\alpha]_2$ est définie en dehors de la demi-droite imaginaire positive, donc il faudra contourner zéro par le bas pour rester dans le domaine d'holomorphic. Ce procédé

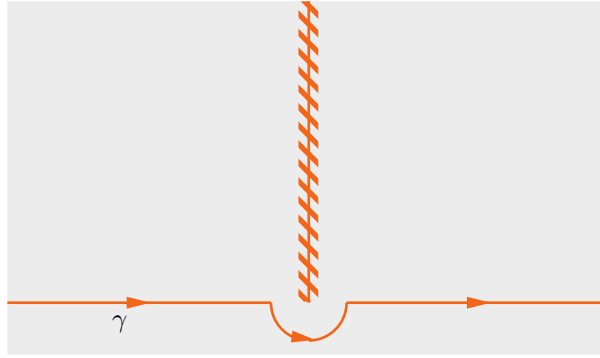


Figure VIII.1 – Contournement dans le plan complexe du point 0

est présenté ici pour des intégrales de dimension 1, mais peut être étendu aux intégrales multiples. Il est très fréquemment utilisé en physique théorique pour les fonctions de Green, les propagateurs, etc. Ainsi, pour résoudre l'équation électromagnétique $\Delta f + k^2 f = g$ par exemple, on peut utiliser la méthode de Fourier présentée en section VII.1. La transformée de Fourier de l'équation est alors :

$$-|\xi|^2 \hat{f}(\xi) + k^2 \hat{f}(\xi) = \hat{g}(\xi) \quad \text{avec} \quad |\xi|^2 = \xi_1^2 + \xi_2^2 + \xi_3^2 \quad (\text{VIII.31})$$

d'où :

$$\hat{f}(\xi) = \frac{\hat{g}(\xi)}{k^2 - |\xi|^2} \quad (\text{VIII.32})$$

La transformée de Fourier inverse de $1/(k^2 - |\xi|^2)$ est appelée *fonction de Green* de l'équation donnée. L'intégrale qui donne cette transformée inverse est :

$$G(x) = \frac{1}{2\pi} \int_{\mathbb{R}^3} \frac{e^{-ix \cdot \xi}}{k^2 - |\xi|^2} d\xi \quad (\text{VIII.33})$$

qui diverge sur les singularités $k^2 - |\xi|^2 = 0$. Pour calculer analytiquement cette fonction de Green, on contourne alors cette singularité en sortant de $\mathbb{R}^3$ par les valeurs complexes de ξ_1, ξ_2, ξ_3 . Grâce à cet expédient, il est assez facile d'obtenir :

$$G(x) = \frac{e^{ikr}}{4\pi r} \quad \text{où} \quad r = |x| = \sqrt{x_1^2 + x_2^2 + x_3^2} \quad (\text{VIII.34})$$

Pour lever la divergence à l'infini, on procède comme dans la section VIII.1, on introduit un facteur régularisant : au lieu de considérer l'intégrale $I_\alpha(x)$, on considère l'intégrale :

$$I_{\alpha,\varepsilon}(x) = \int_{-\infty}^{+\infty} \frac{e^{ix\xi}}{[(i\xi)^\alpha]_2} e^{-\varepsilon|\xi|} d\xi \quad (\text{VIII.35})$$

Le facteur $e^{-\varepsilon|\xi|}$, qui décroît très vite à l'infini, fait converger l'intégrale.

Le premier procédé (contourner la singularité en sortant de l'axe réel) sera donc appliqué lorsque $\alpha > 1$ et le second (introduire un facteur régularisant) lorsque $\alpha < 1$. Pour $\alpha = 1$, il faut faire les deux à la fois.

Le second procédé donne une intégrale convergente tant que $\varepsilon > 0$, mais bien entendu la valeur de l'intégrale divergente $I_\alpha(x)$ sera la limite, dans un sens qu'il faudra préciser, de $I_{\alpha,\varepsilon}(x)$ lorsque ε tend vers zéro. On verra cette fois que la limite au sens usuel (limite dans $\mathbb{R}$ ou $\mathbb{C}$ des nombres $I_{\alpha,\varepsilon}(x)$ pour x fixé) sera insuffisante pour donner un sens consistant à l'intégrale divergente ; même la limite au sens de l'espace $\mathbb{L}_2(\mathbb{R})$ sera insuffisante. Il faudra créer une nouvelle notion de limite.

Ces nouvelles notions de limite ont commencé à être comprises dans les années 1910–1920 par les mathématiciens polonais S. Banach et hongrois F. Riesz. Contrairement à ce qui était le cas dans les siècles précédents, il est difficile de les attribuer entièrement à un auteur précis et elles sont plutôt le résultat d'une lente maturation. Celle que nous utiliserons pour notre exemple a été proposée par L. Schwartz vers 1945.

Supposons donc d'abord $\alpha > 1$ et calculons l'intégrale :

$$I_\alpha(x) = \int_{\Gamma} \frac{e^{ix\xi}}{[(i\xi)^\alpha]_2} d\xi \quad (\text{VIII.36})$$

où Γ est le chemin infini représenté sur la figure VIII.1. Il est bien clair que l'intégrale ne dépend pas du chemin : tous les chemins qui aux grandes distances coïncident avec l'axe réel et qui ne traversent pas la coupure sont homologiquement équivalents. Si on paramètre le chemin par $t \mapsto \xi(t)$, l'intégrale devient :

$$I_\alpha(x) = \int_{-\infty}^{+\infty} \frac{e^{ix\xi(t)}}{[(i\xi(t))^\alpha]_2} \xi'(t) dt \quad (\text{VIII.37})$$

Ⓡ Il n'est pas absolument obligatoire de paramétrer le chemin Γ avec un paramètre allant de $-\infty$ à $+\infty$. Par exemple, la droite peut aussi bien se paramétrer par l'abscisse x elle-même que par $t = \arctan x$ qui parcourt $]-\pi/2; \pi/2[$. Toutefois, si on paramètre sur un domaine fini un chemin infini, la dérivée $\xi'(t)$ deviendra forcément infinie lorsque t atteindra ses valeurs extrémales. Pour fixer les idées, on pourra prendre pour paramètre l'abscisse curviligne le long du chemin, ce qui aura pour avantage d'avoir $|\xi'(t)| = 1$.

Si $x > 0$, la fonction $t \mapsto ix\xi(t)$ paramètre un autre chemin, que nous appellerons Γ et qui se déduit de Γ par une rotation autour de 0 d'angle $+\frac{\pi}{2}$ (multiplication par i) et une

homothétie de rapport x (multiplication par x). Le chemin Γ est alors un chemin qui à grande distance coïncide avec l'axe vertical et qui contourne le point 0 par la droite comme indiqué sur la figure VIII.2. Si on pose $z(t) = ix\xi(t)$, on a évidemment pour la dérivée $z'(t) = ix\xi'(t)$ et on a aussi ⁽¹⁾ $[(z(t))^\alpha]_2 = x^\alpha[(i\xi(t))^\alpha]_2$. Si on reporte cela dans (VIII.37), on voit que :

$$I_\alpha(x) = \frac{x^{\alpha-1}}{i} \int_{-\infty}^{+\infty} \frac{e^{z(t)}}{[z(t)^\alpha]_2} z'(t) dt \quad (\text{VIII.38})$$

Or l'intégrale (VIII.38) n'est rien d'autre que celle qu'on obtiendrait par le paramétrage $z(t)$ à partir de :

$$\int_{\Gamma} \frac{e^z}{[z^\alpha]_2} dz \quad (\text{VIII.39})$$

Par conséquent :

$$I_\alpha(x) = \frac{x^{\alpha-1}}{i} \int_{\Gamma} \frac{e^z}{[z^\alpha]_2} dz \quad (\text{VIII.40})$$

On peut encore faire le constat suivant. Sur des portions de cercles de centre 0 et de rayon R , la fonction sous le signe intégral dans (VIII.40) se paramètre par $z = Re^{i\theta}$, ce qui donne (si on prend θ dans l'intervalle $]-\pi; \pi[$) :

$$\frac{e^z}{[z^\alpha]_2} dz = R^{1-\alpha} \frac{e^{Re^{i\theta}}}{e^{i\alpha\theta}} ie^{i\theta} d\theta \quad (\text{VIII.41})$$

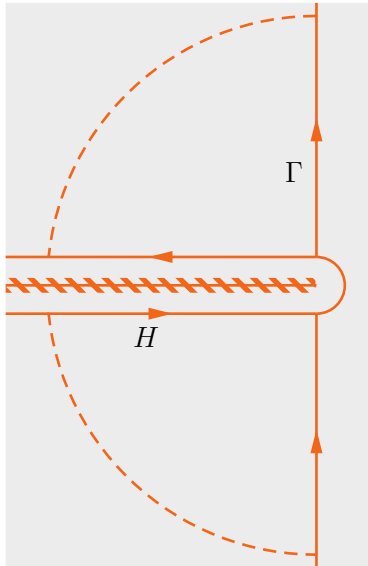


Figure VIII.2 – Chemin H

Si $\cos \theta$ reste ≤ 0 , ce qui est le cas pour $-\pi < \theta \leq -\frac{\pi}{2}$ ou $\frac{\pi}{2} \leq \theta < \pi$, le module de cette fonction est uniformément majoré par $R^{1-\alpha}$ qui tend vers zéro lorsque R tend vers l'infini, puisque nous avons supposé que $\alpha > 1$. Ce qui entraîne que l'intégrale de cette fonction sur n'importe quelle portion de cercle à gauche de l'axe imaginaire tend vers zéro quand R tend vers l'infini et, par suite, que l'intégrale (VIII.40) sur Γ est égale à l'intégrale sur un chemin H du type représenté sur la figure VIII.2.

En définitive :

$$I_\alpha(x) = \frac{x^{\alpha-1}}{i} \int_H \frac{e^z}{[z^\alpha]_2} dz \quad (\text{VIII.42})$$

où on reconnaît l'intégrale de Hankel — section V.6, de sorte que le résultat du calcul est :

$$I_\alpha(x) = \frac{2\pi x^{\alpha-1}}{\Gamma(\alpha)} \quad (\text{VIII.43})$$

Il reste encore à voir le cas plus simple où $x \leq 0$. Revenons à l'expression (VIII.36) de $I_\alpha(x)$. Elle ne dépend pas du détournement choisi, on peut donc prendre pour celui-ci un demi-cercle de rayon arbitraire, qu'on fera tendre vers l'infini. Or, le long d'un tel demi-cercle

(1) L'identité $[a \cdot b^\alpha]_2 = [a^\alpha]_2 \cdot [b^\alpha]_2$ est en général fausse pour a et b complexes mais elle est vraie si a ou b sont réels strictement positifs.

(situé cette fois à droite de l'axe imaginaire) le paramétrage $\xi = Re^{i\theta}$ (avec $\theta \in [-\pi; 0]$ pour le demi-cercle) donnera :

$$I_\alpha(x) = \int_{-\pi}^0 \frac{e^{ixRe^{i\theta}}}{[(iRe^{i\theta})^\alpha]_2} iRe^{i\theta} d\theta \quad (\text{VIII.44})$$

Le module de la fonction qui figure sous le signe intégral est $R^{1-\alpha}e^{-xR\sin\theta}$ où l'on voit aisément que si x est inférieur ou égal à zéro, l'exponentielle restera supérieure à l'unité pour tout θ dans l'intervalle d'intégration $[-\pi; 0]$ et le facteur $R^{1-\alpha}$ tend vers zéro quand R tend vers l'infini. On en conclut que dans ce cas $I_\alpha(x) = 0$. En résumé :

Théorème VIII.1 Pour $\alpha > 1$ on a :

$$\frac{1}{2\pi} \int_{\Gamma} \frac{e^{ix\xi}}{[(i\xi)^\alpha]_2} d\xi = \begin{cases} \frac{x^{\alpha-1}}{\Gamma(\alpha)} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases} \quad (\text{VIII.45})$$

Pour lever la divergence à l'infini, on introduit le facteur régularisant $e^{-\varepsilon|\xi|}$. Ce procédé s'applique pour $\alpha < 1$, de sorte que l'intégrale est convergente en $\xi = 0$. On va donc calculer :

$$I_{\alpha,\varepsilon}(x) = \int_{-\infty}^{+\infty} \frac{e^{ix\xi - \varepsilon|\xi|}}{[(i\xi)^\alpha]_2} d\xi \quad (\text{VIII.46})$$

Cette intégrale peut être décomposée en deux parties, l'une de $-\infty$ à 0, l'autre de 0 à $+\infty$. Dans la première, on peut faire le changement de variable $\xi \mapsto -\xi$ ce qui donne :

$$\int_0^{+\infty} \frac{e^{-ix\xi - \varepsilon\xi}}{[(-i\xi)^\alpha]_2} d\xi \quad (\text{VIII.47})$$

La fonction puissance qui est au dénominateur s'écrit plus explicitement sous la forme trigonométrique :

$$[(i\xi)^\alpha]_2 = \xi^\alpha e^{+i\alpha\pi/2} \quad \text{et} \quad [(-i\xi)^\alpha]_2 = \xi^\alpha e^{-i\alpha\pi/2} \quad (\text{VIII.48})$$

de sorte que :

$$I_{\alpha,\varepsilon}(x) = \int_0^{+\infty} \frac{e^{-(\varepsilon+ix)\xi}}{\xi^\alpha e^{-i\alpha\pi/2}} d\xi + \int_0^{+\infty} \frac{e^{-(\varepsilon-ix)\xi}}{\xi^\alpha e^{+i\alpha\pi/2}} d\xi \quad (\text{VIII.49})$$

On voit que le second terme est conjugué du premier, de sorte qu'on peut tout aussi bien écrire :

$$I_{\alpha,\varepsilon}(x) = 2\Re \left(\int_0^{+\infty} \frac{e^{-(\varepsilon+ix)\xi}}{\xi^\alpha e^{-i\alpha\pi/2}} d\xi \right) \quad (\text{VIII.50})$$

Posons $z = \varepsilon + ix$. Puisque ε a été supposé strictement positif et que x peut être n'importe quel réel, le nombre complexe z pourra être n'importe où dans le demi-plan $\Re(z) > 0$.

Ce que nous cherchons est deux fois la partie réelle de :

$$\int_0^{+\infty} \frac{e^{-z\xi}}{\xi^\alpha e^{-i\alpha\pi/2}} d\xi \quad (\text{VIII.51})$$

Il faut donc calculer (VIII.51). La constante $e^{-i\alpha\pi/2}$ du dénominateur sort de l'intégrale et dans cette dernière, on reconnaît l'intégrale eulérienne de deuxième espèce évoquée en (V.3), au changement de variable $t = z\xi$ près. Ce changement de variable n'est cependant possible que pour z réel; on procédera donc comme d'habitude : on effectue ce changement de variable pour $z > 0$, puis on étend le résultat par prolongement analytique⁽²⁾. Ainsi, pour $z > 0$:

$$\int_0^{+\infty} \frac{e^{-z\xi}}{\xi^\alpha} dt = z^{\alpha-1} \int_0^{+\infty} \frac{e^{-t}}{t^\alpha} d\xi \quad (\text{VIII.52})$$

Les figures qui suivent affichent les fonctions $I_{\alpha,\varepsilon}(x)$. Plus ε est petit, plus les variations au voisinage de zéro sont amples et rapides. On ne peut leur concevoir aucune limite au sens usuel lorsque ε tend vers zéro.

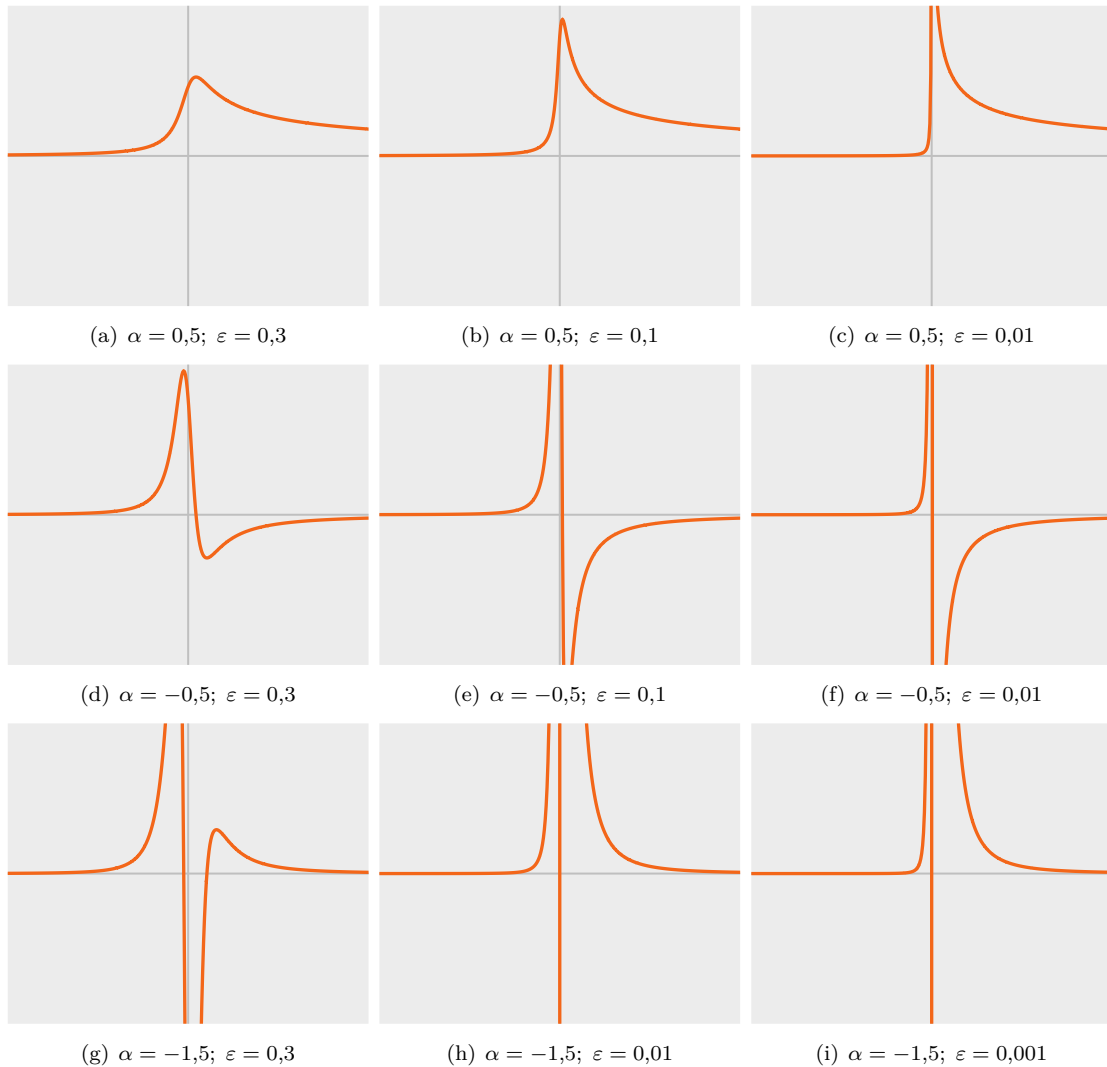


Figure VIII.3 – Évolution de la fonction $I_{\alpha,\varepsilon}(x)$ lorsque ε tend vers zéro.

Sur la figure VIII.3, les fonctions sont représentées toutes à la même échelle avec les caractéristiques suivantes :

⁽²⁾ Les théorèmes généraux garantissent en effet que (VIII.51) dépend analytiquement de z dans le demi-plan $\Re(z) > 0$.

- $\alpha = 0,5$ — pour $x > 0$, la fonction limite est $1/\sqrt{\pi x} = x^{-1/2}/\Gamma(1/2)$;
- $\alpha = -0,5$ — pour $x > 0$, la fonction limite est $-x^{-3/2}/2\sqrt{\pi} = x^{-3/2}/\Gamma(-1/2)$;
- $\alpha = -1,5$ — pour $x > 0$, la fonction limite est $3x^{-5/2}/4\sqrt{\pi} = x^{-5/2}/\Gamma(-3/2)$.

On voit que pour $x < 0$, $I_{\alpha,\varepsilon}(x)$ tend vers zéro mais d'autant plus lentement que x est plus proche de zéro ; pour $x > 0$, $I_{\alpha,\varepsilon}(x)$ tend vers la fonction $x^{\alpha-1}/\Gamma(\alpha)$. Cependant, au voisinage de $x = 0$, le comportement est fortement singulier : pour $\alpha = 0,5$, la fonction passe par un maximum très aigu ; ensuite, pour $\alpha = -0,5$, elle passe d'abord par un maximum très aigu, puis, sur une très courte distance de l'ordre de ε , descend à un minimum, avant de remonter en longeant la courbe $y = x^{-1,5}/\Gamma(-0,5)$. Enfin, pour $\alpha = -1,5$, elle passe par un maximum très aigu, descend à un minimum puis remonte encore à un maximum avant de redescendre en longeant la courbe $y = x^{-2,5}/\Gamma(-1,5)$.

Le voisinage de $x = 0$ présente donc une oscillation d'amplitude énorme entre un maximum et un minimum (de l'ordre de $\varepsilon^{\alpha-1}$) et de période très courte (de l'ordre de ε). Concernant la plus petite valeur de ε , l'oscillation est même devenue invisible car elle est confondue avec l'axe vertical.

Pour d'autres valeurs de α et ε non illustrées, les fonctions limites sont les suivantes :

- $\alpha = -2,5$ — pour $x > 0$, la fonction limite est $-15x^{-7/2}/8\sqrt{\pi} = x^{-7/2}/\Gamma(-5/2)$;
- $\alpha = -3,5$ — pour $x > 0$, la fonction limite est $105x^{-9/2}/16\sqrt{\pi} = x^{-9/2}/\Gamma(-7/2)$;
- $\alpha = -7,5$ — pour $x > 0$, la fonction limite est $2027025x^{-17/2}/256\sqrt{\pi} = x^{-17/2}/\Gamma(-15/2)$.

Le phénomène esquissé sur les figures VIII.3 s'accroît car le nombre d'oscillations augmente avec $|\alpha|$. On ne peut pas voir les courbes complètes : si on comprimait les ordonnées, on ramènerait certes les oscillations centrales dans le cadrage, mais les oscillations les plus extérieures, qui ont une amplitude nettement moindre, seraient alors écrasées sur l'axe horizontal.

La fonction tend bien vers une limite en dehors du voisinage de $x = 0$ mais non au voisinage de $x = 0$; il faudra introduire une nouvelle notion de limite.

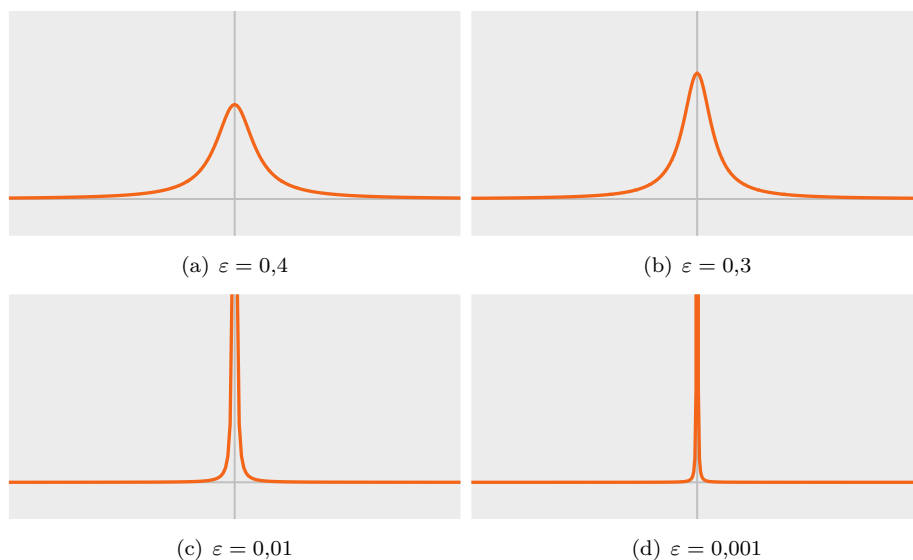


Figure VIII.4 – Évolution de la fonction $I_{0,\varepsilon}(x)$ lorsque ε tend vers zéro. La limite est la distribution δ de Dirac.

Concernant la figure VIII.4 pour $\alpha = 0$, ou plus généralement pour α entier ≤ 0 , le facteur $1/\Gamma(\alpha)$ devient nul, de sorte que la fonction limite pour $x > 0$ est nulle. La fonction $I_{0,\varepsilon}(x)$ tend vers zéro partout en dehors d'un voisinage de zéro et au point $x = 0$, elle tend vers l'infini.

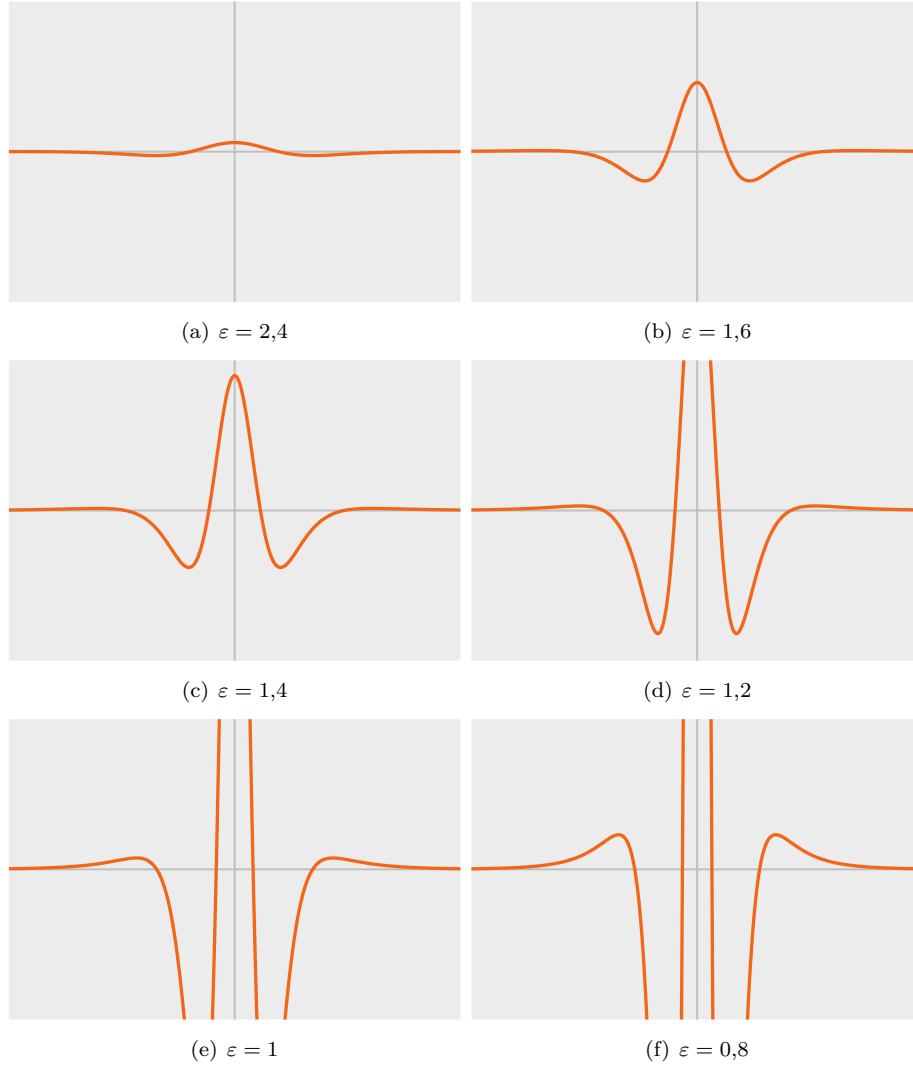


Figure VIII.5 – Évolution de la fonction $I_{-4,\varepsilon}(x)$ lorsque ε tend vers zéro. La limite est la quatrième dérivée de la distribution δ de Dirac. On montre ici l'évolution progressive lorsque ε diminue. Le maximum central et les deux minima s'accroissent ; deux maxima latéraux apparaissent.

L'intégrale eulérienne est égale à $\Gamma(1 - \alpha)$. Le prolongement analytique de $z^{\alpha-1}$ au demi-plan $\Re(z) > 0$ est la fonction $[z^{\alpha-1}]_2$, de sorte que finalement (VIII.51) devient :

$$[z^{\alpha-1}]_2 \Gamma(1 - \alpha) e^{i\alpha\pi/2} \quad (\text{VIII.53})$$

Pour obtenir commodément la partie réelle, écrivons encore la fonction $[z^{\alpha-1}]_2$ sous forme trigonométrique, en introduisant r et θ tels que $z = \varepsilon + ix = re^{i\theta}$. Comme $\Re(z) > 0$, θ sera compris entre $-\frac{\pi}{2}$ et $+\frac{\pi}{2}$, donc $\theta = \arctan(x/\varepsilon)$; d'autre part, on aura :

$$r = \sqrt{\varepsilon^2 + x^2} = \varepsilon \sqrt{1 + (x/\varepsilon)^2} = \varepsilon \sqrt{1 + \tan^2 \theta} = \varepsilon / \cos \theta \quad (\text{VIII.54})$$

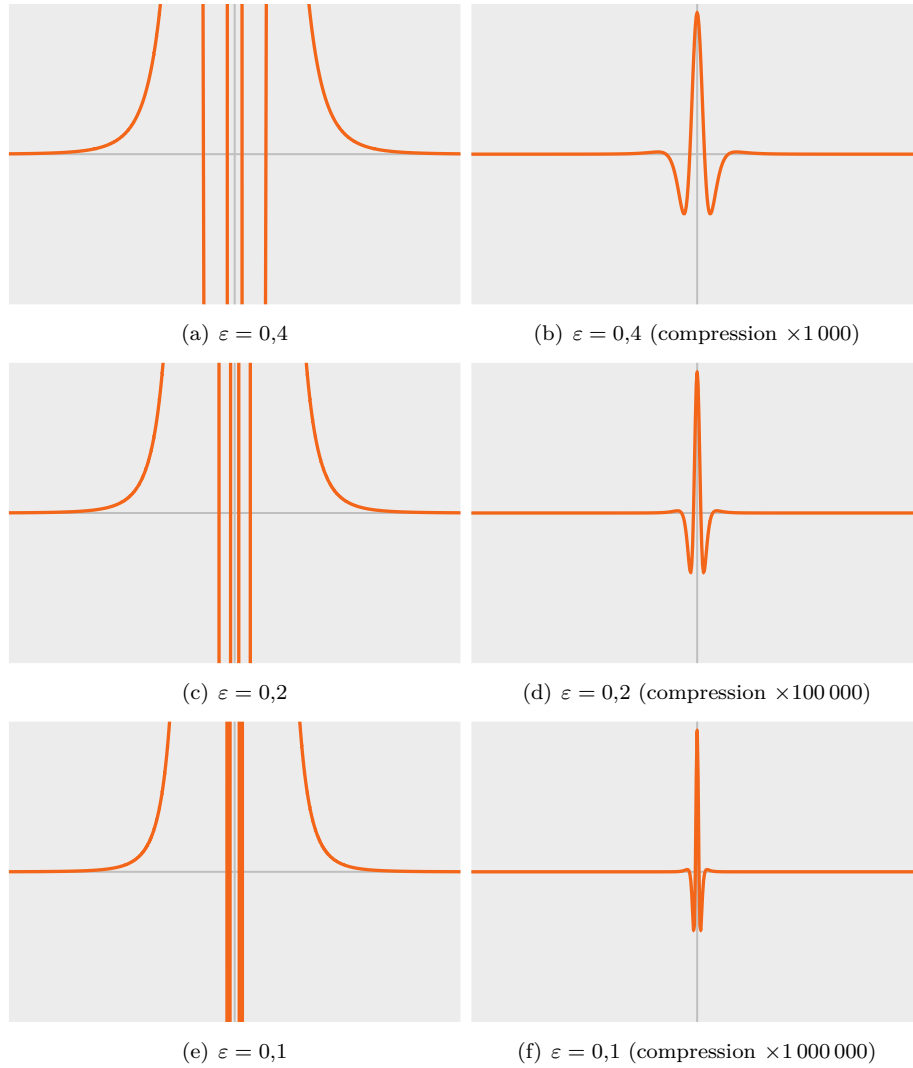


Figure VIII.6 – Compression de la fonction $I_{-4,\varepsilon}$ — À gauche, les deux maxima latéraux sortent du champ ; les portions de courbe qui vont d'un extremum au suivant n'apparaissent plus que comme des droites verticales. À droite, l'ordonnée est comprimée pour ramener les extrema dans le champ, ce qui rend invisibles les maxima latéraux

Par conséquent :

$$[z^{\alpha-1}]_2 = r^{\alpha-1} e^{i(\alpha-1)\theta} \quad (\text{VIII.55})$$

et l'intégrale (VIII.51) s'écrit :

$$r^{\alpha-1} \Gamma(1-\alpha) e^{i\alpha\pi/2 + i(\alpha-1)\theta} \quad (\text{VIII.56})$$

La partie réelle est maintenant évidente :

$$I_{\alpha,\varepsilon}(x) = \frac{2\Gamma(1-\alpha)}{r^{1-\alpha}} \cos\left(\frac{\alpha\pi}{2} - (1-\alpha) \arctan\left(\frac{x}{\varepsilon}\right)\right) \quad (\text{VIII.57})$$

Il ne reste plus qu'à étudier de plus près cette fonction et voir comment elle évolue quand ε tend vers zéro.

Nous avons posé $\theta = \arctan(\frac{x}{\varepsilon})$ et introduisons $\phi = \alpha\frac{\pi}{2} - (1-\alpha) \arctan(\frac{x}{\varepsilon})$, argument de la fonction cos dans (VIII.57). Lorsque x parcourt l'intervalle de $]-\infty; \infty[$, $\arctan(\frac{x}{\varepsilon})$

parcourt l'intervalle $]-\pi/2; \pi/2[$ et $(1 - \alpha) \arctan(\frac{x}{\varepsilon})$ parcourt $-(1 - \alpha)\frac{\pi}{2}; (1 - \alpha)\frac{\pi}{2}[$; par conséquent ϕ parcourt $]\alpha\pi - \frac{\pi}{2}; \frac{\pi}{2}[$. Quand $x \rightarrow -\infty$, $\phi \rightarrow \frac{\pi}{2}$ — donc $\cos \phi \rightarrow 0$ — et quand $x \rightarrow +\infty$, $\phi \rightarrow \alpha\pi - \frac{\pi}{2}$ — donc $\cos \phi \rightarrow \sin(\alpha\pi)$. En utilisant la formule des compléments on voit donc que :

$$\lim_{x \rightarrow +\infty} I_{\alpha, \varepsilon}(x) = \frac{2\pi r^{\alpha-1}}{\Gamma(\alpha)} \quad (\text{VIII.58})$$

Toutefois, $\arctan(\frac{x}{\varepsilon})$ est déjà proche de $\pm\frac{\pi}{2}$ (à 0,1 près) lorsque $|x| \geq 10\varepsilon$. Du fait que ϕ dépend de x par $\arctan(\frac{x}{\varepsilon})$, cela veut dire que lorsque x parcourt l'intervalle de -10ε à $+10\varepsilon$, ϕ parcourt la quasi totalité de l'intervalle $]\alpha\pi - \frac{\pi}{2}; \frac{\pi}{2}[$, dont la longueur est $\frac{1-\alpha}{2} \times 2\pi$.



Figure VIII.7 – Fonctions $I_{-8; \varepsilon}(x)$ — Le nombre d'oscillations a doublé par rapport au cas précédent. Les extrema sont énormes : de l'ordre de 10^{18} pour $\varepsilon = 0,03$ et de 10^{27} pour $\varepsilon = 0,003$.

De façon purement qualitative, on peut donc résumer le comportement de la fonction $x \mapsto I_{\alpha, \varepsilon}(x)$ pour ε donné (en principe petit) :

- pour $x < 0$ et $|x| \gg \varepsilon$: $I_{\alpha, \varepsilon}(x) \simeq 0$;
- pour $|x| \sim \varepsilon$: $I_{\alpha, \varepsilon}(x)$ oscille entre des maxima positifs et des minima négatifs, dont l'amplitude est de l'ordre de $2\Gamma(1 - \alpha)/\varepsilon^{1-\alpha}$;
- pour $x > 0$ et $|x| \gg \varepsilon$: $I_{\alpha, \varepsilon}(x) \simeq 2\pi x^{\alpha-1}/\Gamma(\alpha)$.

Plus ε est petit, plus l'amplitude des oscillations est grande et plus ces oscillations se concentrent dans un petit intervalle (de longueur ε) autour de $x = 0$. Il est clair qu'un tel comportement ne donne aucune limite au sens usuel. Par contre, en dehors d'un voisinage de $x = 0$, la fonction $I_{\alpha, \varepsilon}(x)$ tend *au sens usuel* vers zéro pour $x < 0$ et vers $2\pi x^{\alpha-1}/\Gamma(\alpha)$ pour $x > 0$. De façon précise, on peut dire que lorsque ε tend vers zéro, alors, sur tout intervalle $]-\infty; a]$, $I_{\alpha, \varepsilon}$ tend uniformément vers zéro; et sur tout intervalle $[a; \infty[$, $I_{\alpha, \varepsilon}$ tend uniformément vers $2\pi x^{\alpha-1}/\Gamma(\alpha)$.

Vérifier en détail que la convergence est uniforme est un travail purement technique sans surprise, mais fastidieux : il faut partir de l'équation (VIII.57).

Les figures VIII.3 à VIII.7 montrent les graphiques des fonctions $I_{\alpha, \varepsilon}$ pour différentes valeurs de ε et de α . La particularité la plus intéressante de ces fonctions est leur comportement-limite (lorsque ε tend vers zéro) extrêmement singulier. L'intégrale divergente (VIII.30) pour $\alpha < 1$ est précisément cette limite. Pour obtenir des solutions d'équations aux dérivées partielles sous forme d'intégrales divergentes, J. Hadamard a construit une théorie qui donne

un sens à de telles limites ⁽³⁾. Nous les placerons dans le cadre de la théorie des distributions.

L. Schwartz désigne cette limite par la notation Pf. :

$$I_{\alpha(x)} = \text{Pf.} \left(2\pi \frac{x^{\alpha-1}}{\Gamma(\alpha)} \right) = \frac{2\pi}{\Gamma(\alpha)} \text{Pf.} (x^{\alpha-1}) \quad (\text{VIII.59})$$

Le symbole Pf. se lit indifféremment *pseudo-fonction* (terme introduit par L. Schwartz) ou *partie finie de* (terme employé par J. Hadamard).

Le phénomène qui se produit pour les intégrales $I_{\alpha,\varepsilon}$ lorsque ε tend vers zéro, mérite l'étude détaillée qui lui a été consacrée ici. Dans les applications à la Physique, on rencontre très souvent des intégrales divergentes, mais la cause en est généralement que la description mathématique introduit implicitement des infinis. Par exemple, lorsqu'on décrit un faisceau de laser ou d'électrons, on le fait généralement à l'aide d'une onde plane monochromatique, qui aurait, en toute rigueur, une extension infinie ; dans la réalité, évidemment, le faisceau est limité — ses dimensions spatiales sont simplement grandes par rapport à la longueur d'onde, de sorte que l'infini est une représentation commode. Si on construisait des modèles mathématiques qui respectent mieux les conditions physiques, on ne rencontrerait jamais d'intégrales divergentes, mais les calculs analytiques seraient fortement alourdis : les règles de calcul pour la distribution δ de Dirac sont bien plus simples que les règles équivalentes pour les fonctions $\sqrt{n/\pi} e^{-nx^2}$ par exemple. Toutefois, lorsqu'on utilise des "fonctions" de Dirac, des pseudo-fonctions etc., on ne doit jamais perdre de vue qu'elles représentent *en réalité* des fonctions comme les $I_{\alpha,\varepsilon}(x)$ où ε est simplement petit et non véritablement nul.

Lorsque ε est vraiment petit, les oscillations que nous avons analysées et qu'on peut voir sur les figures VIII.3 à VIII.7, ont certes une amplitude énorme, mais elles ne se produisent que dans un intervalle minuscule (sur les figures où ε est très petit, les oscillations se confondent avec l'axe vertical). En dehors de ce minuscule intervalle, la fonction $I_{\alpha,\varepsilon}(x)$ a une limite (au sens usuel) qui est 0 pour $x < 0$ et $2\pi x^{\alpha-1}/\Gamma(\alpha)$ pour $x > 0$. On peut même dire que la fonction $I_{\alpha,\varepsilon}(x)$ a une limite au sens usuel partout, excepté en $x = 0$. Cette limite n'est pas une fonction intégrable : l'intégrale :

$$\int_0^\infty x^{\alpha-1} dx \quad (\text{VIII.60})$$

diverge ; on notera en particulier qu'elle diverge en $x = 0$ pour $\alpha < 0$. Or, la propriété remarquable qui caractérise les oscillations est que leur moyenne est nulle : les énormes minima négatifs compensent en moyenne les énormes maxima positifs. De sorte que l'intégrale :

$$\int_{-\infty}^{+\infty} I_{\alpha,\varepsilon}(x) dx \quad (\text{VIII.61})$$

converge toujours en $x = 0$ et tend vers une limite quand $\varepsilon \rightarrow 0$, bien qu'à la limite la fonction soit partout, sauf en un seul point, égale à une fonction dont l'intégrale diverge. C'est cette limite que J. Hadamard appelait la *partie finie*.

Voyons cela de plus près. Prenons une fonction $\varphi(x)$ lisse, c'est-à-dire infiniment dérivable ; pour la commodité, on la supposera dans l'espace $\mathcal{S}(\mathbb{R})$. La fonction $I_{\alpha,\varepsilon}(x)$ est la

⁽³⁾ J. Hadamard, *Le problème de Cauchy et les équations aux dérivées partielles linéaires hyperboliques*, éd. Hermann, Paris, 1932, p. 184.

transformée de Fourier de la fonction :

$$\xi \mapsto \frac{e^{-\varepsilon|\xi|}}{[(i\xi)^\alpha]_2} \quad (\text{VIII.62})$$

En utilisant la relation (VII.65), on obtient :

$$\int_{-\infty}^{+\infty} I_{\alpha,\varepsilon}(x) \varphi(x) dx = \int_{-\infty}^{+\infty} \frac{e^{-\varepsilon|\xi|}}{[(i\xi)^\alpha]_2} \widehat{\varphi}(\xi) d\xi \quad (\text{VIII.63})$$

Puisque $\varphi \in \mathcal{S}(\mathbb{R})$, on aura aussi $\widehat{\varphi} \in \mathcal{S}(\mathbb{R})$ d'après les théorèmes VII.1 et VII.2 ; par conséquent, d'après les théorèmes généraux, on peut passer à la limite sous le signe intégral *dans le membre de droite*. On peut donc écrire :

$$\lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{+\infty} I_{\alpha,\varepsilon}(x) \varphi(x) dx = \int_{-\infty}^{+\infty} \frac{\widehat{\varphi}(\xi)}{[(i\xi)^\alpha]_2} d\xi \quad (\text{VIII.64})$$

Cette égalité signifie que, bien que $I_{\alpha,\varepsilon}(x)$ n'ait pas de limite intégrable pour $\varepsilon \rightarrow 0$, son intégrale pondérée par n'importe quelle fonction φ de $\mathcal{S}(\mathbb{R})$ en a cependant une. Ce sont les oscillations moyennisables de la fonction $I_{\alpha,\varepsilon}(x)$ qui rendent cela possible.



IX — Théorie des distributions

IX.1 Modèle mathématique

La théorie mathématique des distributions est une synthèse de tous les procédés qui permettent de donner un sens aux intégrales divergentes, aux dérivées de fonctions non dérivables etc.

Une distribution est un modèle mathématique pour une distribution de charges électriques, d'où le nom. N'importe quelle distribution peut se représenter intuitivement comme une certaine répartition de charges dans l'espace, soit continue, soit discrète. Lorsqu'on écrit l'équation de l'électrostatique :

$$\Delta U = 4\pi\rho \quad (\text{IX.1})$$

le paramètre ρ représente la densité de charges. Si les charges sont discrètes, on doit prendre pour ρ une combinaison de charges ponctuelles dont chacune est représentée par la pseudo-densité δ . La théorie des distributions permet d'englober les deux cas sous un concept unique.

Il est évident que pour le physicien, il n'y a pas de séparation objective entre les densités continues et les distributions de charges discrètes : une charge placée au point $x = 0$ pourrait aussi bien se représenter par une densité continue :

$$\rho(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \quad (\text{IX.2})$$

avec σ suffisamment petit, que par $\delta(x)$. Cependant, les règles de calcul avec δ sont bien plus simples qu'avec ρ .

Si on veut représenter un dipôle (deux charges opposées très proches), on a le choix entre les modèles mathématiques suivants illustrés sur la figure IX.1 :

- une densité continue dipolaire, comme par exemple :

$$\rho(x) = (1/\sqrt{2\pi}\sigma^3) x e^{-x^2/2\sigma^2} \quad (\text{IX.3})$$

cette densité présente un minimum très aigu pour $x = -\sigma$ (charge négative) et un maximum très aigu pour $x = +\sigma$ (charge positive) ;

- une distribution discrète $\frac{1}{2\sigma} [\delta(x - \sigma) - \delta(x + \sigma)]$ (charge positive en $x = \sigma$ et charge négative en $x = -\sigma$) ;
- une nouvelle distribution dite dipolaire, qui est la limite de l'une ou l'autre des deux précédentes lorsque σ tend vers zéro (les deux limites, prises au sens des distributions, sont égales).

La troisième représentation n'est pas plus juste que les deux premières mais les règles de calcul sont bien plus simples avec elle. Les distributions peuvent toujours être dérivées et on verra que la distribution dipolaire est égale à $-\delta'$.

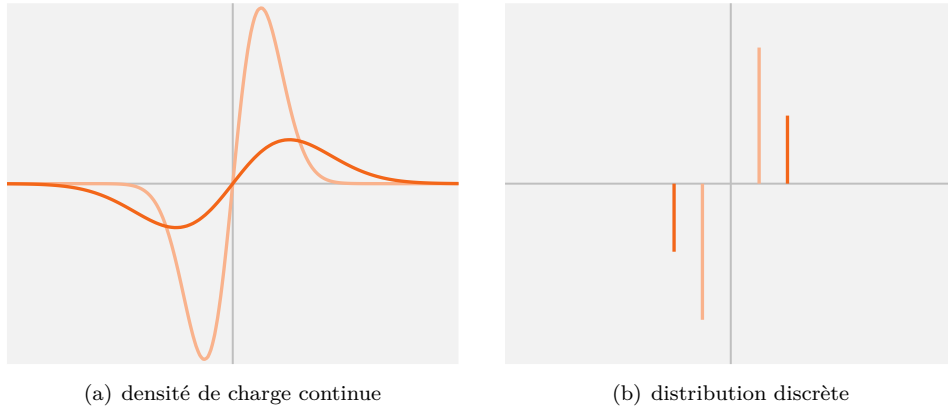


Figure IX.1 – Représentations d'un dipôle. La distance entre les charges est 2σ . En gris, la demi-distance σ entre les deux charges est plus petite et les charges sont donc plus grandes. Lorsque σ tend vers zéro et les charges vers l'infini, le moment dipolaire restant constant, on obtient une limite au sens des distributions

IX.2 Définition des distributions

Nous avons vu à la fin du chapitre précédent que les fonctions $I_{\alpha,\varepsilon}(x)$ n'avaient pas de limite au sens usuel, mais que les intégrales $\int I_{\alpha,\varepsilon}(x) \varphi(x) dx$ en avaient pour toute fonction $\varphi \in \mathcal{S}(\mathbb{R})$. De la même façon, la famille des fonctions (IX.2) n'a pas de limite au sens usuel ⁽¹⁾ mais les intégrales :

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \varphi(x) dx \quad (\text{IX.4})$$

en ont une lorsque $\sigma \rightarrow 0$ pour n'importe quelle fonction continue φ ⁽²⁾.

Il en est de même pour les dipôles considérés de la figure IX.1. Les fonctions (IX.3) n'ont pas de limite quand $\sigma \rightarrow 0$, mais les intégrales :

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma^3} x e^{-x^2/2\sigma^2} \varphi(x) dx \quad (\text{IX.5})$$

⁽¹⁾ Sa limite dans un autre sens à préciser est δ .

⁽²⁾ Il faut évidemment s'assurer que la fonction $\varphi(x)$ ne croisse pas en e^{x^2} à l'infini.

en ont une pour n'importe quelle fonction φ continûment dérivable et raisonnable à l'infini.

On peut calculer facilement la limite de (IX.4) et (IX.5). On remarque que la famille de fonctions (IX.2) tend uniformément vers zéro dans tout intervalle de la forme $] -\infty ; -a]$ ou $[a ; +\infty [$; d'autre part, si φ est continue, elle est pratiquement égale à $\varphi(0)$ au voisinage de $x = 0$, de sorte que (IX.4) se décompose en :

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \varphi(x) dx = \int_{-\infty}^{-a} + \int_{-a}^{+a} + \int_{+a}^{+\infty} \quad (\text{IX.6})$$

où les premier et dernier termes tendent vers zéro et où le deuxième est équivalent à :

$$\varphi(0) \int_{-a}^{+a} \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx \quad (\text{IX.7})$$

Un calcul direct (faire dans l'intégrale le changement de variable $x = \sigma y$) montre que l'intégrale (IX.7) tend vers 1 quand σ tend vers zéro. Autrement dit, la limite de (IX.4) est tout simplement $\varphi(0)$, si φ est continue en $x = 0$.

On remarque que la densité qui intervient dans (IX.5) est au signe près la dérivée de celle qui intervient dans (IX.4) :

$$\frac{1}{\sqrt{2\pi}\sigma^3} x e^{-x^2/2\sigma^2} = -\frac{d}{dx} \left(\frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \right) \quad (\text{IX.8})$$

Par conséquent, en faisant, dans (IX.5), une intégration par parties, on se ramène à (IX.4) et on voit que la limite de (IX.5) est $\varphi'(0)$.

Ces quelques exemples montrent que, selon les cas, il faut que φ soit continue, ou continûment dérivable, ou, pour I_α , suffisamment dérivable. Il fallait aussi qu'elle soit raisonnable à l'infini. Pour être sûr de couvrir tous ces cas, on suppose que φ est dans l'espace $\mathcal{S}(\mathbb{R})$.

Toutes les distributions envisagées jusqu'ici étaient de la forme :

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} f_n(x) \varphi(x) dx \quad \text{ou} \quad \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{+\infty} f_\varepsilon(x) \varphi(x) dx \quad (\text{IX.9})$$

Il est clair que les propriétés de linéarité des intégrales passent à la limite. On généralise donc la notion d'intégrale en disant que les expressions suivantes :

$$\int_{-\infty}^{+\infty} \delta(x) \varphi(x) dx = \varphi(0) \quad (\text{IX.10a})$$

$$\int_{-\infty}^{+\infty} -\delta'(x) \varphi(x) dx = \varphi'(0) \quad (\text{IX.10b})$$

$$\int_{-\infty}^{+\infty} I_\alpha(x) \varphi(x) dx = \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{+\infty} I_{\alpha,\varepsilon}(x) \varphi(x) dx \quad (\text{IX.10c})$$

sont des *fonctionnelles linéaires* ou des *formes linéaires* sur l'espace $\mathcal{S}(\mathbb{R})$. Cela signifie qu'elles dépendent linéairement de la variable φ , qui est une fonction mais aussi un vecteur de l'espace vectoriel $\mathcal{S}(\mathbb{R})$. L'exemple standard de fonctionnelle linéaire est l'intégrale $f \mapsto \int f(x) dx$. La théorie des distributions est une tentative réussie d'étendre la notion usuelle d'intégrale.

Définition IX.1 On appelle distribution une forme linéaire continue sur l'espace $\mathcal{S}(\mathbb{R})$.

Autrement dit, une distribution est une application T de l'espace vectoriel $\mathcal{S}(\mathbb{R})$ dans $\mathbb{R}$ (ou parfois $\mathbb{C}$) :

$$T : \begin{cases} \mathcal{S}(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ \varphi & \longmapsto & T(\varphi) \end{cases} \quad (\text{IX.11})$$

qui est *linéaire* :

$$\begin{aligned} T(\varphi_1 + \varphi_2) &= T(\varphi_1) + T(\varphi_2) \\ T(\lambda\varphi) &= \lambda T(\varphi) \end{aligned} \quad (\text{IX.12})$$

et *continue* :

$$\lim_{n \rightarrow \infty} \varphi_n = \varphi \quad \Rightarrow \quad \lim_{n \rightarrow \infty} T(\varphi_n) = T(\varphi) \quad (\text{IX.13})$$

Au premier abord, il peut paraître curieux qu'on exige la continuité pour une application linéaire mais l'espace $\mathcal{S}(\mathbb{R})$ — tout comme les espaces $\mathbb{L}_1(\mathbb{R})$ et $\mathbb{L}_2(\mathbb{R})$ — est un espace vectoriel de dimension infinie. Sur un tel espace, il peut y avoir des fonctions linéaires discontinues. Il n'y a pas besoin de chercher loin pour en trouver. Prenons la fonctionnelle linéaire :

$$A : \begin{cases} \mathbb{L}_1(\mathbb{R}) \cap \mathbb{L}_2(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \int_{-\infty}^{+\infty} f(x) dx \end{cases} \quad (\text{IX.14})$$

Cette fonctionnelle est simplement l'intégrale. Elle est discontinue si on considère l'espace vectoriel $E = \mathbb{L}_1(\mathbb{R}) \cap \mathbb{L}_2(\mathbb{R})$ avec la notion de limite en moyenne quadratique (mais elle serait continue avec la limite en moyenne tout court). En effet, soit par exemple la suite de fonctions

$$f_n(x) = \begin{cases} \frac{1}{\ln(n+1)} \frac{1}{1+|x|} & \text{si } |x| \leq n \\ 0 & \text{si } |x| > n \end{cases} \quad (\text{IX.15})$$

Un calcul simple donne $A(f_n) = 2$, or :

$$\|f_n\|_2^2 = \int_{-\infty}^{+\infty} \frac{1}{\ln^2(n+1)} \frac{1}{(1+|x|)^2} dx = \frac{2}{\ln^2(n+1)} \left[1 - \frac{1}{n+1} \right] \xrightarrow{n \rightarrow \infty} 0 \quad (\text{IX.16})$$

On a donc ainsi un exemple pour lequel $\|f_n\|_2$ tend vers zéro alors que $A(f_n)$ ne tend pas vers zéro. On voit que la continuité d'une application linéaire n'est pas automatique.

L'exemple choisi est l'intégrale ; c'est l'exemple le moins artificiel possible, puisque l'idée essentielle de la théorie des distributions est d'étendre la notion d'intégrale sous la forme plus générale de fonctionnelle linéaire : $\int f(x) dx$ est une fonctionnelle linéaire de f . Dans cette ordre d'idée, la continuité de ladite fonctionnelle linéaire exprime le passage à la limite sous le signe intégral : les créateurs de la théorie des distributions sont partis du principe que si on veut étendre la notion d'intégrale à des fonctions qui ne sont pas intégrables au sens usuel, ces nouvelles intégrales doivent avoir la propriété du passage à la limite, sinon elles sont peu intéressantes. C'est pour cette raison qu'on exige la continuité.

La limite des $T(\varphi_n)$ ne pose aucun problème, car ils forment une suite numérique : la limite est prise au sens usuel pour les suites de nombres. Par contre, il faut préciser ce qu'on entend par la limite de la suite φ_n , car il s'agit d'une suite de fonctions de l'espace $\mathcal{S}(\mathbb{R})$. Si on adopte l'un des deux sens usuels, à savoir :

1. limite *simple* : pour tout x fixé, la suite numérique $\varphi_n(x)$ tend vers le nombre $\varphi(x)$;
2. limite *uniforme* : la suite numérique $\sup_{x \in \mathbb{R}} |\varphi_n(x) - \varphi(x)|$ tend vers zéro ;

ou bien encore le sens défini pour les espaces $\mathbb{L}_1(\mathbb{R})$ ou $\mathbb{L}_2(\mathbb{R})$, qui s'applique aussi dans l'espace $\mathcal{S}(\mathbb{R})$, alors on ne pourra pas construire une théorie cohérente des distributions. Il a fallu des années de recherche⁽³⁾ pour trouver la bonne notion de limite et pour *démontrer* qu'avec cette bonne notion de limite tout marche bien.

Voici cette notion de limite. On introduit les nombres suivants, appelés les *semi-normes* de l'espace $\mathcal{S}(\mathbb{R})$:

$$\mathcal{N}_{j,k}(\varphi) = \sup_{x \in \mathbb{R}} \left\{ (1 + |x|^k) \left| \frac{d^j}{dx^j} \varphi(x) \right| \right\} \quad (\text{IX.17})$$

où j et k sont des entiers ($\in \mathbb{N}$). La définition même de l'espace de Schwartz garantit que ces nombres sont tous finis⁽⁴⁾. On dit alors que la suite φ_n tend vers zéro dans $\mathcal{S}(\mathbb{R})$ si pour *tout* couple d'entiers j et k positifs ou nuls, la suite numérique $\mathcal{N}_{j,k}(\varphi_n)$ tend vers zéro :

$$\mathcal{S}\text{-}\lim_{n \rightarrow \infty} \varphi_n = 0 \iff \forall j \in \mathbb{N}, \quad \forall k \in \mathbb{N}, \quad \mathbb{R}\text{-}\lim_{n \rightarrow \infty} \mathcal{N}_{j,k}(\varphi_n) = 0 \quad (\text{IX.18})$$

On a introduit les notations $\mathcal{S}\text{-}\lim$ et $\mathbb{R}\text{-}\lim$ pour rendre plus visible la différence. La définition (IX.18) est donnée pour les suites φ_n qui tendent vers zéro, mais cela suffit puisque $\mathcal{S}(\mathbb{R})$ est un espace vectoriel : il revient évidemment au même de dire que $\varphi_n \rightarrow \varphi$ ou que $\varphi_n - \varphi \rightarrow 0$. En outre, pour une forme *linéaire* T , il revient au même de dire que $T(\varphi_n) \rightarrow T(\varphi)$ ou $T(\varphi_n - \varphi) \rightarrow 0$; autrement dit, pour qu'une forme *linéaire* soit continue, il suffit qu'elle soit continue en zéro. Ainsi, une distribution est une forme linéaire sur $\mathcal{S}(\mathbb{R})$ telle que :

$$\forall j \in \mathbb{N}, \quad \forall k \in \mathbb{N}, \quad \lim_{n \rightarrow \infty} \mathcal{N}_{j,k}(\varphi_n) = 0 \implies \lim_{n \rightarrow \infty} T(\varphi_n) = 0 \quad (\text{IX.19})$$

IX.3 Exemples

Pour illustrer cela, essayons cette définition sur les exemples déjà évoqués. Soit d'abord $T(\varphi) = \varphi(0)$ où $T = \delta$ de Dirac :

1. T est définie sur tout $\mathcal{S}(\mathbb{R})$: si φ est élément de $\mathcal{S}(\mathbb{R})$, elle est une fonction définie sur $\mathbb{R}$ et a donc une valeur bien déterminée en $x = 0$;
2. T est linéaire ; en effet, si φ_1 et φ_2 sont deux éléments de l'espace vectoriel $\mathcal{S}(\mathbb{R})$, leur somme $\varphi = \varphi_1 + \varphi_2$ est la fonction telle que $\forall x \in \mathbb{R}, \varphi(x) = \varphi_1(x) + \varphi_2(x)$, donc en particulier pour $x = 0$; de même pour $\lambda \varphi$.

⁽³⁾ L. SCHWARTZ, *Un mathématicien aux prises avec le siècle*, éd. Odile Jacob, Paris, 1997, ch. VI, p. 223–266.

⁽⁴⁾ La logique suivie jusqu'ici dans ce cours peut donner le sentiment que l'introduction de ces semi-normes est naturelle ou évidente compte tenu de la structure de l'espace $\mathcal{S}(\mathbb{R})$; ou encore, le fait qu'elles soient toutes finies peut sembler une coïncidence extraordinaire. Cependant, la voie suivie pour la découverte n'est pas celle adoptée pour ce cours : l'espace de Schwartz a été construit à partir de la nécessité d'avoir ces semi-normes et non l'inverse.

3. T est continue, au sens de (IX.18) : si pour tous j, k on a $\mathcal{N}_{j,k}(\varphi_n) \rightarrow 0$, alors on l'a en particulier pour $j = k = 0$, c'est-à-dire $\sup_{x \in \mathbb{R}} |\varphi_n(x)| \rightarrow 0$; cela veut dire que φ_n converge uniformément vers zéro pour tout x , donc en particulier pour $x = 0$.

Le cas $T(\varphi) = -\varphi'(0)$, où $T = \delta'$ est la dérivée de δ , est analogue :

1. T est définie sur tout $\mathcal{S}(\mathbb{R})$: si φ est élément de $\mathcal{S}(\mathbb{R})$, elle est une fonction définie sur $\mathbb{R}$ et infiniment dérivable, donc φ' a une valeur bien déterminée en $x = 0$;
2. T est linéaire ; en effet, si φ_1 et φ_2 sont deux éléments de l'espace vectoriel $\mathcal{S}(\mathbb{R})$, leur somme $\varphi = \varphi_1 + \varphi_2$ est la fonction telle que $\forall x \in \mathbb{R}$, $\varphi(x) = \varphi_1(x) + \varphi_2(x)$ et on a alors aussi $\varphi'(x) = \varphi_1'(x) + \varphi_2'(x)$ pour tout x , en particulier pour $x = 0$; de même pour $\lambda \varphi$.
3. T est continue : si pour tous j, k on a $\mathcal{N}_{j,k}(\varphi_n) \rightarrow 0$, alors on l'a en particulier pour $j = 1$ et $k = 0$, c'est-à-dire $\sup_{x \in \mathbb{R}} |\varphi_n'(x)| \rightarrow 0$; cela veut dire que φ_n' converge uniformément vers zéro pour tout x , donc en particulier pour $x = 0$.

Le cas des distributions I_α ($\alpha < 1$) est à peine plus compliqué. Souvenons-nous que $I_{\alpha,\varepsilon}(x)$ est la transformée de Fourier de $e^{-\varepsilon|\xi|}/[(i\xi)^\alpha]_2$. Par conséquent, d'après (VII.65), on peut écrire :

$$\int_{-\infty}^{+\infty} I_{\alpha,\varepsilon}(x) \varphi(x) dx = \int_{-\infty}^{+\infty} \frac{e^{-\varepsilon|\xi|}}{[(i\xi)^\alpha]_2} \widehat{\varphi}(\xi) d\xi \quad (\text{IX.20})$$

Posons alors que la distribution $T = I_\alpha$ ($\alpha < 1$) est définie comme étant :

$$T(\varphi) = \mathbb{R}\text{-}\lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{+\infty} I_{\alpha,\varepsilon}(x) \varphi(x) dx \quad (\text{IX.21})$$

D'après (IX.20), cela équivaut à :

$$T(\varphi) = \mathbb{R}\text{-}\lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{+\infty} \frac{e^{-\varepsilon|\xi|}}{[(i\xi)^\alpha]_2} \widehat{\varphi}(\xi) d\xi \quad (\text{IX.22})$$

Or, contrairement à l'intégrale (IX.21) dans laquelle la fonction à intégrer n'a pas de limite quand $\varepsilon \rightarrow 0$, on peut passer à la limite sous le signe intégral dans l'intégrale (IX.22) ⁽⁵⁾.

Par conséquent :

$$T(\varphi) = \int_{-\infty}^{+\infty} \frac{1}{[(i\xi)^\alpha]_2} \widehat{\varphi}(\xi) d\xi \quad (\text{IX.23})$$

On peut donc oublier maintenant (IX.21) et (IX.22) mais considérer que (IX.23) définit $T = I_\alpha$. On constate alors que :

1. T est bien définie sur tout $\mathcal{S}(\mathbb{R})$: si $\varphi \in \mathcal{S}(\mathbb{R})$, alors $\widehat{\varphi} \in \mathcal{S}(\mathbb{R})$ d'après les théorèmes VII.1 et VII.2 ;
2. T est linéaire : si $\varphi = \varphi_1 + \varphi_2$, on aura aussi $\widehat{\varphi} = \widehat{\varphi}_1 + \widehat{\varphi}_2$ et, bien entendu, l'intégrale (IX.23) est linéaire ; de même pour $\lambda \varphi$;

(5) La fonction sous le signe intégral dans (IX.22) satisfait effectivement les conditions des théorèmes généraux grâce à la présence de la fonction $\widehat{\varphi}(\xi)$.

3. T est continue : on peut réécrire (IX.23) sous la forme :

$$T(\varphi) = \int_{-\infty}^{+\infty} \frac{1}{[(i\xi)^\alpha]_2} \frac{1}{1 + \xi^{m+2}} \cdot (1 + \xi^{m+2}) \widehat{\varphi}(\xi) d\xi \quad (\text{IX.24})$$

où m est un entier qu'on choisira judicieusement.

de sorte qu'en appliquant l'inégalité de la moyenne à (IX.24), on obtient :

$$|T(\varphi)| \leq \mathcal{N}_{0,m+2}(\widehat{\varphi}) \cdot \int_{-\infty}^{+\infty} \frac{1}{[(i\xi)^\alpha]_2} \frac{1}{1 + \xi^{m+2}} d\xi \quad (\text{IX.25})$$

Il suffit de choisir m en fonction de α pour assurer la convergence de l'intégrale (IX.25) ⁽⁶⁾. On a alors l'inégalité :

$$|T(\varphi)| \leq M \cdot \mathcal{N}_{0,m+2}(\widehat{\varphi}) \quad (\text{IX.26})$$

où M est la valeur de l'intégrale en question et ne dépend pas de φ . Il ne reste plus qu'à montrer que $\mathcal{S}\text{-lim } \varphi_n = 0 \Leftrightarrow \mathcal{S}\text{-lim } \widehat{\varphi}_n = 0$, ce qui provient d'un théorème de la section suivante.

IX.4 Continuité dans l'espace $\mathcal{S}(\mathbb{R})$

Dans la définition des distributions, il y a trois conditions ; a) T est une *application* de $\mathcal{S}(\mathbb{R})$ dans $\mathbb{R}$; b) cette application est *linéaire* ; c) elle est *continue*. C'est évidemment c) qui est la nouveauté, difficile à comprendre. C'est pourquoi on y consacrera toute la présente section.

Si u_n est une suite de nombres, on dit que u_n tend vers u si $\forall \varepsilon, \exists n_0$ tel que $n \geq n_0 \Rightarrow |u - u_n| \leq \varepsilon$. Cela veut dire que plus n est grand, plus $|u - u_n|$ est petit, c'est-à-dire plus u_n est proche de u ; quand n est *très* grand, u_n est *très* proche de u . La proximité se traduit par la petitesse de $|u - u_n|$, qui est la distance de u à u_n . On voit ainsi que la notion de limite présuppose une notion de proximité.

Lorsqu'on dit qu'une suite de fonctions f_n a pour limite uniforme une fonction f , cela veut dire que plus n est grand, plus le nombre $\sup_x |f_n(x) - f(x)|$ est petit. Cette fois, c'est le nombre $\sup_x |f_n(x) - f(x)|$ qui mesure la distance de f_n à f . C'est-à-dire que f_n est proche de f si $\sup_x |f_n(x) - f(x)|$ est petit ⁽⁷⁾.

Au chapitre VII, nous avons vu deux autres notions de limite ; la limite *en moyenne* et la limite *en moyenne quadratique*. Dans ces deux cas, la distance était respectivement $\|f_n - f\|_1$ et $\|f_n - f\|_2$. La fonction f_n est alors proche de f si $\|f_n - f\|_{1,2}$ est petit. Concrètement, cela signifie que les nombres $f_n(x)$ et $f(x)$ sont proches pour la plupart des valeurs de x ; contrairement à la proximité uniforme, il peut y avoir des valeurs de x pour lesquelles les nombres $f_n(x)$ et $f(x)$ sont éloignés l'un de l'autre mais ces valeurs se regroupent sur des intervalles très courts (d'autant plus courts que n est plus grand).

Dans l'espace $\mathcal{S}(\mathbb{R})$, c'est l'ensemble des semi-normes $\mathcal{N}_{j,k}$ qui définit la proximité. La proximité est plus complexe que dans le cas des proximités uniforme ou en moyenne, car il n'y a plus un nombre unique qui mesure simplement la distance : deux fonctions φ et ψ de

⁽⁶⁾ Par exemple, on prend pour m le plus petit entier supérieur à $1 - \alpha$.

⁽⁷⁾ Cela signifie que, indépendamment de x , les nombres $f_n(x)$ et $f(x)$ sont tous proches, leur distance est petite uniformément en x .

l'espace $\mathcal{S}(\mathbb{R})$ sont proches si tous les nombres $\mathcal{N}_{j,k}(\varphi - \psi)$ sont petits, mais on ne demande pas qu'ils soient tous aussi petits en même temps, c'est-à-dire uniformément en j, k . Ainsi, dire que φ_n tend dans $\mathcal{N}_{j,k}$ vers φ , ce qui en langage mathématique s'écrit $\mathcal{S}\text{-lim } \varphi_n = \varphi$, signifie que plus n est grand, plus les nombres $\mathcal{N}_{j,k}(\varphi_n - \varphi)$ sont petits, autrement dit :

$$\forall j \in \mathbb{N}, \quad \forall k \in \mathbb{N}, \quad \mathbb{R}\text{-lim } \mathcal{N}_{j,k}(\varphi_n - \varphi) = 0 \quad (\text{IX.27})$$

Si on revient à la définition des semi-normes $\mathcal{N}_{j,k}$, l'équation (IX.27) signifie que tous les nombres $\sup_{x \in \mathbb{R}} \{(1 + |x|^k) |\varphi_n^{(j)}(x) - \varphi^{(j)}(x)|\}$ tendent vers zéro quand n tend vers l'infini ($\varphi^{(j)}$ désignant la j^{e} dérivée de φ). Ou encore : que les fonctions $(1 + |x|^k) \varphi_n^{(j)}(x)$ tendent toutes uniformément vers $(1 + |x|^k) \varphi^{(j)}(x)$. Cette notion de limite est donc bien plus forte que la limite uniforme : on ne demande pas seulement que φ_n tende uniformément vers φ , mais, en outre, que toutes les dérivées de φ_n tendent uniformément vers les dérivées correspondantes de φ et aussi que ces dérivées multipliées par n'importe laquelle des fonctions $1 + |x|^k$ tende uniformément vers la limite correspondante. Dire que φ_n tend uniformément vers φ signifie seulement que $\mathcal{N}_{0,0}(\varphi_n - \varphi)$ tend vers zéro. En résumé, $\mathcal{S}\text{-lim } \varphi_n = \varphi$ entraîne $\text{u-lim } \varphi_n = \varphi$ mais non l'inverse.

En utilisant judicieusement l'inégalité de la moyenne, on peut facilement montrer que la convergence dans $\mathcal{S}(\mathbb{R})$ entraîne aussi la convergence dans $\mathbb{L}_1(\mathbb{R})$ ou dans $\mathbb{L}_2(\mathbb{R})$:

$$\varphi = \mathcal{S}\text{-lim } \varphi_n \iff \varphi = \mathbb{L}_1\text{-lim } \varphi_n \quad \text{et} \quad \varphi = \mathbb{L}_2\text{-lim } \varphi_n \quad (\text{IX.28})$$

Il suffit en effet de vérifier que :

$$\begin{aligned} \|\varphi\|_1 &= \int_{-\infty}^{+\infty} |\varphi(x)| dx = \int_{-\infty}^{+\infty} \frac{1}{1+x^2} \cdot [(1+x^2)|\varphi(x)|] dx \\ &\leq \int_{-\infty}^{+\infty} \frac{1}{1+x^2} dx \times \sup_{x \in \mathbb{R}} \{(1+x^2)|\varphi(x)|\} = \pi \mathcal{N}_{0,2}(\varphi) \end{aligned} \quad (\text{IX.29})$$

et de manière analogue :

$$\begin{aligned} \|\varphi\|_2^2 &= \int_{-\infty}^{+\infty} |\varphi(x)|^2 dx = \int_{-\infty}^{+\infty} \frac{1}{(1+|x|)^2} \cdot [(1+|x|)|\varphi(x)|]^2 dx \\ &\leq \int_{-\infty}^{+\infty} \frac{1}{(1+|x|)^2} dx \times \sup_{x \in \mathbb{R}} \{(1+|x|)^2 |\varphi(x)|^2\} = 2 \mathcal{N}_{0,1}(\varphi)^2 \end{aligned} \quad (\text{IX.30})$$

En résumant (IX.29) et en prenant la racine carrée dans (IX.30), on a donc :

$$\|\varphi\|_1 \leq \pi \mathcal{N}_{0,2}(\varphi) \quad \text{et} \quad \|\varphi\|_2 \leq \sqrt{2} \mathcal{N}_{0,1}(\varphi) \quad (\text{IX.31})$$

En conclusion, la convergence d'une suite dans $\mathcal{S}(\mathbb{R})$ entraîne sa convergence uniforme, sa convergence en moyenne et sa convergence en moyenne quadratique.

Voyons maintenant la continuité. Une fonction f de $\mathbb{R}$ dans $\mathbb{R}$ est continue au point x_0 si pour toute suite x_n qui tend vers x_0 , $f(x_n)$ tend vers $f(x_0)$. Par analogie, on dit qu'une fonction $\mathcal{F}$ de $\mathcal{S}(\mathbb{R})$ dans $\mathcal{S}(\mathbb{R})$ est continue au point φ si pour toute suite φ_n qui tend dans $\mathcal{S}(\mathbb{R})$ vers φ , $\mathcal{F}(\varphi_n)$ tend dans $\mathcal{S}(\mathbb{R})$ vers $\mathcal{F}(\varphi)$. Il est un peu gênant d'appeler $\mathcal{F}$ une fonction alors que sa variable et sa valeur sont elles-mêmes des fonctions ; c'est pourquoi on préfère appeler $\mathcal{F}$ une *transformation* ou un *opérateur*. Pour la même raison, on a préféré appeler *fonctionnelle* ou *forme linéaire*, les applications linéaires de $\mathcal{S}(\mathbb{R})$ dans $\mathbb{R}$.

Un exemple de transformation ou d'opérateur est la transformation de Fourier. Nous avons vu au chapitre VII que la transformation de Fourier $\mathcal{F}$ était une application continue de $\mathbb{L}_2(\mathbb{R})$ dans $\mathbb{L}_2(\mathbb{R})$. Il résultait en effet de la relation de Parseval que $\|\widehat{f}\|_2 = 2\pi\|f\|_2$, ce qui entraîne évidemment que si $\|f_n - f\|_2$ tend vers zéro, il en sera de même de $\|\widehat{f_n} - \widehat{f}\|_2$. Nous allons maintenant prouver aussi que :

Théorème IX.1 La transformation de Fourier $\mathcal{F} : f \mapsto \widehat{f}$ est une application linéaire continue de $\mathcal{S}(\mathbb{R})$ dans $\mathcal{S}(\mathbb{R})$.

Autrement dit, si φ_n tend vers zéro dans $\mathcal{S}(\mathbb{R})$, alors $\widehat{\varphi_n}$ tend également vers zéro dans $\mathcal{S}(\mathbb{R})$.

Ⓡ D'après ce qui précède, si φ_n tend vers zéro dans $\mathcal{S}(\mathbb{R})$ alors φ_n tend vers zéro dans $\mathbb{L}_2(\mathbb{R})$; par conséquent, il s'avère aussi que $\widehat{\varphi_n}$ tend vers zéro dans $\mathbb{L}_2(\mathbb{R})$. Cela signifie que la transformation $\mathcal{F}$ est continue si on considère la limite selon les $\mathcal{N}_{j,k}$ dans l'espace de départ et la limite selon $\|\cdot\|_2$ dans l'espace d'arrivée. Par contre, on ne peut pas en conclure la continuité pour la limite selon les $\mathcal{N}_{j,k}$ dans l'espace d'arrivée car la convergence selon les $\mathcal{N}_{j,k}$ entraîne la convergence selon $\|\cdot\|_2$ mais non l'inverse. C'est pourquoi il faut procéder directement.

Preuve (théorème IX.1) On va établir des inégalités, qui toutes seront déduites de l'*inégalité de la moyenne*. Rappelons que la transformée de Fourier de la fonction $x \mapsto D^j f(x)$ (où $D^j f$ est la j^{e} dérivée de f) est la fonction $\xi \mapsto (i\xi)^j \widehat{f}(\xi)$ et que la transformée de Fourier de la fonction $x \mapsto x^k f(x)$ est la fonction $\xi \mapsto -iD^k \widehat{f}(\xi)$ (théorème VII.1). Par conséquent :

$$(-i\xi)^k D^j \widehat{f}(\xi) = \int_{-\infty}^{+\infty} D^k [(ix)^j f(x)] e^{ix\xi} dx \quad (\text{IX.32})$$

D'après la formule de Leibniz pour la dérivation d'un produit, on a :

$$D^k [(ix)^j f(x)] = \sum_{\ell=0}^{\ell=k} M_{j,k,\ell} x^{j-k+\ell} D^\ell f(x) \quad (\text{IX.33})$$

où les $M_{j,k,\ell}$ sont des constantes (qu'on peut exprimer à l'aide de factorielles de $j, k, \ell, j-k, j-\ell$, etc., mais il ne sert à rien ici d'en avoir une expression exacte). En écrivant la fonction sous $\int$ dans (IX.32) sous la forme :

$$D^k [(ix)^j f(x)] = \frac{1}{1+x^2} \cdot (1+x^2) D^k [(ix)^j f(x)] \quad (\text{IX.34})$$

et en appliquant l'inégalité de la moyenne, puis (IX.33), on obtient :

$$\begin{aligned} |\xi|^k |D^j \widehat{f}(\xi)| &\leq \int_{-\infty}^{+\infty} \frac{1}{1+x^2} dx \times \sup_{x \in \mathbb{R}} |(1+x^2) D^k [(ix)^j f(x)]| \\ &\leq \pi \sum_{\ell=0}^{\ell=k} \left[M_{j,k,\ell} \sup_{x \in \mathbb{R}} |x^{j-k+\ell} D^\ell f(x)| + M_{j+2,k,\ell} \sup_{x \in \mathbb{R}} |x^{j+2-k+\ell} D^\ell f(x)| \right] \\ &= \pi \sum_{\ell=0}^{\ell=k} [M_{j,k,\ell} \mathcal{N}_{\ell,j-k+\ell}(f) + M_{j+2,k,\ell} \mathcal{N}_{\ell,j+2-k+\ell}(f)] \end{aligned} \quad (\text{IX.35})$$

Pour $k=0$, l'inégalité (IX.35) devient aussi :

$$|D^j \widehat{f}(\xi)| \leq \pi [M_{j,0,0} \mathcal{N}_{\ell,j}(f) + M_{j+2,0,0} \mathcal{N}_{\ell,j}(f)] \quad (\text{IX.36})$$

Les inégalités (IX.35) et (IX.36) sont vraies pour tout $\xi \in \mathbb{R}$; si on les ajoute membre à membre, alors :

$$\mathcal{N}_{j,k}(\widehat{f}) \leq \pi \times \sum_{j',k'} C^{\text{te}} \cdot \mathcal{N}_{j',k'}(f) \quad (\text{IX.37})$$

où la somme comporte un nombre fini de termes égaux à une constante multipliée par l'une des semi-normes (il importe peu de savoir exactement lesquelles). Si on remplace ci-dessus f par $\varphi_n - \varphi$, on en déduit que :

$$\mathcal{S}\text{-}\lim_{n \rightarrow \infty} (\varphi_n - \varphi) = 0 \Rightarrow \forall j', k', \mathbb{R}\text{-}\lim_{n \rightarrow \infty} \mathcal{N}_{j', k'}(\varphi_n - \varphi) = 0 \Rightarrow \mathbb{R}\text{-}\lim_{n \rightarrow \infty} \mathcal{N}_{j, k}(\widehat{\varphi_n} - \widehat{\varphi}) = 0 \quad (\text{IX.38})$$

et d'après (IX.37), cela est vrai quels que soient les entiers j et k . •

Un autre opérateur très courant est la dérivation : $\varphi \mapsto d\varphi/dx$. Contrairement à la transformation de Fourier, celui-ci n'est pas continu sur $\mathbb{L}_2(\mathbb{R})$: si on prend par exemple $f_n(x) = \frac{1}{n} e^{-\frac{x^2}{2} + inx}$, un calcul simple montre que $\|f_n\|_2 = \frac{1}{n} \pi^{1/4}$ qui tend vers zéro, mais la dérivée est $f'_n(x) = (i - \frac{x}{n}) e^{-\frac{x^2}{2} + inx}$ et le calcul donne $\|f'_n\|_2 = \sqrt{1 + 1/2n^2} \pi^{1/4}$ qui ne tend pas vers zéro. Par contre, la dérivation devient continue pour le type très particulier de limite considéré sur l'espace $\mathcal{S}(\mathbb{R})$:

Théorème IX.2 L'opérateur de dérivation :

$$D : \begin{bmatrix} \mathcal{S}(\mathbb{R}) & \longrightarrow & \mathcal{S}(\mathbb{R}) \\ \varphi & \longmapsto & \frac{d\varphi}{dx} = \varphi' \end{bmatrix} \quad (\text{IX.39})$$

est continu pour la limite définie par les semi-normes $\mathcal{N}_{j, k}$.

Preuve Cela résulte immédiatement de la relation :

$$\mathcal{N}_{j, k}(\varphi') = \mathcal{N}_{j+1, k}(\varphi) \quad (\text{IX.40})$$

qui est évidente compte tenu de la définition des semi-normes. •

IX.5 Intégrales avec poids et extension

Il a déjà été dit en section IX.3 qu'une intégrale est une forme linéaire. Maintenant que nous avons une idée plus claire de la continuité sur les espaces de fonctions, notamment $\mathcal{S}(\mathbb{R})$, nous pouvons mieux discuter cela. On va donc reprendre quelques questions déjà discutées dans les sections IX.3 et IX.4.

Soit la fonctionnelle :

$$\mathcal{I} : \begin{bmatrix} E & \rightarrow & \mathbb{R} \\ f & \mapsto & \int_{-\infty}^{+\infty} f(x) dx \end{bmatrix} \quad (\text{IX.41})$$

où E est l'un des espaces de fonctions déjà étudiés, avec la notion particulière de limite qui le caractérise ; par exemple :

1. $E = \mathbb{L}_1(\mathbb{R})$, où $f_n \rightarrow 0$ signifie $\|f_n\|_1 \rightarrow 0$;
2. $E = \mathbb{L}_1(\mathbb{R}) \cap \mathbb{L}_2(\mathbb{R})$, sous-espace de $\mathbb{L}_2(\mathbb{R})$, où $f_n \rightarrow 0$ signifie $\|f_n\|_2 \rightarrow 0$;
3. $E = \mathcal{S}(\mathbb{R})$, où $f_n \rightarrow 0$ signifie que $\forall j, k \geq 0, \mathcal{N}_{j, k}(f_n) \rightarrow 0$.

La fonctionnelle $\mathcal{I}$ est évidemment linéaire (c'est l'intégrale) ; dans les cas 1 et 3, elle est continue : c'est quasiment tautologique pour 1 et résulte de l'inégalité de la moyenne pour 3. Par contre pour 2, elle est discontinue, comme déjà vu en section IX.3.

Plus généralement, on peut considérer les intégrales avec poids. Une fonction $p(x)$ étant fixée (le poids), on pose :

$$\mathcal{J}_p : \left[\begin{array}{ccc} E & \rightarrow & \mathbb{R} \\ f & \mapsto & \int_{-\infty}^{+\infty} p(x) f(x) dx \end{array} \right] \quad (\text{IX.42})$$

Ceci est une fonctionnelle linéaire ; elle est définie et continue dans les cas suivants :

1. $E = \mathbb{L}_1(\mathbb{R})$ et p est une fonction continue bornée ; cela résulte immédiatement de l'inégalité de la moyenne.
2. $E = \mathbb{L}_2(\mathbb{R})$ et $p \in \mathbb{L}_2(\mathbb{R})$; cela résulte de l'inégalité de Schwartz :

$$\left| \int_{-\infty}^{+\infty} p(x) f(x) dx \right| \leq \|p\|_2 \cdot \|f\|_2. \quad (\text{IX.43})$$

3. $E = \mathcal{S}(\mathbb{R})$ et p est une fonction à croissance polynomiale ; cela résulte de l'inégalité de la moyenne : que p soit à croissance polynomiale signifie qu'il existe une constante M et un entier k tels que $p(x) \leq M(1 + |x|^k)$, donc :

$$\mathcal{J}(f) = \int_{-\infty}^{+\infty} p(x) f(x) dx = \int_{-\infty}^{+\infty} \frac{p(x)}{1 + |x|^k} \cdot (1 + |x|^k) f(x) dx \quad (\text{IX.44})$$

d'où :

$$|\mathcal{J}(f)| \leq M \int_{-\infty}^{+\infty} (1 + |x|^k) |f(x)| dx \leq M \cdot \mathcal{N}_{0,k+2}(f) \cdot \int_{-\infty}^{+\infty} \frac{1 + |x|^k}{1 + |x|^{k+2}} dx \quad (\text{IX.45})$$

L'intégrale (IX.45) ne dépend que de k (donc de p) et non de f , ce qui suffit pour garantir que $\mathcal{N}_{0,k+2}(f) \rightarrow 0 \implies \mathcal{J}(f) \rightarrow 0$.

On pourrait aussi trouver des cas où $\mathcal{J}$ est discontinue.

Voici des conclusions utiles :

Théorème IX.3 Si le poids p est à croissance polynomiale, en particulier si p est un polynôme, la fonctionnelle $\mathcal{J}_p$ définit une distribution.

Preuve Il suffit de remarquer que, conformément à l'étude du cas 3, $\mathcal{J}_p$ est alors continue sur $\mathcal{S}(\mathbb{R})$. •

La réciproque est fautive : toute fonctionnelle linéaire continue sur $\mathcal{S}(\mathbb{R})$ n'est pas forcément une intégrale avec poids. La preuve en est la fonctionnelle introduite plus haut :

$$\delta : \left[\begin{array}{ccc} \mathcal{S}(\mathbb{R}) & \rightarrow & \mathbb{R} \\ f & \mapsto & f(0) \end{array} \right] \quad (\text{IX.46})$$

Si on voulait la représenter sous la forme d'une intégrale avec poids, cela donnerait :

$$f(0) = \int_{-\infty}^{+\infty} \delta(x) f(x) dx \quad (\text{IX.47})$$

ce qui est certes une notation courante, mais dont le sens est donné par la notion de fonctionnelle et non par la notion d'intégrale.

La notation (IX.47) utilise le signe intégral dans un sens *étendu* et non dans son sens usuel (celui de l'intégrale de Riemann ou de Lebesgue). Une intégrale est une limite de sommes

discrètes, tandis qu'une fonctionnelle linéaire est définie abstraitement. Il est tout à fait légitime d'utiliser la notation (IX.47) mais on ne peut pas lui appliquer sans discernement les théorèmes qui s'appliquent aux vraies intégrales, tels que par exemple l'inégalité de la moyenne ou l'inégalité de Schwartz, ou encore, des théorèmes de passage à la limite. Il faudra appliquer des théorèmes qui sont valables pour les fonctionnelles et que nous verrons.

Si T est une distribution, on peut l'écrire comme fonctionnelle $T(\varphi)$ mais aussi comme dans (IX.47) sous la forme :

$$T(\varphi) = \int_{-\infty}^{+\infty} T(x) \varphi(x) dx \quad (\text{IX.48})$$

Une autre notation très commode est la notation *en produit scalaire* $\langle T, \varphi \rangle$. On peut donc indifféremment écrire :

$$T(\varphi) \equiv \langle T, \varphi \rangle \equiv \int_{-\infty}^{+\infty} T(x) \varphi(x) dx \quad (\text{IX.49})$$

Il faut simplement éviter de prendre au sérieux la variable muette x dans la notation en forme d'intégrale.

Si $\langle T, \varphi \rangle$ est une intégrale avec le poids $p(x)$, on peut prendre $p(x)$ pour $T(x)$ et ainsi $T(x)$ sera une véritable fonction au sens usuel et (IX.48), une véritable intégrale. On dit alors que T est une *distribution régulière*.

Par contre, si T est une fonctionnelle qui ne se réduit pas à un poids (comme par exemple δ), on ne pourra pas attribuer une valeur numérique à $T(x)$, du moins pas pour tout x . On dit alors que T est une *distribution singulière*. La distribution T n'étant alors pas une fonction au sens usuel, elle n'a pas de valeur numérique et ne se définit que par rapport aux fonctions φ qui sont dans l'espace $\mathcal{S}(\mathbb{R})$.

On peut revenir à l'analogie faite au début du chapitre avec l'électrostatique : si T représente une densité continue de charges $p(x, y, z)$, le potentiel de ces charges sera :

$$V(X, Y, Z) = \int_{\mathbb{R}^3} \frac{p(x, y, z) dx dy dz}{\sqrt{(x-X)^2 + (y-Y)^2 + (z-Z)^2}} \quad (\text{IX.50})$$

Si les charges deviennent ponctuelles ou encore plus singulières (dipôles, quadrupôles etc.), la densité p cessera d'être une vraie fonction (elle deviendra infinie là où il y a les charges et nulle partout ailleurs) mais le potentiel restera caractéristique de la distribution des charges ; cette dernière sera reconnaissable d'après le potentiel qu'elle crée et qui, lui, est une vraie fonction (quoique ayant des singularités). C'est la même idée que nous suivons ici, mais au lieu de considérer les fonctions $1/\sqrt{(x-X)^2 + (y-Y)^2 + (z-Z)^2}$, qui deviennent infinies en $x = X, y = Y, z = Z$ et qui par conséquent sont peu commodes pour l'analyse on préfère utiliser l'espace $\mathcal{S}(\mathbb{R})$.

Cette idée de généraliser la notion usuelle de fonction ou d'intégrale interdit d'appliquer des théorèmes qui ont été prévus pour les vraies fonctions et les vraies intégrales. Cependant, la théorie des distributions est construite de telle manière à disposer de propriétés commodes, en particulier pour les calculs.

Revenons encore sur la propriété la plus importante des distributions : la continuité en tant que fonctionnelle. Elle signifie que si $\mathcal{S}\text{-lim } \varphi_n = \varphi$, alors $\mathbb{R}\text{-lim } T(\varphi_n) = T(\varphi)$. Si on

interprète la distribution comme une fonction *généralisée* et $\langle T, \varphi_n \rangle$ ou $\langle T, \varphi \rangle$ comme des intégrales généralisées, en utilisant la notation (IX.48), on aura :

$$\mathbb{R}\text{-}\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} T(x) \varphi_n(x) dx = \int_{-\infty}^{+\infty} T(x) \varphi(x) dx \quad (\text{IX.51})$$

Cela exprime la propriété de passage à la limite sous le signe intégral. Ainsi, cette propriété qui pour les vraies intégrales exige une vérification soigneuse de sa validité, est ici partie intégrante de la définition. Le passage à la limite sous le signe intégral est donc automatique pour les distributions et n'exige aucune précaution (sauf évidemment la garantie que T est bien une distribution).

IX.6 À propos de l'espace $\mathcal{S}(\mathbb{R})$

La notion de limite choisie pour l'espace $\mathcal{S}(\mathbb{R})$ (celle définie par les semi-normes $\mathcal{N}_{j,k}$) est extrêmement restrictive : on a vu que la limite uniforme (qui est déjà la plus forte parmi les limites usuelles) est la limite selon la semi-norme $\mathcal{N}_{0,0}$; dans l'espace $\mathcal{S}(\mathbb{R})$ on exige encore une infinité d'autres semi-normes. Pour qu'une suite φ_n de fonctions tende vers zéro dans $\mathcal{S}(\mathbb{R})$, il *faut* qu'elle tende uniformément vers zéro, mais cela ne suffit de loin pas : il faut en plus que toutes ses dérivées tendent uniformément vers zéro et encore que toutes ces dérivées, multipliées par n'importe quel polynôme — ou, ce qui est équivalent, par n'importe quel facteur $1 + |x|^k$ — tendent uniformément vers zéro. Pour en donner une idée un peu plus concrète, voici des exemples.

Exemple IX.1 — $\varphi_n(x) = e^{-n} e^{-x^2/n}$ Cette suite tend vers zéro dans $\mathcal{S}(\mathbb{R})$. Pour le vérifier sans calculs, on procède comme suit : posons $X = x/\sqrt{n}$ et $\phi(X) = e^{-X^2}$. Si $\phi^{(j)}(X)$ est la j^{e} dérivée de ϕ , alors :

$$\frac{d^j}{dx^j} \varphi_n(x) = \frac{e^{-n}}{n^{j/2}} \phi^{(j)}\left(\frac{x}{\sqrt{n}}\right) \quad (\text{IX.52})$$

Le maximum de $|d^j \varphi_n / dx^j|$, c'est-à-dire la semi-norme $\mathcal{N}_{j,0}(\varphi_n)$, est évidemment une expression compliquée, d'autant plus compliquée que j est plus grand. La relation ci-dessus montre que ce maximum est égal à $e^{-n}/n^{j/2}$ fois celui de $|\phi^{(j)}(X)|$, qui est compliqué aussi mais *qui ne dépend pas de n* . Ainsi :

$$\mathcal{N}_{j,0}(\varphi_n) = \frac{e^{-n}}{n^{j/2}} \max \{ |\phi^{(j)}| \} \quad (\text{IX.53})$$

tend bien vers zéro. Avec les facteurs $|x|^k$, on peut écrire :

$$|x|^k \frac{d^j}{dx^j} \varphi_n(x) = \frac{e^{-n} n^{k/2}}{n^{j/2}} \cdot |X|^k \phi^{(j)}(X) \quad (\text{IX.54})$$

Comme avant et bien que cette fois $n^{k/2}$ soit au numérateur, le facteur e^{-n} l'emporte et le coefficient $e^{-n} n^{k/2}/n^{j/2}$ tend vers zéro, tandis que $\max\{|X|^k |\phi^{(j)}(X)|\}$ reste constant, d'où le résultat. $\square$

La vérification est assez longue car il y a beaucoup de semi-normes : il faut vérifier que *chacun* des $\mathcal{N}_{j,k}(\varphi_n)$ tend vers zéro. Voici maintenant un contre-exemple.

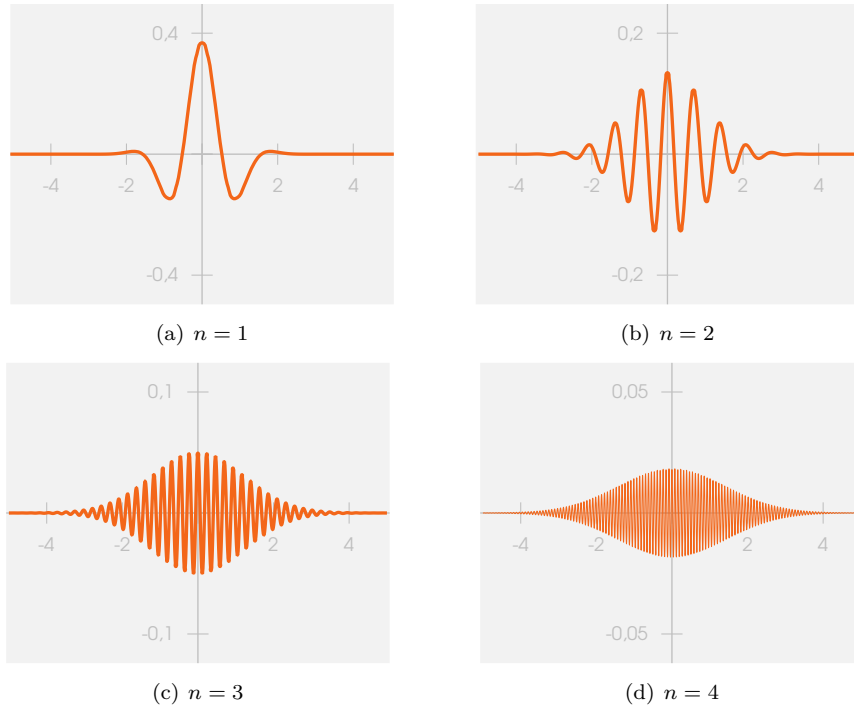


Figure IX.2 – Quatre premières fonctions de la suite $\varphi_n(x) = e^{-n} e^{-x^2/n} \cos(3^n x)$. Ces fonctions sont uniformément de plus en plus petites mais leurs dérivées de plus en plus grandes

Exemple IX.2 — $\varphi_n(x) = e^{-n} \cos(3^n x) e^{-x^2/n}$ On voit, sans calcul, que le maximum de cette fonction est atteint pour $x = 0$ et vaut e^{-n} ; il tend donc vers zéro, donc cette suite de fonctions tend uniformément vers zéro (voir figure IX.2). Par contre sa dérivée est :

$$\frac{d}{dx} \varphi_n(x) = -e^{-n} 3^n \sin(3^n x) e^{-x^2/n} - e^{-n} (2x/n) \cos(3^n x) e^{-x^2/n} \quad (\text{IX.55})$$

on voit que pour $x = 3^{-n}\pi/2$, cette dérivée vaut en valeur absolue $(3/e)^n e(\pi^2/4n3^{2n})$, ce qui tend vers l'infini; donc à plus forte raison son maximum tend vers l'infini. Cela montre que cette suite converge uniformément vers zéro mais pas dans $\mathcal{S}(\mathbb{R})$. $\square$

Exemple IX.3 — $\varphi_n(x) = e^{-n} \cos(1,001^n x) e^{-x^2/n}$ Inutile de faire les calculs en détail mais on devine qu'en dérivant j fois, on aura un facteur $1,001^j e^{-n} = [1,001^j/e]^n$; celui-ci va tendre vers zéro tant que $1,001^j/e < 1$, c'est-à-dire tant que $j < 1/\ln(1,001) \simeq 1\,000,499\,917$ (donc $j \leq 1\,000$), puis tendra vers l'infini quand j sera égal à 1 001 ou plus. On a ainsi une suite de fonctions pour lesquelles $\mathcal{N}_{j,0}(\varphi_n)$ tend vers zéro tant que $j \leq 1\,000$ mais $\mathcal{N}_{1\,001,0}(\varphi_n)$ tend vers l'infini. $\square$

La question qu'on peut se poser est : "pourquoi définir la limite de façon aussi restrictive?" Pour y répondre, notons que plus le type de limite est contraignant, moins il y a de suites convergentes. Il y aura donc d'autant plus de fonctionnelles linéaires continues : en effet, T est continue si pour toutes les suites φ_n qui tendent vers zéro, $\langle T, \varphi_n \rangle$ tend vers zéro; il est évident qu'en réduisant l'ensemble des suites qui tendent vers zéro, on rend moins exigeante la condition que doit satisfaire T . Si par exemple on posait pour la limite d'une suite, la définition la plus contraignante possible, à savoir $\varphi_n \rightarrow 0$ si $\forall n, \varphi_n = 0$,

alors il n'y aurait qu'une seule suite qui tend vers zéro et toutes les fonctionnelles linéaires seraient continues. Cela ne serait pas intéressant car on ne pourrait pas avoir de règles de calcul opératoires. Les *bonnes règles de calcul* sont conditionnées par l'existence effective d'espaces fonctionnels tels que $\mathcal{S}(\mathbb{R})$, qui ont les bonnes propriétés et qu'il faut découvrir.

Après le succès de la théorie des distributions, les mathématiciens ont cherché à pousser plus loin l'idée de *fonctions généralisées* ; dans tous les cas, c'est la nature des limites qui détermine l'efficacité algébrique des nouveaux objets introduits car on calcule principalement par passage à la limite. Or, on n'a pas trouvé mieux que les distributions de Schwartz : ce qui est plus général ne fonctionne qu'avec des règles de calcul compliquées et ce qui est plus simple est insuffisant pour les applications (essentiellement : intégrales divergentes, fonctions de Green). L'espace $\mathcal{S}(\mathbb{R})$ est le meilleur compromis connu.

Cette relation entre le choix d'un espace fonctionnel possédant un type de limite spécifique et la nature et les propriétés des fonctionnelles continues qui en résultent, peut encore être mis en évidence en prenant l'espace $\mathbb{L}_2(\mathbb{R})$. Nous avons vu que la limite dans $\mathbb{L}_2(\mathbb{R})$ est beaucoup moins contraignante que dans $\mathcal{S}(\mathbb{R})$. D'après les remarques précédentes, on s'attend donc à avoir sur $\mathbb{L}_2(\mathbb{R})$ beaucoup moins de fonctionnelles linéaires continues. C'est bien le cas. Un théorème de F. Riesz que l'on admettra dit que toute fonctionnelle linéaire continue sur $\mathbb{L}_2(\mathbb{R})$ est forcément une intégrale avec poids. De façon précise :

Théorème IX.4 Soit

$$T : \left[\begin{array}{ccc} \mathbb{L}_2(\mathbb{R}) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \langle T, f \rangle \end{array} \right] \quad (\text{IX.56})$$

une fonctionnelle linéaire continue, c'est-à-dire telle que

$$\lim_{n \rightarrow \infty} \|f_n\|_2 = 0 \implies \lim_{n \rightarrow \infty} \langle T, f_n \rangle = 0 \quad (\text{IX.57})$$

Alors il existe une fonction p dans $\mathbb{L}_2(\mathbb{R})$ telle que :

$$\forall f \in \mathbb{L}_2(\mathbb{R}), \quad \langle T, f \rangle = \int_{-\infty}^{+\infty} p(x) f(x) dx \quad (\text{IX.58})$$

La démonstration de ce théorème exige des connaissances mathématiques que nous ne verrons qu'au chapitre XI, c'est pourquoi on ne la donne pas ici. Cependant, sa signification est claire : si on prend $\mathbb{L}_2(\mathbb{R})$ comme espace de référence, les fonctionnelles linéaires continues ne fournissent rien de nouveau par rapport aux fonctions usuelles. Pour avoir *plus* que les simples fonctions, il faut avoir *plus* de fonctionnelles et donc il faut abaisser la contrainte de la continuité en renforçant la contrainte sur les suites convergentes. La commodité des règles de calcul sur les fonctionnelles sera évidemment liée aux propriétés de l'espace de référence. Par exemple, le fait que la transformation de Fourier transforme $\mathcal{S}(\mathbb{R})$ en lui-même et y est continue, permet de définir la transformée de Fourier des distributions, avec des règles de calcul simples et naturelles. S'il n'existait aucun espace autre que $\mathbb{L}_2(\mathbb{R})$ sur lequel la transformée de Fourier est continue, on ne pourrait pas étendre la transformation de Fourier aux distributions.

Les distributions sont donc définies abstraitement : on les identifie aux fonctions lorsque (en tant que fonctionnelles) elles sont des intégrales avec poids et on les considère comme des fonctions généralisées dans le cas contraire, celui qui justement ne se produit pas sur $\mathbb{L}_2(\mathbb{R})$. Il ne faudra jamais oublier que les distributions ne sont que des artifices mathématiques ; il n'y a pas plus d'information dans l'expression :

$$\int_{-\infty}^{+\infty} \delta'(x) \varphi(x) dx = \varphi'(0) \quad (\text{IX.59})$$

que dans l'expression :

$$\lim_{n \rightarrow \infty} \frac{2n^{3/2}}{\sqrt{\pi}} \int_{-\infty}^{+\infty} x e^{-nx^2} \varphi(x) dx = \varphi'(0) \quad (\text{IX.60})$$

La différence est qu'il existe des règles de calcul simples pour les expressions comme la première qui évitent d'avoir à traîner la seconde dans les calculs.

IX.7 Dérivation des distributions

L'idée de la théorie des distributions est d'étendre les règles de calcul sur les fonctions telles que dérivation, intégration, transformation de Fourier, etc., aux fonctionnelles linéaires continues. Pour que ces règles soient cohérentes, il faut que, lorsque la fonctionnelle s'identifie à une fonction usuelle (c'est-à-dire lorsqu'elle est une intégrale avec poids), ces opérations coïncident avec leur sens usuel. Par conséquent, l'extension des opérations va toujours partir de l'analogie avec les intégrales.

Voyons d'abord la dérivation. Si p est une fonction poids (à croissance polynomiale) et $\varphi \in \mathcal{S}(\mathbb{R})$, la fonctionnelle associée à la dérivée p' sera :

$$\varphi \mapsto \int_{-\infty}^{+\infty} p'(x) \varphi(x) dx \quad (\text{IX.61})$$

Une intégration par parties montre que ceci est égal à :

$$- \int_{-\infty}^{+\infty} p(x) \varphi'(x) dx \quad (\text{IX.62})$$

le produit $p(x) \varphi(x)$ est en effet nul à l'infini. Autrement dit, si $T_p(\varphi)$ désigne la fonctionnelle $\varphi \mapsto \int p(x) \varphi(x) dx$, on aura la relation :

$$\langle T_{p'}, \varphi \rangle = -\langle T_p, \varphi' \rangle \quad (\text{IX.63})$$

Si p est une fonction poids dérivable et qu'on l'identifie à la distribution T_p , cela signifie qu'on obtient la dérivée d'une distribution T , en faisant opérer T sur $-\varphi'$. On va donc étendre l'opération de dérivation aux fonctionnelles quelconques en posant :

$$\left\langle \frac{dT}{dx}, \varphi \right\rangle = -\langle T, \varphi' \rangle \quad (\text{IX.64})$$

En itérant la définition (IX.64), la dérivée d'ordre j d'une distribution T sera définie par :

$$\left\langle \frac{d^j T}{dx^j}, \varphi \right\rangle = (-1)^j \left\langle T, \frac{d^j \varphi}{dx^j} \right\rangle \quad (\text{IX.65})$$

Pour que cette définition de la dérivée soit cohérente, il faut vérifier que la fonctionnelle $\varphi \mapsto \langle T, \frac{d\varphi}{dx} \rangle$ est linéaire et continue. La linéarité est évidente. La continuité résulte directement du théorème IX.2 : si une suite φ_n tend vers zéro dans $\mathcal{S}(\mathbb{R})$, alors d'après ce théorème, la suite des dérivées φ'_n , tendra aussi vers zéro et donc de même $\langle T, \varphi'_n \rangle$, puisque T était supposée être une distribution. Autrement dit : si T est une fonctionnelle linéaire continue, il en est de même de T' .

Cette définition étendue de la dérivée donne des résultats intéressants si on l'applique à des vraies fonctions au sens usuel, mais non dérivables. L'exemple classique est évidemment la fonction de Heaviside :

$$p(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 & \text{si } x > 0 \end{cases} \quad (\text{IX.66})$$

celle-ci définit bien une intégrale avec poids :

$$\langle \mathcal{H}, \varphi \rangle = \int_{-\infty}^{+\infty} p(x) \varphi(x) dx = \int_0^{\infty} \varphi(x) dx \quad (\text{IX.67})$$

Si on applique la définition (IX.64), cela donne :

$$\langle \mathcal{H}', \varphi \rangle = - \int_0^{\infty} \varphi'(x) dx = \varphi(0) \quad (\text{IX.68})$$

Autrement dit, la dérivée de $\mathcal{H}$ est la distribution δ de Dirac (cf. (IX.10a)).

Une autre exemple très simple de fonction non dérivable est la fonction $p(x) = |x|$. C'est bien une fonction à croissance polynomiale, de sorte que l'intégrale avec poids correspondante est une fonctionnelle continue :

$$\left| \int_{-\infty}^{+\infty} |x| \varphi(x) dx \right| \leq \int_{-\infty}^{+\infty} \frac{|x|}{1+|x|^3} dx \times \mathcal{N}_{0,3}(\varphi) = \frac{4\pi}{3\sqrt{3}} \mathcal{N}_{0,3}(\varphi) \quad (\text{IX.69})$$

La dérivée de cette distribution est :

$$\langle T, \varphi \rangle = - \int_{-\infty}^{+\infty} |x| \varphi'(x) dx = \int_{-\infty}^0 x \varphi'(x) dx - \int_0^{+\infty} x \varphi'(x) dx \quad (\text{IX.70})$$

En intégrant par parties les deux intégrales obtenues :

$$\begin{aligned} \int_{-\infty}^0 x \varphi'(x) dx &= x \varphi(x) \Big|_{-\infty}^0 - \int_{-\infty}^0 \varphi(x) dx - \int_0^{+\infty} x \varphi'(x) dx \\ &= -x \varphi(x) \Big|_0^{+\infty} + \int_0^{+\infty} \varphi(x) dx \end{aligned} \quad (\text{IX.71})$$

Les termes tout intégrés étant nuls, il reste :

$$\langle T, \varphi \rangle = - \int_{-\infty}^0 \varphi(x) dx + \int_0^{+\infty} \varphi(x) dx \quad (\text{IX.72})$$

qui est l'intégrale avec le poids :

$$p(x) = \begin{cases} -1 & \text{si } x < 0 \\ +1 & \text{si } x > 0 \end{cases} \quad (\text{IX.73})$$

La fonction $|x|$ est dérivable pour tout $x \neq 0$ et sa dérivée coïncide alors avec ce résultat ; au point $x = 0$, il n'y a pas de dérivée au sens usuel et au sens des distributions non

plus, puisque les distributions n'ont pas une valeur en chaque point : elles ne sont définies que comme fonctionnelle, ce qui correspond exactement à l'égalité (IX.72). Ceci illustre la compatibilité de l'extension.

On sait que pour les fonctions, seules les constantes ont une dérivée nulle, ce qui a pour conséquence que les primitives sont déterminées à une constante additive près. En est-il de même pour les distributions ? Cette question théorique est importante car les distributions ont été inventées en grande partie pour résoudre les équations différentielles, dont les solutions sont déterminées à une constante d'intégration près. Trouver *toutes* les solutions possibles d'une équation différentielle est un problème qui se ramène de près ou de loin à celui de trouver *toutes* les solutions de $dT/dx = 0$.

La théorie des distributions se construit par analogie à partir des intégrales avec poids ; il est donc naturel de dire que la distribution nulle est la distribution définie par l'intégrale de poids zéro, c'est-à-dire la distribution T telle que $\langle T, \varphi \rangle = 0$ pour toute fonction φ de l'espace $\mathcal{S}(\mathbb{R})$. De même la distribution correspondant à la constante C doit être l'intégrale de poids $p(x) = C$, soit :

$$\langle C, \varphi \rangle = C \int_{-\infty}^{+\infty} \varphi(x) dx \quad (\text{IX.74})$$

La relation (IX.63) identifie la dérivée au sens des distributions à la dérivée au sens usuel, de sorte que, dans la théorie des distributions aussi, les constantes ont une dérivée nulle. Réciproquement, on sait que les fonctions de dérivée nulle sont constantes, mais qu'est-ce qui prouve qu'il n'y aurait pas des distributions singulières à dérivée nulle ? Il ne s'agit pas d'une simple évidence et il faut voir cela de plus près :

Théorème IX.5 Toute distribution dont la dérivée est nulle est une distribution du type (IX.74).

Preuve Dire que $T' = 0$ équivaut à dire que pour toute fonction $\varphi \in \mathcal{S}(\mathbb{R})$, on a $\langle T, \varphi' \rangle = 0$. On ne peut en déduire que $T = 0$ car les fonctions φ' , dérivées d'une fonction de $\mathcal{S}(\mathbb{R})$, ne peuvent pas être n'importe quelle fonction de $\mathcal{S}(\mathbb{R})$, soit, en langage mathématique : l'application $\varphi \mapsto \varphi'$ de $\mathcal{S}(\mathbb{R})$ dans $\mathcal{S}(\mathbb{R})$ n'est pas surjective. Cela se voit très bien sur la fonction e^{-x^2} qui appartient à $\mathcal{S}(\mathbb{R})$ mais n'est la dérivée d'aucune fonction de $\mathcal{S}(\mathbb{R})$; une primitive en est $\text{erf}(x) = \int_{-\infty}^x e^{-t^2} dt$ qui tend rapidement vers zéro pour $x \rightarrow -\infty$ mais non pour $x \rightarrow +\infty$, où elle tend vers $\sqrt{\pi}$. La fonction $\text{erf}(x) - \sqrt{\pi}$ a la même dérivée et tend rapidement vers zéro pour $x \rightarrow +\infty$ mais non plus pour $x \rightarrow -\infty$. On ne peut trouver aucune constante qui, ajoutée à $\text{erf}(x)$, donne une fonction de $\mathcal{S}(\mathbb{R})$, c'est-à-dire une fonction qui tend rapidement vers zéro à la fois pour $x \rightarrow +\infty$ et pour $x \rightarrow -\infty$.

Il est cependant facile de vérifier que si $\varphi \in \mathcal{S}(\mathbb{R})$, alors la primitive $\phi_1(x) = \int_{-\infty}^x \varphi(t) dt$ tend rapidement vers zéro pour $x \rightarrow -\infty$ et la primitive $\phi_2(x) = \int_{+\infty}^x \varphi(t) dt$ tend rapidement vers zéro pour $x \rightarrow +\infty$ (il suffit d'appliquer l'inégalité de la moyenne et les inégalités $|\varphi(t)| \leq M |t|^{-k}$), mais ces deux primitives ne sont pas forcément égales ; toutefois elles diffèrent évidemment d'une constante, puisque toutes deux ont pour dérivée φ : cette constante est $\phi_1 - \phi_2 = \int_{-\infty}^{\infty} \varphi(t) dt$. On voit ainsi que les fonctions qui sont la dérivée d'une fonction de $\mathcal{S}(\mathbb{R})$ sont caractérisées par le fait que les deux primitives ϕ_1 et ϕ_2 sont, justement, égales, car alors elles tendent rapidement vers zéro à la fois pour $x \rightarrow -\infty$ et $x \rightarrow +\infty$; ou encore autrement dit, ces fonctions sont caractérisées par le fait que $\int_{-\infty}^{\infty} \varphi(t) dt = 0$.

Prenons maintenant une fonction φ_0 particulière de $\mathcal{S}(\mathbb{R})$ telle que $\int_{-\infty}^{\infty} \varphi_0(x) dx = 1$; par exemple $\varphi_0(x) = (1/\sqrt{\pi}) e^{-x^2}$. Étant donnée n'importe quelle fonction φ de $\mathcal{S}(\mathbb{R})$, posons $A = \int_{-\infty}^{\infty} \varphi(x) dx$. Si φ est la dérivée d'une autre fonction ϕ de $\mathcal{S}(\mathbb{R})$, alors $A = 0$ d'après ce qui précède. Si tel n'est pas le cas, alors la fonction $\varphi_1 = \varphi - A\varphi_0$ vérifiera la condition $\int_{-\infty}^{\infty} \varphi_1(x) dx = 0$ et sera, par conséquent, la dérivée

d'une fonction $\phi_1 \in \mathcal{S}(\mathbb{R})$, de sorte qu'on aura $\varphi = \phi_1' + A\varphi_0$. Puisque T est une fonctionnelle *linéaire*, on pourra écrire que :

$$\langle T, \varphi \rangle = \langle T, \phi_1' \rangle + A\langle T, \varphi_0 \rangle = -\langle T', \phi_1 \rangle + A\langle T, \varphi_0 \rangle \quad (\text{IX.75})$$

Le terme $\langle T', \phi_1 \rangle$ est nul puisque $T' = 0$, il reste donc :

$$\langle T, \varphi \rangle = A\langle T, \varphi_0 \rangle = \langle T, \varphi_0 \rangle \int_{-\infty}^{+\infty} \varphi(x) dx \quad (\text{IX.76})$$

ce qui montre que T est l'intégrale avec le poids constant $C = \langle T, \varphi_0 \rangle$ (c'est un nombre indépendant de φ qui ne dépend donc que de la distribution T donnée). •

IX.8 Transformation de Fourier des distributions

Voyons maintenant l'autre opération classique sur les distributions, qui est la transformation de Fourier. C'est toujours l'intégrale avec poids qui sert de base pour l'analogie. Si p est le poids, supposé ici appartenir à $\mathbb{L}_1(\mathbb{R})$ et $\varphi \in \mathcal{S}(\mathbb{R})$ on a

$$\int_{-\infty}^{+\infty} \widehat{p}(x) \varphi(x) dx = \int_{-\infty}^{+\infty} p(x) \widehat{\varphi}(x) dx \quad (\text{IX.77})$$

ce qui conduit à définir la transformée de Fourier d'une distribution comme :

$$\langle \widehat{T}, \varphi \rangle = \langle T, \widehat{\varphi} \rangle \quad (\text{IX.78})$$

Pour que cette définition soit cohérente, il faut vérifier que $\varphi \mapsto \langle T, \widehat{\varphi} \rangle$ est une fonctionnelle linéaire continue. Ceci résulte du théorème IX.1 : $\varphi \mapsto \widehat{\varphi}$ est linéaire et continue relativement à la limite dans l'espace $\mathcal{S}(\mathbb{R})$. Si donc la suite φ_n tend vers zéro dans $\mathcal{S}(\mathbb{R})$, il en sera de même de $\widehat{\varphi}_n$ et donc aussi de $\langle T, \widehat{\varphi} \rangle$, puisque T est supposée être une fonctionnelle linéaire continue.

La définition de la transformée de Fourier d'une distribution a évidemment été choisie telle que, lorsque T est une intégrale avec poids p , c'est-à-dire lorsque T s'identifie à une fonction p , on retrouve la transformée de Fourier usuelle.

IX.8.1 Transformée de Fourier de 1

La constante 1 est une fonction au sens usuel, mais non intégrable de sorte que l'intégrale $\int e^{ix\xi} dx$ diverge ⁽⁸⁾ ; par conséquent la transformation de Fourier ne peut être considérée au sens usuel. Cependant, la fonction $p(x) = 1$ définit bien un poids. La fonctionnelle :

$$\langle 1, \varphi \rangle = \int_{-\infty}^{+\infty} \varphi(x) dx \quad (\text{IX.79})$$

est continue puisque d'après l'inégalité de la moyenne :

$$|\langle 1, \varphi \rangle| \leq \mathcal{N}_{0,2}(\varphi) \int_{-\infty}^{+\infty} \frac{1}{1+x^2} dx \quad (\text{IX.80})$$

D'après la définition (IX.78) :

$$\langle \widehat{1}, \varphi \rangle = \int_{-\infty}^{+\infty} \widehat{\varphi}(x) dx \quad (\text{IX.81})$$

(8) Cette intégrale divergente a déjà été étudiée au chapitre VIII : c'est l'intégrale I_α pour $\alpha = 0$.

Or, d'après la formule d'inversion, on a :

$$\int_{-\infty}^{+\infty} e^{-ixt} \widehat{\varphi}(x) dx = 2\pi \varphi(t) \quad (\text{IX.82})$$

Il suffit donc de prendre $t = 0$ pour avoir :

$$\langle \widehat{1}, \varphi \rangle = 2\pi \varphi(0) \quad (\text{IX.83})$$

ce qui prouve que :

$$\widehat{1} = 2\pi \delta \quad (\text{IX.84})$$

IX.8.2 Transformée de Fourier d'un polynôme

N'importe quel polynôme $P(x) = \sum_j a_j x^j$ est un poids possible, puisqu'on peut déduire de l'inégalité de la moyenne l'inégalité suivante :

$$|\langle P, \varphi \rangle| \leq \mathcal{N}_{0,k+2}(\varphi) \int_{-\infty}^{+\infty} \frac{|P(x)|}{1+|x|^{k+2}} dx \quad (\text{IX.85})$$

où k est le degré du polynôme. D'après la définition (IX.78) :

$$\langle \widehat{P}, \varphi \rangle = \int_{-\infty}^{+\infty} P(x) \widehat{\varphi}(x) dx \quad (\text{IX.86})$$

On sait, d'après le théorème VII.1, que $x^j \widehat{\varphi}(x)$ est la transformée de Fourier de $i^j \frac{d^j \varphi}{dx^j}$. D'après la formule d'inversion, on a alors :

$$\int_{-\infty}^{+\infty} x^j \widehat{\varphi}(x) dx = 2\pi i^j \frac{d^j \varphi}{dx^j}(0) \quad (\text{IX.87})$$

et donc pour le polynôme $P(x) = \sum_j a_j x^j$ on aura :

$$\int_{-\infty}^{+\infty} P(x) \widehat{\varphi}(x) dx = 2\pi \sum_{j=0}^k a_j i^j \frac{d^j \varphi}{dx^j}(0) \quad (\text{IX.88})$$

Cela montre que :

$$\widehat{P} = 2\pi \sum_{j=0}^k a_j i^j (-1)^j \frac{d^j \delta}{dx^j} = 2\pi P\left(-i \frac{d}{dx}\right) \delta \quad (\text{IX.89})$$

L'expression $P(-i \frac{d}{dx})$ est ce qu'on appelle un *opérateur différentiel* : on remplace formellement dans le polynôme $P(X)$ la variable X par $-i \frac{d}{dx}$.

On voit dans ces deux exemples que les opérations sur les distributions (ou, ici, sur les fonctions non intégrables) consistent à toujours se ramener aux opérations correspondantes sur les fonctions φ , pour lesquelles elles sont légitimes ; l'espace $\mathcal{S}(\mathbb{R})$ a été spécialement fabriqué pour cela.

IX.8.3 Transformée de Fourier de δ et de ses dérivées

Soit $\delta^{(j)}$ la j^{e} dérivée de δ ; par définition on a $\langle \delta^{(j)}, \varphi \rangle = (-1)^j \varphi^{(j)}(0)$; par conséquent la transformée de Fourier sera définie par :

$$\langle \widehat{\delta^{(j)}}, \varphi \rangle = (-1)^j [\widehat{\varphi}]^{(j)}(0) \quad (\text{IX.90})$$

Or, la j^e dérivée de $\widehat{\varphi}$ est la transformée de Fourier de $(ix)^j \varphi(x)$, de sorte que :

$$\langle \widehat{\delta^{(j)}}, \varphi \rangle = \int_{-\infty}^{+\infty} (-ix)^j \varphi(x) dx \quad (\text{IX.91})$$

qui est une intégrale avec poids $p(x) = (-ix)^j$. Autrement dit, la transformée de Fourier de $\delta^{(j)}$ est, en tant que fonctionnelle, l'intégrale avec poids $(-ix)^j$, donc, en identifiant :

$$\widehat{\delta^{(j)}} = (-ix)^j \quad (\text{IX.92})$$

IX.8.4 Transformée de Fourier de $1/[ix^\alpha]_2$

Pour $\alpha < 1$, cette transformée de Fourier est par définition la distribution T telle que $\forall \varphi \in \mathcal{S}(\mathbb{R})$:

$$\langle T, \varphi \rangle = \int_{-\infty}^{+\infty} \frac{\widehat{\varphi}(\xi)}{[i\xi^\alpha]_2} d\xi \quad (\text{IX.93})$$

On reconnaît que T est la distribution I_α introduite (pour $\alpha < 1$) à la section IX.4 et sur laquelle on reviendra. Pour $\alpha \geq 1$, l'intégrale ci-dessus diverge en $\xi = 0$; au chapitre VIII, nous avons traité ce cas en contournant la singularité dans le plan complexe, mais ci-dessus on ne peut faire cela car les fonctions φ ne sont pas forcément analytiques ; la valeur de l'intégrale ne serait plus indépendante du chemin choisi. En calculant l'intégrale selon cet artifice, on trouvait (pour $\alpha > 1$) $I_\alpha(x) = 2\pi x^{\alpha-1}/\Gamma(\alpha)$, ce qui est un poids acceptable, même pour $\alpha > 0$. On va donc poser :

$$\langle T, \varphi \rangle = \begin{cases} \int_{-\infty}^{+\infty} \frac{2\pi x^{\alpha-1}}{\Gamma(\alpha)} \varphi(x) dx & \text{si } \alpha > 0 \\ \int_{-\infty}^{+\infty} \frac{\widehat{\varphi}(\xi)}{[i\xi^\alpha]_2} d\xi & \text{si } \alpha < 1 \end{cases} \quad (\text{IX.94})$$

On vérifiera que les deux expressions sont compatibles quand $0 < \alpha < 1$ en utilisant le théorème VII.3 avec un facteur régularisant.

IX.8.5 Transformée de Fourier de e^{ix^2}

C'est la distribution T définie par :

$$\langle T, \varphi \rangle = \int_{-\infty}^{+\infty} e^{ix^2} \widehat{\varphi}(x) dx \quad (\text{IX.95})$$

Cette intégrale est la limite, lorsque z tend dans $\mathbb{C}$ vers $-i$ en restant de partie réelle positive, de $\int_{-\infty}^{+\infty} e^{-zx^2} \widehat{\varphi}(x) dx$. Or, tant que $\Re(z) > 0$, la fonction e^{-zx^2} , contrairement à e^{ix^2} , est intégrable et sa transformée de Fourier au sens usuel des fonctions est $\sqrt{\pi/z}^1 e^{-x^2/4z}$. Ici, $\sqrt{\pi/z}^1$ désigne la détermination $N^\circ 1$ de la racine carrée, analytique sur $\Omega_1 = \mathbb{C} -]-\infty; 0]$. Par conséquent, pour $\Re(z) > 0$:

$$\int_{-\infty}^{+\infty} e^{-zx^2} \widehat{\varphi}(x) dx = \sqrt{\frac{\pi}{z}}^1 \int_{-\infty}^{+\infty} e^{-x^2/4z} \varphi(x) dx \quad (\text{IX.96})$$

Dans l'intégrale de droite, on peut aussi passer à la limite sous le signe intégral puisque les conditions pour cela sont satisfaites, ce qui donne :

$$\langle T, \varphi \rangle = \sqrt{\pi} e^{i\pi/4} \int_{-\infty}^{+\infty} e^{-ix^2/4} \varphi(x) dx \quad (\text{IX.97})$$

qui est une intégrale avec le poids :

$$p(x) = \sqrt{\pi} e^{i\pi/4} e^{-ix^2/4} \quad (\text{IX.98})$$

La distribution T est donc une distribution régulière.

Ainsi, la transformée de Fourier au sens des distributions de la fonction non intégrable e^{ix^2} est la fonction (IX.98), également non intégrable.

IX.9 Limites de distributions

Nous avons très fréquemment utilisé le passage à la limite pour obtenir des formules intégrales, comme dans le chapitre V qui concerne les fonctions eulériennes. Ces passages à la limite sous le signe intégral exigent que soient satisfaites des conditions bien précises, qu'il faut vérifier rigoureusement. Lorsqu'on interprète les distributions comme des extensions de la notion d'intégrale avec poids, le passage à la limite sur les φ (au sens défini par les semi-normes $\mathcal{N}_{j,k}$) est automatique et fait partie de la définition même des distributions.

Néanmoins, que se passe-t-il lors du passage à la limite sur les poids pour un φ fixé ? Soit par exemple p_n , une suite de poids. D'après les théorèmes généraux classiques, si p_n converge uniformément sur tout intervalle borné vers une limite p , alors on aura :

$$\mathbb{R}\text{-}\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} p_n(x) \varphi(x) dx = \int_{-\infty}^{+\infty} p(x) \varphi(x) dx \quad (\text{IX.99})$$

et cela pour toute $\varphi \in \mathcal{S}(\mathbb{R})$. Cependant, nous avons vu aussi en section IX.4 que les intégrales avec poids :

$$\int_{-\infty}^{+\infty} I_{\alpha,\varepsilon}(x) \varphi(x) dx \quad (\text{IX.100})$$

avaient pour toute $\varphi \in \mathcal{S}(\mathbb{R})$ une limite quand $\varepsilon \rightarrow 0$ (pour $\alpha < 1$), alors que les poids $I_{\alpha,\varepsilon}(x)$ eux-mêmes n'en avaient pas — voir le chapitre VIII. Nous avons défini la distribution I_α en disant que pour toute $\varphi \in \mathcal{S}(\mathbb{R})$, $\langle I_\alpha, \varphi \rangle$ était cette limite. Il serait donc naturel de dire que la distribution I_α est la limite, pour $\varepsilon \rightarrow 0$, des distributions $I_{\alpha,\varepsilon}$.

Définition IX.2 On dit qu'une suite de distributions T_n a pour limite la distribution T si pour toute $\varphi \in \mathcal{S}(\mathbb{R})$ on a :

$$\lim_{n \rightarrow \infty} \langle T_n, \varphi \rangle = \langle T, \varphi \rangle \quad (\text{IX.101})$$

De même, on dira que la famille de distributions T_ε , paramétrée par ε , tend vers la distribution T si pour toute $\varphi \in \mathcal{S}(\mathbb{R})$, on a :

$$\lim_{\varepsilon \rightarrow 0} \langle T_\varepsilon, \varphi \rangle = \langle T, \varphi \rangle \quad (\text{IX.102})$$

Au contraire de la limite définie sur l'espace $\mathcal{S}(\mathbb{R})$, la limite ainsi définie est très peu contraignante : selon cette définition, les fonctions $I_{\alpha,\varepsilon}$ (pour $\alpha < 1$) ont une limite pour $\varepsilon \rightarrow 0$! La limite selon les semi-normes $\mathcal{N}_{j,k}$ était la plus contraignante de toutes celles que nous connaissons ; à l'inverse, la limite selon la définition IX.2 est la moins contraignante de toutes. C'est pourquoi on l'appelle la limite *faible*.

Ⓡ Cette limite faible suffit pour tous les cas pratiques qu'on peut rencontrer mais cet avantage n'est pas une évidence. Dans la définition IX.2, on a pris soin de préciser que la limite T est une distribution ; c'est-à-dire que les T_n et aussi T sont supposées *a priori* être des fonctionnelles linéaires continues. Lorsque nous voudrions montrer que $T = \lim T_n$ nous devrions montrer que $\forall \varphi \in \mathcal{S}(\mathbb{R})$, $\lim \langle T_n, \varphi \rangle = \langle T, \varphi \rangle$, mais aussi, indépendamment, que $\varphi \mapsto \langle T, \varphi \rangle$ est une fonctionnelle continue. C'est ce que nous avons fait pour I_α en section IX.4.

Toutefois on peut démontrer⁽⁹⁾ que si une suite de distributions T_n est telle que pour toute $\varphi \in \mathcal{S}(\mathbb{R})$, la suite numérique $\langle T_n, \varphi \rangle$ est convergente (dans $\mathbb{R}$), alors la fonctionnelle $\varphi \mapsto \lim \langle T_n, \varphi \rangle$ est forcément continue. Cela résulte d'un théorème obtenu par le mathématicien polonais S. BANACH (1892–1945)⁽¹⁰⁾, appelé théorème d'équi-continuité, qui s'applique ici parce que l'espace $\mathcal{S}(\mathbb{R})$ vérifie les conditions requises. S'attarder sur ces questions théoriques est évidemment sans intérêt pour une formation d'ingénieur et cette subtilité n'est mentionnée ici que pour attirer l'attention sur le caractère non évident de la chose.

Dans les cas concrets, il y aura toujours un moyen direct et simple, en général par l'inégalité de la moyenne, de montrer la continuité de la fonctionnelle limite, de sorte qu'on peut entièrement se passer du théorème de Banach. Ce dernier fournit malgré tout une garantie que les passages à la limite fonctionnent toujours sans problèmes ; s'il n'en était pas ainsi, certains passages à la limite pourraient poser des problèmes compliqués et les distributions n'auraient certainement pas le maniement simple qui explique leur succès.

Pour illustrer cela, revenons à nos exemples qui ont été choisis parce qu'ils sont les plus courants dans les applications. En section IX.3, on a vu que les intégrales avec poids suivantes avaient une limite pour tout $\varphi \in \mathcal{S}(\mathbb{R})$ (mais la limite ne passe pas sous le signe intégral) :

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma^3} x e^{-x^2/2\sigma^2} \varphi(x) dx; \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma^3} x e^{-x^2/2\sigma^2} dx; \int_{-\infty}^{+\infty} I_{\alpha,\varepsilon} \varphi(x) dx \quad (\text{IX.103})$$

Appelons ρ_σ , le poids de la première et ρ'_σ — dérivée du précédent, le poids de la deuxième, identifiés à la distribution correspondante. En utilisant le nouveau langage introduit, on peut dire que δ (distribution de Dirac) est la limite faible de ρ_σ quand $\sigma \rightarrow 0$; que δ' (dérivée de la distribution de Dirac) est la limite faible de ρ'_σ et que la distribution I_α (la transformée de Fourier de $1/[ix^\alpha]_2$, pour $\alpha < 1$) est la limite faible de $I_{\alpha,\varepsilon}$ quand $\varepsilon \rightarrow 0$.

Les exemples de distributions de charges de la section IX.2 s'interprètent naturellement : les deux distributions qui représentaient le dipôle, la densité :

$$\rho'_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma^3} x e^{-x^2/2\sigma^2} \quad (\text{IX.104})$$

et la distribution discrète :

$$\frac{1}{2\sigma} [\delta(x - \sigma) - \delta(x + \sigma)] \quad (\text{IX.105})$$

sont toutes deux proches (lorsque σ est petit) de la distribution dipolaire δ' , au sens de la limite faible. C'est ce qui est écrit ci-dessus pour ρ'_σ . Il est facile de le vérifier pour la distribution discrète. Précisons d'abord les notations : $\delta_{\{a\}}$ désignera la charge unité placée au point $x = a$, de sorte que :

$$\begin{aligned} \langle \delta_{\{a\}}, \varphi \rangle &= \int_{-\infty}^{+\infty} \delta_{\{a\}}(x) \varphi(x) dx = \int_{-\infty}^{+\infty} \delta(x - a) \varphi(x) dx = \varphi(a) \\ \left\langle \frac{1}{2\sigma} [\delta_{\{\sigma\}} - \delta_{\{-\sigma\}}], \varphi \right\rangle &= \frac{1}{2\sigma} [\varphi(\sigma) - \varphi(-\sigma)] \end{aligned} \quad (\text{IX.106})$$

(8) L. SCHWARTZ, *Théorie des distributions*, Hermann, Paris, 1966, p. 74.

(9) S. BANACH, *Théorie des opérations linéaires*, 1932.

Cela tend bien vers $\varphi'(0)$ lorsque $\sigma \rightarrow 0$, c'est-à-dire que la distribution $\frac{1}{2\sigma} [\delta_{\{+\sigma\}} - \delta_{\{-\sigma\}}]$ tend faiblement vers $-\delta'$.

De la même façon, en reprenant la sous-section IX.8.5, on peut dire que la distribution régulière e^{ix^2} est la limite faible des distributions e^{-zx^2} lorsque z tend dans $\mathbb{C}$ vers $-i$ en restant dans le demi-plan $\Re(z) > 0$ ⁽⁹⁾.

Si une suite T_n (ou une famille T_ε) de distributions tend faiblement vers la limite T , alors la suite de leurs dérivées tendra vers la dérivée de T et la suite de leurs transformées de Fourier tendra vers la transformée de Fourier de T . En effet, puisque $\varphi \in \mathcal{S}(\mathbb{R}) \implies \varphi' \in \mathcal{S}(\mathbb{R})$, on peut écrire la chaîne logique :

$$\begin{aligned} \forall \varphi \in \mathcal{S}(\mathbb{R}), \lim_{n \rightarrow \infty} \langle T_n, \varphi \rangle &= \langle T, \varphi \rangle \\ \Downarrow \\ \forall \varphi \in \mathcal{S}(\mathbb{R}), \lim_{n \rightarrow \infty} \langle T_n, \varphi' \rangle &= \langle T, \varphi' \rangle \\ \Updownarrow \\ \forall \varphi \in \mathcal{S}(\mathbb{R}), \lim_{n \rightarrow \infty} \langle T'_n, \varphi \rangle &= \langle T', \varphi \rangle \end{aligned} \quad (\text{IX.107})$$

et de même, puisque $\varphi \in \mathcal{S}(\mathbb{R}) \iff \widehat{\varphi} \in \mathcal{S}(\mathbb{R})$:

$$\begin{aligned} \forall \varphi \in \mathcal{S}(\mathbb{R}), \lim_{n \rightarrow \infty} \langle T_n, \varphi \rangle &= \langle T, \varphi \rangle \\ \Downarrow \\ \forall \varphi \in \mathcal{S}(\mathbb{R}), \lim_{n \rightarrow \infty} \langle T_n, \widehat{\varphi} \rangle &= \langle T, \widehat{\varphi} \rangle \\ \Updownarrow \\ \forall \varphi \in \mathcal{S}(\mathbb{R}), \lim_{n \rightarrow \infty} \langle \widehat{T}_n, \varphi \rangle &= \langle \widehat{T}, \varphi \rangle \end{aligned} \quad (\text{IX.108})$$

Ainsi, les théorèmes élémentaires de passage à la limite sous le signe intégral, détaillés dans le chapitre I, permettent de montrer que $\forall \varphi \in \mathcal{S}(\mathbb{R})$:

$$\lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{+\infty} \frac{e^{-\varepsilon|\xi|}}{[(i\xi)^\alpha]_2} \varphi(x) \, d\xi = \int_{-\infty}^{+\infty} \frac{1}{[(i\xi)^\alpha]_2} \varphi(x) \, d\xi \quad (\text{IX.109})$$

c'est-à-dire que la limite faible des poids $e^{-\varepsilon|\xi|}/[(i\xi)^\alpha]_2$ est le poids $1/[(i\xi)^\alpha]_2$.

On peut alors déduire de ce qui précède que pour $\alpha < 1$, la limite des distributions $I_{\alpha,\varepsilon}$ est la distribution I_α mais cela n'est plus possible par les voies élémentaires et nécessite la théorie qui a été développée dans ce chapitre.

C'est donc la théorie des distributions qui donne un sens rigoureux à l'intégrale divergente I_α du chapitre VIII. On peut écrire des formules telles que :

$$\int_{-\infty}^{+\infty} e^{ix\xi} \, dx = 2\pi \delta(\xi) \quad (\text{IX.110})$$

ou :

$$\int_{-\infty}^{+\infty} e^{ix\xi} [(i\xi)^\alpha]_2 \, d\xi = I_\alpha(x) \quad (\text{IX.111})$$

Cependant il ne faudra pas oublier que $\delta(\xi)$ ou $I_\alpha(x)$, malgré la notation abusive, ne sont pas soumises aux règles de calcul sur les fonctions, mais aux règles de calcul sur les distributions.

⁽⁹⁾ Cette condition est nécessaire : si $\Re(z) < 0$, la densité e^{-zx^2} n'est plus à croissance polynomiale et ne peut pas être un poids, ni même une distribution.

- Ⓡ La famille des distributions I_α a des propriétés remarquables que nous étudierons dans les chapitres suivants. Pour $\alpha = -n$ entier positif, le poids $1/[(i\xi)^\alpha]_2$ devient $(i\xi)^n$ — pour α entier, la fonction cesse d'être multiforme. Or, on a vu que la transformée de Fourier de ce poids est la distribution $2\pi \delta^{(n)}$ (la n^e dérivée de $2\pi \delta$). Par conséquent :

$$I_{-n} = 2\pi \delta^{(n)} \quad (\text{IX.112})$$

D'autre part, pour toute $\varphi \in \mathcal{S}(\mathbb{R})$, la fonction :

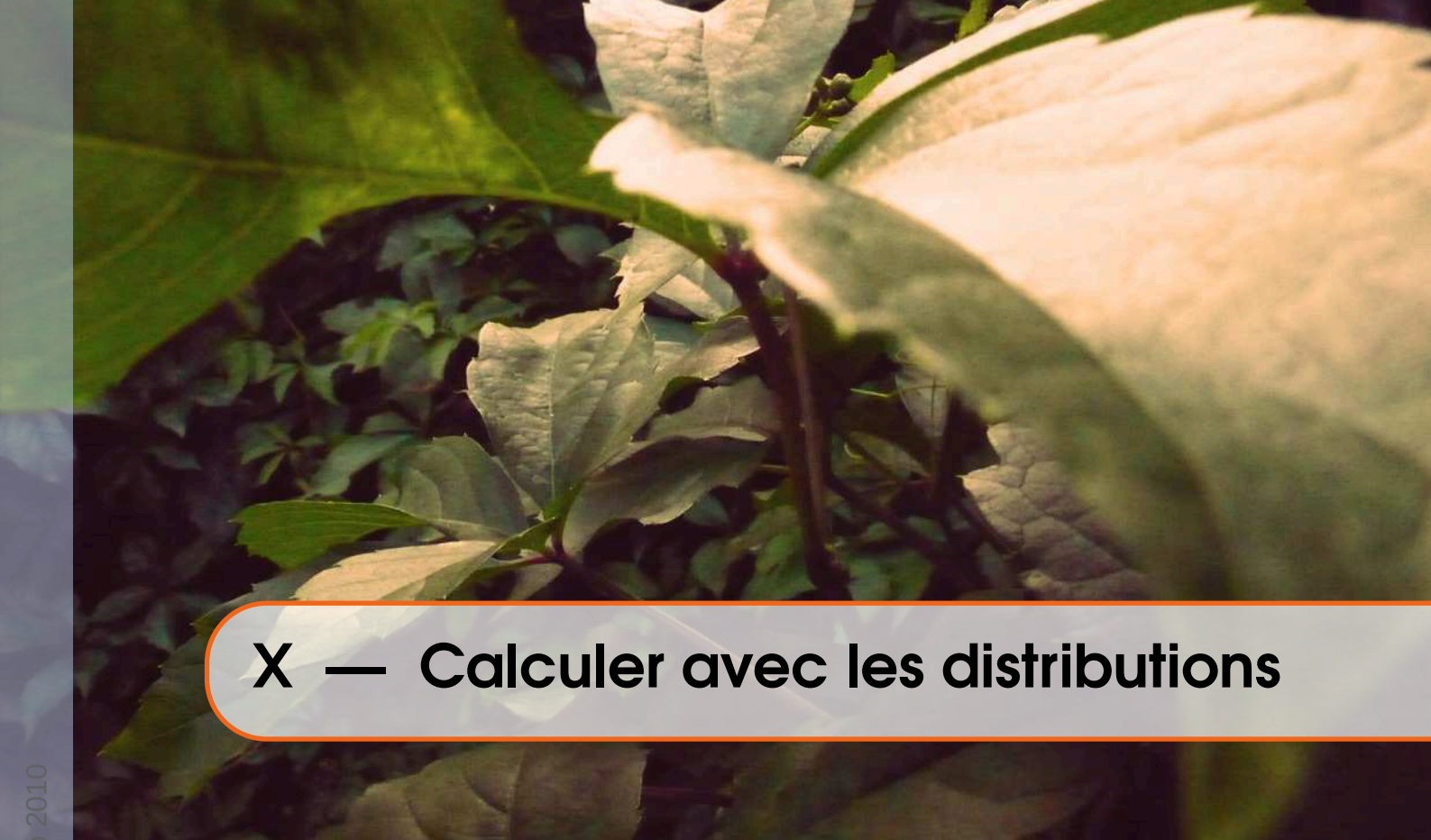
$$\alpha \mapsto \int_{-\infty}^{+\infty} \frac{1}{[(i\xi)^\alpha]_2} \varphi(x) \, d\xi \quad (\text{IX.113})$$

est une fonction analytique dans tout le plan complexe, sans point singulier ni coupure ; cela se montre en appliquant à l'intégrale ci-dessus un théorème élémentaire. Or, cette intégrale, si on y remplace φ par $\widehat{\varphi}$, n'est autre que $\langle I_\alpha, \varphi \rangle$. Ceci signifie que $\forall \varphi \in \mathcal{S}(\mathbb{R})$, la fonction :

$$\alpha \mapsto \langle I_\alpha, \varphi \rangle = \begin{cases} \int_{-\infty}^{+\infty} \frac{2\pi x^{\alpha-1}}{\Gamma(\alpha)} \varphi(x) \, dx & \text{si } \Re(\alpha) > 0 \\ \int_{-\infty}^{+\infty} \frac{\widehat{\varphi}(\xi)}{[i\xi^\alpha]_2} \, d\xi & \text{si } \Re(\alpha) < 1 \end{cases} \quad (\text{IX.114})$$

est analytique dans tout le plan complexe. Cela aura, comme nous le verrons, l'avantage que toute égalité vérifiée par la famille I_α sur un ensemble non discret de valeurs de α , par exemple l'intervalle $]0; 1[$, se prolongera analytiquement à toutes les valeurs de α .

La famille de distributions I_α intervient dans beaucoup de problèmes comme l'étude des équations différentielles et le traitement du signal. Elle permet de définir une dérivation d'ordre non entier : en effet, puisque I_{-n} est la n^e dérivée de δ , on peut dire que, même lorsque r n'est pas entier, I_{-r} est la r^e dérivée de δ . La *convolution* permet alors de *dériver* une fonction ou une distribution quelconque à l'ordre r . Indépendamment de cela, les distributions les plus courantes sont des membres de cette famille : δ et ses dérivées, ainsi que les pseudo-fonctions de J. Hadamard, en sont des cas particuliers.



X — Calculer avec les distributions

X.1 Dérivées de fonctions non dérivables

Toutes les fonctions ne sont pas dérivables. Par contre, d'après la définition (IX.64), toutes les distributions ont une dérivée. Or, nous avons vu qu'une fonction quelconque à croissance polynomiale (pouvant donc être un poids dans une intégrale avec une fonction à décroissance rapide) s'identifiait à une distribution : on identifie le poids p qui est une fonction, à la fonctionnelle T_p qui est l'intégrale avec poids. Si le poids p est une fonction dérivable, quel rapport y a-t-il entre p' (la dérivée de p au sens usuel) et la dérivée de la distribution T_p ?

Pour le voir, il suffit d'entreprendre un calcul. Par la définition (IX.64) :

$$\langle T'_p, \varphi \rangle = -\langle T_p, \varphi' \rangle = -\int_{-\infty}^{+\infty} p(x) \varphi'(x) dx \quad (\text{X.1})$$

Cette intégrale peut être intégrée par parties, ce qui donne :

$$-p(x) \varphi(x) \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} p'(x) \varphi(x) dx \quad (\text{X.2})$$

La partie intégrée est nulle car la fonction $p(x) \varphi(x)$ est nulle à l'infini (p est à croissance polynomiale et φ à décroissance rapide). Il reste :

$$\langle T'_p, \varphi \rangle = \int_{-\infty}^{+\infty} p'(x) \varphi(x) dx = \langle T_{p'}, \varphi \rangle \quad (\text{X.3})$$

On obtient donc un résultat qui montre la compatibilité entre la dérivation au sens des distributions et la dérivation usuelle lorsque cette dernière est possible.

On a déjà vu des exemples de ce qui se passe pour des fonctions non dérivables : en (IX.68), on a obtenu la dérivée au sens des distributions de la fonction de Heaviside,

qui est δ . En (IX.72), la dérivée au sens des distributions de la fonction $|x|$, qui est le poids égal à la constante -1 pour $x < 0$ et à la constante $+1$ pour $x > 0$. Au point $x = 0$, ce poids n'a pas besoin d'être défini, car la fonctionnelle correspondante (l'intégrale) n'est pas affectée par ce qui se passe en un seul point ; en général, on ne change pas la valeur d'une intégrale si on change la valeur que prend la fonction en un nombre fini (ou même dénombrable) de points. Un poids n'a donc pas à être défini partout mais seulement presque partout ⁽¹⁾. Entre ces deux exemples, il y a une différence qualitative : la fonction de Heaviside comme la fonction $|x|$ sont toutes deux des fonctions au sens usuel mais si on les dérive au sens des distributions, on obtient, dans le premier cas une distribution singulière et dans le second, une distribution régulière. Cela provient de la discontinuité : la première est discontinue, la seconde continue.

On peut généraliser ces exemples comme suit.

Théorème X.1 Soit $p(x)$, une fonction poids dérivable par morceaux, c'est-à-dire qu'il existe un nombre fini de points $a_0, a_1, a_2, \dots, a_n$ tels que p soit dérivable sur chacun des intervalles $]-\infty; a_0[,]a_0; a_1[,]a_1; a_2[,]a_2; a_3[, \dots,]a_{n-2}; a_{n-1}[,]a_{n-1}; a_n[,]a_n; +\infty[$, mais pas aux points $a_0, a_1, \dots, a_n$. On suppose aussi que la fonction a en chacun de ces points une limite à gauche et une limite à droite finies, mais non nécessairement égales, de sorte qu'il y a un saut de discontinuité s_i au point a_i . Alors la dérivée au sens des distributions de cette fonction $p(x)$ est :

$$T'_p = T_{p'} + \sum_i s_i \delta_{\{a_i\}} \quad (\text{X.4})$$

où p' est la dérivée de p au sens usuel (définie en dehors des points a_i) et $\delta_{\{a_i\}}$ le pic de Dirac au point a_i : $\langle \delta_{\{a_i\}}, \varphi \rangle = \varphi(a_i)$

Preuve Du fait de la définition (IX.64), on a :

$$\langle T'_p, \varphi \rangle = - \int_{-\infty}^{+\infty} p(x) \varphi'(x) dx = - \int_{-\infty}^{a_0} - \sum_{i=0}^{n-1} \int_{a_i}^{a_{i+1}} - \int_{a_n}^{+\infty} \quad (\text{X.5})$$

Intégrons par parties chacune des intégrales de la somme ci-dessus :

$$\begin{aligned} - \int_{-\infty}^{a_0} p(x) \varphi'(x) dx &= -p(x) \varphi(x) \Big|_{-\infty}^{a_0} + \int_{-\infty}^{a_0} p'(x) \varphi(x) dx \\ - \int_{a_i}^{a_{i+1}} p(x) \varphi'(x) dx &= -p(x) \varphi(x) \Big|_{a_i}^{a_{i+1}} + \int_{a_i}^{a_{i+1}} p'(x) \varphi(x) dx \\ - \int_{a_n}^{+\infty} p(x) \varphi'(x) dx &= -p(x) \varphi(x) \Big|_{a_n}^{+\infty} + \int_{a_n}^{+\infty} p'(x) \varphi(x) dx \end{aligned} \quad (\text{X.6})$$

En sommant tout, on obtient :

- pour la somme des intégrales, $\int_{-\infty}^{+\infty} p'(x) \varphi(x) dx$;
- pour la somme des parties intégrées : $-p^-(a_0) \varphi(a_0) + p^+(a_0) \varphi(a_0) - p^-(a_1) \varphi(a_1) + p^+(a_1) \varphi(a_1) - p^-(a_2) \varphi(a_2) + p^+(a_2) \varphi(a_2) - \dots - p^-(a_n) \varphi(a_n) + p^+(a_n) \varphi(a_n)$, les termes $p(\pm\infty) \varphi(\pm\infty)$ étant nuls. Conformément aux hypothèses, $p^+(a_i)$ n'est pas forcément égal à $p^-(a_i)$, mais φ est continue. En regroupant les termes deux par deux, on obtient bien la somme des $[p^+(a_i) - p^-(a_i)] \varphi(a_i) = s_i \varphi(a_i)$. •

(1) La théorie de l'intégrale de Lebesgue donne un sens précis à cette expression. Voir chapitre II, p. 17.

Il existe des fonctions qui sont si irrégulières qu'on ne peut découper leur domaine en intervalles sur chacun desquels elle est dérivable. Par exemple, il existe des fonctions qui ne sont dérivables en aucun point. D'ailleurs, il n'est même pas correct de dire qu'il en existe, car les fonctions qui ne sont dérivables en aucun point sont infiniment plus nombreuses que les fonctions dérivables par morceaux et on les rencontre bien plus souvent dans la nature. On est habitué en mathématique à utiliser des fonctions construites avec les opérations arithmétiques, ce qui donne généralement des fonctions analytiques. On finit alors par croire que ces fonctions sont la règle, alors qu'elles sont l'exception. La trajectoire d'un grain soumis au mouvement brownien ou un bruit blanc donnent une idée des fonctions qui ne sont nullement dérivables. C'est pourquoi il ne serait pas sans intérêt de calculer les dérivées au sens des distributions de telles fonctions. Certaines applications reposent même sur de tels calculs (géométrie fractale, analyse du bruit). Cela nécessiterait un chapitre à part entière.

Bien entendu, on obtiendrait, pour les dérivées de telles fonctions, des distributions *vraiment singulières*, alors que les théorèmes X.1 et X.2 prédisent que si on se limite aux fonctions *normales*, on ne rencontrera rien d'autre en les dérivant que des combinaisons des trois sortes de distributions suivantes :

1. les distributions régulières (ou fonctions usuelles) ;
2. les pics de Dirac ;
3. les pseudo-fonctions.

La formule (X.4) donne la dérivée d'une fonction n'ayant que des discontinuités dites de première espèce, c'est-à-dire lorsque les limites à gauche et à droite de la discontinuité existent et sont finies. On va maintenant étudier les distributions qu'on obtient en dérivant une fonction ayant une discontinuité de seconde espèce mais intégrable, car il doit s'agir d'un poids. Pour simplifier l'énoncé, on considérera le cas d'une seule discontinuité, contrairement à ce qui a été fait au théorème X.1 ; le passage au cas plus général où il y aurait n discontinuités de seconde espèce, ou un mélange de discontinuités de première et de seconde espèce, est alors une complication purement technique.

Théorème X.2 Soit $p(x)$ une fonction dérivable sur $\mathbb{R} - \{0\}$, et ayant en $x = 0$ une discontinuité de seconde espèce, mais intégrable ; c'est-à-dire que l'une au moins des deux limites à droite ou à gauche est infinie mais de sorte que l'intégrale $\int |p(x)| dx$ converge en $x = 0$. La distribution T'_p est alors définie par :

$$\langle T'_p, \varphi \rangle = \lim_{\varepsilon \rightarrow 0} \left\{ [p(+\varepsilon) - p(-\varepsilon)] \varphi(0) + \int_{|x| > \varepsilon} p'(x) \varphi(x) dx \right\} \quad (\text{X.7})$$

Ⓡ La limite dans (X.7) existe toujours ; pourtant la fonction $p'(x)$ n'est pas forcément intégrable en $x = 0$; par exemple $1/\sqrt{x}$ est intégrable en $x = 0$, mais sa dérivée $-1/x^{3/2}$ ne l'est pas. Si p' n'est pas intégrable l'intégrale dans (X.7) tendra vers l'infini, et le terme $p(+\varepsilon) - p(-\varepsilon)$ aussi, les deux infinis se compensant. Le mathématicien J. Hadamard appelait la limite (X.7) la *partie finie* de l'intégrale $\int p'(x) \varphi(x) dx$, qui, elle diverge.

Preuve La démonstration du théorème est une simple intégration par parties. •

L'exemple type pour illustrer le théorème X.2 est la fonction :

$$I_\alpha(x) = \begin{cases} \frac{x^{\alpha-1}}{\Gamma(\alpha)} & \text{si } x > 0 \\ 0 & \text{si } x < 0 \end{cases} \quad 0 < \alpha < 1 \quad (\text{X.8})$$

Cette fonction est en effet intégrable en $x = 0$ et y a une discontinuité de seconde espèce.

Ⓡ Pour $\alpha = 1$, il s'avère que la discontinuité est de première espèce alors que pour $\alpha > 1$, la fonction est continue.

Laissant de côté le facteur de normalisation $1/\Gamma(\alpha)$, nous devons dériver *au sens des distributions* le poids :

$$p(x) = \begin{cases} 0 & \text{si } x < 0 \\ x^{\alpha-1} & \text{si } x > 0 \end{cases} \quad 0 < \alpha < 1 \quad (\text{X.9})$$

Par définition de la dérivée d'une distribution, la dérivée de p est la distribution :

$$\langle T, \varphi \rangle = - \int_0^\infty p(x) \varphi'(x) dx - \lim_{\varepsilon \rightarrow 0} \int_\varepsilon^\infty p(x) \varphi'(x) dx \quad (\text{X.10})$$

En intégrant par parties, on obtient :

$$\langle T, \varphi \rangle = \lim_{\varepsilon \rightarrow 0} \left\{ p(\varepsilon) \varphi(\varepsilon) + \int_\varepsilon^\infty p'(x) \varphi(x) dx \right\} \quad (\text{X.11})$$

Il est facile de voir que le terme $p(\varepsilon) \varphi(\varepsilon)$ est équivalent, lorsque $\varepsilon \rightarrow 0$, à :

$$p(\varepsilon) \varphi(\varepsilon) \sim p(\varepsilon) \varphi(0) \sim \frac{p(\varepsilon)}{\varepsilon} \int_0^\varepsilon \varphi(x) dx \quad (\text{X.12})$$

ce qui montre que si on introduit le poids :

$$q_\varepsilon(x) = \begin{cases} 0 & \text{si } x < 0 \\ p(\varepsilon)/\varepsilon & \text{si } 0 < x < \varepsilon \\ p'(x) & \text{si } x > \varepsilon \end{cases} \quad (\text{X.13})$$

on peut écrire :

$$\langle T, \varphi \rangle = \lim_{\varepsilon \rightarrow 0} \left\{ \int_{-\infty}^{+\infty} q_\varepsilon(x) \varphi(x) dx \right\}, \quad \forall \varphi \in \mathcal{S}(\mathbb{R}) \quad (\text{X.14})$$

ce qui signifie que la distribution T que nous cherchons est la limite faible des fonctions q_ε lorsque ε tend vers zéro. (voir figure X.1).

On voit que, si $p(\varepsilon)$ tend vers l'infini lorsque ε tend vers zéro, $p(\varepsilon)/\varepsilon$ tendra encore plus vite vers l'infini ; par contre la dérivée $p'(x)$, forcément négative puisque $p(x)$, partant de $+\infty$, ne peut que décroître, tend vers $-\infty$ quand $x \rightarrow 0^+$. C'est bien ce qu'on peut voir sur la figure X.1. Il faut donc se représenter la distribution T comme la limite faible des fonctions du type q_ε . Ceci est évidemment à rapprocher des distributions I_α , qui sont obtenues comme limite des fonctions $I_{\alpha,\varepsilon}$: pour $p(x) = x^{\alpha-1}$ ($0 < \alpha < 1$), on a $T = \Gamma(\alpha) \cdot I_\alpha$.

Un autre exemple illustrant le théorème X.2 est la distribution de poids $\ln(|x|)$ (figure X.2). La fonction $\ln(|x|)$ a bien une singularité intégrable en $x = 0$; elle diffère de

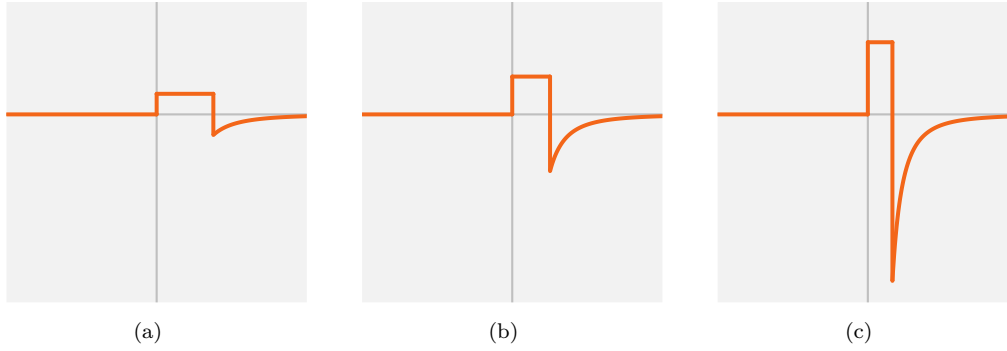


Figure X.1 – Suite de fonctions qui tendent faiblement vers la distribution $I_{-1/2}$. Le calcul a été effectué pour un poids $p(x)$ défini en (X.9). Le calcul est le même pour n'importe quel poids du même type

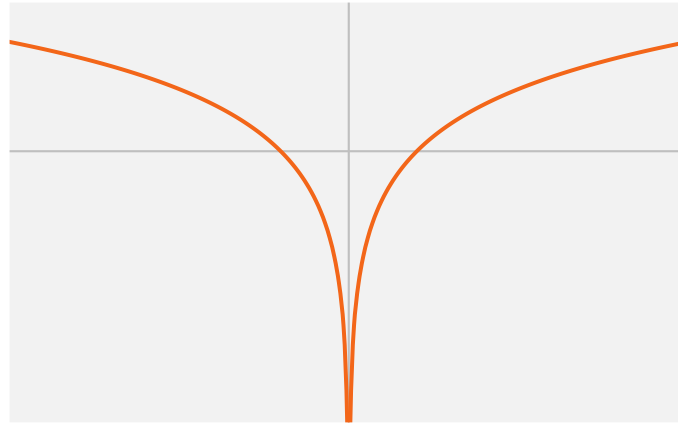


Figure X.2 – Fonction poids $p(x) = \ln(|x|)$

l'exemple I_α surtout par le fait qu'elle tend vers l'infini des deux côtés, alors que I_α ne tendait vers l'infini que du côté $x > 0$. Si on applique le théorème X.2, on constate que dans (X.7), le terme $[p(+\varepsilon) - p(-\varepsilon)]\varphi(0)$ est nul, puisque $p(x) = \ln(|x|)$ est une fonction paire. La dérivée de cette fonction au sens des distributions est une distribution classique qu'on rencontre dans des applications, et qu'il faut connaître ; on l'appelle *valeur principale* de $\frac{1}{x}$ et on note $\text{VP}(\frac{1}{x})$. D'après le théorème X.2 :

$$\left\langle \text{VP}\left(\frac{1}{x}\right), \varphi \right\rangle = \lim_{\varepsilon \rightarrow 0} \left\{ \int_{-\infty}^{-\varepsilon} \frac{1}{x} \varphi(x) dx + \int_{+\varepsilon}^{+\infty} \frac{1}{x} \varphi(x) dx \right\} \quad (\text{X.15})$$

Dans la relation (X.7), le terme $[p(+\varepsilon) - p(-\varepsilon)]\varphi(0)$ doit compenser l'infini de l'intégrale ; ici, sa disparition est liée au fait que la fonction $\frac{1}{x}$, qui tend vers $-\infty$ pour $x < 0$ et vers $+\infty$ pour $x > 0$, compense déjà deux infinis opposés : les deux intégrales de (X.15) divergent toutes les deux, mais avec des signes opposés. La fonction $\frac{1}{x}$ n'est donc pas un poids, car la singularité n'est pas intégrable. C'est pour souligner cette particularité qu'on note $\text{VP}(\frac{1}{x})$ et non $\frac{1}{x}$. Cette notation sert à rappeler que l'intégrale :

$$\int_{-\infty}^{+\infty} \frac{1}{x} \varphi(x) dx \quad (\text{X.16})$$

n'est pas une intégrale au sens usuel, et qu'on ne peut pas lui appliquer par exemple l'inégalité de la moyenne ou les théorèmes généraux de passage à la limite sous le signe intégral : il

faut remplacer ces derniers par de nouveaux théorèmes de passage à la limite, prévus pour les distributions, et qui sont essentiellement les suivants :

- on peut passer à la limite sous le signe $\langle T, \varphi_n \rangle$ selon φ_n , si φ_n tend vers une limite dans $\mathcal{S}(\mathbb{R})$, donc au sens défini par les semi-normes $\mathcal{N}_{j,k}$;
- on peut passer à la limite sous le signe $\langle T_n, \varphi \rangle$ selon T_n , si T_n tend *faiblement* vers une distribution T .

Il est donc absurde, pour passer à la limite sur φ dans l'intégrale (X.16), d'invoquer la convergence uniforme comme s'il s'agissait d'une intégrale au sens usuel.

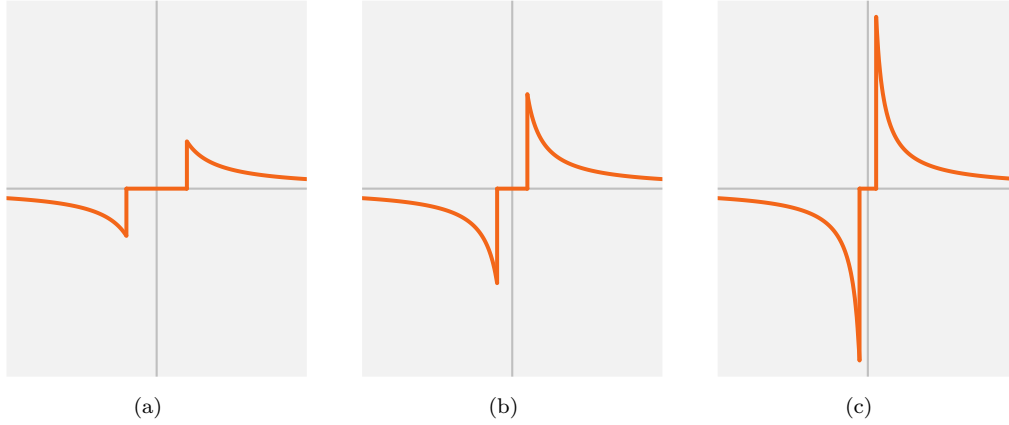


Figure X.3 – Suite de fonctions qui tendent faiblement vers la distribution $\text{VP}(\frac{1}{x})$

Comme toujours, on peut interpréter (X.15) en disant que la distribution $\text{VP}(\frac{1}{x})$ est la limite faible, quand ε tend vers zéro, des poids

$$p_\varepsilon(x) = \begin{cases} \frac{1}{x} & \text{si } |x| \geq \varepsilon \\ 0 & \text{si } |x| < \varepsilon \end{cases} \quad (\text{X.17})$$

Les graphiques de ces poids sont donnés dans la figure X.3.

Lorsqu'on définit une distribution, il faut toujours vérifier avec soin qu'il s'agit bien d'une fonctionnelle linéaire *continue*. En principe, cela demande de jongler un peu avec les semi-normes, comme nous avons vu au chapitre IX. En jonglant justement avec ces semi-normes, on se rendra aisément compte que :

- une intégrale avec poids est une distribution si le poids est localement intégrable (*i.e.* si toutes ses singularités sont intégrables) et s'il croît polynomialement à l'infini. En effet, l'intégrale :

$$\int_{-\infty}^{+\infty} p(x) \varphi(x) dx \quad (\text{X.18})$$

ne peut être convergente pour toute $\varphi \in \mathcal{S}(\mathbb{R})$ que si la fonction p est dépourvue de singularités qui la rendraient divergente, et si la décroissance rapide de $\varphi(x)$ compense la croissance de $p(x)$. Pour que la fonctionnelle linéaire soit continue, il faudra aussi pouvoir appliquer l'inégalité de la moyenne comme ceci :

$$\left| \int_{-\infty}^{+\infty} p(x) \varphi(x) dx \right| \leq \int_{-\infty}^{+\infty} \frac{|p(x)|}{1 + |x|^m} dx \times \mathcal{N}_{0,m}(\varphi) \quad (\text{X.19})$$

et il faudra donc pouvoir choisir un m tel que le premier facteur soit fini.

- la dérivée d'une distribution est toujours une distribution ⁽²⁾ ;
- la transformée de Fourier d'une distribution est toujours une distribution ⁽³⁾.

Par conséquent toute fonctionnelle qui se déduit d'une intégrale avec poids par dérivation ou transformation de Fourier sera forcément une distribution. De même, toute fonctionnelle qui se déduit par dérivation ou transformation de Fourier d'une distribution déjà connue comme telle, sera forcément aussi une distribution. La vérification directe par des inégalités de semi-normes ne sera donc nécessaire que dans les cas où on ne peut pas se ramener aux cas précédents ; cela pourra malheureusement arriver parfois, pour des distributions définies comme produit ou comme convolution car il n'existe pas, pour ces opérations, de critère aussi simple que pour la dérivation ou la transformation de Fourier.

Ainsi il n'est pas nécessaire de vérifier directement que $\text{VP}(\frac{1}{x})$ est bien une fonctionnelle continue, puisqu'elle est la dérivée du poids $\ln(|x|)$.

Calculons encore la dérivée de la distribution $\text{VP}(\frac{1}{x})$. Par définition, ce sera la distribution T telle que $\forall \varphi \in \mathcal{S}(\mathbb{R})$:

$$\langle T, \varphi \rangle = -\langle \text{VP}(\frac{1}{x}), \varphi' \rangle = \lim_{\varepsilon \rightarrow 0} \left\{ -\int_{-\infty}^{-\varepsilon} \frac{1}{x} \varphi'(x) dx - \int_{+\varepsilon}^{+\infty} \frac{1}{x} \varphi'(x) dx \right\} \quad (\text{X.20})$$

En intégrant par parties les deux intégrales ci-dessus, on obtient :

$$\langle T, \varphi \rangle = \lim_{\varepsilon \rightarrow 0} \left\{ \frac{1}{\varepsilon} [\varphi(-\varepsilon) + \varphi(+\varepsilon)] - \int_{-\infty}^{-\varepsilon} \frac{1}{x^2} \varphi(x) dx - \int_{+\varepsilon}^{+\infty} \frac{1}{x^2} \varphi(x) dx \right\} \quad (\text{X.21})$$

Lorsque ε tend vers zéro, le terme $[\varphi(-\varepsilon) + \varphi(+\varepsilon)]/\varepsilon$ a la même limite que $2\varphi(0)/\varepsilon$, ou encore que :

$$\int_{-\varepsilon}^{+\varepsilon} \frac{1}{\varepsilon^2} \varphi(x) dx \quad (\text{X.22})$$

En fin de compte on voit que :

$$\langle T, \varphi \rangle = \lim_{\varepsilon \rightarrow 0} \int_{-\infty}^{+\infty} f_{\varepsilon}(x) \varphi(x) dx \quad (\text{X.23})$$

où f_{ε} est la fonction définie par :

$$f_{\varepsilon}(x) = \begin{cases} -\frac{1}{x^2} & \text{si } x < -\varepsilon \\ \frac{1}{\varepsilon^2} & \text{si } 0 < |x| < \varepsilon \\ -\frac{1}{x^2} & \text{si } x > \varepsilon \end{cases} \quad (\text{X.24})$$

Ainsi, la distribution que nous cherchons est la limite faible des fonctions f_{ε} (figure X.4).

Dans la théorie des distributions, on montre que toute distribution est limite faible de fonctions ⁽⁴⁾. Le meilleur moyen de représenter graphiquement une distribution est donc de dessiner le graphe d'une fonction proche (au sens de la limite faible) de cette distribution.

⁽²⁾ En effet, si T est une distribution et que $\mathcal{S}\text{-lim } \varphi_n = 0$, alors $\mathcal{S}\text{-lim } \varphi'_n = 0$, donc $\langle T', \varphi \rangle = -\langle T, \varphi' \rangle$ tend vers zéro.

⁽³⁾ On fait appel au même argument : si T est une distribution et que $\mathcal{S}\text{-lim } \varphi_n = 0$, alors $\mathcal{S}\text{-lim } \widehat{\varphi}_n = 0$, donc $\langle \widehat{T}, \varphi \rangle = \langle T, \widehat{\varphi} \rangle$ tend vers zéro.

⁽⁴⁾ L. SCHWARTZ, *Théorie des distributions*, p. 75 et 166 ; aussi le théorème X.7 dans ce document.

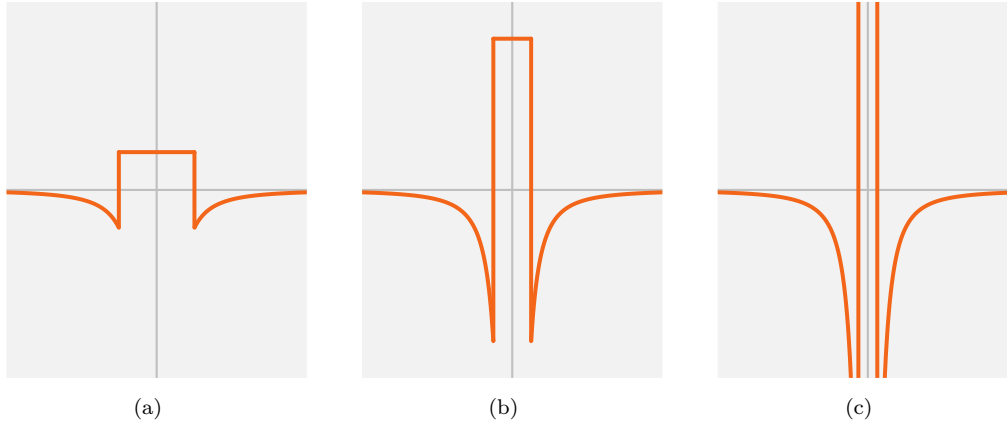


Figure X.4 – Suite de fonctions qui tendent faiblement vers la distribution $\frac{d}{dx} \text{VP}(\frac{1}{x})$

X.2 Multiplication et convolution des distributions

Les distributions étant une extension des fonctions, les opérations usuelles sur les fonctions doivent s'étendre aux distributions. On a déjà vu cela pour la dérivation et la transformation de Fourier. Toutefois, la dérivation et la transformation de Fourier sont possibles pour n'importe quelle distribution car l'espace $\mathcal{S}(\mathbb{R})$ a été construit spécialement pour cela. Les deux nouvelles opérations que nous allons examiner maintenant ne sont pas toujours possibles.

La multiplication de deux fonctions f et g est simplement la fonction fg dont la valeur en x est le produit des deux nombres $f(x)$ et $g(x)$:

$$\forall x \in \mathbb{R}, fg(x) = f(x) \cdot g(x) \quad (\text{X.25})$$

Cette opération a toujours un sens puisqu'elle se ramène au produit de deux nombres (aussi bien lorsque f et g prennent leurs valeurs dans $\mathbb{R}$ ou dans $\mathbb{C}$). Mais les distributions singulières ne sont pas sensées avoir une valeur numérique pour tout x , donc on ne peut pas utiliser (X.25). La théorie des distributions étant essentiellement construite par analogie à partir des intégrales avec poids, il faut regarder à quoi correspond le produit de deux poids. Or l'intégrale de poids $p(x)q(x)$ peut s'écrire de trois façons :

$$\int_{-\infty}^{+\infty} [p(x)q(x)]\varphi(x) dx = \int_{-\infty}^{+\infty} p(x)[q(x)\varphi(x)] dx = \int_{-\infty}^{+\infty} q(x)[p(x)\varphi(x)] dx \quad (\text{X.26})$$

Ces trois écritures sont équivalentes à cause de l'associativité du produit. En écriture fonctionnelle, ces trois formes de l'intégrale deviennent :

$$\langle T_{pq}, \varphi \rangle = \langle T_p, q\varphi \rangle = \langle T_q, p\varphi \rangle \quad (\text{X.27})$$

Pour que dans l'une ou l'autre des deux dernières variantes, T_p ou T_q puisse être considérée comme une fonctionnelle opérant sur $\mathcal{S}(\mathbb{R})$, il faut que (respectivement) $q\varphi$ ou $p\varphi$ soit dans $\mathcal{S}(\mathbb{R})$.

Or, si p est infiniment dérivable et à croissance polynomiale ainsi que toutes ses dérivées, alors $\varphi \in \mathcal{S}(\mathbb{R}) \Rightarrow p\varphi \in \mathcal{S}(\mathbb{R})$. De même pour q . Par analogie on est donc conduit à poser :

Définition X.1 Si T est une distribution et p une fonction infiniment dérivable et à croissance polynomiale ainsi que toutes ses dérivées, le produit pT est la fonctionnelle :

$$\varphi \mapsto \langle pT, \varphi \rangle = \langle T, p\varphi \rangle \quad (\text{X.28})$$

Il faut comme toujours vérifier la cohérence de cette définition en s'assurant que si T est continue, alors pT est continue — la linéarité ne pose évidemment aucun problème. Or, si p est à croissance polynomiale ainsi que toutes ses dérivées, on aura :

$$\frac{d^j[p\varphi]}{dx^j} = \sum_{\ell=0}^{\ell=j} \binom{j}{\ell} \frac{d^{j-\ell}p}{dx^{j-\ell}} \frac{d^\ell\varphi}{dx^\ell} \quad (\text{X.29})$$

et chacun des facteurs $\frac{d^{j-\ell}p}{dx^{j-\ell}}$ sera majoré par une expression du type $M_{j,\ell}(1 + |x|^{n_{j\ell}})$; de sorte que :

$$\mathcal{N}_{j,k}(p\varphi) \leq \sum_{\ell=0}^{\ell=j} \binom{j}{\ell} M_{j,\ell} \mathcal{N}_{\ell,n_{j\ell}} \quad (\text{X.30})$$

ce qui prouve bien la continuité.

La définition X.1 du produit pT est donnée sous une condition assez restrictive : p est supposé être une fonction infiniment dérivable et à croissance polynomiale ainsi que toutes ses dérivées. Nous venons de voir que cette condition est suffisante pour garantir, en vertu de l'argument simple ci-dessus, que si T est une fonctionnelle continue sur $\mathcal{S}(\mathbb{R})$, il en sera automatiquement de même de pT . Toutefois cette condition suffisante n'est de loin pas toujours nécessaire ; certaines distributions particulières peuvent être multipliées par des fonctions plus générales, par exemple des fonctions non infiniment dérivables.

Exemple X.1 La distribution δ peut être multipliée par n'importe quelle fonction continue en $x = 0$: si $f(x)$ est une fonction aussi irrégulière qu'on voudra, qui peut même être discontinue en tout point autre que $x = 0$, le produit $f \cdot \delta$ a un sens, c'est la fonctionnelle $\langle f\delta, \varphi \rangle = f(0)\varphi(0)$. Par contre si f est discontinue en $x = 0$, on ne peut donner un sens cohérent au produit $f \cdot \delta$, même si f est aussi régulière qu'on voudra en dehors de $x = 0$. Par exemple, si $H(x)$ est la fonction de Heaviside, on ne peut pas faire des calculs sensés avec le produit $H \cdot \delta$; $\square$

Exemple X.2 Si T est l'intégrale de poids p_0 , on peut la multiplier par n'importe quelle fonction p localement intégrable et à croissance polynomiale, même partout discontinue, car dans ce cas le produit $p \cdot p_0$ est le produit usuel des fonctions, et si p et p_0 sont tous deux à croissance polynomiale, il en sera de même de leur produit. $\square$

Une des faiblesses classiques de la théorie des distributions est l'inexistence d'une multiplication qui serait possible sous des conditions à la fois simples et générales. Beaucoup de mathématiciens ont tenté de définir un produit de deux distributions arbitraires mais aucun n'y est parvenu sans imposer des bases théoriques d'une complexité exorbitante. Il faut donc retenir que le produit est une affaire qui doit se régler dans chaque cas particulier, lorsque c'est possible. La définition (X.28) peut servir de modèle mais il ne faudra jamais

hésiter à sortir de son cadre (nous le ferons à l'occasion). Dans les applications, il arrive en effet assez souvent que la condition invoquée pour (X.28) ne soit pas vérifiée, mais que le cas particulier considéré permette néanmoins la multiplication.

Une autre opération très courante et soumise aux mêmes aléas est la convolution. On appelle produit de convolution de deux fonctions f et g la nouvelle fonction $h = f * g$ définie par :

$$h(x) = \int_{-\infty}^{+\infty} f(x-y) g(y) dy = \int_{-\infty}^{+\infty} f(y) g(x-y) dy \quad (\text{X.31})$$

les deux intégrales ci-dessus sont convergentes si f et g sont toutes deux dans l'espace $\mathbb{L}_1(\mathbb{R})$; leur égalité signifie que $f * g = g * f$, autrement dit que la convolution est une opération commutative entre éléments de l'espace $\mathbb{L}_1(\mathbb{R})$. La fonction $h = f * g$ est alors elle aussi dans l'espace $\mathbb{L}_1(\mathbb{R})$. En effet, d'après l'inégalité de la moyenne appliquée à (X.31), on aura pour tout $x \in \mathbb{R}$:

$$|h(x)| \leq \int_{-\infty}^{+\infty} |f(x-y)| |g(y)| dy \quad (\text{X.32})$$

et donc en intégrant cela par rapport à la variable x :

$$\int_{-\infty}^{+\infty} |h(x)| dx \leq \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |f(x-y)| |g(y)| dy dx \quad (\text{X.33a})$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |f(x')| |g(y)| dy dx' \quad (\text{X.33b})$$

$$= \int_{-\infty}^{+\infty} |f(x')| dx' \cdot \int_{-\infty}^{+\infty} |g(y)| dy \quad (\text{X.33c})$$

Le passage de (X.33a) à (X.33b) résulte du changement de variable $(x, y) \mapsto (x-y, y)$ et le passage de (X.33b) à (X.33c) de la factorisation de l'intégrale. En notation plus condensée :

$$\|f * g\|_1 \leq \|f\|_1 \cdot \|g\|_1 \quad (\text{X.34})$$

La convolution a beaucoup d'applications en traitement du signal (filtrage). Une propriété essentielle, relative à la transformation de Fourier, est la suivante :

Théorème X.3 Si f et g sont deux fonctions de $\mathbb{L}_1(\mathbb{R})$, alors :

$$\widehat{f * g} = \widehat{f} \cdot \widehat{g} \quad (\text{X.35})$$

Autrement dit, la transformée de Fourier de la convolution est le produit des transformées de Fourier.

L'égalité (X.35) est vraie pour la transformation de Fourier $f \mapsto \widehat{f}$. Si on utilise d'autres versions de la transformation de Fourier (voir (VII.116), (VII.117), et (VII.119)) il faudra modifier (X.35) par un coefficient multiplicatif. Pour la transformation (VII.116) :

$$\mathcal{F}_1 f(\xi) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(x) e^{ix\xi} dx$$

on aura :

$$\mathcal{F}_1(f * g) = \sqrt{2\pi} \mathcal{F}_1(f) \cdot \mathcal{F}_1(g) \quad (\text{X.36a})$$

Pour (VII.117) :

$$\Phi_1 f(\xi) = \int_{-\infty}^{+\infty} f(x) e^{2\pi i x \xi} dx$$

on aura :

$$\Phi_1(f * g) = \Phi_1(f) \cdot \Phi_1(g) \quad (\text{X.36b})$$

et enfin pour (VII.119), qui est la transformation de Fourier de la mécanique quantique :

$$\mathcal{F}_\hbar f(\xi) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} f(x) e^{-\frac{i}{\hbar} x \xi} dx$$

on aura :

$$\mathcal{F}_\hbar(f * g) = \sqrt{2\pi\hbar} \mathcal{F}_\hbar(f) \cdot \mathcal{F}_\hbar(g) \quad (\text{X.36c})$$

Si on utilise une autre variante, il faudra faire les modifications convenables ; par exemple pour la transformation $f \mapsto \tilde{f}$ (inverse de $f \mapsto \hat{f}$), on aura

$$\widetilde{f * g} = 2\pi \tilde{f} \cdot \tilde{g} \quad (\text{X.36d})$$

Preuve (théorème X.3) Puisque $f * g \in \mathbb{L}_1(\mathbb{R})$, la transformée de Fourier de $f * g$ est tout simplement l'intégrale :

$$\begin{aligned} \int_{-\infty}^{+\infty} e^{ix\xi} f * g(x) dx &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{ix\xi} f(x-y) g(y) dx dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{i(x'+y)\xi} f(x') g(y) dx' dy \\ &= \int_{-\infty}^{+\infty} e^{ix'\xi} f(x') dx' \cdot \int_{-\infty}^{+\infty} e^{iy\xi} g(y) dy \end{aligned} \quad (\text{X.37})$$

d'où l'énoncé. •

Il s'agit maintenant d'étendre la convolution aux distributions. On procède toujours par analogie à partir des intégrales avec poids. Si p et q sont deux fonctions de $\mathbb{L}_1(\mathbb{R})$, et $\varphi \in \mathcal{S}(\mathbb{R})$, l'intégrale de poids $p * q$ est :

$$\int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} p(x-y) q(y) dy \right] \varphi(x) dx = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(x) q(y) \varphi(x+y) dx dy \quad (\text{X.38a})$$

$$= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} p(x-y) \varphi(x) dx \right] q(y) dy \quad (\text{X.38b})$$

$$= \int_{-\infty}^{+\infty} p(x) \left[\int_{-\infty}^{+\infty} q(y-x) \varphi(y) dy \right] dx \quad (\text{X.38c})$$

On a fait dans l'intégrale double (X.38a), les changements de variable $(x, y) \mapsto (x-y, y)$ pour obtenir (X.38b), et $(x, y) \mapsto (x, y-x)$ pour obtenir (X.38c). Dans les notations fonctionnelles, cela se résume comme suit :

$$\langle p * q, \varphi \rangle = \langle q, \check{p} * \varphi \rangle = \langle p, \check{q} * \varphi \rangle \quad (\text{X.39})$$

où on a introduit la notation $\check{p}, \check{q}$ qui représente la fonction $\check{p}(x) = p(-x)$. D'autre part, puisque $\widehat{p * q} = \widehat{p} \cdot \widehat{q}$, on peut aussi écrire :

$$\langle p * q, \varphi \rangle = \langle \widehat{p} \cdot \widehat{q}, \widetilde{\varphi} \rangle = \langle \widehat{p}, \widehat{q} \cdot \widetilde{\varphi} \rangle = \langle \widehat{q}, \widehat{p} \cdot \widetilde{\varphi} \rangle \quad (\text{X.40})$$

en supposant qu'on puisse donner un sens élargi à $\langle \widehat{q}, \widehat{p} \cdot \widetilde{\varphi} \rangle$, ce qui ramène la convolution au produit. On va donc étendre la convolution aux distributions par la définition suivante :

Définition X.2 Étant données deux distributions S et T , la convolution $S * T$ est indifféremment une des fonctionnelles suivantes :

- si la transformée de Fourier au sens des distributions de S est une fonction $p(x)$ infiniment dérivable à croissance polynomiale :

$$\varphi \longmapsto \langle \widehat{T}, p \cdot \widetilde{\varphi} \rangle \quad (\text{X.41})$$

- si S est une fonction $p(x)$ localement intégrable à croissance polynomiale (c'est-à-dire si en tant que fonctionnelle S est l'intégrale de poids p) :

$$\varphi \longmapsto \langle T, \check{p} * \varphi \rangle \quad (\text{X.42})$$

Pour rappel, une fonction localement intégrable est, pour nous, une fonction continue par morceaux, qui peut devenir infinie en certains points singuliers, mais qui est alors intégrable en ces points. Pour que cette définition soit cohérente, il faut que les conditions suivantes soient satisfaites :

1. l'expression (X.41) ou (X.42) doit avoir une valeur finie pour toute $\varphi \in \mathcal{S}(\mathbb{R})$;
2. cette même expression doit tendre vers zéro lorsque φ tend vers zéro dans $\mathcal{S}(\mathbb{R})$.

Ces conditions ne sont pas forcément satisfaites : cela dépendra des distributions S et T ; comme pour le produit, on ne peut pas donner une définition qui satisfasse automatiquement ces conditions, tout en couvrant l'ensemble des cas intéressants. On peut certes donner de la convolution de deux distributions une définition plus générale que celle ci-dessus ; cela exigerait encore un supplément de théorie. La définition ci-dessus couvre à peu près les cas qu'on rencontre en pratique, mais il faudra vérifier cas par cas si les conditions 1 et 2 ci-dessus sont bien satisfaites.

Il est d'ailleurs logique que les difficultés soient les mêmes pour les produits et pour les convolutions, puisque la transformation de Fourier transforme l'un en l'autre. Il reste donc la charge de prouver la légitimité de l'opération dans chaque cas particulier.

Dans (X.42), on a défini la fonctionnelle $\langle p * T, \varphi \rangle$ comme étant égale à $\langle T, \check{p} * \varphi \rangle$. Cela peut paraître à première vue contraire à la nature des distributions, qui veut qu'en tant que fonctionnelles, elles opèrent sur les fonctions infiniment dérivables ; mais en réalité, la fonction $\check{p} * \varphi$ est bien infiniment dérivable, car :

Théorème X.4 Dans la convolution de deux fonctions, si l'une, p , est localement intégrable et à croissance polynomiale mais non nécessairement dérivable, et l'autre, φ , dans $\mathcal{S}(\mathbb{R})$,

alors leur convolution sera aussi infiniment dérivable (mais pas forcément à décroissance rapide) et on aura :

$$(p * \varphi)' = p * \varphi' \quad (\text{X.43})$$

Preuve Pour s'assurer que $p * \varphi$ est bien infiniment dérivable, il suffit de remarquer que les théorèmes généraux de dérivation sous le signe intégral s'appliquent à l'intégrale :

$$\int_{-\infty}^{+\infty} p(y)\varphi(x-y) dy \quad (\text{X.44})$$

L'hypothèse que p est à croissance polynomiale et φ à décroissance rapide garantit les conditions pour que ces théorèmes généraux s'appliquent. Si on dérive sous le signe intégral, seules les dérivées de φ interviennent, puisque $p(y)$ ne dépend pas de x .

Pour voir que $p * \varphi$ n'est pas forcément à décroissance rapide, on examine le contre-exemple $H * \varphi$, où H est la fonction de Heaviside :

$$H * \varphi(x) = \int_0^{\infty} \varphi(x-y) dy = \int_{-\infty}^x \varphi(z) dz \quad (\text{X.45})$$

Il est clair que si $\int_{-\infty}^{\infty} \varphi(z) dz$ n'est pas nulle, $H * \varphi(x)$ ne tend pas vers zéro quand $x \rightarrow +\infty$. •

Pour calculer avec la convolution des distributions, il faut connaître la propriété importante suivante :

Théorème X.5 Étant données trois distributions R, S, T , si les convolutions $S * T, R * S, R * (S * T), (R * S) * T$ sont définies, on a nécessairement $R * (S * T) = (R * S) * T$ (associativité). On note donc $R * S * T$. On a aussi pour les dérivées $(R * S)' = R' * S = R * S'$, $(R * S * T)' = R' * S * T = R * S' * T = R * S * T'$, etc.

Preuve Vérifications élémentaires mais fastidieuses à partir de (X.41) et (X.42). •

X.3 Exemples et applications des produits et convolutions

X.3.1 Convolution par les distributions de Dirac

La transformée de Fourier de δ est la constante 1. En effet, d'après la définition (IX.78), on doit avoir :

$$\langle \widehat{\delta}, \varphi \rangle = \langle \delta, \widehat{\varphi} \rangle = \widehat{\varphi}(0) = \int_{-\infty}^{+\infty} \varphi(x) dx \quad (\text{X.46})$$

On reconnaît bien l'intégrale de poids 1.

La distribution $\varphi \mapsto \langle \delta, \varphi \rangle = \varphi(0)$ est le *pic* de Dirac placé en $x = 0$; le *pic* de Dirac placé en un point quelconque $x = a$ est, en tant que fonctionnelle, $\langle \delta_a, \varphi \rangle = \varphi(a)$. La transformée de Fourier est alors :

$$\langle \widehat{\delta}_a, \varphi \rangle = \langle \delta_a, \widehat{\varphi} \rangle = \widehat{\varphi}(a) = \int_{-\infty}^{+\infty} e^{iax} \varphi(x) dx \quad (\text{X.47})$$

C'est l'intégrale de poids e^{iax} .

On constate que ces fonctions, 1 et e^{iax} , sont infiniment dérivables et à croissance polynomiale. On peut donc appliquer la définition (X.41), ce qui donne pour une distribution arbitraire T : $\langle \delta * T, \varphi \rangle = \langle \widehat{T}, 1 \cdot \widehat{\varphi} \rangle = \langle T, \varphi \rangle$. D'où le résultat :

$$\delta * T = T \quad (\text{X.48})$$

Ⓡ D'après les conditions de validité de la convolution discutées en section X.2, il faut vérifier dans chaque cas particulier que le résultat obtenu est bien une fonctionnelle continue ; mais ici, c'est évident.

Voyons le cas de δ_a . D'après la formule (X.41), T étant une distribution arbitraire alors $\langle \delta_a * T, \varphi \rangle = \langle \widehat{T}, e^{iax} \cdot \widetilde{\varphi} \rangle = \langle T, \psi_a \rangle$, où la fonction ψ_a est la transformée de Fourier de $e^{iax} \cdot \widetilde{\varphi}$, c'est-à-dire :

$$\psi_a(\xi) = \int_{-\infty}^{+\infty} e^{ix\xi} e^{ixa} \widetilde{\varphi}(x) dx = \varphi(\xi + a) \quad (\text{X.49})$$

Si T était un poids $p(x)$, alors $\langle T, \psi_a \rangle$ serait l'intégrale :

$$\int_{-\infty}^{+\infty} p(x) \psi_a(x) dx = \int_{-\infty}^{+\infty} p(x - a) \varphi(x) dx \quad (\text{X.50})$$

c'est-à-dire que la convolution par δ_a équivaut à une translation. C'est pourquoi la distribution $\delta_a * T$ est aussi appelée la translatée de T .

Preuve Pour vérifier que la fonctionnelle $\varphi \mapsto \langle T, \psi_a \rangle$ est continue, on observera d'abord que $\forall j \, d^j \psi_a(x) = d^j \varphi(x + a)$ (D désigne la dérivation), et par conséquent $\mathcal{N}_{j,k}(\psi_a) = \sup_x \{(1 + |x|^k) d^j \varphi(x + a)\} = \sup_x \{(1 + |x - a|^k) d^j \varphi(x)\}$. D'autre part :

$$|x - a|^k = \left| \sum_{\ell=0}^k \binom{k}{\ell} (-a)^{k-\ell} x^\ell \right| \leq \sum_{\ell=0}^k \binom{k}{\ell} |a|^{k-\ell} |x|^\ell \quad (\text{X.51})$$

et :

$$\begin{aligned} \sum_{\ell=0}^k \binom{k}{\ell} |a|^{k-\ell} (1 + |x|^\ell) &= \sum_{\ell=0}^k \binom{k}{\ell} |a|^{k-\ell} + \sum_{\ell=0}^k \binom{k}{\ell} |a|^{k-\ell} |x|^\ell \\ &= (|a| + 1)^k + \sum_{\ell=0}^k \binom{k}{\ell} |a|^{k-\ell} |x|^\ell \geq 1 + \sum_{\ell=0}^k \binom{k}{\ell} |a|^{k-\ell} |x|^\ell \end{aligned} \quad (\text{X.52})$$

d'où :

$$1 + |x - a|^k \leq \sum_{\ell=0}^k \binom{k}{\ell} |a|^{k-\ell} (1 + |x|^\ell) \quad (\text{X.53})$$

Par conséquent :

$$\mathcal{N}_{j,k}(\psi_a) \leq \sum_{\ell=0}^k \binom{k}{\ell} |a|^{k-\ell} \mathcal{N}_{j,\ell}(\varphi) \quad (\text{X.54})$$

Cette dernière inégalité montre que si φ tend vers zéro dans $\mathcal{S}(\mathbb{R})$, il en sera de même pour ψ_a . •

On a déjà vu que la convolution est commutative et associative ; la distribution δ est donc un élément neutre pour cette opération. Cet élément neutre n'est pas dans l'espace $\mathbb{L}_1(\mathbb{R})$, sur lequel la convolution était cependant partout définie.

X.3.2 Convolution par les dérivées de δ

La transformée de Fourier de $\delta^{(j)}$ — j^{e} dérivée de δ — est la fonction $(-i\xi)^j$. Celle-ci est infiniment dérivable et à croissance polynomiale ; on se réfère donc à (X.41) :

$$\langle \delta^{(j)} * T, \varphi \rangle = \langle \widehat{T}, (-i\xi)^j \cdot \widetilde{\varphi} \rangle = \left\langle \widehat{T}, (-1)^j \frac{d^j \varphi}{dx^j} \right\rangle = \left\langle T, (-1)^j \frac{d^j \varphi}{dx^j} \right\rangle \quad (\text{X.55})$$

Autrement dit, la convolution par $\delta^{(j)}$ équivaut à la dérivation :

$$\delta^{(j)} * T = \frac{d^j T}{dx^j} \quad (\text{X.56})$$

On ne trouvera rien de vraiment nouveau dans la convolution par les $\delta_a^{(j)}$; ce serait simplement la composition de la dérivation et de la translation.

La vérification que le résultat de la convolution est bien une fonctionnelle continue est dans ce cas évidente : on sait déjà que les dérivées d'une distribution sont des distributions.

X.3.3 Régularisation

Théorème X.6 La convolution d'une distribution T par une fonction infiniment dérivable ϕ , lorsqu'elle est possible, est une fonction infiniment dérivable. C'est-à-dire qu'elle s'identifie, en tant que fonctionnelle, à une intégrale avec un poids infiniment dérivable.

Preuve Par la définition (X.42), on a $\langle \check{\phi} * T, \varphi \rangle = \langle T, \phi * \varphi \rangle$. Prenons :

$$\varphi(x) = \rho_{n,a}(x) = \sqrt{\frac{n}{\pi}} e^{-n(x-a)^2} \quad (\text{X.57})$$

Lorsque n tend vers l'infini, $\rho_{n,a}$ tend vers δ_a selon la limite faible, donc $\langle \check{\phi} * T, \rho_{n,a} \rangle = \langle T, \phi * \rho_{n,a} \rangle$ va tendre vers $\langle T, \phi_a \rangle$, où ϕ_a est la fonction traduite $\phi_a(x) = \phi(x+a)$. Il est facile de voir que la fonction $F : a \mapsto \langle T, \phi_a \rangle$ est dérivable : puisque T est une fonctionnelle *linéaire*, on a :

$$\frac{F(a+h) - F(a)}{h} = \langle T, \frac{\phi_{a+h} - \phi_a}{h} \rangle \quad (\text{X.58})$$

Or $\mathcal{S}\text{-}\lim_{h \rightarrow 0} [\phi_{a+h} - \phi_a]/h = \phi'_a$, où $\phi'_a(x) = \phi'(x+a)$. Comme T est aussi une fonctionnelle *continue*, on peut passer à la limite quand h tend vers zéro, ce qui montre que $(F(a+h) - F(a))/h$ a bien une limite, qui est $\langle T, \phi'_a \rangle$. Puisque la fonction ϕ est infiniment dérivable, on peut recommencer indéfiniment la procédure, ce qui prouve que la fonction $F : a \mapsto \langle T, \phi_a \rangle$ est infiniment dérivable.

On remarquera que cette démonstration utilise d'une manière essentielle la continuité de la fonctionnelle : il faut pouvoir passer à la limite sous le signe $\langle , \rangle$ lorsque $h \rightarrow 0$.

Il reste à vérifier que $\langle \check{\phi} * T, \varphi \rangle$ est identique à l'intégrale de poids F . L'idée est la même que pour la dérivée : on joue sur la nature de l'intégrale, qui est une limite de sommes finies (les sommes de Riemann). On fait passer la somme finie sous le signe $\langle , \rangle$ grâce à la linéarité, puis on passe des sommes finies à leur limite (l'intégrale) grâce à la continuité. Le détail est long et fastidieux, car il faut vérifier que l'intégrale est la limite des sommes finies dans $\mathcal{S}(\mathbb{R})$ mais l'idée est simple ⁽⁵⁾. On voit ici encore à quel point la continuité est essentielle. •

Nous avons déjà vu, d'après les expressions (IX.4), (IX.6) et (IX.7), que les poids $\rho_n(x) = \sqrt{\frac{n}{\pi}} e^{-nx^2}$ formaient une suite qui converge faiblement vers la distribution δ . On s'attend donc à ce que pour n'importe quelle distribution T , la suite $\rho_n * T$ converge faiblement vers T . C'est bien le cas :

Théorème X.7 Si T est une distribution quelconque, les convolutions $\rho_n * T$ forment une suite de fonctions de $\mathcal{S}(\mathbb{R})$ qui converge faiblement vers T . Par conséquent, toute distribution est la limite (faible) d'une suite de fonctions infiniment dérivables.

On peut donc approcher les distributions par des fonctions très régulières. On appelle ce procédé la *régularisation* de la distribution. Nous avons vu par exemple que les distributions I_α étaient les limites des poids $I_{\alpha,\varepsilon}$. Le théorème X.7 nous dit que n'importe quelle

(5) L. SCHWARTZ, *Théorie des distributions*, p. 166

distribution est toujours une limite faible de *vraies* fonctions. C'est bien ce que nous avons constaté empiriquement sur les exemples déjà étudiés. Imaginer les distributions comme des fonctions extrêmement irrégulières est donc une représentation correcte. La régularisation est un procédé essentiel en traitement du signal (filtre passe bas). Elle généralise aussi la méthode du *facteur régularisant* introduite au chapitre VIII pour les intégrales divergentes. Ici nous avons choisi le filtre gaussien pour effectuer les calculs mais rien n'imposait ce choix.

Preuve Dire que $\chi_n = \rho_n * T$ converge faiblement vers T signifie, par définition, que :

$$\forall \varphi \in \mathcal{S}(\mathbb{R}), \langle \chi_n, \varphi \rangle \longrightarrow \langle T, \varphi \rangle \quad (\text{X.59})$$

Mais on a aussi, par définition de la convolution (voir la relation (X.42)) :

$$\langle \chi_n, \varphi \rangle = \langle \rho_n * T, \varphi \rangle = \langle T, \rho_n * \varphi \rangle \quad (\text{X.60})$$

on notera que $\check{\rho}_n = \rho_n$, puisque ρ_n est une fonction paire. Comme T est une distribution, elle est continue en tant que fonctionnelle, donc il suffit de vérifier que $\rho_n * \varphi$ converge dans $\mathcal{S}(\mathbb{R})$ vers φ . En résumé : il s'agit de montrer que $\forall \varphi \in \mathcal{S}(\mathbb{R})$, la suite numérique $n \mapsto \langle \rho_n * T, \varphi \rangle = \langle T, \rho_n * \varphi \rangle$ tend vers $\langle T, \varphi \rangle$. Le fait que $\langle \rho_n * T, \varphi \rangle = \langle T, \rho_n * \varphi \rangle$ a pour conséquence qu'il est exactement équivalent de dire que $\rho_n * T$ tend faiblement vers T ou que $\forall \varphi \in \mathcal{S}(\mathbb{R})$, la suite $\rho_n * \varphi$ tend dans $\mathcal{S}(\mathbb{R})$ vers φ .

La transformation de Fourier étant un opérateur continu et bijectif sur $\mathcal{S}(\mathbb{R})$, il est encore équivalent de montrer que $\forall \varphi \in \mathcal{S}(\mathbb{R})$, la suite $\widehat{\rho}_n \cdot \varphi$ tend dans $\mathcal{S}(\mathbb{R})$ vers φ . C'est sous cette dernière forme que la démonstration sera techniquement la plus aisée. Notons que $\widehat{\rho}_n(x) = e^{-x^2/4n}$. En posant $\varepsilon = 1/4n$, on doit donc montrer que les semi-normes $\mathcal{N}_{j,k}(\varphi - e^{-\varepsilon x^2} \varphi)$ tendent toutes vers zéro lorsque ε tend vers zéro. Il va donc falloir majorer les expressions de la forme $(1 + |x|^k) d^j [(1 - e^{-\varepsilon x^2}) \varphi(x)]$, où d^j désigne la dérivation d'ordre j .

Or, en utilisant la formule de Leibniz pour la dérivée d'un produit, on peut écrire :

$$d^j [(1 - e^{-\varepsilon x^2}) \varphi(x)] = (1 - e^{-\varepsilon x^2}) d^j \varphi(x) - \sum_{\ell=0}^{j-1} \binom{j}{\ell} d^{j-\ell} e^{-\varepsilon x^2} \cdot d^\ell \varphi(x) \quad (\text{X.61})$$

Il faut donc commencer par majorer les expressions $d^{j-\ell} e^{-\varepsilon x^2}$. Pour cela, on remarque d'abord qu'en posant $X = \sqrt{\varepsilon} x$, on a :

$$\frac{d}{dx} = \sqrt{\varepsilon} \frac{d}{dX} \quad (\text{X.62})$$

donc on est ramené à majorer les dérivées $\frac{d}{dX} e^{-X^2}$. Comme e^{-X^2} est une fonction analytique dans tout le plan complexe, on peut utiliser les inégalités de Cauchy du corollaire III.6.1 :

$$\frac{d^n}{dx^n} e^{-X^2} = \frac{n!}{r^n} M_r \quad (\text{X.63})$$

où M_r désigne le maximum de la fonction $z \mapsto |e^{-z^2}| = e^{-\Re(z^2)}$ sur le cercle $|z - X| = r$. Un calcul élémentaire montre que $M_r \leq e^{r^2 - X^2/2}$, de sorte que si on choisit $r = 1$:

$$\frac{d^n}{dx^n} e^{-X^2} \leq n! e \cdot e^{-X^2/2} \quad (\text{X.64})$$

Le choix $r = 1$ est loin de donner la majoration la plus serrée, mais c'est celui qui donne l'expression la plus simple. Par conséquent on aura :

$$\frac{d^n}{dx^n} e^{-\varepsilon x^2} \leq n! e \cdot \varepsilon^{n/2} \cdot e^{-\frac{1}{2}\varepsilon x^2} \quad (\text{X.65})$$

En reportant cela dans (X.61), compte tenu de (X.62) et de l'inégalité dite du triangle, on obtient :

$$\left| d^j [(1 - e^{-\varepsilon x^2}) \varphi(x)] \right| \leq (1 - e^{-\varepsilon x^2}) |d^j \varphi(x)| + \sum_{\ell=0}^{j-1} \binom{j}{\ell} (j-\ell)! e \varepsilon^{(j-\ell)/2} \cdot e^{-\frac{1}{2}\varepsilon x^2} \cdot |d^\ell \varphi(x)| \quad (\text{X.66})$$

Pour avoir toutes les semi-normes, il faut encore multiplier cela par les facteurs $1 + |x|^k$, ce qui donne :

$$(1 + |x|^k) \left| d^j [(1 - e^{-\varepsilon x^2}) \varphi(x)] \right| \leq (1 - e^{-\varepsilon x^2}) (1 + |x|^k) |d^j \varphi(x)| + \sum_{\ell=0}^{j-1} \binom{j}{\ell} (j - \ell)! \varepsilon^{(j-\ell)/2} \cdot e^{-\frac{1}{2}\varepsilon x^2} \cdot (1 + |x|^k) |d^\ell \varphi(x)| \quad (\text{X.67})$$

Dans le premier terme à droite ci-dessus, on a le produit $(1 - e^{-\varepsilon x^2}) (1 + |x|^k)$; il faut utiliser le fait que $1 - e^{-\varepsilon x^2}$ tend vers zéro quand ε tend vers zéro, mais de manière à retrouver la convergence au sens des semi-normes. Pour cela, on remarque que pour tout x , $1 - e^{-\varepsilon x^2} \leq \varepsilon x^2$; d'où :

$$(1 - e^{-\varepsilon x^2}) (1 + |x|^k) \leq \varepsilon (x^2 + |x|^{k+2}) \leq \varepsilon [(1 + x^2) + (1 + |x|^{k+2})] \quad (\text{X.68})$$

Dans les termes sous $\sum$ de (X.67), on a aussi le facteur $e^{-\frac{\varepsilon x^2}{2}}$, qui est toujours ≤ 1 de sorte que (X.67) se majore comme suit :

$$(1 + |x|^k) \left| d^j [(1 - e^{-\varepsilon x^2}) \varphi(x)] \right| \leq \varepsilon (1 + |x|^2) |d^j \varphi(x)| + \quad (\text{X.69a})$$

$$+ \varepsilon (1 + |x|^{k+2}) |d^j \varphi(x)| + \sum_{\ell=0}^{j-1} \binom{j}{\ell} (j - \ell)! e^{(j-\ell)/2} \cdot (1 + |x|^k) |d^\ell \varphi(x)| \leq \quad (\text{X.69b})$$

$$\leq \varepsilon \mathcal{N}_{j,2}(\varphi) + \varepsilon \mathcal{N}_{j,k+2}(\varphi) + \sum_{\ell=0}^{j-1} \binom{j}{\ell} (j - \ell)! e^{(j-\ell)/2} \cdot \mathcal{N}_{\ell,k}(\varphi) \quad (\text{X.69c})$$

Sur cette dernière inégalité, on voit clairement ce qui se passe : à droite on a une somme de j termes, dont chacun contient un facteur ε^α , la plus petite valeur prise par α étant $\frac{1}{2}$. Les coefficients de ε^α sont formés de factorielles qui ne dépendent que des indices j, k, ℓ , et de semi-normes $\mathcal{N}_{j,2}(\varphi)$, $\mathcal{N}_{j,k+2}(\varphi)$, $\mathcal{N}_{\ell,k}(\varphi)$ qui sont toutes finies et indépendantes de ε . Le tout tend bien vers zéro et, cela, quels que soient les indices j, k choisis. •

X.3.4 Résolution d'équations différentielles

Cette application est la plus importante ; sa mise en œuvre utilise tout ce qui a été vu jusqu'ici, y compris les considérations théoriques (notamment la limite faible). On peut, même si on ne maîtrise pas la théorie, suivre les calculs présentés par un auteur, en lui faisant confiance pour ce qui est de leur validité. Mais lorsqu'on est livré à soi-même, une connaissance correcte de la théorie est nécessaire. En fait, la méthode des distributions ne devient vraiment puissante qu'en plusieurs dimensions (équations aux dérivées partielles). Cependant, nous étudierons, ici, un exemple en dimension un, puisque nous avons détaillé toute la théorie des distributions dans ce cadre. même s'il peut aussi être résolu sans les distributions, par des méthodes élémentaires, justement parce qu'il est de dimension un. Toutefois, la résolution par les distributions d'un problème à une dimension donnera une idée de ce qu'on peut faire en dimension supérieure.

Soit par exemple, l'équation :

$$u''(x) + k^2 u(x) = 1 + x^2 \quad (\text{X.70})$$

Cette équation est à coefficients constants : on est donc tenté de la résoudre par la méthode de Fourier (cf. section VII.1). L'ennui est que toutes les intégrales de Fourier qu'on va rencontrer seront divergentes. C'est pourquoi on va faire appel à la transformation de Fourier *au sens des distributions*. On cherche donc une distribution $T = \hat{u}$, telle que :

$$-\xi^2 \cdot T + k^2 T = 2\pi (\delta - \delta'') \quad (\text{X.71})$$

Le terme $-\xi^2 \cdot T$ est le produit de T par une fonction infiniment dérivable et à croissance polynomiale (cf. (X.28)). Il provient de la transformation de Fourier appliquée à $u'' = \delta'' * u$. La transformée de Fourier de la convolution étant le produit des transformées, on obtient en effet $\widehat{u''} = \widehat{\delta'' * u} = \widehat{\delta''} \cdot \widehat{u} = -\xi^2 \cdot T$. Le second membre de (X.70) se transforme comme suit : $\widehat{1} = 2\pi \delta$ et $\widehat{x^2} = -2\pi \delta''$. C'est ainsi qu'on obtient (X.71).

Pour résoudre (X.71), il suffit de diviser par $k^2 - \xi^2$:

$$T = \frac{2\pi(\delta - \delta'')}{-\xi^2 + k^2} \quad (\text{X.72})$$

On obtient alors la fonction u en prenant la transformée de Fourier inverse de ce résultat. Comme ce résultat est le produit de $2\pi(\delta - \delta'')$ — la transformée de Fourier du second membre de (X.70) — et de la fonction $g(\xi) = 1/(k^2 - \xi^2)$, on voit qu'il suffit de trouver la transformée de Fourier inverse de g au sens des distributions.

L'intégrale de Fourier au sens usuel :

$$G(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{-ix\xi}}{k^2 - \xi^2} d\xi \quad (\text{X.73})$$

est divergente en $\xi = \pm k$. Cela n'exclut pas que la fonction g puisse avoir une transformée au sens des distributions. Si c'est le cas, celle-ci devra être la limite (faible) des transformées de Fourier des fonctions régularisées :

$$g_z(\xi) = -\frac{1}{z^2 + \xi^2} \quad (\text{X.74})$$

où le dénominateur ne s'annule plus car on prend pour z un nombre complexe de partie réelle > 0 . Il est facile de vérifier que lorsque z tend vers ik , g_z tend faiblement vers $g_{ik} = g$ — théorèmes généraux de passage à la limite sous le signe intégral. Donc la transformée de Fourier de g_{ik} sera la limite faible des transformées de Fourier des g_z lorsque $z \rightarrow ik$ avec $\Re(z) > 0$. Mais on connaît déjà la transformée de Fourier inverse de g_z pour z réel > 0 : c'est la fonction $G_z(x) = -e^{-z|x|}/2z$. Par prolongement analytique, puisque tant que $\Re(z) > 0$, l'intégrale de Fourier dépend analytiquement de z en vertu des théorèmes généraux, ceci reste vrai dans tout le demi-plan $\Re(z) > 0$. Puis par continuité cela reste encore vrai en passant à la limite faible $z \rightarrow ik$, de sorte que :

$$G_{ik}(x) = -\frac{e^{-ik|x|}}{2ik} \quad (\text{X.75})$$

On appelle cette fonction *la fonction de Green* de l'équation (X.70). On voit comment la théorie des distributions permet un calcul rigoureux, alors que par la méthode de Fourier classique, toutes les intégrales seraient divergentes.

On obtient donc la solution de (X.70) sous la forme :

$$u(x) = G_{ik} * (1 + x^2) = -\frac{i}{2k} \int_{-\infty}^{+\infty} e^{-ik|x-y|} [1 + y^2] dy \quad (\text{X.76})$$

On remarquera que la fonction de Green G_{ik} donne la solution par convolution avec le second membre, quel que soit ce second membre, pourvu évidemment que cette convolution soit bien définie. Ainsi l'équation $u'' + k^2 u = f$ aura pour solution $u = G_{ik} * f$. On remarquera

que c'est aussi le résultat qu'on aurait obtenu par la méthode élémentaire de variation des constantes. C'est pour cette dernière raison que la méthode décrite ci-dessus est surtout intéressante en dimension supérieure car alors elle n'a pas de rivale.

X.4 Famille Y_α

La famille de distributions Y_α est, à des détails près, la famille déjà rencontrée sous le nom I_α . Il s'agit des distributions définies comme suit pour $\alpha \in \mathbb{C}$:

$$\langle Y_\alpha, \varphi \rangle = \begin{cases} \int_0^{+\infty} \frac{x^{\alpha-1}}{\Gamma(\alpha)} \varphi(x) dx & \text{si } \Re(\alpha) > 0 \\ \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\widehat{\varphi}(\xi)}{[(i\xi)^\alpha]_2} d\xi & \text{si } \Re(\alpha) < 1 \end{cases} \quad (\text{X.77})$$

Cette définition comporte deux cas : si $\Re(\alpha) > 0$, on voit que Y_α est simplement le poids nul pour $x < 0$ et égal à $x^{\alpha-1}/\Gamma(\alpha)$ pour $x > 0$. Si $\Re(\alpha) < 1$, on reconnaît que Y_α est la transformée de Fourier inverse du poids $1/[(i\xi)^\alpha]_2$. Dans les deux cas, le poids considéré satisfait bien aux conditions requises pour définir une distribution, c'est-à-dire que l'intégrale correspondante soit une fonctionnelle continue sur $\mathcal{S}(\mathbb{R})$: la fonction est localement intégrable et à croissance polynomiale. La fonctionnelle est donc bien définie dans les deux cas.

Lorsque $0 < \Re(\alpha) < 1$, les deux cas se recouvrent. Il faut donc vérifier que les deux définitions donnent le même résultat, c'est-à-dire que :

$$\int_0^{+\infty} \frac{x^{\alpha-1}}{\Gamma(\alpha)} \varphi(x) dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\widehat{\varphi}(\xi)}{[(i\xi)^\alpha]_2} d\xi \quad (\text{X.78})$$

On ne peut malheureusement pas faire la vérification directe par le calcul intégral ordinaire, en utilisant par exemple le théorème VII.3, car les fonctions $x^{\alpha-1}$ et $1/[(+i\xi)^\alpha]_2$ ne sont pas dans $\mathbb{L}_1(\mathbb{R})$ (pour $0 < \Re(\alpha) < 1$, elles sont toutes deux intégrables en $x = 0$ ou $\xi = 0$, mais pas à l'infini). On va donc utiliser le biais suivant : pour $\varepsilon > 0$, on a l'intégrale eulérienne de seconde espèce :

$$\int_0^{+\infty} \frac{x^{\alpha-1}}{\Gamma(\alpha)} e^{-(\varepsilon-i\xi)x} dx = \frac{1}{[(\varepsilon-i\xi)^\alpha]_2} \quad (\text{X.79})$$

Quand $\varepsilon \rightarrow 0$, on ne peut pas prendre la limite sous le signe intégral dans (X.79) ; mais (X.79) signifie néanmoins que $1/[(\varepsilon-i\xi)^\alpha]_2$ est la transformée de Fourier de la fonction égale à 0 pour $x < 0$ et à $x^{\alpha-1} e^{-\varepsilon x}/\Gamma(\alpha)$ pour $x > 0$, qui, elle, est dans $\mathbb{L}_1(\mathbb{R})$; d'après le théorème VII.3, on peut donc écrire, pour $\varphi \in \mathcal{S}(\mathbb{R})$:

$$\int_0^{+\infty} \frac{x^{\alpha-1} e^{-\varepsilon x}}{\Gamma(\alpha)} \varphi(x) dx = \int_{-\infty}^{+\infty} \frac{\widetilde{\varphi}(\xi)}{[(\varepsilon-i\xi)^\alpha]_2} d\xi \quad (\text{X.80})$$

D'autre part, $\widetilde{\varphi}(\xi) = \widehat{\varphi}(-\xi)/2\pi$; en substituant et en effectuant le changement de variable $\xi \mapsto -\xi$, on obtient :

$$\int_0^{+\infty} \frac{x^{\alpha-1} e^{-\varepsilon x}}{\Gamma(\alpha)} \varphi(x) dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\widehat{\varphi}(\xi)}{[(\varepsilon+i\xi)^\alpha]_2} d\xi \quad (\text{X.81})$$

Cette fois (grâce à la présence du facteur φ ou $\widehat{\varphi}$), on peut passer à la limite sous le signe intégral lorsque $\varepsilon \rightarrow 0$, ce qui donne (X.78) et prouve ainsi la cohérence de la définition (X.77).

Les théorèmes généraux garantissent aussi que l'intégrale :

$$\int_0^{+\infty} \frac{x^{\alpha-1}}{\Gamma(\alpha)} \varphi(x) dx \quad (\text{X.82})$$

est analytique dans tout le demi-plan $\Re(\alpha) > 0$ et que l'autre intégrale :

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\widehat{\varphi}(\xi)}{[(i\xi)^\alpha]_2} d\xi \quad (\text{X.83})$$

est analytique dans tout le demi-plan $\Re(\alpha) < 1$. Cela prouve donc que pour toute $\varphi \in \mathcal{S}(\mathbb{R})$, la fonction :

$$\alpha \mapsto \langle Y_\alpha, \varphi \rangle \quad (\text{X.84})$$

est analytique dans $\mathbb{C}$ tout entier.

La famille Y_α étant maintenant bien définie, on peut énoncer ses principales propriétés :

Théorème X.8 Pour tout $\alpha \in \mathbb{C}$, on a la relation $Y'_\alpha = Y_{\alpha-1}$; pour tous $\alpha, \beta \in \mathbb{C}$, on a $Y_\alpha * Y_\beta = Y_{\alpha+\beta}$; et enfin, pour n , entier supérieur ou égal à zéro, $Y_{-n} = \delta^{(n)}$.

On notera que la première relation se déduit des deux autres. La seconde relation signifie que la famille Y_α est un groupe pour la convolution. La troisième relation montre que la convolution d'une distribution ou d'une fonction par Y_α peut être interprétée comme une dérivée d'ordre non entier. Il est facile de voir que Y_α est égale à $\frac{1}{2\pi} I_\alpha$, où I_α est la distribution déjà introduite. La famille Y_α renferme donc les distributions singulières les plus courantes : δ et ses dérivées, ainsi que les pseudofonctions puissance.

Preuve Puisque $\alpha \mapsto \langle Y_\alpha, \varphi \rangle$ est, comme nous avons vu plus haut, analytique, il suffit de vérifier la relation $Y'_\alpha = Y_{\alpha-1}$ pour des valeurs de α d'un domaine non discret. On peut donc choisir un domaine où la vérification est particulièrement aisée, par exemple $\alpha \in]1; \infty[$. Dans ce cas Y_α s'identifie au poids égal à zéro pour $x < 0$ et à $x^{\alpha-1}/\Gamma(\alpha)$ pour $x > 0$, ce qui est dérivable par morceaux et continu (plus précisément, dérivable partout sauf au point $x = 0$ où la discontinuité est nulle). Dans ce cas on a vu que la dérivée de la distribution s'identifie à la dérivée usuelle, d'où le résultat.

Sachant que la transformée de Fourier d'une convolution donne un produit, la relation $Y_\alpha * Y_\beta = Y_{\alpha+\beta}$ semble évidente si on remarque qu'après transformation de Fourier elle devient :

$$[(i\xi)^{-\alpha}]_2 \cdot [(i\xi)^{-\beta}]_2 = [(i\xi)^{-\alpha-\beta}]_2 \quad (\text{X.85})$$

Toutefois cette dernière égalité n'est évidente que pour $\Re(\alpha)$ et $\Re(\beta)$ négatifs, car dans le cas contraire, les facteurs du produit ne sont pas de véritables fonctions, ce sont des pseudofonctions ; l'égalité est bien sûr vraie, mais n'est pas justifiée par la simple évidence, il faut des arguments supplémentaires faciles à trouver : supposons que $\Re(\alpha)$ et $\Re(\beta)$ ne soient pas tous deux négatifs. Il est alors possible de toujours trouver des entiers m et n tels que $\alpha - m$ et $\alpha - n$ soient tous deux de partie réelle négative ; de sorte que $Y_{\alpha-m} * Y_{\beta-n} = Y_{\alpha+\beta-m-n}$. Or, d'après la première relation, cela équivaut à $Y_\alpha^{(m)} * Y_\beta^{(n)} = Y_{\alpha+\beta}^{(m+n)}$. D'après les propriétés de la convolution, cela signifie que la $(m+n)^{\text{e}}$ dérivée de $Y_\alpha * Y_\beta$ est égale à la $(m+n)^{\text{e}}$ dérivée de $Y_{\alpha+\beta}$, donc que $Y_\alpha * Y_\beta$ diffère de $Y_{\alpha+\beta}$ par un polynôme. Ce polynôme est forcément nul puisque les distributions Y_α sont nulles sur $] -\infty; 0[$.

La troisième relation est facile à prouver : la transformée de Fourier de Y_α est $[(i\xi)^{-\alpha}]_2$; pour $\alpha = -n$ cette fonction cesse d'être multivoque et se réduit à $(i\xi)^n$, dont on sait que c'est la transformée de Fourier de $\delta^{(n)}$ (la n^{e} dérivée de δ). •

Nous savons déjà que, pour $\alpha < 1$, la distribution Y_α est la transformée de Fourier inverse de la fonction $\xi \mapsto 1/[(i\xi)^\alpha]_2$, puisque cela résulte directement de sa définition (X.77). Cette fonction est localement intégrable (la question se pose en $\xi = 0$) et à croissance polynomiale et définit bien un poids. Il n'en est plus de même pour $\alpha \leq 1$ ou plus généralement $\Re(\alpha) \leq 1$; dans ce cas, la fonction $\xi \mapsto 1/[(i\xi)^\alpha]_2$ n'est plus intégrable en $\xi = 0$ et par conséquent ne peut plus être un poids dans une intégrale — c'est pourquoi la définition (X.77) distingue deux cas. Nous avons vu au chapitre VIII qu'on régularise l'intégrale divergente correspondante en la réinterprétant comme intégrale le long d'un chemin contournant la singularité. Cette réinterprétation doit, pour être cohérente avec tout le reste, redonner la distribution Y_α (pour $\alpha > 1$ cette fois) ; c'est bien le cas si on compare le résultat du calcul avec la définition (X.77). La régularisation par contournement doit donc aussi pouvoir s'interpréter par la théorie des distributions. C'est ce que nous allons voir maintenant. Il s'agit donc de voir de plus près de quelle nature est la fonctionnelle associée à la fonction $1/[(i\xi)^\alpha]_2$ lorsque celle-ci n'est pas un poids, c'est-à-dire lorsque $\alpha > 1$.

Lorsque $\alpha > 0$, la distribution Y_α est régulière, de poids $p(x)$, égal à 0 pour $x < 0$ et à $x^{\alpha-1}/\Gamma(\alpha)$ pour $x > 0$. Ce poids est localement intégrable et à croissance polynomiale, mais n'est pas dans l'espace $\mathbb{L}_1(\mathbb{R})$. Donc sa transformée de Fourier n'est pas définie au sens usuel, et n'est donc pas non plus une vraie fonction. En tant que fonctionnelle, par contre, sa transformée de Fourier est définie par :

$$\langle \widehat{Y}_\alpha, \varphi \rangle = \langle Y_\alpha, \widehat{\varphi} \rangle = \frac{1}{\Gamma(\alpha)} \int_{-\infty}^{+\infty} x^{\alpha-1} \widehat{\varphi} dx \quad (\text{X.86})$$

Comme toujours, le meilleur moyen de se représenter visuellement cette distribution est de l'approcher (au sens de la limite faible) par des fonctions. Pour y parvenir, on remarque que les poids :

$$p_\varepsilon(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{x^{\alpha-1} e^{-\varepsilon x}}{\Gamma(\alpha)} & \text{si } x > 0 \end{cases} \quad (\text{X.87})$$

tendent faiblement (lorsque $\varepsilon \rightarrow 0$) vers Y_α ; en effet, cela signifie simplement que $\forall \varphi \in \mathcal{S}(\mathbb{R})$, on peut passer à la limite sous le signe intégral dans l'intégrale :

$$\int_0^\infty \frac{x^{\alpha-1}}{\Gamma(\alpha)} e^{-\varepsilon x} \widehat{\varphi}(x) dx \quad (\text{X.88})$$

ce qui, grâce à la présence du facteur $\widehat{\varphi}(x)$, est garanti par les théorèmes généraux.

Donc, d'après la théorie, les transformées de Fourier des p_ε vont aussi tendre (faiblement) vers $\widehat{Y}_\alpha$; or les p_ε , eux, sont des fonctions de $\mathbb{L}_1(\mathbb{R})$, et on peut calculer leurs transformées par le calcul intégral usuel. Ce qui donne :

$$\widehat{p}_\varepsilon(\xi) = \int_0^{+\infty} \frac{x^{\alpha-1} e^{-\varepsilon x}}{\Gamma(\alpha)} e^{ix\xi} dx \quad (\text{X.89})$$

On reconnaît l'intégrale eulérienne, et on obtient :

$$\widehat{p}_\varepsilon(\xi) = \frac{1}{[(\varepsilon - i\xi)^\alpha]_2} \quad (\text{X.90})$$

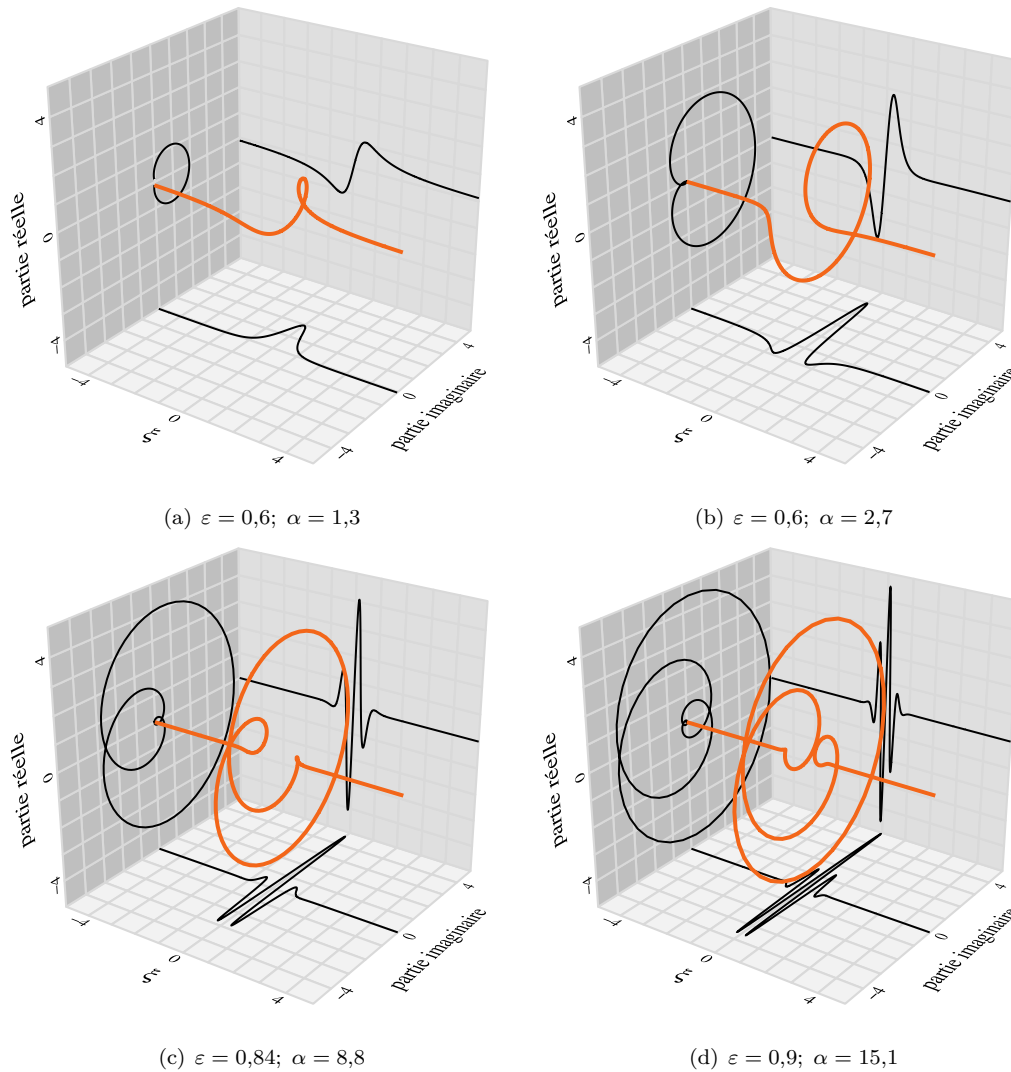


Figure X.5 – Vue en perspective du graphique de la fonction $[(\varepsilon - i\xi)^{-\alpha}]_2$ pour différents couples (ε, α) . La perspective est nécessaire puisque les valeurs sont complexes. La fonction reste très proche de zéro en dehors d’une région de largeur $\sim \varepsilon$, mais à l’intérieur de cette région elle décrit des orbes d’amplitude $\sim \varepsilon^{-\alpha}$.

ce qui fournit une approximation (selon la limite faible) de la distribution cherchée (voir figure X.5).

La méthode qui a été suivie pour ce calcul peut se résumer ainsi : pour calculer la transformée de Fourier d’une fonction $p(x)$ qui n’est pas intégrable, on commence par la multiplier par un facteur régularisant (ici $e^{-\varepsilon x}$), ce qui donne une famille p_ε de fonctions intégrables ; puis :

1. on vérifie que p_ε tend faiblement vers p , ce qui est possible par les passages à la limite élémentaires ;
2. on en déduit que $\widehat{p}_\varepsilon$ tend faiblement vers $\widehat{p}$, ce qui résulte de la théorie ;
3. en utilisant le fait que les p_ε sont des fonctions intégrables, on calcule leur transformée de Fourier par le calcul intégral classique ;
4. on obtient ainsi une approximation (au sens de la limite faible) de la distribution

cherchée par des fonctions.

Cette méthode a été suivie ici pour calculer les transformées de Fourier des fonctions $x^{\alpha-1}$, avec le facteur régularisant $e^{-\varepsilon x}$. La même méthode avait été suivie dans les sous-sections IX.8.4 et IX.8.5 pour calculer les transformées de Fourier des fonctions $[(i\xi)^{-\alpha}]_2$ et e^{ix^2} ; les facteurs régularisants étaient respectivement $e^{-\varepsilon|x|}$ et $e^{-\varepsilon x^2}$. Le choix du facteur régularisant est essentiel : il faut choisir celui qui rendra le calcul des intégrales $\int p_\varepsilon(x) e^{ix\xi} dx$ le plus simple possible. *C'est la théorie qui garantit que le résultat ne dépend pas du choix du facteur régularisant* : $\hat{p}$ est une distribution bien définie, qui sera forcément la limite faible de *n'importe quelle* suite $\hat{p}_\varepsilon$, pourvu que p soit bien la limite faible de p_ε .

On peut aussi déduire de (X.90) par la formule d'inversion :

$$p_\varepsilon(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{-ix\xi}}{[(\varepsilon - i\xi)^\alpha]_2} d\xi \quad (\text{X.91})$$

En posant $z = \varepsilon - i\xi$, on voit que la fonction sous le signe intégral est une fonction analytique de z en dehors de la demi-droite $\{z < 0\}$, et on peut interpréter l'intégrale ci-dessus comme intégrale prise selon z le long du chemin $\Re(z) = \varepsilon$, qu'on peut donc déformer sans changer la valeur de l'intégrale :

$$\int_{-\infty}^{+\infty} \frac{e^{-ix\xi}}{[(\varepsilon - i\xi)^\alpha]_2} d\xi = i e^{-\varepsilon x} \int_\gamma \frac{e^{zx}}{[z^\alpha]_2} dz \quad (\text{X.92})$$

En particulier, on peut prendre pour γ un chemin qui suit l'axe $\Re(z) = 0$ excepté autour de $z = 0$, où le chemin fait un détour pour éviter la singularité.

Ainsi se trouve justifié le procédé de régularisation d'intégrale divergente introduit en section VIII.3.



XI — Espaces de Hilbert

XI.1 Espaces euclidiens de dimension infinie

En dimension finie, un espace euclidien est un espace vectoriel avec un produit scalaire. Les espaces euclidiens de même dimension finie étant tous isomorphes, l'espace $\mathcal{E}_n$ des polynômes de degré $\leq n$ fournit l'exemple générique des espaces de dimension $n + 1$ (on aurait tout aussi bien pu prendre pour exemple $\mathbb{R}^{n+1}$).

Sur $\mathcal{E}_n$, on peut définir une infinité de produits scalaires différents ; en voici trois, notés $\langle P | Q \rangle_k$ avec $P(x) = \sum a_j x^j$ et $Q(x) = \sum b_j x^j$:

$$\begin{aligned}\langle P | Q \rangle_1 &= \sum_{j=0}^{j=n} a_j b_j \\ \langle P | Q \rangle_2 &= \int_{-1}^1 P(x) Q(x) dx \\ \langle P | Q \rangle_3 &= \int_{-\infty}^{+\infty} e^{-x^2} P(x) Q(x) dx\end{aligned}\tag{XI.1}$$

Si on identifie un polynôme au vecteur de $\mathbb{R}^{n+1}$ formé par ses coefficients, $\langle | \rangle_1$ est le produit scalaire usuel mais les deux autres sont différents.

Les polynômes $1, x, x^2, x^3, \dots, x^n$ sont deux à deux orthogonaux pour le produit scalaire $\langle | \rangle_1$ mais ils ne le sont plus pour les produits scalaires $\langle | \rangle_2$ et $\langle | \rangle_3$. Des polynômes orthogonaux pour le produit scalaire $\langle | \rangle_2$ sont, par exemple, les polynômes de Legendre :

$$\begin{aligned}P_0(x) &= 1; & P_1(x) &= x; & P_2(x) &= \frac{3x^2 - 1}{2}; & P_3(x) &= \frac{5x^3 - 3x}{2}; \\ P_4(x) &= \frac{35x^4 - 30x^2 + 3}{8}; & P_5(x) &= \frac{63x^5 - 70x^3 + 15x}{8}; & \dots\end{aligned}\tag{XI.2}$$

Pour le produit scalaire $\langle | \rangle_3$ ce seraient les polynômes d'Hermite :

$$\begin{aligned} H_0(x) &= 1; & H_1(x) &= 2x; & H_2(x) &= 4x^2 - 2; & H_3(x) &= 8x^3 - 12x; \\ H_4(x) &= 16x^4 - 48x^2 + 12; & H_5(x) &= 32x^5 - 160x^3 + 120x; & \dots \end{aligned} \quad (\text{XI.3})$$

Les familles de polynômes $\{1, x, x^2, x^3, \dots, x^n\}$, $\{P_0, P_1, P_2, \dots, P_n\}$ et $\{H_0, H_1, H_2, \dots, H_n\}$ sont des bases de l'espace $\mathcal{E}_n$, orthogonales pour (respectivement) les produits scalaires $\langle | \rangle_1$, $\langle | \rangle_2$ et $\langle | \rangle_3$. Pour les rendre orthonormées, il suffit de diviser chaque polynôme par sa norme.

Si, au lieu de considérer l'espace $\mathcal{E}_n$ des polynômes de degré inférieur ou égal à n , on considère l'espace $\mathcal{E}_\infty$ de tous les polynômes quel que soit leur degré, on obtient un espace de dimension infinie. Rien n'est changé en ce qui concerne les produits scalaires, ni l'orthogonalité; les familles de polynômes $\{1, x, x^2, \dots, x^n, \dots\}$, $\{P_0, P_1, P_2, \dots, P_n, \dots\}$ et $\{H_0, H_1, H_2, \dots, H_n, \dots\}$ sont toujours orthogonales (pour le produit scalaire correspondant), mais sont maintenant infinies.

L'infini introduit des propriétés nouvelles, la plus importante étant la suivante : tout polynôme est une combinaison linéaire *finie* des polynômes de base $\{1, x, x^2, \dots, x^n\}$ (ou $\{P_0, P_1, P_2, \dots, P_n\}$, ou encore $\{H_0, H_1, H_2, \dots, H_n\}$), mais on peut envisager aussi une série convergente *infinie* des polynômes de base.

Or, puisqu'un polynôme est une somme *finie* du type $\sum a_j x^j$, $\sum a_j P_j(x)$ ou $\sum a_j H_j(x)$, c'est donc qu'une somme *infinie* de ce type *n'est pas* un polynôme.

La question est alors : « que représente une somme infinie (série convergente) ? » Pour que la réponse ait un sens, il faut d'abord préciser ce qu'on entend par série convergente; or, dans un espace euclidien, la notion de convergence qui s'impose naturellement est celle liée à la distance euclidienne : on dit que la série $\sum f_j$ converge vers f si $\mathbb{R}\text{-}\lim_n \|f - \sum_{j \leq n} f_j\|_k = 0$ avec $\|g\|_k = \sqrt{\langle g | g \rangle_k}$. Ici, nous ne savons pas à l'avance ce que peut être f ; il nous faut donc une définition *intrinsèque* de la convergence. Une telle définition intrinsèque est fournie par le critère de Cauchy :

Définition XI.1 Dans un espace euclidien E , on dit qu'une série $\sum_{n \geq 0} f_n$ est intrinsèquement convergente si :

$$\mathbb{R}\text{-}\lim_{n \rightarrow \infty} \left\{ \sup_{p \geq n} \left\| \sum_{j=n}^p f_j \right\| \right\} = 0 \quad (\text{XI.4})$$

avec toujours $\|g\| = \sqrt{\langle g | g \rangle}$.

On vérifie immédiatement par l'inégalité du triangle (qui s'applique à la norme $\| \cdot \|$) que ce critère est satisfait si la série est normalement convergente, c'est-à-dire si la série des normes $\sum \|f_n\|$ est convergente dans $\mathbb{R}$. Il s'agit là d'une condition suffisante et non nécessaire. Le point essentiel est cependant qu'une série peut être intrinsèquement convergente sans pour autant avoir une somme dans le même espace.

Dans un espace euclidien de dimension infinie, une série intrinsèquement convergente n'est pas forcément convergente.

On peut donner immédiatement un contre-exemple. La série :

$$\sum_{n \geq 0} \frac{x^n}{n} \quad (\text{XI.5})$$

est intrinsèquement convergente dans l'espace euclidien $\mathcal{E}_\infty$ de tous les polynômes muni du produit scalaire $\langle | \rangle_1$, puisque, en utilisant le fait que la famille x^j est orthonormée, on peut écrire pour n'importe quelle somme partielle :

$$\left\| \sum_{j=n}^{j=p} \frac{1}{n} x^n \right\|^2 = \sum_{j=n}^{j=p} \frac{1}{n^2} \quad (\text{XI.6})$$

et la série de terme général $1/n^2$ est convergente dans $\mathbb{R}$. Or, la série (XI.5) ne peut pas être elle-même un polynôme, puisque ses coefficients ne sont pas nuls à partir d'un certain rang ; si cette série devait d'une façon ou d'une autre avoir une somme, ce serait la fonction $f(x) = -\ln(1-x)/x$, qui n'est pas un polynôme.

Ce phénomène est semblable à celui qu'on rencontre avec les séries ou les suites de nombres rationnels. Les séries :

$$\sum_{n \geq 1} \frac{1}{n^2} \quad \text{ou} \quad \sum_{n \geq 0} \frac{(-1)^n}{2n+1} \quad (\text{XI.7})$$

sont intrinsèquement convergentes puisqu'elles satisfont au critère de Cauchy. Toutefois, elles n'ont pas de somme dans le même espace : cette somme n'est pas un nombre rationnel puisque c'est respectivement $\pi^2/6$ et $\pi/4$. On dit alors que l'espace $\mathbb{Q}$ n'est pas complet, tandis que l'espace $\mathbb{R}$ l'est. Dans la présente discussion, il apparaît donc que l'espace euclidien $\mathcal{E}_\infty$ des polynômes n'est pas complet.

En conclusion, dans un espace euclidien de dimension infinie, rien ne garantit qu'une série intrinsèquement convergente ait une somme appartenant à cet espace ; sa somme peut être en dehors. Autrement dit, rien ne garantit qu'un tel espace soit complet, bien que ce soit un espace vectoriel sur $\mathbb{R}$. Par contre, en dimension finie, un espace euclidien sur $\mathbb{R}$ est forcément complet.

La question posée plus haut : « que représente une somme infinie de la forme $\sum a_j x^j$? » doit alors être précisée ainsi : « que représentent les séries infinies *intrinsèquement convergentes* de la forme $\sum a_j x^j$? » Dans les trois cas considérés en exemple, la réponse est alors la suivante :

1. les séries infinies de la forme $\sum a_j x^j$, $\sum a_j P_j(x)$ ou $\sum a_j H_j(x)$ sont intrinsèquement convergentes pour la distance euclidienne $\|g\|_1 = \sqrt{\langle g | g \rangle_1}$: on obtient un élément de l'espace ℓ^2 , qui est l'espace de toutes les suites $\{c_n\}_{n \geq 0}$ telles que la série $\sum_n c_n^2$ soit convergente ;
2. les séries infinies de la forme $\sum a_j x^j$, $\sum a_j P_j(x)$ ou $\sum a_j H_j(x)$ sont intrinsèquement convergentes pour la distance euclidienne $\|g\|_2 = \sqrt{\langle g | g \rangle_2}$: on obtient un élément de l'espace $\mathbb{L}_2([-1; 1])$ des fonctions de carré intégrable sur $[-1; 1]$;
3. les séries infinies de la forme $\sum a_j x^j$, $\sum a_j P_j(x)$ ou $\sum a_j H_j(x)$ sont intrinsèquement convergentes pour la distance euclidienne $\|g\|_3 = \sqrt{\langle g | g \rangle_3}$: on obtient un élément de

l'espace $\mathbb{L}_2(e^{-x^2/2}\mathbb{R})$ des fonctions $f(x)$ sur $\mathbb{R}$ telles que $\int |f(x)|^2 e^{-x^2} dx$ soit convergente — *i.e.* l'espace des fonctions définies sur $\mathbb{R}$ telles que $e^{-x^2/2} f(x)$ soit de carré intégrable.

Ces trois espaces ℓ^2 , $\mathbb{L}_2([-1; 1])$ et $\mathbb{L}_2(e^{-x^2/2}\mathbb{R})$ sont complets.

Les phénomènes nouveaux introduits par l'infini sont donc les suivants :

1. une série infinie, contrairement aux sommes finies, peut avoir une somme extérieure à l'espace considéré ;
2. ces sommes (de séries infinies intrinsèquement convergentes) forment une espace *qui dépend du produit scalaire considéré*.

Notons que ce n'est pas le choix de la base qui détermine le résultat, c'est bien celui du produit scalaire, parce que les séries intrinsèquement convergentes ne sont pas les mêmes selon la notion de limite choisie.

On pourrait aussi se demander quel espace on obtiendrait pour les séries infinies convergentes au sens de la limite uniforme déjà connue⁽¹⁾. Réponse : l'espace correspondant est alors l'ensemble des limites uniformes de polynômes et cet espace n'est autre que celui des fonctions continues sur $[-1; 1]$.

Bien entendu, ces résultats ne sont pas évidents et les démonstrations sont difficiles. Pour chaque produit scalaire particulier, l'ensemble des séries convergentes selon la métrique qu'il induit, forme un espace spécifique qui est complet ; il faut dire aussi que les trois exemples de produits scalaires considérés, $\langle | \rangle_1$, $\langle | \rangle_2$ et $\langle | \rangle_3$, ont été choisis parce que l'espace correspondant est spécialement intéressant.

XI.2 Espaces de Hilbert

On appelle *espace de Hilbert*, un espace euclidien de dimension infinie qui est complet. Nous admettrons que les espaces suivants sont complets.

- ℓ^2 : espace de toutes les suites infinies $\{c_n\}_{n \in \mathbb{N}}$ de nombres réels telles que la série $\sum c_n^2$ converge ; c'est donc un espace de Hilbert pour le produit scalaire $\langle \{a_n\} | \{b_n\} \rangle_{n \in \mathbb{N}} = \sum a_n b_n$;
- $\mathbb{L}_2(p, [a; b])$: espace de toutes les fonctions $f(x)$ définies sur l'intervalle $[a; b]$ telles que l'intégrale $\int_a^b p(x) f(x)^2 dx$ converge ; c'est donc un espace de Hilbert pour le produit scalaire :

$$\langle f | g \rangle = \int_a^b p(x) f(x) g(x) dx \quad (\text{XI.8})$$

La fonction $p(x)$ est appelée le *poids*. Si le poids est 1, on écrit simplement $\mathbb{L}_2([a; b])$ au lieu de $\mathbb{L}_2(p, [a; b])$.

- $\mathbb{L}_2(p, [0; \infty[)$ ou $\mathbb{L}_2(p,]-\infty; \infty[)$: on devine ;

Les espaces suivants sont complets, mais ne sont pas des espaces de Hilbert, car leur métrique n'est pas définie par un produit scalaire.

(1) On notera que la limite uniforme n'est pas liée à une distance euclidienne ; il n'y a pas de produit scalaire qui conduit à la limite uniforme.

- $\mathbb{L}_1(p, [a; b])$, $\mathbb{L}_1(p, [0; \infty[)$, $\mathbb{L}_1(p,]-\infty; \infty[)$: leur métrique est définie par la norme $\|f\|_{1,p} = \int p(x) |f(x)| dx$, l'intégrale portant sur le domaine $[a; b]$, $[0; \infty[$ ou $]-\infty; \infty[$;
- $\mathbb{C}_0([a; b])$: espace des fonctions f continues sur $[a; b]$ avec la métrique définie par la norme $\|f\|_\infty = \max\{|f(x)|\}$ (c'est la métrique de la limite uniforme);
- $\mathbb{C}_0([0; \infty[)$ ou $\mathbb{C}_0(]-\infty; \infty[)$: espace des fonctions f continues sur $[0; \infty[$ ou $]-\infty; \infty[$ et nulles à l'infini⁽²⁾, également avec la métrique de la limite uniforme.

Afin d'illustrer les propriétés des espaces de Hilbert, nous allons étudier de plus près l'espace $\mathbb{L}_2([-1; 1])$ des fonctions dont le carré est intégrable sur l'intervalle $[-1; 1]$. Cet espace contient $\mathbb{C}_0([-1; 1])$ comme sous-espace vectoriel; mais pour la métrique euclidienne de $\mathbb{L}_2([-1; 1])$, $\mathbb{C}_0([-1; 1])$ n'est pas complet (il l'est pour la métrique uniforme). $\mathbb{C}_0([-1; 1])$ contient à son tour le sous-espace vectoriel $\mathcal{T}$ des polynômes trigonométriques de la forme :

$$P(x) = a_0 + \sum_{j=1}^n a_j \cos nx + b_j \sin nx \quad (\text{XI.9})$$

où n est appelé le degré du polynôme (si a_n ou b_n est $\neq 0$). Pour éviter les confusions, appelons $\mathcal{T}_\infty$ l'espace de tous les polynômes trigonométriques quel que soit leur degré et $\mathcal{T}_N$, l'espace de tous les polynômes trigonométriques de degré inférieur ou égal à N . Il est clair que $\mathcal{T}_\infty$ est de dimension infinie et $\mathcal{T}_N$ de dimension finie $2N + 1$.

Pour résumer, $\mathcal{T}_N$ est un sous-espace vectoriel de $\mathcal{T}_\infty$, qui est lui-même un sous-espace vectoriel de $\mathbb{C}_0([-1; 1])$, lui-même un sous-espace vectoriel de $\mathbb{L}_2([-1; 1])$.

Considérons la fonction $f_0(x) = |x|$; en tant que fonction définie sur $[-1; 1]$, elle est continue et donc appartient à l'espace $\mathbb{C}_0([-1; 1])$ même si elle n'appartient pas au sous-espace $\mathcal{T}_\infty$: f_0 n'est pas un polynôme trigonométrique (un polynôme trigonométrique serait dérivable, alors que $|x|$ ne l'est pas en $x = 0$).

Développons f_0 en série de Fourier. Puisqu'on considère l'intervalle $[-1; 1]$, on prendra une série de la forme $a_0 + a_1 \cos \pi x + a_2 \cos 2\pi x + a_3 \cos 3\pi x + \dots$. D'après le théorème de Fourier, les coefficients sont :

$$c_n = \int_{-1}^1 f_0(x) \cos n\pi x dx \quad (\text{XI.10})$$

Cette intégrale est aussi l'expression du produit scalaire de l'espace de Hilbert $\mathbb{L}_2([-1; 1])$ dans lequel se trouvent toutes les fonctions considérées ici. Les coefficients de Fourier peuvent donc s'écrire aussi $c_n = \langle f_0 | \cos_n \rangle$, où $\cos_n$ désigne la fonction $x \mapsto \cos n\pi x$. Il est facile de calculer les c_n : la fonction $f_0(x) \cos n\pi x$ étant paire, l'intégrale (XI.10) est égale à :

$$c_n = 2 \int_0^1 x \cos n\pi x dx = \begin{cases} 0 & \text{si } n \text{ est pair} \\ -\frac{4}{(n\pi)^2} & \text{si } n \text{ est impair} \end{cases} \quad (\text{XI.11})$$

Les sommes partielles $S_N = \sum_0^N c_n \cos_n$ sont des polynômes trigonométriques de degré N (ou $N - 1$ si N est pair) et sont donc dans le sous-espace $\mathcal{T}_N$.

On remarquera que les fonctions $\cos_n$ sont orthonormées :

(2) Si on néglige cette condition, il peut se produire des horreurs.

– si $n \neq m$:

$$\begin{aligned}\langle \cos_n | \cos_m \rangle &= 2 \int_0^1 \cos n\pi x \cos m\pi x \, dx \\ &= \int_0^1 \cos(n+m)\pi x + \cos(n-m)\pi x \, dx \\ &= \frac{\sin(n+m)\pi x}{n+m} + \frac{\sin(n-m)\pi x}{n-m} \, dx = 0\end{aligned}\quad (\text{XI.12a})$$

– si $n = m$:

$$\langle \cos_n | \cos_n \rangle = 2 \int_0^1 \cos^2 n\pi x \, dx = \int_0^1 1 + \cos 2n\pi x \, dx = 1 \quad (\text{XI.12b})$$

Calculons la distance $\|f_0 - S_{2p-1}\|$. On va d'abord développer son carré. En utilisant la bilinéarité du produit scalaire puis le fait que les fonctions $\cos_n$ sont orthonormées, on obtient successivement :

$$\begin{aligned}\|f_0 - S_N\|^2 &= \langle f_0 - \sum_{n=0}^{2p-1} c_n \cos_n | f_0 - \sum_{n=0}^{2p-1} c_n \cos_n \rangle \\ &= \langle f_0 | f_0 \rangle - 2 \sum_{n=0}^{2p-1} c_n \langle f_0 | \cos_n \rangle + \sum_{n=0}^{2p-1} \sum_{m=0}^{2p-1} c_n c_m \langle \cos_n | \cos_m \rangle \\ &= \|f_0\|^2 - 2 \sum_{n=0}^{2p-1} c_n^2 + \sum_{n=0}^{2p-1} c_n^2 \\ &= \|f_0\|^2 - \sum_{n=0}^{2p-1} c_n^2\end{aligned}\quad (\text{XI.13})$$

Calculons aussi la distance $f_0 - T_{2p-1}$ de f_0 à un élément *quelconque* T_{2p-1} du sous-espace $\mathcal{T}_{2p-1}$. Le polynôme trigonométrique T_{2p-1} ne s'exprime pas seulement avec les fonctions $\cos_n$, mais aussi avec les fonctions $\sin_n$; il est cependant facile de vérifier que l'ensemble des fonctions $\cos_n$ et $\sin_n$ forment une famille orthonormée. Posons donc :

$$T_{2p-1} = a_0 + \sum_{n=1}^{2p-1} a_n \cos_n + b_n \sin_n \quad (\text{XI.14})$$

et calculons :

$$\begin{aligned}\|f_0 - T_N\|^2 &= \left\langle f_0 - \sum_{n=0}^{2p-1} a_n \cos_n + b_n \sin_n \left| f_0 - \sum_{n=0}^{2p-1} a_n \cos_n + b_n \sin_n \right. \right\rangle \\ &= \langle f_0 | f_0 \rangle - 2 \sum_{n=0}^{2p-1} a_n \langle f_0 | \cos_n \rangle + b_n \langle f_0 | \sin_n \rangle + \\ &\quad \sum_{n=0}^{2p-1} \sum_{m=1}^{2p-1} a_n a_m \langle \cos_n | \cos_m \rangle + a_n b_m \langle \cos_n | \sin_m \rangle + b_n a_m \langle \sin_n | \cos_m \rangle + b_n b_m \langle \sin_n | \sin_m \rangle \\ &= \langle f_0 | f_0 \rangle - 2 \sum_{n=0}^{2p-1} a_n c_n + \sum_{n=0}^{2p-1} \sum_{m=1}^{2p-1} a_n^2 + b_n^2 \\ &= \|f_0\|^2 - \sum_{n=0}^{2p-1} c_n^2 + \sum_{n=0}^{2p-1} (a_n - c_n)^2 + b_n^2 \\ &= \|f_0 - S_{2p-1}\|^2 + \sum_{n=0}^{2p-1} (a_n - c_n)^2 + b_n^2\end{aligned}\quad (\text{XI.15})$$

Pour obtenir l'avant-dernière ligne, on a utilisé l'identité $a_n^2 - 2a_n c_n = -c_n^2 + (a_n - c_n)^2$. On voit d'après la dernière ligne que la distance $\|f_0 - T_{2p-1}\|^2$, égale à $\|f_0 - S_{2p-1}\|^2$ plus une somme de carrés, est donc *toujours* supérieure ou égale à $\|f_0 - S_{2p-1}\|^2$. Elle devient *égale* à $\|f_0 - S_{2p-1}\|^2$ lorsque cette somme de carrés devient nulle, c'est-à-dire lorsque pour tout $n \leq 2p-1$, on a $a_n = c_n$ et $b_n = 0$; autrement dit, lorsque $T_{2p-1} = S_{2p-1}$.

On peut donc dire que, parmi tous les polynômes trigonométriques T de degré $\leq 2p-1$, S_{2p-1} est celui pour lequel la distance $f_0 - T$ est minimale. En général, les coefficients de Fourier d'une fonction sont ceux qui rendent minimum cette distance. La somme partielle de Fourier S_{2p-1} est appelée *la projection orthogonale de f_0 sur le sous-espace $\mathcal{T}_{2p-1}$* .

La projection orthogonale sur un sous-espace n'existe pas toujours. Cherchons en effet le minimum de $f_0 - T$ lorsque T parcourt, non $\mathcal{T}_{2p-1}$ mais $\mathcal{T}_\infty$. Pour cela, nous utilisons le fait supposé connu que la somme de la série de Fourier complète est égale à f_0 , c'est-à-dire que :

$$\lim_{p \rightarrow \infty} \|f_0 - S_{2p-1}\| = 0 \quad (\text{XI.16})$$

Cela signifie que la borne inférieure de la distance de f_0 à $\mathcal{T}_\infty$ est nulle. Mais ce n'est pas un minimum. S'il existait un polynôme trigonométrique S_∞ qui réalise ce minimum, il ne pourrait être qu'égal à f_0 : on aurait $\|f_0 - S_\infty\| = \min \{\|f_0 - S_{2p-1}\|\} = 0$, d'où $f_0 = S_\infty$. Or, c'est impossible puisque f_0 n'est pas un polynôme trigonométrique.

On voit ainsi la manifestation d'une propriété impensable en dimension finie : la distance de f_0 au sous-espace vectoriel $\mathcal{T}_\infty$ est nulle et pourtant f_0 n'est pas dans ce sous-espace. L'élément f_0 de l'espace $\mathbb{C}_0([-1; 1])$ est la limite d'une suite de polynômes trigonométriques comme un nombre irrationnel est la limite d'une suite de fractions. C'est le même phénomène qu'en section XI.1 avec la fonction $-\ln(1-x)/x$, qui n'est pas un polynôme mais qui est la limite, selon le produit scalaire $\langle | \rangle_1$, d'une suite de polynômes.

On peut donc dire que l'espace $\mathcal{T}_\infty$ n'est pas complet pour la métrique issue du produit scalaire $\langle | \rangle_2$. L'espace $\mathbb{C}_0([-1; 1])$ n'est pas complet non plus pour cette métrique. Seul l'espace $\mathbb{L}_2([-1; 1])$, qui les englobe tous, l'est, du fait du théorème de Fischer-Riesz.

Pour se rendre compte que $\mathbb{C}_0([-1; 1])$ n'est pas complet, on peut prendre la suite de fonctions continues :

$$f_n(x) = \begin{cases} -1 & \text{pour } x < -\frac{1}{n} \\ nx & \text{pour } -\frac{1}{n} \leq x \leq +\frac{1}{n} \\ 1 & \text{pour } x > \frac{1}{n} \end{cases} \quad (\text{XI.17})$$

Cette suite converge vers la fonction discontinue :

$$f(x) = \begin{cases} -1 & \text{pour } x < 0 \\ 1 & \text{pour } x > 0 \end{cases} \quad (\text{XI.18})$$

pour la métrique issue du produit scalaire $\langle | \rangle_2$. En effet, $f_n - f$ est une fonction nulle en dehors de l'intervalle $[-1/n; 1/n]$, où elle est plus petite que 1, donc l'intégrale de son carré tend vers zéro. On a ainsi exhibé dans $\mathbb{L}_2([-1; 1])$ une fonction discontinue qui est limite de fonctions continues, donc $\mathbb{C}_0([-1; 1])$ n'est pas complet.

Il ne faut cependant pas en conclure hâtivement que les sous-espaces de dimension infinie d'un espace de Hilbert sont forcément non complets. Par exemple, on peut considérer le sous-espace $\mathcal{H}_+$ de $\mathbb{L}_2([-1; 1])$ des fonctions paires et le sous-espace $\mathcal{H}_-$ des fonctions impaires. Il est facile de voir que toute fonction définie sur l'intervalle $[-1; 1]$ s'écrit d'une manière unique comme la somme d'une fonction paire et d'une fonction impaire : si $f \in \mathbb{L}_2([-1; 1])$, on pose $f_+(x) = \frac{1}{2}(f(x) + f(-x))$ et $f_-(x) = \frac{1}{2}(f(x) - f(-x))$; on a alors $f = f_+ + f_-$ et l'unicité se vérifie aisément. Les deux sous-espaces $\mathcal{H}_+$ et $\mathcal{H}_-$ sont complets en vertu du fait que la limite d'une suite de fonctions paires est forcément aussi une fonction paire (et de même pour les fonctions impaires) : si f_n est une suite intrinsèquement convergente (ou suite de Cauchy) de fonctions paires, elle a une limite f dans $\mathbb{L}_2([-1; 1])$ car cet espace est complet et d'autre part, cette limite est paire car c'est une limite de fonctions paires, donc $f \in \mathcal{H}_+$.

Ⓡ Cette argumentation a conduit à introduire indépendamment de la notion d'espace complet, celle de sous-espace fermé : un sous-espace fermé est défini par la propriété que s'il contient une suite convergente, il contient aussi la limite de cette suite ; ainsi les sous-espaces $\mathcal{H}_+$ et $\mathcal{H}_-$ sont des sous-espaces fermés de $\mathbb{L}_2([-1; 1])$: si une suite f_n de $\mathcal{H}_-$ converge vers un élément f de $\mathbb{L}_2([-1; 1])$, cet élément f sera forcément dans $\mathcal{H}_-$. L'argument ci-dessus peut donc se résumer ainsi : un sous-espace fermé d'un espace complet est complet ; mais un sous-espace fermé d'un espace non complet n'est pas forcément complet : par exemple, le sous-espace $\mathcal{T}_+$ de $\mathcal{T}_\infty$ des fonctions continues paires n'est pas complet, bien qu'il soit fermé dans $\mathcal{T}_\infty$; il suffit pour s'en convaincre de se souvenir que la fonction $f_0(x) = |x|$ est limite de polynômes trigonométriques formés de fonctions $\cos_n$, donc pairs.

La différence entre *fermé* et *complet* est que le premier terme est relatif et le second absolu ou intrinsèque : on ne dira jamais qu'un espace E est fermé, mais qu'un sous-espace F de E est fermé dans E ; par contre E ou F sera dit complet ou non indépendamment de ce qui se passe en dehors. Ainsi $\mathcal{T}_+$ est fermé dans $\mathcal{T}_\infty$ mais non complet.

À l'intérieur d'un espace complet, fermé équivaut à complet : tout sous-espace fermé est complet et tout sous-espace non fermé est non complet. Dans un espace non complet, par contre, on peut dire que tout sous-espace complet est fermé, mais non l'inverse. Par exemple les sous-espaces de dimension finie sont complets et fermés ; mais le sous-espace $\mathbb{C}_{0+}([-1; 1])$ des fonctions continues paires est fermé dans $\mathbb{C}_0([-1; 1])$ et non complet.

Enfin, insistons encore sur le rôle de la métrique. Le fait qu'un espace vectoriel normé ou semi-normé soit complet ou non dépend de la norme ou des semi-normes : $\mathcal{S}(\mathbb{R})$ est complet pour les semi-normes $\mathcal{N}_{j,k}$ mais non complet pour la norme $\mathcal{N}_{0,0}$ de la limite uniforme ou pour la norme $\|\cdot\|_2$; $\mathbb{C}_0([-1; 1])$ est non complet pour la norme $\|\cdot\|_2$, mais complet pour la limite uniforme ; $\mathbb{L}_2([-1; 1])$ est complet pour la norme $\|\cdot\|_2$, mais non complet pour la norme $\|\cdot\|_1$ etc.

XI.3 Bases orthonormées

Les propriétés connues des séries de Fourier nous ont permis de constater que la fonction $f_0(x) = |x|$ pouvait être approchée d'aussi près qu'on veut par des polynômes trigonométriques. Approchée au sens de la distance euclidienne définie par la norme $\|\cdot\|_2$.

Le théorème classique de Weierstrass⁽³⁾ dit que toute fonction continue f sur un intervalle borné $[a; b]$ — donc appartenant à l'espace $\mathbb{C}_0([a; b])$ — peut être approchée uniformément par des polynômes trigonométriques de $\sin\left(\frac{2\pi}{b-a}x\right)$ et $\cos\left(\frac{2\pi}{b-a}x\right)$. Ce théorème s'applique donc en particulier à l'intervalle $[-1; 1]$. Cela peut s'écrire en langage plus mathématique :

$$\forall \varepsilon > 0, \quad \exists P \in \mathcal{T}_\infty, \quad \max_{-1 \leq x \leq 1} |f(x) - P(x)| < \varepsilon \quad (\text{XI.19})$$

(3) K. WEIERSTRASS, *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin* 1885.

Ⓡ Ce théorème de Weierstrass ne signifie pas que la série de Fourier de la fonction f converge uniformément vers f ; l'assertion (XI.19) signifie que pour tout ε on peut trouver un P mais rien ne dit que ce P est précisément une somme partielle de la série de Fourier ; et ce serait faux ! Ce qui est vrai est dit par le théorème de Féjer : Soient $S_n(x) = a_0 + \sum_1^n a_j \cos n\pi x + b_j \sin n\pi x$, les sommes partielles de la série de Fourier de $f(x)$ et $F_n(x) = \frac{1}{n} \sum_1^n S_j$, la moyenne des n premières de ces sommes partielles. Si f est une fonction périodique (de période 2) continue, alors la suite F_n converge uniformément vers f . Mais la suite S_n ne converge pas toujours uniformément vers f ; c'est cependant le cas dans des cas particuliers, par exemple pour la fonction f_0 considérée ci-dessus.

La norme $\max |f(x) - P(x)|$ est la distance *uniforme* entre P et f . On la désigne souvent par $\|f - P\|_\infty$. D'après l'inégalité de la moyenne, on a pour toute fonction $g \in \mathbb{C}_0([-1; 1])$ la majoration :

$$\|g\|_2^2 \leq \int_{-1}^1 [g(x)]^2 dx \leq 2 \|g\|_\infty^2 \quad (\text{XI.20})$$

d'où :

$$\|g\|_2 \leq \sqrt{2} \|g\|_\infty \quad (\text{XI.21})$$

Ce qui montre que la convergence uniforme implique la convergence euclidienne (pour la norme $\| \cdot \|_2$). Sur un intervalle $[a; b]$ quelconque on aurait $\|f\|_2 \leq \sqrt{b-a} \|f\|_\infty$ et sur un intervalle non borné tel que $]0; \infty[$ ou $]-\infty; \infty[$ on ne peut pas avoir une inégalité du type $\|f\|_2 \leq A \|f\|_\infty$; de ce fait, sur un intervalle non borné la convergence uniforme *n'implique pas* la convergence euclidienne.

Ainsi, le théorème de Weierstrass garantit que toute fonction de l'espace $\mathbb{C}_0([-1; 1])$ peut être approchée d'aussi près qu'on veut pour la distance $\| \cdot \|_2$ par un polynôme trigonométrique. Il va donc se produire pour n'importe quelle fonction $f \in \mathbb{C}_0([-1; 1])$ la même chose que pour la fonction $f_0(x) = |x|$: la borne inférieure de la distance de f à un polynôme trigonométrique P , lorsque celui-ci parcourt le sous-espace $\mathcal{T}_\infty$, est zéro.

Un autre théorème de Weierstrass dit aussi que toute fonction continue sur $\mathbb{C}_0([a; b])$ peut être approchée uniformément d'aussi près qu'on veut par un polynôme algébrique (non trigonométrique) : si on désigne par $\mathcal{P}_\infty$ l'espace vectoriel des polynômes de la forme $p(x) = \sum a_n x^n$ quel que soit leur degré, alors pour $f \in \mathbb{C}_0([-1; 1])$:

$$\forall \varepsilon > 0, \quad \exists p \in \mathcal{P}_\infty, \quad \max_{-1 \leq x \leq 1} |f(x) - p(x)| < \varepsilon \quad (\text{XI.22})$$

Pour la même raison que plus haut, on peut donc conclure que la borne inférieure de la distance euclidienne $\|f - p\|_2$ de f à un polynôme p lorsque celui-ci parcourt l'espace $\mathcal{P}_\infty$, est zéro ; mais ce n'est pas un minimum.

On exprime ces propriétés en disant que les sous-espaces $\mathcal{T}_\infty$ et $\mathcal{P}_\infty$ sont *denses* dans $\mathbb{C}_0([-1; 1])$.

Définition XI.2 Un sous-espace E d'un espace euclidien F est dense dans F si tout élément f de F est la limite d'une suite f_n d'éléments de E .

Le fait que les sous-espaces $\mathcal{T}_\infty$ et $\mathcal{P}_\infty$ sont *denses* dans $\mathbb{C}_0([-1; 1])$ résulte de ces théorèmes de Weierstrass qui ne sont pas évidents.

Par ailleurs, il se trouve que $\mathbb{C}_0([-1; 1])$ lui-même est dense dans $\mathbb{L}_2([-1; 1])$. Cela se prouve par le même type d'argument que le théorème X.7. Soit $f \in \mathbb{L}_2([-1; 1])$; si

$\rho_n(x) = \sqrt{\frac{n}{\pi}} e^{-nx^2}$ est le filtre régularisant introduit à cette occasion, les fonctions $f_n = \rho_n * f$ sont continues et tendent — selon la norme $\| \cdot \|_2$ — vers f . La démonstration en sera également donnée à la section suivante.

Ces propriétés de densité, à savoir, que $\mathcal{T}_\infty$, l'ensemble des polynômes trigonométriques et $\mathcal{P}_\infty$, l'ensemble des polynômes algébriques, sont denses dans $\mathbb{L}_2([-1; 1])$, ont la conséquence très importante que voici :

Théorème XI.1 Si $\{\varphi_n\}_{n \in \mathbb{N}}$ est une famille orthonormée de fonctions, alors n'importe quelle fonction f de $\mathbb{L}_2([-1; 1])$ est égale à la somme de la série $\sum c_n \varphi_n$ avec $c_n = \langle f | \varphi_n \rangle_2$, la série étant convergente dans $\mathbb{L}_2([-1; 1])$, c'est-à-dire pour la norme $\| \cdot \|_2$. En langage plus mathématique :

$$\forall f \in \mathbb{L}_2([-1; 1]), \quad \mathbb{R}\text{-}\lim_{n \rightarrow \infty} \left\| f - \sum_{j=0}^{j=n} c_j \varphi_j \right\|_2 = 0 \quad (\text{XI.23})$$

Preuve Soit en effet Φ_N le sous-espace vectoriel engendré par les fonctions $\{\varphi_0, \varphi_1, \varphi_2, \dots, \varphi_N\}$; si $\varphi_n = P_n / \|P_n\|_2$, ce sous-espace sera $\mathcal{P}_N$; si $\varphi_0 = 1$, $\varphi_{2n} = \cos_n$ et $\varphi_{2n-1} = \sin_n$, on aura $\Phi_{2N+1} = \mathcal{T}_N$. La projection orthogonale de f sur Φ_N est $f_N = \sum_{j=0}^N c_j \varphi_j$. Si la réunion Φ_∞ des Φ_N est dense dans $\mathbb{L}_2([-1; 1])$, on aura :

$$\forall \varepsilon > 0, \quad \exists n_\varepsilon \in \mathbb{N}, \quad \exists P \in \Phi_{n_\varepsilon}, \quad \|f - P\|_2 \leq \varepsilon \quad (\text{XI.24})$$

Mais puisque $f_N = \sum_{j=0}^N c_j \varphi_j$ est la projection orthogonale de f sur Φ_N , qui rend la distance *minimum*, on aura $\|f - f_{n_\varepsilon}\|_2 \leq \|f - P\|_2$, donc $\|f - f_{n_\varepsilon}\|_2 \leq \varepsilon$; cela suffit pour garantir que $\mathbb{L}_2\text{-}\lim f_n = f$, car la série $\sum c_n \varphi_n$ est intrinsèquement convergente. •

Ce raisonnement montre que le produit scalaire joue un rôle essentiel. Pour la métrique uniforme (celle de la norme $\| \cdot \|_\infty$), qui ne dérive pas d'un produit scalaire, on ne peut pas construire une suite de projections orthogonales comme f_n , bien qu'il existe une suite g_n de fonctions de Φ_∞ qui tende vers f ; on peut certes construire aussi cette suite g_n et c'est d'ailleurs ce que l'on fait pour démontrer les théorèmes de Weierstrass mais cette construction est plus complexe. Le grand avantage des séries de Fourier est qu'on dispose d'une formule simple pour calculer leurs coefficients. Cet avantage n'est pas particulier aux fonctions trigonométriques mais est commun à toutes les bases hilbertiennes.

XI.4 Exemples de bases orthonormées

XI.4.1 Polynômes de Legendre

Le premier exemple est déjà connu : sur $\mathbb{L}_2([-1; 1])$, c'est la famille trigonométrique $1, \cos, \sin, \cos_2, \sin_2, \cos_3, \sin_3, \cos_4, \sin_4, \dots$. Dans le même espace, une autre base est formée par les *polynômes de Legendre* que nous présentons brièvement.

L'expression $(x^2 - 1)^n$ est un polynôme de degré $2n$; si on la dérive n fois, on obtiendra donc un polynôme de degré n . Le n^{e} polynôme de Legendre est :

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n [(x^2 - 1)^n]}{dx^n} \quad (\text{XI.25})$$

Le coefficient numérique $1/2^n n!$ est un facteur de normalisation conventionnel. Nous prenons la formule (XI.25) comme définition des polynômes de Legendre mais ce n'est pas comme cela qu'ils ont été introduits à l'origine. Dans la littérature, (XI.25) est connue sous le nom de *formule de Rodrigues*.

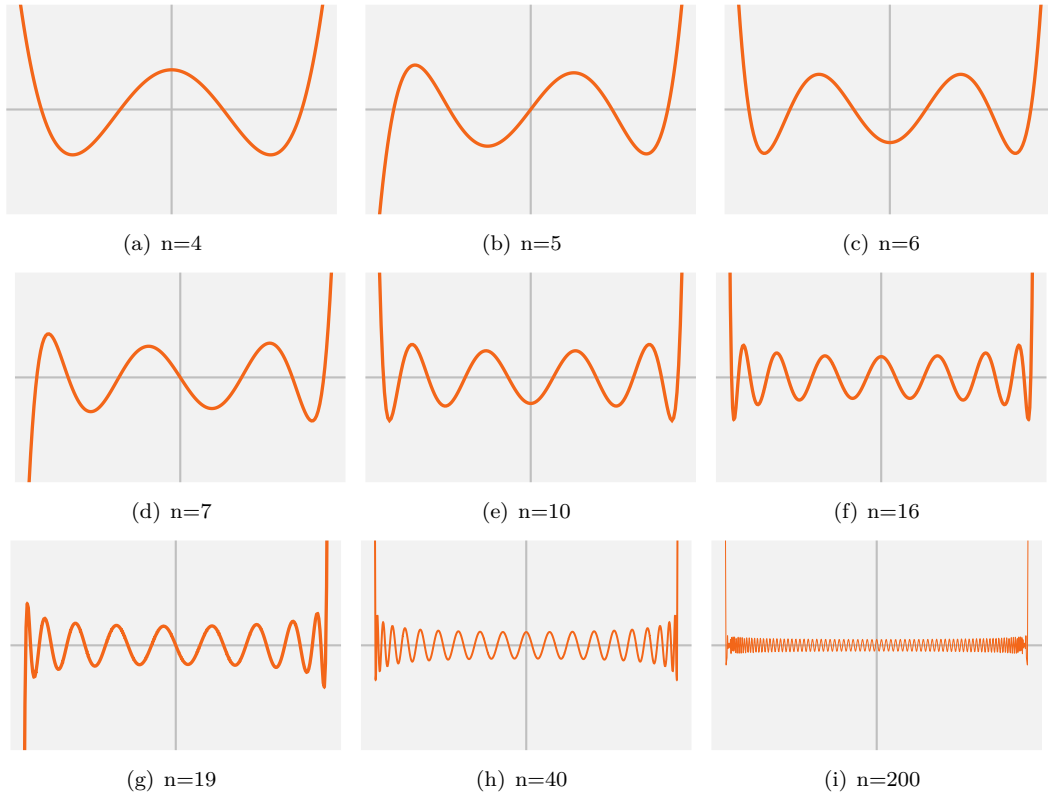


Figure XI.1 – Polynômes de Legendre

Sur la figure XI.1, les polynômes de Legendre $P_n(x)$ sont les coefficients de Fourier de la fonction $\theta \mapsto 1/|2 \sin \frac{\theta}{2}|$. À partir de $n = 40$, on a utilisé l'approximation asymptotique :

$$P_n(\cos \theta) \simeq \sqrt{\frac{2}{\pi n \sin \theta}} \cos \left[\left(n + \frac{1}{2} \right) \theta - \frac{\pi}{4} \right] \quad (\text{XI.26})$$

la différence étant graphiquement imperceptible. Cependant, pour le calcul numérique, mieux vaut préférer une relation de récurrence, par exemple :

$$P_{n+1}(x) = \frac{2n+1}{n+1} x P_n(x) - \frac{2n+1}{n+1} P_{n-1}(x) \quad (\text{XI.27})$$

Les polynômes de Legendre prennent des valeurs comprises entre -1 et 1 lorsque $x \in [-1; 1]$. Il est donc possible d'avoir un cadrage identique pour toutes les valeurs de n . Par contre, dès que x franchit la limite 1 ou -1 , $P_n(x)$ croît très rapidement, d'autant plus rapidement que n est grand. Les racines et les oscillations sont entièrement incluses dans l'intervalle $[-1; 1]$.

Ces polynômes se rencontrent dans toutes sortes de problèmes, par exemple la théorie du spin de l'électron. Ils sont solution de l'équation différentielle du second ordre :

$$(1 - x^2) y''(x) - 2x y'(x) + n(n+1)y(x) = 0 \quad (\text{XI.28})$$

Puisque chaque polynôme $P_n(x)$ est de degré n , ils forment une base algébrique de $\mathcal{P}_\infty$, l'espace vectoriel de tous les polynômes. Or, nous avons vu que $\mathcal{P}_\infty$ est dense dans $\mathbb{L}_2([-1; 1])$. Montrons maintenant qu'ils sont orthogonaux (les polynômes $P_n(x)$ ne sont pas normés, mais on en déduit une famille normée en prenant $p_n = P_n/\|P_n\|$). Le coefficient de normalisation $1/2^n n!$ dans (XI.25) ne joue aucun rôle dans l'orthogonalité ; il suffit donc de vérifier que les intégrales :

$$\int_{-1}^1 \frac{d^n([x^2 - 1]^n)}{dx^n} \frac{d^m([x^2 - 1]^m)}{dx^m} dx \quad (\text{XI.29})$$

sont nulles pour $m \neq n$. Pour cela effectuons une intégration par parties : l'intégrale (XI.29) est égale à :

$$\left. \frac{d^{n-1}([x^2 - 1]^n)}{dx^{n-1}} \frac{d^m([x^2 - 1]^m)}{dx^m} \right|_{-1}^{+1} - \int_{-1}^1 \frac{d^{n-1}([x^2 - 1]^n)}{dx^{n-1}} \frac{d^{m+1}([x^2 - 1]^m)}{dx^{m+1}} dx \quad (\text{XI.30})$$

L'un des facteurs du terme intégré est la $(n-1)^e$ dérivée de $(x^2 - 1)^n$, qui est factorisable par $(x^2 - 1)$; plus généralement, $(n-k)^e$ dérivée de $(x^2 - 1)^n$ serait factorisable par $(x^2 - 1)^k$. Par conséquent, ce facteur est nul aussi bien pour $x = -1$ que pour $x = +1$, de sorte que le terme intégré disparaît. Il reste une intégrale qu'on peut à nouveau intégrer par parties et ainsi de suite, les termes intégrés étant à chaque fois nuls pour la même raison. En définitive :

$$\int_{-1}^1 \frac{d^n([x^2 - 1]^n)}{dx^n} \frac{d^m([x^2 - 1]^m)}{dx^m} dx = (-1)^n \int_{-1}^1 [x^2 - 1]^n \frac{d^{m+n}([x^2 - 1]^m)}{dx^{m+n}} dx \quad (\text{XI.31})$$

Le second facteur sous le signe intégral est la dérivée $(m+n)^e$ de $[x^2 - 1]^m$, qui est un polynôme de degré $2m$; si $n > m$ on aura donc dérivé un nombre de fois supérieur au degré, ce qui donne zéro. Avec $m > n$, on aurait procédé dans l'autre sens. Ainsi, on a bien $\langle P_n | P_m \rangle = 0$ si $n \neq m$. L'égalité (XI.31) peut aussi être utilisée si $n = m$; dans ce cas le second facteur sous le signe intégral est la $2n^e$ dérivée de $[x^2 - 1]^n$, qui est égale à la $2n^e$ dérivée du terme du plus haut degré ; ce dernier étant x^{2n} , sa dérivée $2n^e$ est $(2n)!$. En tenant compte cette fois des coefficients de normalisation, on en déduit :

$$\langle P_n | P_n \rangle = \frac{(-1)^n}{(2^n n!)^2} \int_{-1}^1 [x^2 - 1]^n (2n)! dx = \frac{(-1)^n (2n)!}{2^{2n} n!^2} \int_{-1}^1 [x^2 - 1]^n dx \quad (\text{XI.32})$$

Il reste à calculer l'intégrale ; celle-ci se ramène aux intégrales eulériennes de première espèce (cf. chapitre V) par le changement de variable $t = \frac{1}{2}(1+x)$:

$$\begin{aligned} (-1)^n \int_{-1}^1 [x^2 - 1]^n dx &= 2^{2n+1} \int_0^1 t^n (1-t)^n dt = 2^{2n+1} \beta(n+1, n+1) \\ &= 2^{2n+1} \frac{\Gamma(n+1)\Gamma(n+1)}{\Gamma(2n+2)} = \frac{2^{2n+1} n!^2}{(2n+1)!} \end{aligned} \quad (\text{XI.33})$$

Substituant cela dans (XI.32), on obtient :

$$\langle P_n | P_n \rangle = \frac{2}{2n+1} \quad \text{ou} \quad \|P_n\| = \sqrt{\frac{2}{2n+1}} = \frac{1}{\sqrt{n + \frac{1}{2}}} \quad (\text{XI.34})$$

Ainsi la famille des fonctions $p_n(x) = \sqrt{n + \frac{1}{2}} P_n(x)$ est orthonormée. Comme elle engendre algébriquement le sous-espace $\mathcal{P}_\infty$ qui est dense dans $\mathbb{L}_2([-1; 1])$, elle est donc une base

hilbertienne de $\mathbb{L}_2([-1; 1])$. Ce qui signifie concrètement que *n'importe* quelle fonction f de $\mathbb{L}_2([-1; 1])$ est égale à la somme de la série $\sum c_n p_n$, avec $c_n = \langle f | p_n \rangle$, cette série étant convergente pour la norme $\| \cdot \|$:

$$f = \mathbb{L}_2\text{-}\lim_{n \rightarrow \infty} \sum_{n=0}^{\infty} c_n p_n \iff \mathbb{R}\text{-}\lim_{n \rightarrow \infty} \left\| f - \sum_{n=0}^{\infty} c_n p_n \right\|_2 = 0 \quad (\text{XI.35})$$

Bien entendu, (XI.35) ne donne aucun renseignement concernant la convergence uniforme ou point par point de la série. Nous verrons que dans les applications, notamment à la mécanique quantique, la convergence en moyenne quadratique est la plus significative et la mieux adaptée.

XI.4.2 Fonctions et polynômes d'Hermite

On va maintenant construire une base de l'espace de Hilbert des fonctions de carré intégrable sur tout $\mathbb{R}$: $\mathbb{L}_2(]-\infty; \infty[)$. Ces fonctions sont les fonctions propres de l'oscillateur quantique ⁽⁴⁾.

Dérivons la fonction e^{-x^2} :

$$\frac{d e^{-x^2}}{dx} = -2x e^{-x^2}; \quad \frac{d^2 e^{-x^2}}{dx^2} = (4x^2 - 1) e^{-x^2}; \quad \frac{d^3 e^{-x^2}}{dx^3} = (-8x^3 + 12x) e^{-x^2} \quad (\text{XI.36})$$

On voit que ces dérivées sont de la forme polynôme $\times e^{-x^2}$. Cela est confirmé par la récurrence :

$$H_{n+1}(x) e^{-x^2} = \frac{d H_n(x) e^{-x^2}}{dx} = [H'_n(x) - 2x H_n(x)] e^{-x^2} \quad (\text{XI.37})$$

d'où :

$$H_{n+1}(x) = H'_n(x) - 2x H_n(x) \quad (\text{XI.38})$$

qui montre que si H_n est un polynôme de degré n , H_{n+1} sera un polynôme de degré $n + 1$. Cela montre aussi que si l'entier n est pair, le polynôme H_n sera pair et si l'entier n est impair, le polynôme H_n sera impair. Cette même récurrence montre enfin que pour tout n , le terme du plus haut degré du polynôme H_n est $(-2x)^n$, donc sa n^{e} dérivée sera $(-2)^n n!$

Les polynômes $(-1)^n H_n(x)$ sont appelés les *polynômes d'Hermite*.

Les polynômes $H_n(x)/n!$ sont les coefficients du développement en série entière de la fonction $z \mapsto e^{2xz - z^2}$. On obtient aussi les polynômes d'Hermite $H_n(x)$ en dérivant la fonction e^{-x^2} . On a en effet :

$$\frac{d^n e^{-x^2}}{dx^n} = (-1)^n H_n(x) e^{-x^2} \quad (\text{XI.39})$$

Sur la figure XI.2, si on voulait représenter les graphiques pour n beaucoup plus grand, on devrait élargir peu à peu le cadre pour rendre visibles toutes les oscillations. On remarque aussi que les oscillations sont d'amplitude très inégale : si on veut garder les oscillations latérales dans le cadre, on rend les oscillations centrales à peine perceptibles. Ce phénomène

(4) E. SCHRÖDINGER, *Der stetige Übergang von der Mikro- zur Makromechanik*, Die Naturwissenschaften, 1926 vol. 28, p. 664-666. Traduction française de A. PROCA dans le recueil *Mémoires sur la mécanique ondulatoire* (p. 65-70), réédité en 1988 aux éditions Jacques Gabay.

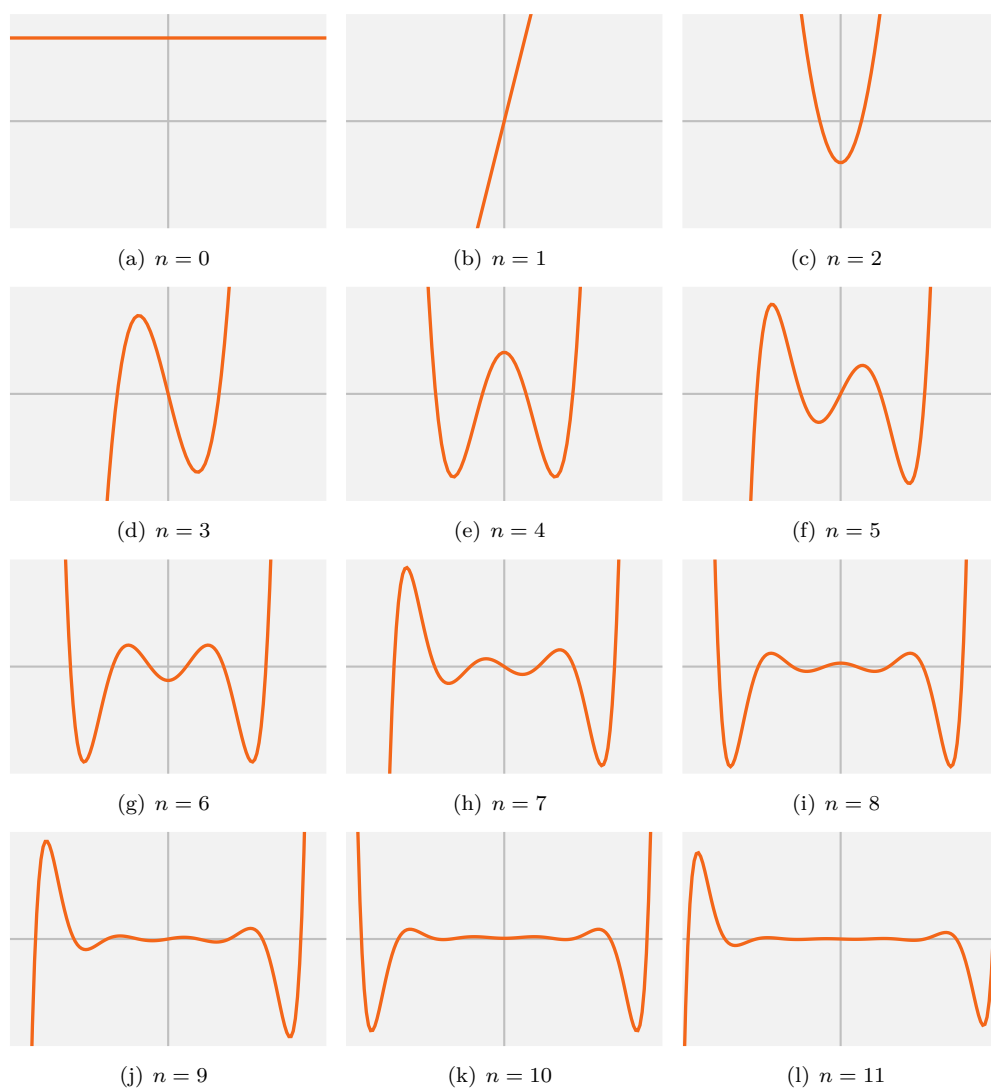


Figure XI.2 – Polynômes d'Hermite $H_n(x)$. La fenêtre représente les abscisses $x \in [-3,83; 3,83]$ et les ordonnées $y \in [-1,3; 1,3]$. Ce cadrage est choisi pour montrer la région où les polynômes ont leurs racines. En dehors de ce cadre, ils tendent rapidement vers l'infini. Les graphiques représentent, en fait, les polynômes $H_n(x)/n!$ qui ne sortent pas de ce cadre lorsque n augmente.

s'accroît quand n augmente. En multipliant $H_n(x)$ par la fonction $e^{-x^2/2}$, on arrive à compenser cette disproportion, comme indiqué sur la figure XI.3.

Les polynômes d'Hermite se rencontrent dans de nombreuses applications, dont l'oscillateur quantique. Ils sont notamment solution de l'équation du second ordre :

$$y''(x) - 2x y'(x) + 2n y(x) = 0 \quad (\text{XI.40})$$

et on les rencontre donc dans la résolution d'équations différentielles similaires.

Pour tout $\alpha > 0$, les fonctions $h_n(\sqrt{\alpha}x)$, illustrées sur la figure XI.3 forment une base de $\mathbb{L}_2(\mathbb{R})$; avec $\alpha = \sqrt{mk}/\hbar$, elles représentent les états liés de l'oscillateur quantique de masse m et d'élasticité k .

Les fonctions $h_n(x) = H_n(x) e^{-x^2/2}$ sont orthogonales pour le produit scalaire de l'espace $\mathbb{L}_2(]-\infty; \infty[)$ pour la même raison que les polynômes de Legendre puisque :

$$\begin{aligned} \langle h_n | h_m \rangle &= \int_{-\infty}^{+\infty} h_n(x) h_m(x) dx = \int_{-\infty}^{+\infty} e^{-x^2} H_n(x) H_m(x) dx \\ &= \int_{-\infty}^{+\infty} \frac{d^n e^{-x^2}}{dx^n} H_m(x) dx \end{aligned} \quad (\text{XI.41})$$

puis intégrer n fois de suite par parties ; les termes tout intégrés sont nuls à cause de la décroissance très rapide à l'infini du facteur e^{-x^2} , donc il reste :

$$\langle h_n | h_m \rangle = (-1)^n \int_{-\infty}^{+\infty} e^{-x^2} \frac{d^n H_m(x)}{dx^n} dx \quad (\text{XI.42})$$

On constate à nouveau que l'un des facteurs sous l'intégrale est un polynôme de degré m dérivé n fois, ce qui donne zéro si $m < n$. Pour trouver la norme des h_n on utilise également (XI.42), mais avec $n = m$. La n^{e} dérivée de H_n se réduit à celle du terme de degré n , qui est $(-2x)^n$, soit $(-2)^n n!$; de sorte que :

$$\langle h_n | h_n \rangle = 2^n n! \int_{-\infty}^{+\infty} e^{-x^2} dx = 2^n n! \sqrt{\pi} \quad \Rightarrow \quad \|h_n\| = \pi^{1/4} 2^{n/2} \sqrt{n!} \quad (\text{XI.43})$$

Par conséquent, la famille des fonctions :

$$\varphi_n(x) = \frac{H_n(x) e^{-x^2/2}}{\pi^{1/4} 2^{n/2} \sqrt{n!}} \quad (\text{XI.44})$$

est orthonormée. N'importe quelle fonction de carré intégrable sur l'intervalle $]-\infty; \infty[$ peut donc être développée en *série d'Hermite* comme une fonction de période 2π peut l'être en série de Fourier.

XI.4.3 Fonctions et polynômes de Laguerre

Le troisième exemple nous conduit maintenant à l'espace $\mathbb{L}_2([0; \infty[)$. Cet espace de Hilbert se rencontre surtout dans les problèmes à deux ou trois dimensions, où il apparaît par suite de la factorisation des coordonnées polaires ou sphériques : l'intervalle $[0; \infty[$ est en effet le domaine naturel de la coordonnée radiale. La base hilbertienne que nous allons construire correspond cette fois à la partie radiale des fonctions propres d'un électron dans un champ

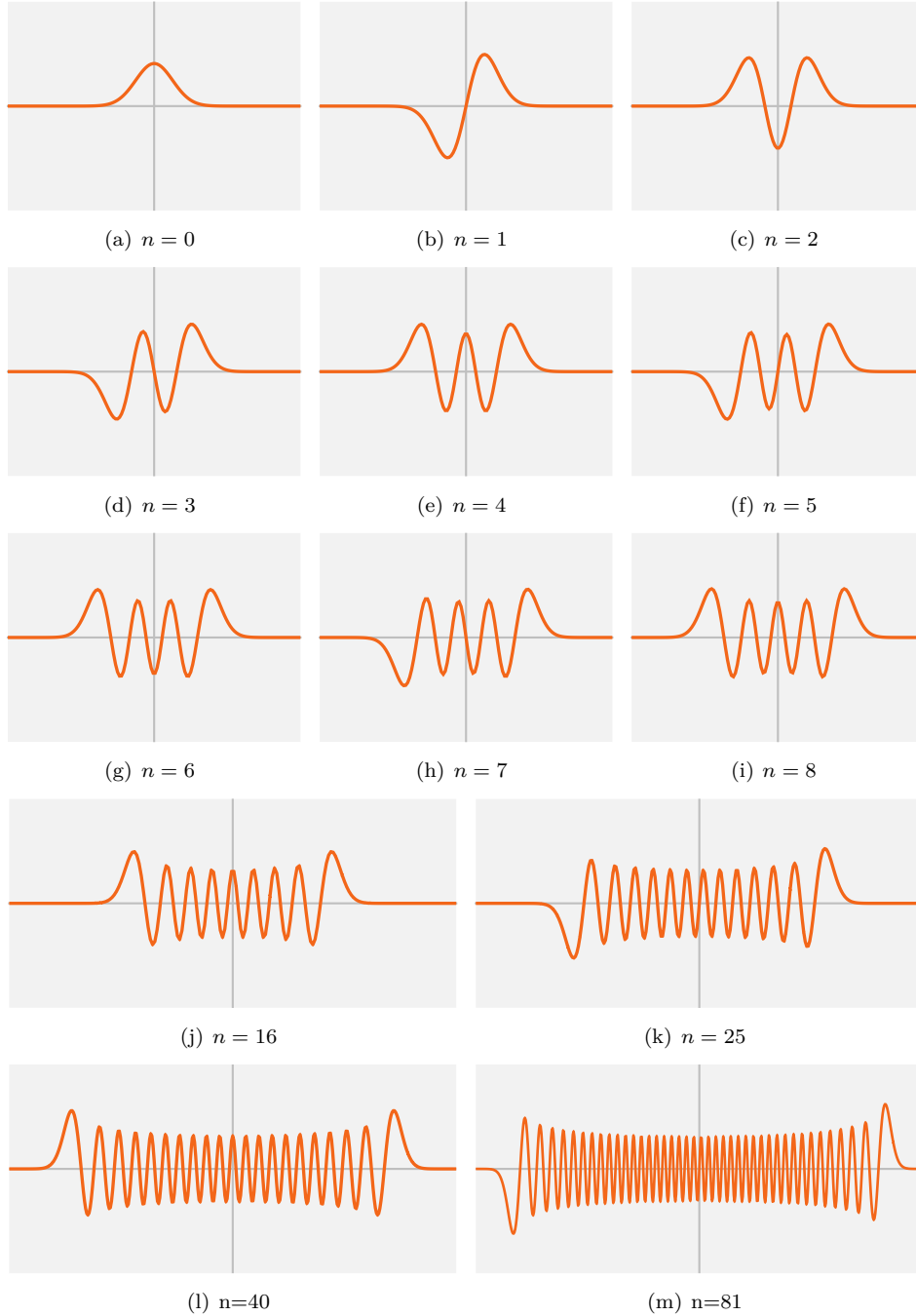


Figure XI.3 – Fonctions d’Hermite $h_n(x) = c_n/n! H_n(x) e^{-x^2/2}$. Le coefficient de normalisation c_n a été pris égal à $(-1)^n (n/2)^{n/2} e^{-n/2} \sqrt{1+n/2}$, qui donne un graphique visible pour tout n sans varier l’échelle verticale. Dans les petites cases, le cadrage est le même que pour les polynômes d’Hermite de la figure XI.2. Quand n augmente, il faut étendre les abscisses.



Figure XI.4 – Fonctions de Laguerre $q_n^0(x) = \frac{1}{n!} e^{-x/2} Q_n^0(x)$. Les petites fenêtres graphiques ($n = 0$ à 11) correspondent aux abscisses positives $x \in [0, 57; 5]$ et aux ordonnées $y \in [-1, 3; 1, 3]$. Les fenêtres élargies sont à la même échelle et dévoilent simplement davantage d'abscisses : $x \in [0, 84; 3]$ pour $n = 13$, $x \in [0; 100]$ pour $n = 18$ et $x \in [0; 200]$ pour $n = 41$.

coulombien, problème célèbre résolu par E. Schrödinger et un des premiers succès de la mécanique quantique ⁽⁵⁾.

Les fonctions de la figure XI.4 (pour $n = 0, 1, 2, \dots, \infty$) forment une base de l'espace $\mathbb{L}_2([0; \infty[)$.

On définit d'abord les polynômes de Laguerre d'une manière analogue aux polynômes d'Hermite : pour ces derniers, on dérivait la fonction e^{-x^2} , ici on dérive n fois les fonctions $x^{n+\ell} e^{-x}$. Ces polynômes sont nommés d'après le mathématicien Laguerre (1834–1886) :

$$\begin{aligned}\frac{dx^{\ell+1} e^{-x}}{dx} &= [(\ell+1) - x] x^\ell e^{-x} \\ \frac{d^2 x^{\ell+2} e^{-x}}{dx^2} &= [(\ell+1)(\ell+1) - 2(\ell+1)x + x^2] x^\ell e^{-x}\end{aligned}\tag{XI.45}$$

Pour le cas général, le mieux est d'utiliser la formule de Leibniz :

$$\frac{d^n(fg)}{dx^n} = \sum_{j=0}^n \binom{n}{j} \frac{d^{n-j}f}{dx^{n-j}} \cdot \frac{d^j g}{dx^j}\tag{XI.46}$$

avec $f(x) = x^{\ell+n}$ et $g(x) = e^{-x}$. On obtient :

$$\begin{aligned}\frac{d^n x^{\ell+n} e^{-x}}{dx^n} &= \sum_{j=0}^n \binom{n}{j} \frac{(n+\ell)!}{(j+\ell)!} x^{j+\ell} \cdot (-1)^j e^{-x} \\ &= \left[\sum_{j=0}^n (-1)^j \binom{n}{j} \frac{(n+\ell)!}{(j+\ell)!} x^j \right] \cdot x^\ell e^{-x}\end{aligned}\tag{XI.47}$$

ce qui montre que la n^e dérivée de la fonction $x^{n+\ell} e^{-x}$ est égale à un polynôme de degré n , multiplié par $x^\ell e^{-x}$. Ce polynôme est :

$$Q_n^\ell(x) = \sum_{j=0}^n (-1)^j \binom{n}{j} \frac{(n+\ell)!}{(j+\ell)!} x^j\tag{XI.48}$$

Ainsi :

$$\frac{d^n x^{\ell+n} e^{-x}}{dx^n} = Q_n^\ell(x) x^\ell e^{-x}\tag{XI.49}$$

où les Q_n^ℓ sont les polynômes de Laguerre. Comme ils dépendent de deux indices, on peut fabriquer plusieurs familles de fonctions orthogonales à partir de ces polynômes, selon la manière de combiner les indices. En voici deux :

$$q_n^\ell(x) = x^{\ell/2} e^{-x/2} Q_n^\ell(x); \quad f_{n\ell}(x) = x^{\ell+1/2} e^{-x/2} Q_{n-\ell-1}^{2\ell+1}(x)\tag{XI.50}$$

C'est la seconde famille, $f_{n\ell}$, qui intervient dans le problème du champ coulombien ; $f_{n\ell}$ n'est définie que pour ℓ entier compris entre 0 et $n-1$.

(5) E. SCHRÖDINGER, *Quantisierung als Eigenwertproblem*, Annalen der Physik, 1926, vol. 80, p. 437. Traduction française de A. PROCA sous le titre *Mémoires sur la Mécanique ondulatoire*, rééditée en 1988 aux éditions Jacques Gabay.



Figure XI.5 – Fonctions $f_{n\ell}(x) = \sqrt{(\ell+1)/[(n-\ell-1)!(n+\ell)!]} x^\ell e^{-x/2} Q_{n-\ell-1}^{2\ell+1}(x)$. La fenêtre graphique correspond aux abscisses $x \in [0, 46]$ et aux ordonnées $y \in [-1; 1]$. Le coefficient numérique n'a pas d'autre signification que de faire correspondre tous les graphiques à ce cadre. Pour chaque n entier fixé, les n fonctions $f_{n\ell}(2r/n)$ pour lesquelles ℓ est compris entre 0 et $n-1$ représentent la partie radiale des états liés d'énergie $E_n = -m\alpha^2/2\hbar^2 n^2$ de l'électron de masse m dans un champ coulombien de potentiel $U = -\alpha/r$.

Exercice XI.1

1. montrer que pour tout ℓ fixé, les fonctions q_n^ℓ et q_m^ℓ sont orthogonales si $n \neq m$;
2. montrer que si $n \neq m$, les $f_{n\ell}$ sont orthogonales aux f_{mk} , indépendamment du doublet ℓ, k mais avec $0 \leq \ell \leq n-1$ et $0 \leq k \leq m-1$;
3. calculer les normes de ces fonctions. ■

XI.5 Théorèmes de Weierstrass

Dans cette section, on donne les démonstrations des théorèmes de Weierstrass utilisés à la section XI.3 pour établir que certains systèmes orthonormés sont des bases hilbertiennes. Les démonstrations que nous reproduisons ici sont celles, devenues classiques, de Bernstein ⁽⁶⁾.

Ces théorèmes sont essentiels puisqu'ils garantissent la possibilité de développer n'importe quelle fonction de carré intégrable en séries de fonctions orthonormées analogues aux séries de Fourier. Cela permet d'étendre la *méthode de Fourier* expliquée au chapitre VII à toutes sortes d'équations aux dérivées partielles, pourvu qu'on sache construire la base hilbertienne adéquate ; nous en verrons deux exemples de mécanique quantique : l'oscillateur harmonique et l'électron dans un champ coulombien.

Ces théorèmes de Weierstrass sont donc deux des sept piliers de l'analyse fonctionnelle.

Soit f une fonction continue sur l'intervalle $[0; 1]$, à valeurs réelles. On pose :

$$B_n(x) = \sum_{j=0}^n \binom{n}{j} f\left(\frac{j}{n}\right) x^j (1-x)^{n-j} \quad (\text{XI.51})$$

On appelle B_n , le n^{e} polynôme de Bernstein de la fonction f . On a choisi l'intervalle $[0; 1]$ pour alléger l'écriture, mais sur un intervalle $[a; b]$ quelconque, on aurait pris :

$$B_n(x) = \sum_{j=0}^n \binom{n}{j} f\left(a + (b-a)\frac{j}{n}\right) \left(\frac{x-a}{b-a}\right)^j \left(\frac{b-x}{b-a}\right)^{n-j} \quad (\text{XI.52})$$

Ⓡ Les polynômes de Bernstein ne dépendent pas précisément de la *fonction* f mais seulement des valeurs qu'elle prend aux points $a + (b-a)j/n$. Étant donné une famille discrète de points du plan, de coordonnées $(x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)$, tels que les x_j soient équidistants, on peut également définir leur polynôme de Bernstein par :

$$B_n(x) = \sum_{j=0}^n \binom{n}{j} y_j \left(\frac{x-x_0}{x_1-x_0}\right)^j \left(\frac{x_n-x}{x_1-x_0}\right)^{n-j} \quad (\text{XI.53})$$

Cette formule se prête particulièrement bien à une traduction algorithmique et cet avantage est fréquemment utilisé dans les logiciels graphiques pour tracer sur écran digital des courbes, dites *splines*, qui suivent de près une série de points donnés. On utilise aussi sa version bi-dimensionnelle en image de synthèse, pour dessiner des surfaces.

Théorème XI.2 Si f est continue sur $[0; 1]$, la suite (XI.51) des polynômes B_n converge uniformément sur $[0; 1]$ vers f . Plus généralement, si f est continue sur $[a; b]$, ce sera la suite (XI.52) des polynômes B_n qui convergera uniformément sur $[0; 1]$ vers f .

(6) Mémoires de l'Académie de Belgique, 1912.

Preuve La deuxième partie du théorème se ramène évidemment à la première par le changement de variable $x \mapsto (x - a)/(b - a)$. Rappelons la formule du binôme de Newton :

$$\sum_{j=0}^n \binom{n}{j} x^j y^{n-j} = (x + y)^n \quad (\text{XI.54})$$

En la dérivant par rapport à x on obtient les identités suivantes :

$$\sum_{j=0}^n \binom{n}{j} j x^j y^{n-j} = n x (x + y)^{n-1} \quad (\text{XI.55})$$

$$\sum_{j=0}^n \binom{n}{j} j(j-1) x^j y^{n-j} = n(n-1) x^2 (x + y)^{n-2}$$

En prenant $y = 1 - x$ dans ces identités, on obtient :

$$\sum_{j=0}^n \binom{n}{j} x^j (1-x)^{n-j} = 1 \quad (\text{XI.56a})$$

$$\sum_{j=0}^n \binom{n}{j} j x^j (1-x)^{n-j} = n x \quad (\text{XI.56b})$$

$$\sum_{j=0}^n \binom{n}{j} j(j-1) x^j (1-x)^{n-j} = n(n-1) x^2 \quad (\text{XI.56c})$$

Puis, en les combinant :

$$\begin{aligned} \sum_{j=0}^n \binom{n}{j} (j - nx)^2 x^j (1-x)^{n-j} &= \sum_{j=0}^n \binom{n}{j} (n^2 x^2 - 2jnx + j^2) x^j (1-x)^{n-j} \\ &= n^2 x^2 \cdot 1 - 2nx \cdot nx + n(n-1) x^2 = nx(1-x) \end{aligned} \quad (\text{XI.57})$$

Grâce à (XI.56a), on peut écrire :

$$|f(x) - B_n(x)| = \left| \sum_{j=0}^n \binom{n}{j} \left[f(x) - f\left(\frac{j}{n}\right) \right] x^j (1-x)^{n-j} \right| \quad (\text{XI.58})$$

Dans le membre de droite ci-dessus, on va distinguer les termes pour lesquels j/n est proche de x . Soit donc $\alpha > 0$. En appliquant l'inégalité du triangle dans le second membre de (XI.58) :

$$|f(x) - B_n(x)| \leq \sum_{|x-j/n| \leq \alpha} \binom{n}{j} \left| f(x) - f\left(\frac{j}{n}\right) \right| x^j (1-x)^{n-j} + \sum_{|x-j/n| > \alpha} \dots \quad (\text{XI.59})$$

Pour la première partie (somme pour $|x - j/n| \leq \alpha$) on peut, ε étant donné, choisir α de sorte que $|f(x) - f(j/n)| \leq \varepsilon$: cela provient du fait que f est par hypothèse continue et donc aussi *uniformément* continue sur $[0; 1]$. La somme pour $|x - j/n| \leq \alpha$ est donc majorée par ε , multiplié par la somme des $\binom{n}{j} x^j (1-x)^{n-j}$ (somme de termes tous ≥ 0) qui d'après (XI.56a) est ≤ 1 . La deuxième partie de (XI.59) se majore en utilisant (XI.57). Soit M le maximum de f sur $[0; 1]$. On a $|f(x) - f(j/n)| \leq 2M$, d'où :

$$\left| \sum_{|x-j/n| > \alpha} \binom{n}{j} \left[f(x) - f\left(\frac{j}{n}\right) \right] x^j (1-x)^{n-j} \right| \leq 2M \sum_{|x-j/n| > \alpha} \binom{n}{j} x^j (1-x)^{n-j} \quad (\text{XI.60})$$

De plus, l'inégalité $1 \leq (x - j/n)^2 / \alpha^2 = (j - nx)^2 / n^2 \alpha^2$ est partout vraie sur le domaine de sommation :

$$\sum_{|x-j/n| > \alpha} \binom{n}{j} x^j (1-x)^{n-j} \leq \frac{1}{n^2 \alpha^2} \sum_{j=0}^n (j - nx)^2 \binom{n}{j} x^j (1-x)^{n-j} \quad (\text{XI.61})$$

D'après (XI.57) cette dernière expression est égale à $x(1-x)/n\alpha^2$; et d'autre part le maximum de $x(1-x)$ sur $[0; 1]$ est $1/4$, de sorte qu'en définitive la deuxième partie de (XI.59) est majorée par $M/2n\alpha^2$. En regroupant tous ces résultats :

$$\forall \varepsilon > 0, \quad \exists \alpha, \quad |f(x) - B_n(x)| \leq \varepsilon + \frac{M}{2n\alpha^2} \quad (\text{XI.62})$$

Pour conclure, il faut encore un argument. On ne peut pas prendre simplement $\varepsilon = \frac{1}{n}$. En effet, si la fonction était par exemple $f(x) = x$, on aurait $\alpha = \varepsilon = 1/n$ et donc le rapport $M/2n\alpha^2$ ne tendrait pas vers zéro, mais vers l'infini. Il faut donc, pour une fonction continue donnée f , choisir ε en fonction de n , mais de telle sorte que le α correspondant ne tende pas trop vite vers zéro. D'autre part, on ne peut pas non plus garder α constant, car alors ce serait ε qui ne tendrait pas vers zéro. Il faut donc choisir le meilleur compromis. Par exemple si $f(x) = \sqrt{x}$, un tel compromis sera $\varepsilon = \sqrt{\alpha} = n^{-r}$ avec $0 < r < \frac{1}{4}$: en effet, il faut que ε tende vers zéro, donc que $r > 0$, mais aussi que $n\alpha^2 = n^{4r+1}$ tende vers l'infini, donc $r < \frac{1}{4}$. Si $f(x) = e^{-1/x^2}$ (prolongée en $x = 0$ par $f(0) = 0$), un bon compromis est $\varepsilon = e^{-\sqrt{n}}$; quand n tend vers l'infini, cela tend bien vers zéro et $\alpha^2 = 1/\ln(\frac{1}{\varepsilon}) = 1/\sqrt{n}$, donc $n\alpha^2 = \sqrt{n}$ tend bien vers l'infini. Ce compromis existe pour n'importe quelle fonction continue, puisque $\alpha \rightarrow 0 \Rightarrow \varepsilon \rightarrow 0$; il suffit donc de prendre $\alpha = 1/\sqrt{n}$ pour qu'obligatoirement $\varepsilon = \sup_x \sup_{|h| < \alpha} f(x+h)$ tende lui aussi vers zéro. •

Ce théorème s'applique sur les intervalles bornés, ce qui est logique puisque sa démonstration utilise de manière essentielle le fait que la fonction f est *uniformément* continue. On peut donc en déduire directement que toute fonction continue sur $[-1; 1]$ est limite uniforme d'une suite de polynômes, puis, comme cela a été fait à la section XI.3, en déduire que les polynômes sont denses dans $\mathbb{C}_0([-1; 1])$ pour la métrique euclidienne du produit scalaire $\langle f | g \rangle = \int f(x)g(x)dx$. Par contre, on ne peut pas appliquer *directement* le théorème sur $\mathbb{C}_0(-\infty; \infty)$ ou sur $\mathbb{C}_0([0; \infty[)$, car ces espaces regroupent les fonctions continues sur un intervalle non borné. Or, pour montrer que les fonctions d'Hermite forment une base hilbertienne de $\mathbb{L}_2(-\infty; \infty)$, ou que les fonctions de Laguerre forment une base hilbertienne de $\mathbb{L}_2([0; \infty[)$, il faut vérifier que les combinaisons linéaires de ces fonctions sont denses dans $\mathbb{C}_0(-\infty; \infty)$ et $\mathbb{C}_0([0; \infty[)$ pour la métrique euclidienne.

Théorème XI.3 Toute fonction continue de $\mathbb{L}_2(-\infty; \infty)$ est la limite dans cet espace d'une suite de fonctions de la forme polynôme $\times e^{-x^2/2}$.

Toute fonction continue de $\mathbb{L}_2([0; \infty[)$ est la limite dans cet espace d'une suite de fonctions de la forme polynôme $\times e^{-x}$.

Mais elle n'est pas la limite uniforme d'une telle suite !

Preuve On la donne pour $\mathbb{L}_2(-\infty; \infty)$ et on la laisse en exercice pour $\mathbb{L}_2([0; \infty[)$. On ne peut utiliser le théorème de Weierstrass sur $]-\infty; \infty[$ mais on peut l'utiliser sur $[-A; A]$ et faire tendre ensuite A vers l'infini. Soit donc f une fonction de carré intégrable et continue. La fonction $g(x) = e^{+x^2/2} f(x)$ est alors elle aussi continue sur $[-A; A]$ et par conséquent elle est la limite uniforme d'une certaine suite B_n de polynômes. Ce qui entraîne que :

$$\forall \varepsilon, \quad \exists n(A, \varepsilon), \quad \forall x \in [-A; A], \quad |g(x) - B_n(x)| \leq \varepsilon \quad (\text{XI.63})$$

et par conséquent aussi :

$$\forall \varepsilon, \quad \exists n(A, \varepsilon), \quad \forall x \in [-A; A], \quad |f(x) - B_n(x)e^{-x^2/2}| \leq \varepsilon e^{-x^2/2} \quad (\text{XI.64})$$

Par ailleurs, pour tout ε et tout n on peut choisir A tel que :

$$\int_{x \geq A} |f(x) - B_n(x)e^{-x^2/2}|^2 dx \leq \varepsilon \quad (\text{XI.65})$$

puisque $f(x)$ et $B_n(x)e^{-x^2/2}$ sont des fonctions de carré intégrable. En choisissant A de la sorte, on aura :

$$\int_{-\infty}^{+\infty} |f(x) - B_n(x)e^{-x^2/2}|^2 dx \leq \int_{-A}^A \varepsilon^2 e^{-x^2} dx + \varepsilon \leq \varepsilon^2 \sqrt{\pi} + \varepsilon \quad (\text{XI.66})$$

Théorème XI.4 Si f est une fonction continue périodique de période 2π , alors il existe une suite T_n de polynômes trigonométriques qui converge, uniformément sur tout $\mathbb{R}$, vers f .

Ⓡ Il est équivalent de dire que T_n converge uniformément sur tout $\mathbb{R}$, ou sur une seule période, par exemple $[0; 2\pi]$ ou $[-\pi; \pi]$. Par contre, il n'est pas équivalent de dire que f est une fonction continue périodique, ou que f est une fonction continue sur $[0; 2\pi]$ ou sur $[-\pi; \pi]$. En effet, f pourrait être continue sur $[0; 2\pi]$, mais prendre des valeurs différentes en 0 et en 2π ; en ce cas, elle ne pourrait pas se prolonger en une fonction continue et périodique, ni être la limite uniforme d'une suite de polynômes trigonométriques.

Preuve Le théorème XI.4 est un simple corollaire du théorème XI.2. Supposons d'abord que f est paire; on peut la considérer seulement sur l'intervalle $[0; \pi]$: sur $[-\pi; 0]$, elle s'en déduit par symétrie et au-delà, par périodicité. La fonction $g(y) = f(\arccos(y))$ est alors une fonction continue sur $[-1; 1]$, à laquelle on peut appliquer le théorème XI.2. Donc pour tout ε , il existe un polynôme $P(y)$ tel que

$$\forall y \in [-1; 1], \quad |g(y) - P(y)| \leq \varepsilon \quad (\text{XI.67})$$

si cela est vrai $\forall y \in [-1; 1]$, c'est vrai aussi en remplaçant y par $\cos x$, x pouvant être *n'importe* quel nombre réel :

$$\forall x \in \mathbb{R}, \quad |g(\cos x) - P(\cos x)| \leq \varepsilon \quad (\text{XI.68})$$

Pour $x \in [0; \pi]$ on a $g(\cos x) = f(x)$ par définition de g ; mais puisque f est paire cela est encore vrai pour $x \in [-\pi; \pi]$ et puisque f est périodique, c'est encore vrai pour tout $x \in \mathbb{R}$.

Notons maintenant que $P(\cos x)$ est un polynôme algébrique de $\cos x$, mais on sait que toute puissance de $\cos x$, disons $\cos^n x$, est combinaison linéaire des fonctions 1, $\cos x$, $\cos 2x$, ..., $\cos nx$. Donc $P(\cos x)$ peut s'écrire sous forme d'une combinaison linéaire de ces fonctions.

Si f est impaire, on va montrer que pour tout $\varepsilon > 0$, il existe un polynôme algébrique $Q(y)$ tel que $\forall x \in [0; 2\pi]$, $|f(x) - \sin x Q(\cos x)| \leq \varepsilon$. Les expressions de la forme $\sin x \cos^{n-1} x$ pouvant toutes s'écrire comme une combinaison linéaire des fonctions $\sin x$, $\sin 2x$, $\sin 3x$, ..., $\sin nx$, on aura ainsi prouvé le théorème également pour les fonctions impaires. Si f est impaire, introduisons à nouveau la fonction $g(y) = f(\arccos(y))$. Puisque $f(x)$ est impaire, on a $f(0) = 0$; on doit aussi avoir $f(-\pi) = f(\pi)$ puisque f doit être continue et périodique; mais en outre $f(-\pi) = -f(\pi)$ puisque f est impaire; cela implique que $f(-\pi) = f(\pi) = 0$. Donc $g(y)$ est nul pour $y = 1$ et pour $y = -1$. Le polynôme de Bernstein qui sert à notre approximation de $g(y)$ est donné par la formule (XI.52), avec $a = -1$ et $b = 1$, ce qui donne ici :

$$B_n(y) = \sum_{j=0}^n \binom{n}{j} g\left(-1 + \frac{2j}{n}\right) \left(\frac{y+1}{2}\right)^j \left(\frac{1-y}{2}\right)^{n-j} \quad (\text{XI.69})$$

Du fait que $g(\pm 1) = 0$, on voit qu'il est nul également pour $y = +1$ et pour $y = -1$; on voit même sur cette expression qu'il est factorisable par $(1-y)(1+y) = (1-y^2)$: en effet chaque terme de la somme (XI.69) contient en facteur $g(-1 + 2j/n)(y+1)^j(1-y)^{n-j}$; si $0 < j < n$, ceci est divisible par $(y+1)(1-y)$ et si $j = 0$ ou $j = n$, c'est le coefficient $g(-1 + \frac{2j}{n})$ qui est nul. Ainsi, le polynôme de Bernstein B_n de notre fonction impaire est divisible par $1 - y^2$, donc on peut le factoriser sous la forme $(1 - y^2) A_n(y)$. Puisque le polynôme B_n est de degré n , le polynôme A_n sera de degré $n - 2$.

Ainsi les polynômes $B_n = (1 - y^2) A_n(y)$ convergent uniformément vers $g(y)$ sur $[-1; 1]$, ce qui implique que $\sin^2 x A_n(\cos x)$ va converger uniformément vers $g(\cos x)$ sur tout $\mathbb{R}$; pour n assez grand on aura :

$$\forall x \in \mathbb{R}, \quad |g(\cos x) - \sin^2 x A_n(\cos x)| \leq \frac{1}{2} \varepsilon \quad (\text{XI.70})$$

Cependant, $g(\cos x)$ est une fonction paire et n'est égale à $f(x)$ que sur les intervalles $[0; \pi]$, $[2\pi; 3\pi]$ etc. Dans $[-\pi; \pi]$, on peut donc seulement en déduire que $\forall x \in [0; \pi]$, $|f(x) - \sin^2 x A_n(\cos x)| \leq \frac{1}{2} \varepsilon$.

Notons cependant que la fonction $h(y) = \sqrt{1 - y^2} A_n(y)$ est continue; ce n'est pas un polynôme à cause du facteur $\sqrt{1 - y^2}$, mais c'est une fonction continue. Par conséquent on peut elle-même l'approcher

uniformément par un polynôme de Bernstein ; de sorte qu'il existera un polynôme $Q(y)$ tel que $\forall y \in [-1; 1]$, $|h(y) - Q(y)| \leq \frac{1}{2}\varepsilon$, ou, ce qui est équivalent, $\forall x \in [0; \pi]$, $|h(\cos x) - Q(\cos x)| \leq \frac{1}{2}\varepsilon$. Pour $x \in [0; \pi]$, $\sin x = \sqrt{1 - y^2}$, donc $\sin^2 x A_n(\cos x) = \sin x h(\cos x)$. Ce qui, puisque $|\sin x| \leq 1$, entraîne que $\forall x \in [0; \pi]$, $|\sin^2 x A_n(\cos x) - \sin x Q(\cos x)| \leq \frac{1}{2}\varepsilon$. Or, d'après ce que nous avons vu avant, $|g(\cos x) - \sin^2 x A_n(\cos x)|$ est aussi inférieur à $\frac{1}{2}\varepsilon$ pour tout $x \in [0; \pi]$. L'inégalité triangulaire conduit alors à la conclusion que :

$$\forall x \in [0; \pi], \quad |g(\cos x) - \sin x Q(\cos x)| = |f(x) - \sin x Q(\cos x)| \leq \frac{1}{2}\varepsilon \quad (\text{XI.71})$$

Cette fois on a approché $g(\cos x)$ par un polynôme trigonométrique impair, de sorte que cette inégalité est vraie non seulement pour $x \in [0; \pi]$, mais aussi pour $x \in [-\pi; \pi]$.

Enfin, si f est quelconque, elle se décompose en somme d'une fonction paire et d'une fonction impaire, à chacune desquelles on peut appliquer ce qui précède. ●



XII — Opérateurs

XII.1 Dimension infinie

Un *opérateur* sur un espace vectoriel est une transformation linéaire de cet espace dans lui-même. Si l'espace est de dimension finie, on retrouve la notion d'endomorphisme. En fixant une base, l'opérateur est entièrement défini par sa matrice dans cette base.

Dans un espace vectoriel de dimension infinie, les opérateurs sont évidemment plus difficiles à manipuler. De même que pour les fonctionnelles linéaires, un opérateur linéaire n'est pas forcément continu. Nous ferons une place à part aux opérateurs continus, mais, ayant en vue la mécanique quantique, nous ne pourrions pas éviter de considérer aussi des opérateurs discontinus, car les plus importants (impulsion, moment cinétique, hamiltonien) sont malheureusement discontinus.

Parmi les espaces de dimension infinie, ceux qui se rapprochent le plus des espaces de dimension finie sont les espaces de Hilbert. En mécanique quantique, les fonctions d'onde (ou vecteurs d'état) sont justement des vecteurs évoluant dans des espaces du type $\mathbb{L}_2$, c'est-à-dire des espaces de Hilbert. Comme nous avons en vue la mécanique quantique, nous éviterons d'envisager les opérateurs sur d'autres espaces que les espaces de Hilbert. La perte de généralité qui en résulte est sans conséquence pour des physiciens et aura l'avantage d'éviter bon nombre de complications.

XII.2 Opérateurs continus et opérateurs discontinus fermés

Dans tout le chapitre, on appellera $\mathcal{H}$, l'espace de Hilbert, $\langle | \rangle$, le produit scalaire et $\| \cdot \|$, la norme. On ne précisera la nature de $\mathcal{H}$ que dans les exemples.

La continuité d'un opérateur linéaire se réduit à la continuité en 0. En effet, si l'opérateur $A : f \mapsto Af$ est continu à l'origine de l'espace $\mathcal{H}$, on aura par définition de la continuité :

$$\forall \varepsilon > 0, \quad \exists \eta > 0, \quad \|f\| < \eta \quad \Rightarrow \quad \|Af\| < \varepsilon \quad (\text{XII.1})$$

Si on remplace f par $f - f_0$ cela donne, puisque $A(f - f_0) = Af - Af_0$:

$$\forall \varepsilon > 0, \quad \exists \eta > 0, \quad \|f - f_0\| < \eta \quad \Rightarrow \quad \|Af - Af_0\| < \varepsilon \quad (\text{XII.2})$$

ce qui est l'expression de la continuité au *point* f_0 .

La continuité d'un opérateur linéaire se traduit aussi par l'existence d'une constante $M > 0$ telle que :

$$\forall f \in \mathcal{H}, \quad \|Af\| \leq M\|f\| \quad (\text{XII.3})$$

Prenons en effet $\varepsilon = 1$ dans (XII.1), on obtient alors $\exists \eta > 0, \|f\| < \eta \Rightarrow \|Af\| < 1$, ce qui, puisque $A(\eta g) = \eta Ag$, équivaut à :

$$\exists \eta > 0, \quad \|f\| < 1 \quad \Rightarrow \quad \|Af\| < \frac{1}{\eta} \quad (\text{XII.4})$$

Mais pour *tout* $f \in \mathcal{H}$, on peut dire en posant $h = (1/\|f\|)f$, que $\|h\| \leq 1$ d'où :

$$\|Ah\| < \frac{1}{\eta} \quad \text{c'est-à-dire} \quad \|Af\| < \frac{1}{\eta}\|f\| \quad (\text{XII.5})$$

ce qui est bien (XII.3) en prenant $M = \frac{1}{\eta}$. On exprime souvent l'inégalité (XII.3) en disant que l'opérateur A est *borné* ; c'est pourquoi on trouve souvent dans la littérature l'expression *opérateur borné*, pour dire *opérateur continu*.

La plus petite possible de ces constantes M s'appelle la *norme* de l'opérateur A .

Un exemple d'opérateur continu est la *transformation de Fourier* sur $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$. Cet opérateur est continu du fait de la formule de Plancherel :

$$\|\hat{f}\| = \sqrt{2\pi} \|f\| \quad (\text{XII.6})$$

l'inégalité (XII.3) est donc vérifiée avec $M = \sqrt{2\pi}$.

Un autre exemple d'usage très courant est l'opérateur de projection sur un sous-espace fermé. Un tel opérateur vérifie par nature l'inégalité $\|Pf\| \leq \|f\|$. En voici deux cas :

- dans $\mathcal{H} = \mathbb{L}_2([-1; 1])$, la projection P_N sur le sous-espace de dimension finie $\mathcal{T}_N$ des polynômes trigonométriques de degré $\leq N$;
- dans $\mathcal{H} = \mathbb{L}_2([-1; 1])$ ou $\mathcal{H} = \mathbb{L}_2(]-\infty; \infty[) = \mathbb{L}_2(\mathbb{R})$, la projection P_+ sur le sous-espace de dimension infinie (mais fermé) $\mathcal{H}_+$ des fonctions paires.

On rencontre très souvent aussi l'*opérateur de translation* T_a ; si $f \in \mathbb{L}_2(\mathbb{R})$, la fonction $T_a f$ est $f(x - a)$. On voit que $\|T_a f\| = \|f\|$, puisque par un changement de variable, on a :

$$\int_{-\infty}^{+\infty} |f(x - a)|^2 dx = \int_{-\infty}^{+\infty} |f(x)|^2 dx \quad (\text{XII.7})$$

Toutefois les opérateurs les plus importants de la mécanique quantique sont discontinus. Voici les principaux :

multiplication par x — c'est l'opérateur qui à une fonction $f(x)$, associe la fonction $xf(x)$.

On l'appelle aussi *opérateur de position* ;

dérivation — c'est l'opérateur qui à une fonction f , associe sa dérivée df/dx . L'opérateur d'impulsion est $i\hbar d/dx$;

multiplication par une fonction quelconque $V(x)$ — c'est l'opérateur qui, à une fonction $f(x)$, associe la fonction $V(x)f(x)$; on a aussi l'opérateur d^n/dx^n ; ainsi l'opérateur *hamiltonien* est :

$$Hf = -\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + V \cdot f \quad (\text{XII.8})$$

Non seulement ces opérateurs sont discontinus, mais ils ne sont même pas partout définis : sur l'espace $\mathbb{L}_2(\mathbb{R})$, il y a des fonctions non dérivables, pour lesquelles on ne peut pas définir l'opérateur de dérivation ; il y a aussi des fonctions qu'on peut certes multiplier par x ou par $V(x)$, mais sans que la fonction produit soit elle aussi dans $\mathbb{L}_2(\mathbb{R})$, de sorte que l'opérateur de multiplication ne peut pas être défini en tant qu'opérateur de $\mathbb{L}_2(\mathbb{R})$ dans lui-même.

Pour de tels opérateurs, on prendra en compte leur domaine de définition : au lieu d'avoir un opérateur A qui transforme *toute* fonction de l'espace de Hilbert $\mathcal{H}$ en une autre fonction de $\mathcal{H}$, on aura un opérateur A défini sur un sous-espace vectoriel $\mathcal{D}_A$ de l'espace de Hilbert $\mathcal{H}$, qui transforme toute fonction de $\mathcal{D}_A$ en une fonction de $\mathcal{H}$.

Ⓡ On pourrait aussi convenir que l'opérateur A serait défini sur l'espace de Hilbert $\mathcal{H}$ tout entier, mais qu'en revanche l'espace image serait plus gros que $\mathcal{H}$: ainsi l'opérateur de multiplication par x , au lieu de n'être défini que sur le sous-espace $\mathcal{D}_x$ des fonctions $f \in \mathbb{L}_2(\mathbb{R})$ telles que $xf \in \mathbb{L}_2(\mathbb{R})$, serait défini sur $\mathbb{L}_2(\mathbb{R})$, mais l'espace-image serait un espace plus grand que $\mathbb{L}_2(\mathbb{R})$, contenant aussi les fonctions $f(x)$ telles que $f(x)/x$ soit de carré intégrable. La pratique montre que c'est une mauvaise solution : les produits scalaires et autres commodités des espaces de Hilbert (bases etc.) sont utilisables sur les sous-espaces $\mathcal{D}_A$, mais ne s'étendent pas forcément à des espaces plus gros.

Ⓡ Comme déjà remarqué pour les fonctionnelles linéaires, la continuité dépend de la métrique. Nous ne considérons ici que des opérateurs sur des espaces de Hilbert, où la métrique est fixée par le produit scalaire. Il est facile de vérifier que par exemple l'opérateur de multiplication par x , défini sur $\mathcal{D}_x$, est discontinu. En effet, les fonctions :

$$f_n(x) = \frac{1}{\sqrt{n}} e^{-x^2/2n} \quad (\text{XII.9})$$

ont pour norme $(\pi/n^3)^{1/4}$, donc la suite f_n tend vers zéro dans $\mathbb{L}_2(\mathbb{R})$; d'autre part les fonctions $xf_n(x)$ sont de carré intégrable (et donc, les f_n sont toutes dans $\mathcal{D}_x$), de norme $(n\pi)^{1/4}/\sqrt{2}$ qui tend vers l'infini, ce qui prouve ainsi la discontinuité de l'opérateur. Si on avait pris sur $\mathcal{D}_x$ la métrique :

$$\|f\| = \sqrt{\|f\|^2 + \|xf\|^2} \quad (\text{XII.10})$$

tout en conservant la métrique hilbertienne sur l'espace image, le même opérateur serait alors continu : en effet, avec *n'importe quelle* suite f_n de $\mathcal{D}_x$, le fait que $\|f_n\|$ tend vers zéro implique automatiquement que $\|xf_n\|_2$ tend vers zéro. Toutefois, la nouvelle norme $\|f\|$ dépend de l'opérateur, est fabriquée *ad hoc*, ne dérive d'aucun produit scalaire et (pour ce qui concerne la mécanique quantique) ne correspond à rien de physique. Il s'agit d'un exemple scolaire qui sert uniquement ici à rappeler que la discontinuité n'est pas intrinsèque et dépend de la métrique.

Pour de tels opérateurs, non partout définis et discontinus, on introduit une nouvelle notion : celle d'opérateur *fermé*.

Définition XII.1 Soit A un opérateur sur l'espace de Hilbert $\mathcal{H}$, ayant pour domaine de définition le sous-espace vectoriel dense ⁽¹⁾ $\mathcal{D}_A$. On dit que A est fermé si pour toute suite f_n dans $\mathcal{D}_A$, telle que f_n et Af_n soient toutes deux convergentes, on a $\mathcal{H}\text{-lim } f_n \in \mathcal{D}_A$ et $\mathcal{H}\text{-lim } Af_n = A \mathcal{H}\text{-lim } f_n$.

On dit que A est préfermé s'il est défini sur un sous-espace dense $\mathcal{D}$ de $\mathcal{H}$ et si pour toute suite f_n de $\mathcal{D}$ qui tend (dans $\mathcal{H}$) vers zéro et telle que, la suite Af_n soit convergente dans $\mathcal{H}$, on ait nécessairement $\mathcal{H}\text{-lim } Af_n = 0$.

On vérifie en une ligne qu'un opérateur continu sur $\mathcal{H}$ est forcément fermé et qu'un opérateur continu sur un sous-espace dense $\mathcal{D}$ de $\mathcal{H}$ est forcément préfermé. L'intérêt de cette nouvelle notion est de donner une possibilité de prolongement pour les opérateurs discontinus, analogue au prolongement par continuité. Imaginons en effet un opérateur qui serait défini *a priori* sur un sous-espace dense de $\mathcal{H}$ et qui serait continu : un tel opérateur se prolonge alors par continuité à $\mathcal{H}$ tout entier (tout comme une fonction continue sur $\mathbb{Q}$ se prolonge par continuité à $\mathbb{R}$).

De la même façon, imaginons un opérateur A qui serait discontinu et défini sur un sous-espace dense $\mathcal{D}$. Si l'opérateur A est préfermé, alors on peut le prolonger à un domaine plus grand $\mathcal{D}_A$, qui est défini comme étant l'ensemble des éléments de $\mathcal{H}$ qui sont des limites de suites f_n de $\mathcal{D}$ telles que f_n et Af_n soient toutes deux convergentes. Si $f \in \mathcal{D}_A$, il existe une suite f_n de $\mathcal{D}$ qui tend vers f et telle que Af_n tende elle aussi vers une limite g : on prolonge alors A en posant $Af = g$. La définition XII.1 garantit l'unicité et donc la cohérence de ce prolongement : si on avait pris une autre suite $f_n^{(1)}$ qui tend vers f , telle que $Af_n^{(1)}$ tende vers $g^{(1)}$, on aurait $g^{(1)} = g$.

Contrairement au prolongement par continuité, qui permet de prolonger l'opérateur du domaine dense $\mathcal{D}$ à $\mathcal{H}$ tout entier, le prolongement par fermeture permet de prolonger l'opérateur du domaine dense $\mathcal{D}$ à un domaine plus grand (et donc dense aussi) $\mathcal{D}_A$, mais qui *n'est pas* $\mathcal{H}$ tout entier. Ce domaine étendu $\mathcal{D}_A$ est alors le domaine de définition maximal de l'opérateur ; on peut montrer que si ce domaine maximal est $\mathcal{H}$ tout entier, l'opérateur est continu. Autrement dit :

Théorème XII.1 — Théorème de Banach Un opérateur fermé dont le domaine est $\mathcal{H}$ tout entier est continu.

L'adhérence $\overline{K}$ d'un sous-ensemble K de $\mathcal{H}$ est définie comme l'ensemble des éléments de $\mathcal{H}$ qui sont des limites de suites convergentes f_n incluses dans K . Un prolongement par continuité s'effectue de K à $\overline{K}$; si $K = \mathcal{D}$ qui est dense, l'adhérence est $\mathcal{H}$ et si A est continu, il se prolonge donc de $\mathcal{D}$ à $\mathcal{H}$.

Cependant, $\mathcal{D}_A$ n'est pas l'ensemble des limites de *n'importe quelle* suite convergente de $\mathcal{D}$; on ne considère que les suites f_n pour lesquelles Af_n est elle aussi convergente : il y a donc moins de suites admises et c'est pourquoi $\mathcal{D}_A$ n'a en général aucune raison d'être $\mathcal{H}$

(1) La plupart des auteurs omettent cette restriction ; toutefois, les opérateurs qu'on rencontre en mécanique quantique ont toujours un domaine dense, de sorte que la considération de domaines non denses ne ferait que nous compliquer la vie, déjà assez difficile comme cela.

tout entier ⁽²⁾.

On comprendra encore mieux le rapprochement entre l'adhérence $\overline{\mathcal{D}}$ et l'extension $\mathcal{D}_A$ considérant à nouveau la norme $\|f\| = \sqrt{\|f\|^2 + \|Af\|^2}$. Une suite f_n de $\mathcal{D}$ qui converge dans $\mathcal{H}$, ainsi que Af_n , est une suite intrinsèquement convergente (ou suite de Cauchy) pour la norme $\|\cdot\|$. Ainsi l'adhérence $\overline{\mathcal{D}}$ de $\mathcal{D}$ est le complété du sous-espace $\mathcal{D}$ pour la norme $\|\cdot\|$, tandis que l'extension $\mathcal{D}_A$ de $\mathcal{D}$ est le complété du sous-espace $\mathcal{D}$ pour la norme $\|\cdot\|$. $\overline{\mathcal{D}} = \mathcal{H}$ est complet pour la norme $\|\cdot\|$, tandis que $\mathcal{D}_A$ est complet pour la norme $\|\cdot\|$.

On peut encore dire que l'ensemble $\mathcal{G}$ des *points* (f, Af) lorsque f parcourt $\mathcal{D}$ est un sous-espace vectoriel de l'espace de Hilbert $\mathcal{H} \times \mathcal{H}$ ⁽³⁾ : c'est le graphe de l'opérateur A . Dire que A est fermé équivaut à dire que le graphe de A est un sous-espace vectoriel fermé de $\mathcal{H} \times \mathcal{H}$. Dire que A est préfermé équivaut à dire que l'adhérence $\overline{\mathcal{G}}$ du graphe $\mathcal{G}$ dans $\mathcal{H} \times \mathcal{H}$ est un graphe d'opérateur : l'opérateur A est discontinu en tant que transformation linéaire de $\mathcal{H}$ dans $\mathcal{H}$, mais continu en tant que transformation linéaire de $\mathcal{G}$ dans $\mathcal{H}$; ainsi le prolongement par fermeture est un prolongement par continuité déguisé.

Pour résumer, on peut dire que toute $f \in \mathcal{D}_A$ est la limite d'une suite f_n telle que :

- $\forall n, f_n \in \mathcal{D}$;
- $\exists g \in \mathcal{H}, g = \mathcal{H}\text{-lim } Af_n$;

pour prolonger A à $\mathcal{D}_A$ on pose $Af = g$; sur le domaine étendu $\mathcal{D}_A$, A sera un opérateur fermé.

La notion d'opérateur préfermé est utile pour la raison suivante : comme cela a été dit plus haut, on ne peut malheureusement pas se restreindre à l'étude des seuls opérateurs continus, car ceux de la mécanique quantique sont discontinus. Par contre, les opérateurs de la mécanique quantique sont préfermés et même fermés si on convient de toujours considérer leur prolongement maximal. C'est ce que nous ferons dans la suite : les opérateurs (dérivation etc.) seront définis *a priori* sur des domaines non nécessairement maximaux, mais (après avoir dûment vérifié qu'ils sont préfermés) on les supposera toujours implicitement prolongés à leur domaine maximal.

Voici une série d'exercices ayant pour but de montrer que les opérateurs de dérivation et de multiplication par une fonction sont bien des opérateurs fermés ou du moins préfermés, ainsi que de préciser le domaine maximal sur lequel on peut les prolonger pour avoir un opérateur fermé.

Exercice XII.1 On pose $f_n(x) = e^{-nx^2/2}$; ces fonctions forment une suite sur l'espace $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$.

1. Calculer leurs normes et en déduire que la suite f_n tend vers zéro dans $\mathcal{H}$.
2. Calculer les normes des dérivées f'_n , constater que la suite f'_n ne tend pas vers zéro dans $\mathcal{H}$ et en conclure que l'opérateur $f \mapsto f'$ est discontinu. ■

Exercice XII.2 Soit $V(x)$ une fonction que l'on supposera définie et continue sur $]-\infty; \infty[$.

1. On définit le domaine $\mathcal{D}_V = \{f \in \mathcal{H} \mid V \cdot f \in \mathcal{H}\}$; vérifier que c'est un sous-espace vectoriel dense de $\mathcal{H}$ (indication : pour $f \in \mathcal{H}$ considérer la suite f_n des fonctions définies par $f_n(x) = f(x)$ si $|x| \leq n$ et $f_n(x) = 0$ si $|x| > n$).

(2) Lorsque cela arrive, c'est que l'opérateur A est continu, d'après le théorème XII.1.

(3) L'espace de Hilbert $\mathcal{H} \times \mathcal{H}$ est muni du produit scalaire $\langle (f_1, g_1), (f_2, g_2) \rangle = \langle f_1, f_2 \rangle + \langle g_1, g_2 \rangle$.

2. Sur le domaine $\mathcal{D}_V$ introduit en a) on définit l'opérateur A_V (multiplication par V) : $A_V f$ est la fonction $V(x)f(x)$. Montrer que A_V est fermé. Indication : faire jouer l'inégalité de la moyenne dans les intégrales sur le domaine $|x| \leq M$ en utilisant le fait que (pour tout M fini) V doit avoir un maximum sur l'intervalle $\{|x| \leq M\}$. ■

Exercice XII.3 On étudie maintenant l'opérateur de dérivation.

1. On prend le domaine $\mathcal{D}$ égal à l'ensemble des fonctions de $\mathcal{H}$ qui sont continûment dérivables et dont la dérivée est dans $\mathcal{H}$ (ce domaine n'est pas maximal). Vérifier que $\mathcal{D}$ est un sous-espace vectoriel dense dans $\mathcal{H}$;
2. Soit f_n une suite de fonctions continûment dérivables, qui tend vers 0 dans $\mathcal{H}$ et telle que la suite des dérivées f'_n ait une limite g (il s'agira de montrer que $g = 0$). Montrer que la fonction :

$$G(x) = \int_0^x g(t) dt \quad (\text{XII.11})$$

est bien définie et que pour tout x on a :

$$|G(x) - f_n(x)| \leq |f_n(0)| + \sqrt{|x|} \|g - f'_n\| \quad (\text{XII.12})$$

(utiliser l'inégalité de Schwarz dans l'intégrale $\int g(t) - f'_n(t) dt$). Montrer aussi que $f_n(0)$ tend vers zéro (utiliser le fait que les f_n , qui tendent vers zéro dans $\mathcal{H}$, sont dérivables et leur appliquer habilement la formule des accroissements finis) ;

3. Prouver (en utilisant b) que pour tout $M > 0$ on a l'inégalité :

$$\int_{-M}^M |G(x) - f_n(x)|^2 dx \leq M |f_n(0)| + M^2 \|g - f'_n\|^2 \quad (\text{XII.13})$$

et en déduire que $G = 0$;

4. Conclure. Caractériser le domaine maximal de l'opérateur. ■

XII.3 Valeurs propres et spectre d'un opérateur

En dimension finie, on dit que λ est valeur propre de l'endomorphisme A s'il existe un vecteur non nul V tel que $AV = \lambda V$. Une manière un peu plus sophistiquée de dire cela est que *le noyau de $\lambda I - A$ n'est pas réduit à 0* (I représentant l'endomorphisme *identité*). Ce noyau est appelé le *sous-espace propre* associé à la valeur propre λ et sa dimension est appelée la *multiplicité* de la valeur propre λ .

On démontre⁽⁴⁾ que l'espace vectoriel $\mathcal{E}$ (de dimension finie) sur lequel A opère est la somme directe des sous-espace propres ; sa dimension est donc la somme des multiplicités des valeurs propres. Le nombre de valeurs propres ne peut donc pas excéder la dimension de l'espace.

Si λ n'est égal à aucune des valeurs propres, le noyau de $\lambda I - A$ est réduit à zéro et l'endomorphisme est injectif, donc inversible. L'ensemble fini des valeurs propres, qui est un sous-ensemble du corps $\mathbb{C}$, est appelé *le spectre* de l'endomorphisme A . L'application $\lambda \mapsto (\lambda I - A)^{-1}$, appelée *résolvante* de A , est définie sur le complémentaire du spectre et prend ses valeurs dans l'espace $\text{End}(\mathcal{E})$.

En dimension infinie, la situation est plus compliquée. Étant donné un opérateur A sur un espace de Hilbert $\mathcal{H}$, on peut toujours considérer $\lambda I - A$ qui est aussi un opérateur sur

⁽⁴⁾ J. LELONG-FERRAND & J. M. ARNAUDIÈS, *Algèbre*, Dunod, Paris, 1971, chapitre XI.

$\mathcal{H}$ et si son noyau est de dimension $\alpha > 0$, on dit que λ est une valeur propre de multiplicité α . En dehors de ce cadre, il y a les phénomènes nouveaux qui ne peuvent pas se produire en dimension finie :

- le noyau peut être réduit à 0, sans que pour autant $\lambda I - A$ soit inversible⁽⁵⁾ et le spectre peut donc contenir autre chose que des valeurs propres ;
- le noyau peut être réduit à 0 et $\lambda I - A$ être inversible, mais cet inverse peut être discontinu ;
- l'opérateur A n'est lui-même pas forcément continu et peut n'être défini que sur un sous-espace dense $\mathcal{D}_A$: si tel est le cas, l'opérateur $(\lambda I - A)^{-1}$ ne peut pas être surjectif⁽⁶⁾ même s'il est bien défini et continu.

Le dernier cas fait bien comprendre pourquoi on a intérêt à considérer des opérateurs fermés, c'est-à-dire ayant un domaine maximal. Imaginons un opérateur A défini sur un sous-espace dense $\mathcal{D}$, mais préfermé et non fermé et soit A_0 son extension fermée, définie sur le domaine maximal $\mathcal{D}_0$. Il est clair que si $\lambda I - A_0$ est inversible, il *ne pourra pas* en être de même de $\lambda I - A$, car les éléments de $\mathcal{D}_0 - \mathcal{D}$ n'auront pas d'image. Ainsi, pour une opération linéaire telle que la dérivation, le fait pour $\lambda I - A$ d'être inversible ou non dépendrait du choix du domaine et ne serait pas *intrinsèque*.

On supposera donc *toujours* dans la suite que les opérateurs sont fermés. Pour faire la théorie spectrale de ces opérateurs, on fixe alors les définitions suivantes.

Définition XII.2 On appelle ensemble résolvant l'ensemble des valeurs complexes de λ pour lesquelles $\lambda I - A$ possède un inverse continu.

Si λ est résolvant, $\lambda I - A$ applique le domaine maximal $\mathcal{D}_A$ bijectivement sur $\mathcal{H}$ et son inverse $(\lambda I - A)^{-1}$ applique bijectivement $\mathcal{H}$ sur $\mathcal{D}_A$ et en outre de façon continue. Ce critère n'est correct que pour des opérateurs fermés.

Définition XII.3 On appelle spectre de A le complémentaire (dans $\mathbb{C}$) de l'ensemble résolvant.

Le spectre n'est donc pas (contrairement à ce qui se passe en dimension finie) formé que de valeurs propres ; il contient aussi les valeurs λ pour lesquelles $\lambda I - A$ n'est pas surjectif et même les valeurs de λ pour lesquelles $(\lambda I - A)^{-1}$ existe, mais n'est pas continu ; ces cas ne se produisent pas quand la dimension est finie.

Définition XII.4 On appelle indice de nullité de λ , noté $\text{nul}_A(\lambda)$, la dimension du noyau de $\lambda I - A$ et indice de défaut de λ , noté $\text{def}_A(\lambda)$, la codimension de l'image de $\lambda I - A$.

Si λ est valeur propre, $\text{nul}_A(\lambda)$ est donc sa multiplicité. La codimension d'un sous-espace vectoriel est la dimension de son complémentaire orthogonal (ici, dans $\mathcal{H}$). En dimension finie, on aurait *toujours* $\text{nul}_A(\lambda) = \text{def}_A(\lambda)$, mais en dimension infinie ces deux paramètres peuvent prendre indépendamment l'un de l'autre n'importe quelle valeur entière ou infinie.

(5) On rappelle qu'en dimension finie, un endomorphisme injectif est automatiquement surjectif, mais cela n'est pas vrai en dimension infinie.

(6) C'est-à-dire avoir pour image $\mathcal{H}$ tout entier

Définition XII.5 On appelle spectre essentiel de A l'ensemble des valeurs de λ pour lesquelles $\text{nul}_A(\lambda) = \text{def}_A(\lambda) = \infty$.

Le spectre est une partie du plan complexe, qui est formée de différentes composantes : il y a les valeurs propres, pour lesquelles $\text{nul}_A(\lambda) > 0$ (y compris $\text{nul}_A(\lambda) = \infty$), le spectre essentiel, mais aussi des composantes intermédiaires, correspondant à différentes valeurs des indices. On remarque que les valeurs propres de multiplicité finie sont par nature toujours hors du spectre essentiel. Le spectre essentiel tire son nom du fait qu'il constitue la partie la plus stable du spectre lorsqu'on perturbe l'opérateur.

Théorème XII.2 Le spectre est toujours un ensemble fermé.

Preuve On utilise le fait qu'un opérateur proche de l'identité est forcément inversible et son inverse continu ; soit A un opérateur borné et ε un nombre complexe ; alors, pour ε assez petit, $I + \varepsilon A$ est inversible et son inverse est continu. Prouvons d'abord ce lemme technique.

Puisque A est borné, il existe une constante M telle que pour toute $f \in \mathcal{H}$, $\|Af\| \leq M\|f\|$. Cela implique que :

$$\|(I + \varepsilon A)f\| \geq \|f\| - |\varepsilon| \|Af\| \geq (1 - |\varepsilon|M)\|f\| \quad (\text{XII.14})$$

Cette inégalité prouve déjà que si $|\varepsilon|M < 1$, $I + \varepsilon A$ est injective.

Voyons la surjectivité. Si g est dans $\mathcal{H}$, soit f la somme de la série $\sum_{n \geq 0} (-1)^n \varepsilon^n A^n g$, où A^n est l'opérateur A itéré n fois. On a donc $\|A^n g\| \leq M^n \|g\|$. Chaque terme de la série peut donc être majoré en norme comme ceci : $\|(-1)^n \varepsilon^n A^n g\| \leq (|\varepsilon|M)^n \|g\|$. Cela montre que si $|\varepsilon|M < 1$, la série est normalement convergente et donc convergente dans $\mathcal{H}$ (puisque $\mathcal{H}$ est complet). Sa somme définit donc un élément f de $\mathcal{H}$. En appliquant l'opérateur $I + \varepsilon A$ à f , on obtient g , donc g est l'image de f . On a ainsi prouvé que si $|\varepsilon|M < 1$, $I + \varepsilon A$ est surjective.

Le fait que l'inverse soit continu résulte immédiatement de l'inégalité (XII.14). En effet, en posant $g = (I + \varepsilon A)f$ ou, ce qui est équivalent, $f = (I + \varepsilon A)^{-1}g$, cette inégalité donne :

$$\|f\| \leq (1/(1 - |\varepsilon|M)) \|g\| \quad (\text{XII.15})$$

En conclusion, si A est un opérateur borné de norme M , alors l'opérateur $I + \varepsilon A$ est inversible si $|\varepsilon| < 1/M$.

Pour montrer que le spectre est un ensemble fermé dans $\mathbb{C}$, on montre que son complémentaire, l'ensemble résolvant, est ouvert. L'argument consiste à montrer que si λ_0 est résolvant, tout λ assez proche de λ_0 est résolvant aussi. Ou encore : si $\lambda_0 I - A$ est inversible et son inverse continu il en sera de même de $\lambda I - A$ si λ est assez proche de λ_0 . Or on peut écrire :

$$\lambda I - A = (\lambda - \lambda_0)I + \lambda_0 I - A = (\lambda_0 I - A)(I + (\lambda - \lambda_0)(\lambda_0 I - A)^{-1}) \quad (\text{XII.16})$$

Si $\lambda_0 I - A$ est inversible et d'inverse continu, appelons M la norme de cet inverse ; d'après le lemme, l'opérateur $(I + (\lambda - \lambda_0)(\lambda_0 I - A)^{-1})$ sera inversible et d'inverse continu si $|\lambda - \lambda_0| < 1/M$, donc aussi $\lambda I - A$. •

On pourrait montrer par des arguments du même type que le spectre essentiel est aussi un ensemble fermé⁽⁷⁾. Étudions quelques exemples.

Exemple XII.1 — Multiplication par une fonction Soit $V(x)$, une fonction à valeurs complexes, continûment dérivable et l'opérateur A qui, à toute fonction $f(x)$ de $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$, associe la fonction $V(x)f(x)$.

Cet opérateur est continu si V est une fonction continue bornée (cela résulte de l'inégalité de la moyenne, car $\|Af\| \leq \sup_{x \in \mathbb{R}} |V(x)| \cdot \|f\|$), mais discontinu dans le cas contraire. Il s'agit d'examiner l'inversion de $\lambda I - A$. Pour cela on considère l'équation $\lambda f - Af = g$,

(7) T. KATO, *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1966, p. 235 et 242–243.

qui se traduit par $\forall x, \lambda f(x) - V(x)f(x) = g(x)$. On obtient l'inverse par division : $f(x) = g(x)/[\lambda - V(x)]$. λ est résolvant si et seulement si l'application $g \mapsto g/(\lambda - V)$ est définie pour tout $g \in \mathcal{H}$ et continue. On va voir que le spectre est l'adhérence de $\mathfrak{S}(V)$, l'ensemble des valeurs prises par la fonction V . Supposons en effet que λ ne soit pas adhérente à $\mathfrak{S}(V)$; il existe alors une distance minimum de λ aux points de $\overline{\mathfrak{S}(V)}$, soit α . Le dénominateur de $f = g/(\lambda - V)$ ne s'annule jamais et en outre on a toujours $|f(x)| \leq \frac{1}{\alpha}|g(x)|$, donc d'après l'inégalité de la moyenne $\|f\| \leq \frac{1}{\alpha}\|g\|$, ce qui prouve la continuité. Inversement, il faut prouver que si λ est adhérente à $\mathfrak{S}(V)$, la transformation $f \mapsto \lambda f - V \cdot f$ ne peut pas être inversée. Nous savons déjà que le spectre est un ensemble fermé, donc il suffit de voir le cas où $\lambda \in \mathfrak{S}(V)$. Soit x_0 tel que $\lambda = V(x_0)$. Nous avons supposé que V est continûment dérivable (sans cette hypothèse supplémentaire, l'argument deviendrait très compliqué) ; alors, dans un voisinage de x_0 on aura, d'après la formule des accroissements finis, $|V(x_0) - V(x)| \leq M|x - x_0|$, d'où $|\lambda f(x) - V(x)f(x)| \leq M|x - x_0||f(x)|$ (dans un voisinage de x_0). Cela suffit à prouver que la transformation $f \mapsto \lambda f - V \cdot f$ ne peut pas être inversée : si cela était, une fonction $g \in \mathbb{L}_2(\mathbb{R})$ qui reste par exemple égale à M dans un voisinage de x_0 serait l'image d'une fonction $f(x)$, minorée par $1/|x - x_0|$ dans ce voisinage ; or une telle fonction ne peut pas être de carré intégrable. $\square$

Exemple XII.2 — Dérivation L'opérateur est maintenant la dérivation $B \cdot f = f'$. Il est possible de se ramener au cas précédent par transformation de Fourier avec $V(\xi) = i\xi$. En effet, l'équation $g = \lambda f - Bf = \lambda f - f'$ est équivalente à l'équation $\widehat{g} = \lambda \widehat{f} + i\xi \widehat{f}$. Comme la transformée de Fourier est un isomorphisme de l'espace $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$, elle conserve les propriétés qui décident si λ est, ou non, une valeur résolvante : $f \mapsto \lambda f - f'$ est inversible si et seulement si $f \mapsto \lambda f + i\xi f$ l'est ; et son inverse est continu si et seulement si $g \mapsto g/(\lambda + i\xi)$ l'est. On en déduit que le spectre de B est l'axe imaginaire. $\square$

Ⓡ L'opérateur $f \mapsto \lambda f - f'$, tout comme l'opérateur $f \mapsto \lambda f - xf$ ou $f \mapsto \lambda f + i\xi f$, est *toujours* injectif : si $\forall x \in \mathbb{R}, (\lambda - x)f(x) = 0$ ou si $\forall \xi \in \mathbb{R}, (\lambda + i\xi)f(\xi) = 0$, alors $\forall x \in \mathbb{R}, f(x) = 0$. Le spectre ne contient donc *aucune* valeur propre. Par contre, si la fonction $V(x)$ est constante sur tout un intervalle $[a; b]$, alors les valeurs de λ égales à cette constante sont, pour l'opérateur $f \mapsto Vf$, des valeurs propres de multiplicité infinie (la vérification de cette affirmation est un bon exercice ; déterminer aussi les sous-espaces propres correspondants).

Exemple XII.3 — Transformation de Fourier L'opérateur est $\mathcal{F} : f \mapsto \widehat{f}$. Il s'agit de voir pour quelles valeurs de λ la transformation $\lambda I - \mathcal{F}$ possède un inverse continu.

On va commencer par montrer que les fonctions d'Hermite $\varphi_n(x) = H_n(x)e^{-x^2/2}$ (voir sous-section XI.4.2) sont des fonctions propres de $\mathcal{F}$: $\widehat{\varphi_n} = \sqrt{2\pi}i^n \varphi_n$. Cela se prouve par récurrence. On sait que la transformée de Fourier de $e^{-x^2/2}$ est la fonction $\sqrt{2\pi}e^{-\xi^2/2}$, vu à l'exemple VII.1, c'est-à-dire $\widehat{\varphi_0} = \sqrt{2\pi}\varphi_0$. Pour les suivants, on procède par récurrence : supposons que la relation $\widehat{\varphi_n} = \sqrt{2\pi}i^n \varphi_n$ soit vraie pour n et prouvons qu'elle sera vraie aussi pour $n + 1$. Pour cela, on va utiliser la relation de récurrence des polynômes d'Hermite (XI.38) : celle-ci conduit à la relation $\varphi_{n+1} = \varphi'_n - x\varphi_n$ pour les φ_n . En lui appliquant la transformation de Fourier, on obtient :

$$\widehat{\varphi_{n+1}} = -i\xi \widehat{\varphi_n} + i \frac{d}{d\xi} \widehat{\varphi_n} \quad (\text{XII.17})$$

En utilisant maintenant l'hypothèse de récurrence, on parvient à :

$$\widehat{\varphi_{n+1}}(\xi) = \sqrt{2\pi} i^{n+1} \left(-\xi \varphi_n(\xi) + \frac{d}{d\xi} \varphi_n(\xi) \right) = \sqrt{2\pi} i^{n+1} \varphi_{n+1}(\xi) \quad (\text{XII.18})$$

ce qui prouve la récurrence.

Ceci montre que la transformation de Fourier $\mathcal{F}$ sur l'espace $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$ a des valeurs propres ; ce qui, rappelons-le, n'était pas le cas des opérateurs $f \mapsto Vf$ ou $f \mapsto f'$. Ces valeurs propres sont $\sqrt{2\pi} i^n$: comme i^n ne prend en réalité que les quatre valeurs 1, i , -1 et $-i$, on peut dire que $\mathcal{F}$ a pour valeurs propres les quatre nombres complexes $\sqrt{2\pi}$, $\sqrt{2\pi} i$, $-\sqrt{2\pi}$ et $-\sqrt{2\pi} i$. Les sous-espaces propres correspondants sont de dimension infinie ; ils sont engendrés (respectivement) par les familles $\varphi_0, \varphi_4, \varphi_8, \varphi_{12}, \varphi_{16}, \dots$, $\varphi_1, \varphi_5, \varphi_9, \varphi_{13}, \varphi_{17}, \dots$, $\varphi_2, \varphi_6, \varphi_{10}, \varphi_{14}, \varphi_{18}, \dots$ et $\varphi_3, \varphi_7, \varphi_{11}, \varphi_{15}, \varphi_{19}, \dots$. Ainsi ces valeurs propres sont de multiplicité infinie.

Une autre question se pose : le spectre est-il composé uniquement de ces valeurs propres, ou contient-il encore autre chose que les valeurs propres ? Le fait que λ soit valeur propre signifie que l'opérateur $\lambda I - \mathcal{F}$ est non injectif. Contrairement à ce qui se passe en dimension finie, λ peut être dans le spectre même si $\lambda I - \mathcal{F}$ est injectif : il suffit pour cela que $\lambda I - \mathcal{F}$ soit non surjectif (c'est ce qui arrivait dans le spectre des opérateurs Vf ou f' , voir exemples XII.1 et XII.2), ou même, que $\lambda I - \mathcal{F}$, quoique injectif et surjectif, ait un inverse discontinu. Il faut donc contrôler ce qui arrive lorsque λ n'est aucune des quatre valeurs $\sqrt{2\pi}$, $\sqrt{2\pi} i$, $-\sqrt{2\pi}$ et $-\sqrt{2\pi} i$.

Pour cela, on va utiliser le fait que la famille φ_n est une base orthogonale de $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$, ou, mieux, que $\phi_n = \varphi_n / \|\varphi_n\| = \varphi_n / \pi^{1/4} 2^{n/2} \sqrt{n!}$ est une base orthonormée. Il s'agit d'étudier l'inversion de l'opérateur $\lambda I - \mathcal{F}$, autrement dit, étant donné g quelconque dans $\mathcal{H}$, trouver $f \in \mathcal{H}$ telle que $\lambda f - \widehat{f} = g$. Or, les fonctions de $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$ sont des sommes de *séries d'Hermite* de la forme $\sum c_n \phi_n$, telles que les séries numériques $\sum |c_n|^2$ soient convergentes (cf. théorème XI.1). On peut donc traduire le problème ainsi : étant donnée une série $g = \sum b_n \phi_n$ telle que $\sum |b_n|^2$ soit convergente, trouver une série $f = \sum a_n \phi_n$ telle que $\sum |a_n|^2$ soit convergente et que $\lambda f - \widehat{f} = g$.

Or $\lambda f - \widehat{f} = \sum a_n (\lambda \phi_n - \widehat{\phi_n}) = \sum a_n (\lambda - \sqrt{2\pi} i^n) \phi_n$ et puisque φ_n est une base, l'équation $\lambda f - \widehat{f} = g$ se traduit par :

$$\forall n \geq 0, \quad b_n = (\lambda - \sqrt{2\pi} i^n) a_n \quad \Leftrightarrow \quad \forall n \geq 0, \quad a_n = \frac{b_n}{\lambda - \sqrt{2\pi} i^n} \quad (\text{XII.19})$$

Il est facile de voir que la condition $\sum |b_n|^2 < \infty \Rightarrow \sum |a_n|^2 < \infty$ sera satisfaite pour tout λ autre que les quatre valeurs propres, car on peut minorer le dénominateur $\lambda - \sqrt{2\pi} i^n$: si λ est distinct des quatre valeurs propres, il existe un $\alpha > 0$ tel que $\forall n, |\lambda - \sqrt{2\pi} i^n| \geq \alpha$, d'où :

$$\forall n \geq 0, \quad |a_n| = \frac{|b_n|}{|\lambda - \sqrt{2\pi} i^n|} \leq \frac{1}{\alpha} |b_n| \quad (\text{XII.20})$$

L'opérateur $\lambda I - \mathcal{F}$ est donc bien inversible. En outre :

$$\|f\|^2 = \sum_{n \geq 0} |a_n|^2 \leq \frac{1}{\alpha^2} \sum_{n \geq 0} |b_n|^2 = \frac{1}{\alpha^2} \|g\|^2 \quad (\text{XII.21})$$

ce qui prouve la continuité de l'opérateur $(\lambda I - \mathcal{F})^{-1}$.

Conclusion : le spectre de $\mathcal{F}$ est formé uniquement des quatre valeurs propres $\sqrt{2\pi}$, $\sqrt{2\pi}i$, $-\sqrt{2\pi}$ et $-\sqrt{2\pi}i$, toutes de multiplicité infinie. $\square$

Exercice XII.4 vérifier que l'ensemble de ces quatre valeurs propres est le *spectre essentiel* de $\mathcal{F}$. (Calculer les indices de défaut et de nullité.) $\blacksquare$

XII.4 Oscillateur quantique

Dans cette section, on va illustrer la théorie des opérateurs développée en sections XII.2 et XII.3, en montrant comment elle s'applique à la mécanique quantique. L'idée essentielle a été avancée pour la première fois par E. SCHRÖDINGER⁽⁸⁾. Dans ce mémoire, la mécanique quantique est construite en partant de l'analogie entre la mécanique newtonienne et l'Optique géométrique, selon le postulat, appelé *principe de correspondance*, que *la mécanique quantique serait à la mécanique classique ce que l'Optique ondulatoire est à l'Optique géométrique*. Voici un passage du début de ce mémoire :

« L'existence d'une relation intime entre la théorie de Hamilton et le phénomène de propagation des ondes est loin d'être quelque chose de nouveau. Hamilton lui-même le connaissait très bien et l'avait même choisi comme point de départ de sa théorie de la dynamique, laquelle n'était d'ailleurs qu'un prolongement de son optique des milieux non homogènes.

Le principe de variation de Hamilton (principe de moindre action) peut être considéré comme un principe de Fermat pour une propagation d'ondes dans l'espace de configuration et son équation aux dérivées partielles n'exprime autre chose que le principe d'Huyghens pour cette même propagation. »

et plus loin :

« L'hypothèse à laquelle je crois pouvoir attacher un haut degré de certitude est la suivante : la manière correcte de concevoir ou de représenter les phénomènes mécaniques consiste à les rattacher à une propagation d'ondes dans l'espace des points représentatifs et non à un mouvement de points représentatifs. L'étude du mouvement des points représentatifs qui forme l'objet de la mécanique classique, n'est qu'un procédé d'approximation et son emploi est tout aussi peu justifié que l'emploi de l'optique géométrique, ou optique de rayons, dans le cas de phénomènes lumineux réels. »

La méthode de Hamilton mentionnée par Schrödinger est la suivante : dans un milieu non homogène (c'est-à-dire d'indice de réfraction variable $\mathcal{N}(x, y, z)$), la propagation d'ondes a pour équation :

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} + k^2 \mathcal{N}(x, y, z)^2 f = 0 \quad (\text{XII.22})$$

où $k = 2\pi/\lambda$ est le nombre d'onde. Lorsque l'onde f se propage, les *surfaces d'onde*, d'équation $W(x, y, z) = C^{\text{te}}$ vérifient l'équation dite *équation eikonale* :

$$\left| \frac{\partial W}{\partial x} \right|^2 + \left| \frac{\partial W}{\partial y} \right|^2 + \left| \frac{\partial W}{\partial z} \right|^2 = |\text{grad } W|^2 = k^2 \mathcal{N}(x, y, z)^2 \quad (\text{XII.23})$$

(8) *Quantisierung als Eigenwertproblem II*, Annalen der Physik, vol. 79, 1926, p. 489. La traduction française de 1933 a été rééditée récemment : Erwin Schrödinger *Mémoires sur la mécanique ondulatoire* Editions Jacques Gabay, Paris, 1988. La théorie de l'oscillateur quantique à une dimension, qui est présentée ici avec le formalisme des opérateurs, se trouve aux pages 49 à 53 de cette édition.

qui décrit l'onde selon l'optique géométrique ; les *rayons lumineux* sont alors les lignes orthogonales aux surfaces d'onde.

Hamilton fait remarquer que dans un champ de force dérivant d'un potentiel $V(x, y, z)$, les trajectoires des points matériels d'énergie E sont également les lignes orthogonales à des surfaces d'équation $W(x, y, z) = C^{\text{te}}$, si on prend pour W une solution de l'équation :

$$|\text{grad } W|^2 = 2(E - V) \quad (\text{XII.24})$$

dite équation de Hamilton-Jacobi. Cette dernière est formellement semblable à l'équation eikonale (XII.23). C'est pourquoi, dans l'ouvrage cité par Schrödinger, Hamilton fait remarquer que les trajectoires des points matériels d'énergie E soumis au champ de potentiel $V(x, y, z)$ sont identiques aux rayons lumineux dans un milieu non homogène d'indice de réfraction $\mathcal{N}(x, y, z)$, si on pose :

$$\mathcal{N}(x, y, z) = \sqrt{2[E - V(x, y, z)]} \quad (\text{XII.25})$$

Ainsi, les trajectoires seraient les *rayons lumineux* du phénomène ondulatoire correspondant à l'équation des ondes :

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} + 2k^2 [E - V] f = 0 \quad (\text{XII.26})$$

qu'on obtient à partir de (XII.22) en remplaçant $\mathcal{N}(x, y, z)^2$ par $2[E - V(x, y, z)]$, conformément à (XII.23). Il reste à deviner quelle valeur donner à k ; après une courte évaluation de ce que doit être l'ordre de grandeur de k d'après les données physiques connues, Schrödinger fait le rapprochement avec l'hypothèse (formulée trois ans avant) par de Broglie et conclut que k doit être $\sqrt{m}/\hbar$, m étant la masse de la particule. Voici encore un passage :

« On retrouve ainsi un théorème que M. De Broglie avait établi en faisant largement appel à la théorie de la Relativité, pour les ondes de phase d'un électron, dans ses belles recherches⁽⁹⁾ qui ont formé le point de départ du présent travail. On voit qu'il s'agit là au fond d'un théorème beaucoup plus général qui ne découle pas nécessairement de la théorie de la Relativité, mais qui est valable pour n'importe quel système conservatif de la mécanique classique. »

En remplaçant ainsi k par $\sqrt{m}/\hbar$, l'équation (XII.26) devient :

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \right) + V f = E f \quad (\text{XII.27})$$

qui est restée sous le nom d'*équation de Schrödinger*.

Si on interprète le premier membre de cette équation comme un opérateur :

$$A : f \longmapsto -\frac{\hbar^2}{2m} \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \right) + V f \quad (\text{XII.28})$$

l'équation (XII.27) s'interprète comme l'équation aux valeurs propres de l'opérateur A : la solution de (XII.27), si elle existe, est le vecteur propre associé à la valeur propre E .

Schrödinger a proposé de poser comme un principe de la mécanique quantique que :

(9) L. De Broglie, *Annales de Physique*, vol. 3, 1925, page 22

« Les niveaux quantiques sont déterminés tous à la fois, comme valeurs propres de l'équation (XII.27) qui contient en elle-même ses propres conditions aux limites. »

Les conditions aux limites dont il est question implicitement sont les conditions à l'infini : pour que E puisse être une valeur propre de l'opérateur A , il faut que le vecteur propre associé soit une fonction de carré intégrable, c'est-à-dire un élément de l'espace de Hilbert $\mathbb{L}_2(\mathbb{R}^3)$. Ceci est lié à l'interprétation (effectuée au même moment par Max Born) de la fonction f , dont le carré doit représenter une densité de probabilité. Notons en passant que les valeurs et vecteurs propres concernent les états stationnaires ou états liés, mais les valeurs du spectre autres que les valeurs propres correspondent aux processus de collision.

L'idée initiale de Schrödinger, qui par l'intermédiaire des mathématiques conduisait à des conséquences expérimentalement justes, a été développée et systématisée. La synthèse a été faite par le mathématicien John von Neumann et présentée dans son livre paru en 1943 aux états-unis *Mathematische Grundlagen der Quantenmechanik* (Dover Publ. N.Y.) Le formalisme adopté aujourd'hui encore universellement pour la mécanique quantique est pour l'essentiel celui proposé par von Neumann et est entièrement basé sur la théorie des opérateurs dans les espaces de Hilbert. Nous l'appellerons le *formalisme standard* de la mécanique quantique. Pour l'Histoire, il est intéressant de signaler que la théorie des espaces de Hilbert et des opérateurs est l'œuvre quasi exclusive de l'école *mathématique de Göttingen*, dont Johann von Neumann est issu ; elle a été développée entre 1900 et 1920 (en gros), donc *avant la mécanique quantique*. En 1933, cette école a déménagé à Princeton (New Jersey) ; seuls les membres les plus âgés, dont David Hilbert, sont restés en Allemagne.

Les principes de ce *formalisme standard* de la mécanique quantique se résument ainsi :

1. l'état de préparation d'un système quantique est décrit par un élément d'un espace de Hilbert $\mathcal{H}$, qui est toujours un espace $\mathbb{L}_2(X)$, mais X , l'espace de configuration, dépend de la nature du système et de son environnement : il ne peut être déterminé qu'à partir du problème classique correspondant ;
2. les grandeurs qu'on peut mesurer physiquement sont représentées par un opérateur auto-adjoint A sur l'espace $\mathcal{H}$; on ne peut mesurer simultanément, avec une précision arbitraire, que des grandeurs correspondant à des opérateurs qui commutent ;
3. les valeurs observées sont nécessairement des valeurs propres de ces opérateurs et pour un système *préparé* dans l'état $\psi \in \mathcal{H}$, la probabilité d'avoir la valeur a_n (valeur propre de A) est le carré du produit scalaire $\langle \psi | f_n \rangle$, où f_n est la fonction propre normée associée à la valeur propre a_n .

Le troisième principe sous-entend que, en vertu du principe des probabilités totales qui veut que la somme des probabilités de toutes les possibilités soit 1, on doit avoir pour toute $\psi \in \mathcal{H}$ de norme 1, la relation :

$$\sum_n |\langle \psi | f_n \rangle|^2 = 1 \quad (\text{XII.29})$$

la somme étant étendue à toutes les fonctions propres. Cela suppose, comme nous l'avons vu au chapitre XI, que les fonctions propres de A forment une base. Si on a l'esprit de

système (c'était le cas du mathématicien J. von Neumann), on peut ajouter cette exigence à la définition d'une grandeur physique, mais cela complique l'axiomatique sans bénéfices pratiques. Nous nous contenterons de la constater dans les cas *intéressants*.

Pour illustrer son approche, le premier exemple présenté par E. Schrödinger est celui de l'oscillateur quantique à une dimension. Nous allons examiner de près ce problème.

L'équation classique de l'oscillateur est :

$$m\ddot{x} = -Kx \quad (\text{XII.30})$$

où K est le paramètre d'élasticité. Le potentiel est $V(x) = \frac{1}{2}Kx^2$, donc l'équation de Schrödinger qui correspond à (XII.27) dans ce cas particulier est :

$$-\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + \frac{1}{2}Kx^2 f = Ef \quad (\text{XII.31})$$

Nous devons en chercher des solutions de carré intégrable.

Si on interprète ce problème dans le cadre du formalisme *standard*, l'opérateur A correspond à la grandeur *énergie* (A est appelé l'*hamiltonien* du système). Nous allons calculer les valeurs propres E_n (appelées *niveaux d'énergie* du système) et les vecteurs propres normalisés f_n correspondants. D'après le principe 3, les E_n sont les seules valeurs qui peuvent être observées dans un processus de mesure et si le système a été préparé dans l'état ψ , elles seront observées avec la probabilité $p_n = |\langle \psi | f_n \rangle|^2$; cela signifie que si on reproduit mille fois le processus en préparant chaque fois le système de la même façon, on observera environ $1\,000 p_1$ fois l'énergie E_1 , $1\,000 p_2$ fois l'énergie E_2 , $1\,000 p_3$ fois l'énergie E_3 etc.

On procède comme pour l'étude de l'opérateur $\mathcal{F}$ à la section XII.3, en introduisant les fonctions d'Hermite $\varphi_n(x) = H_n(x) e^{-x^2/2}$. Appliquons d'abord l'opérateur $A_0 = -\frac{d^2}{dx^2} + x^2$ à ces fonctions. Nous avons déjà vu que les opérateurs $f \mapsto -f''$ et $f \mapsto x^2 f$, quoique discontinus, sont fermés si on les définit sur un domaine maximal. Il n'est pas difficile de vérifier qu'il en sera de même de leur somme A_0 (considérer d'abord A_0 sur le sous-espace dense $\mathcal{S}(\mathbb{R})$ et vérifier qu'il est préfermé). On obtient en dérivant :

$$\frac{d^2 \varphi_n}{dx^2} = \frac{d^2}{dx^2} (H_n(x) e^{-x^2/2}) = (H_n''(x) - 2xH_n'(x) + (x^2 - 1)H_n(x)) e^{-x^2/2} \quad (\text{XII.32})$$

ce qui montre que :

$$A_0 \varphi_n(x) = (-H_n''(x) + 2xH_n'(x) + H_n(x)) e^{-x^2/2} \quad (\text{XII.33})$$

Or, on sait que les polynômes d'Hermite vérifient la relation :

$$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0 \quad (\text{XII.34})$$

de sorte que :

$$A_0 \varphi_n(x) = (2n + 1) H_n(x) e^{-x^2/2} = (2n + 1) \varphi_n(x) \quad (\text{XII.35})$$

Cette relation signifie que les fonctions φ_n sont des vecteurs propres de l'opérateur A_0 , les valeurs propres correspondantes étant $2n + 1$. On voit que (contrairement à ce qui se passait pour $\mathcal{F}$), ces valeurs propres sont toutes distinctes; ainsi chacune a pour unique vecteur propre associé φ_n et par conséquent a pour multiplicité 1.

Exercice XII.5 — polynômes d'Hermite Démontrer la relation (XII.34) par récurrence à partir de la relation :

$$H_{n+1}(x) = H'_n(x) - 2xH_n(x) \quad (\text{XII.36})$$

qui, elle, résulte directement de la définition des polynômes d'Hermite. ■

On vérifie par le même procédé que pour $\mathcal{F}$, que le spectre de A_0 ne contient que ces valeurs propres : si $g \in \mathbb{L}_2(\mathbb{R})$, on a $g = \sum b_n \varphi_n$ avec $\sum |b_n|^2$ convergente, donc la solution de l'équation $\lambda f - A_0 f = g$ est $f = \sum a_n \varphi_n$, où $a_n = b_n/(\lambda - 2n - 1)$. On peut donc énoncer :

Théorème XII.3 Le spectre de l'opérateur $A_0 = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + x^2$ sur l'espace $\mathbb{L}_2(\mathbb{R})$ est formé de la famille des valeurs propres simples $\lambda_n = 2n + 1$, les vecteurs propres correspondants étant les fonctions d'Hermite $\varphi_n(x) = H_n(x)e^{-x^2/2}$.

On se ramène ensuite à l'opérateur hamiltonien :

$$A = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}Kx^2 \quad (\text{XII.37})$$

de l'équation (XII.31) par un changement de variable linéaire $x \mapsto y = \alpha x$. En effet, $\frac{d^2}{dy^2} = \alpha^2 \frac{d^2}{dx^2}$, donc l'équation (XII.31) se transforme en :

$$-\frac{\hbar^2}{2m} \frac{1}{\alpha^2} \frac{d^2 f}{dy^2} + \frac{1}{2}K\alpha^2 y^2 f = E f \quad (\text{XII.38})$$

ce qui, en divisant par $\frac{1}{2}K\alpha^2$, donne :

$$-\frac{\hbar^2}{mK\alpha^4} \frac{d^2 f}{dy^2} + y^2 f = \frac{2E}{K\alpha^2} f \quad (\text{XII.39})$$

On voit qu'en choisissant $\alpha = [\hbar^2/(mK)]^{1/4}$, cette équation devient $A_0 f = \lambda f$, avec $E = \frac{1}{2}K\alpha^2 \lambda$. Ainsi, les valeurs propres de l'opérateur A sont $E_n = \frac{1}{2}K\alpha^2(2n + 1) = \hbar\omega(n + \frac{1}{2})$, où $\omega = \sqrt{K/m}$, qui est la fréquence (ou plutôt la pulsation) de l'oscillateur classique.

En conclusion, l'équation (XII.31), postulée par E. Schrödinger en vertu de son *principe de correspondance*, n'a de solutions dans $\mathbb{L}_2(\mathbb{R})$ que pour les valeurs de E égales à $\hbar\omega(n + 1/2)$ et ces solutions sont alors les fonctions $\varphi_n(x [Km]^{1/4}/\sqrt{\hbar})$. La signification physique est la suivante :

1. l'oscillateur, en changeant d'état, ne peut émettre ou absorber de l'énergie que par multiples entiers de $\hbar\omega$;
2. lorsqu'il est dans l'état d'énergie $E_n = \hbar\omega(n + 1/2)$, la densité de présence de la particule dans l'espace est $H_n(x [Km]^{1/4}/\sqrt{\hbar})^2 e^{-x^2\sqrt{Km}/\hbar}$.

XII.5 Opérateurs auto-adjoints et unitaires

La notion d'*adjoint* d'un opérateur généralise celle d'*endomorphisme adjoint* sur les espaces de dimension finie, ou (une base étant fixée) de *matrice adjointe*. Si $\mathcal{E}$ est un espace hermitien

de dimension finie n et si $A \in \text{End}(\mathcal{E})$, l'endomorphisme adjoint A^* de A est défini par la relation :

$$\forall X \in \mathcal{E}, \quad \forall Y \in \mathcal{E}, \quad \langle AX \mid Y \rangle = \langle X \mid A^*Y \rangle \quad (\text{XII.40})$$

Cette relation définit entièrement A^* , car il est clair que pour tout $Y \in \mathcal{E}$, $f : X \mapsto f(X) = \langle AX \mid Y \rangle$ est une forme linéaire sur $\mathcal{E}$, à laquelle correspond par dualité hermitienne un élément unique Z tel que $\forall X \in \mathcal{E}$, $f(X) = \langle X \mid Z \rangle$. Ce vecteur Z dépend linéairement de X et l'endomorphisme adjoint est la transformation $X \mapsto Z$; on pose $Z = A^*X$.

Dans un espace de Hilbert $\mathcal{H}$ de dimension infinie, la situation est plus subtile, car l'existence d'un vecteur Z tel que $f(X) = \langle X \mid Z \rangle$ n'est garantie que pour les formes linéaires *continues* (théorème de F. Riesz). Si l'opérateur A est lui-même continu, $f : X \mapsto f(X) = \langle AX \mid Y \rangle$ est continue également et tout se passe bien comme en dimension finie ; mais si A est discontinu et défini seulement sur un domaine dense, on ne peut plus procéder de cette façon. On utilise alors le biais suivant : l'opérateur adjoint ne sera pas défini partout, mais seulement sur un sous-espace $\mathcal{D}_{A^*}$ (comme A lui-même). Plus précisément, on prend pour $\mathcal{D}_{A^*}$ l'ensemble des $h \in \mathcal{H}$ pour lesquels la fonctionnelle linéaire $f \mapsto \langle Af \mid h \rangle$ est continue. Cette fonctionnelle n'est bien définie *a priori* que sur $\mathcal{D}_A$ (toujours supposé dense dans $\mathcal{H}$), mais puisqu'elle est supposée continue et que $\mathcal{D}_A$ est dense, elle se prolonge par continuité à $\mathcal{H}$ tout entier.

On voit qu'il est essentiel de supposer $\mathcal{D}_A$ dense ; sinon, l'opérateur adjoint A^* ne peut pas être défini univoquement sur la foi du théorème de Riesz.

Le nouveau domaine introduit, $\mathcal{D}_{A^*}$, est un sous-espace vectoriel de $\mathcal{H}$; cela se voit immédiatement : si $f \mapsto \langle Af \mid h_1 \rangle$ et $f \mapsto \langle Af \mid h_2 \rangle$ sont continues, il en sera de même de leur somme, qui est $f \mapsto \langle Af \mid h_1 + h_2 \rangle$, donc $h_1 + h_2 \in \mathcal{D}_{A^*}$ et de même pour λh .

Il faudrait pouvoir s'assurer que $\mathcal{D}_{A^*}$ est dense dans $\mathcal{H}$. Pour cela, on peut prouver que :

Théorème XII.4 Si A est un opérateur fermé, le domaine $\mathcal{D}_{A^*}$ est dense et l'opérateur adjoint A^* est alors lui aussi un opérateur fermé. En outre, l'adjoint de l'adjoint, $(A^*)^*$ (qui est alors bien défini) est identique à A .

Preuve Le meilleur moyen de prouver cela est de considérer les graphes, qui sont des sous-espaces fermés de $\mathcal{H} \times \mathcal{H}$. Soit en effet $G_A = \{(f, Af) \mid f \in \mathcal{D}_A\}$ le graphe de A ; par hypothèse il est fermé. Soit aussi le sous-espace $\widehat{G}_A = \{(-Af, f) \mid f \in \mathcal{D}_A\}$, qui est l'image de G_A par la transformation $(f, g) \mapsto (-g, f)$ et qui évidemment est fermé aussi. La relation $\forall f \in \mathcal{H}, \forall g \in \mathcal{H}, \langle Af \mid g \rangle = \langle f \mid A^*g \rangle$ signifie que A^* a pour graphe le supplémentaire orthogonal de $\widehat{G}_A$ dans $\mathcal{H} \times \mathcal{H}$ (le fait que ce supplémentaire orthogonal est bien un graphe d'opérateur résulte de l'existence, déjà signalée, de l'opérateur adjoint). Pour montrer que $\mathcal{D}_{A^*}$ est dense dans $\mathcal{H}$, il suffit de montrer qu'il n'existe dans $\mathcal{H}$ aucun vecteur non nul f qui soit orthogonal à $\mathcal{D}_{A^*}$; or s'il en existait un, le couple $(f, 0)$ serait orthogonal dans $\mathcal{H} \times \mathcal{H}$ à tout couple (g, h) tel que $g \in \mathcal{D}_{A^*}$, donc en particulier à tout couple de G_{A^*} : on aurait $\forall g \in \mathcal{D}_{A^*}, (f, 0) \perp (g, A^*g)$. Puisque le graphe G_{A^*} de A^* est le supplémentaire orthogonal de $\widehat{G}_A$ (qui est fermé), cela impliquerait que $(f, 0)$ est un élément de $\widehat{G}_A$, donc que $(0, -f)$ est un élément de G_A . Autrement dit, $-f$ serait l'image par A de 0, ce qui est absurde. Par conséquent $\mathcal{D}_{A^*}$ est dense. •

Ⓡ Pour définir l'adjoint, il n'est même pas nécessaire que A soit fermé ; *préfermé* suffit : en effet, si A est préfermé et si A_0 est son prolongement fermé, le graphe G_A a pour adhérence le graphe G_{A_0} ; le supplémentaire orthogonal de G_A est le même que celui de G_{A_0} (si $\langle f \mid g \rangle = 0, \forall f \in G_A$, on peut passer à la limite en faisant tendre f vers un élément quelconque de G_{A_0}). Dans les espaces de

Hilbert, l'adhérence d'un sous-espace V est $(V^\perp)^\perp$, le supplémentaire du supplémentaire. De même, le supplémentaire orthogonal de $\widehat{G}_A$ est le même que celui de $\widehat{G}_{A_0}$. On peut dire que si A est préfermé, son prolongement fermé est $(A^*)^*$, l'adjoint de l'adjoint.

L'existence de l'adjoint A^* n'exige pas que A soit fermé ou préfermé, mais seulement que le domaine $\mathcal{D}_A$ sur lequel A est défini soit dense ; par contre, si A n'est pas préfermé, le domaine $\mathcal{D}_{A^*}$ de l'adjoint ainsi obtenu ne sera pas dense et par conséquent $(A^*)^*$ ne sera pas défini. En ajoutant quelques détails à la démonstration du théorème XII.4, on peut montrer que si $\mathcal{D}_A$ est dense, alors A est préfermé si et seulement si $\mathcal{D}_{A^*}$ est dense.

Bien entendu, ces remarques deviennent triviales pour les opérateurs continus : dans ce cas, l'adjoint existe toujours et est continu et les domaines sont $\mathcal{H}$ tout entier. Rappelons qu'un opérateur fermé dont le domaine est $\mathcal{H}$ tout entier est forcément continu (théorème de Banach).

Exercice XII.6 Sur un espace de dimension finie, on peut caractériser les endomorphismes par leur matrice : si l'endomorphisme A a pour matrice $\{a_{ij}\}$, l'endomorphisme adjoint a pour matrice $\{\bar{a}_{ji}\}$. Montrer par un procédé direct et calculatoire que cette relation entre les matrices est bien équivalente à la relation entre les graphes $G_{A^*} \perp \widehat{G}_A$. Examiner plus particulièrement le cas des dimensions 1 et 2. ■

Définition XII.6 On dit qu'un opérateur A (dont le domaine est toujours supposé dense) est auto-adjoint, si $A = A^*$.

Si A est auto-adjoint, le supplémentaire orthogonal de G_A dans $\mathcal{H} \times \mathcal{H}$ est égal à $\widehat{G}_A$. Voici quelques exemples :

Exemple XII.4 La multiplication par une fonction réelle (disons continûment dérivable) $V(x)$, définie sur le domaine $\mathcal{D}_V = \{f \in \mathcal{H} \mid Vf \in \mathcal{H}\}$. Pour déterminer le domaine $\mathcal{D}_{V^*}$ de l'adjoint, on cherche, conformément à la définition, les éléments g pour lesquels la fonctionnelle $f \mapsto \langle Vf \mid g \rangle = \int V(x)f(x)\overline{g(x)}dx$ est continue ; or d'après le théorème de Riesz, il en est ainsi si et seulement si Vg est de carré intégrable, ce qui signifie bien que $\mathcal{D}_{V^*} = \mathcal{D}_V$. □

Exemple XII.5 L'opérateur $\mathcal{P} = i\hbar \frac{d}{dx}$, sur le domaine $\mathcal{D}_{\mathcal{P}} = \{f \in \mathcal{H} \mid \xi \hat{f} \in \mathcal{H}\}$. (même argument que précédemment, mais appliqué aux transformées de Fourier). □

Exemple XII.6 L'opérateur $A = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2}Kx^2$. Pour alléger l'écriture on prend $A_0 = -\frac{d^2}{dx^2} + x^2$. En utilisant la base hilbertienne ϕ_n formée par les fonctions d'Hermite normalisées, on peut caractériser le domaine de A_0 comme étant l'ensemble des éléments f de $\mathcal{H}$ dont les coefficients d'Hermite $a_n = \langle f \mid \phi_n \rangle$ sont tels que la série numérique $\sum_{n \geq 0} (2n+1)^2 |a_n|^2$ converge. Le domaine de l'opérateur adjoint A_0^* est alors l'ensemble des $g = \sum b_n \phi_n \in \mathcal{H}$ tels que la fonctionnelle $f \mapsto \langle A_0 f \mid g \rangle = \sum (2n+1) a_n b_n$ est continue, ce qui est le cas si et seulement si la série numérique $\sum (2n+1)^2 |b_n|^2$ est convergente : cela montre bien que $\mathcal{D}_{A_0^*} = \mathcal{D}_{A_0}$. □

Exemple XII.7 Plus généralement, l'opérateur $A = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x)$, qui est l'opérateur hamiltonien usuel en mécanique quantique non relativiste. Dans le cas particulier précédent, on pouvait utiliser une base de vecteurs propres ; cela n'est plus possible en général. Interprétons le premier terme de l'opérateur comme un dérivation au sens des distributions, qui

existe toujours ⁽¹⁰⁾ ; posons alors $\mathcal{D}_A = \{f \in \mathbb{L}_2(\mathbb{R}) \mid -\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + Vf \in \mathbb{L}_2(\mathbb{R})\}$. La fonctionnelle $f \mapsto \int \left[-\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2}(x) + V(x)f(x) \right] g(x) dx$ est bien définie pour $f \in \mathcal{D}_A$. Si V est une fonction infiniment dérivable à croissance polynômiale et si f appartient à $\mathcal{S}(\mathbb{R})$, qui est un sous-espace de $\mathcal{D}_A$, on peut écrire $\langle \bar{g}, -\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + Vf \rangle = \langle f, -\frac{\hbar^2}{2m} \frac{d^2 \bar{g}}{dx^2} + V\bar{g} \rangle$ en considérant $\bar{g}$ comme une distribution. Mais cela revient à $\langle -\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + Vf \mid g \rangle = \langle f \mid -\frac{\hbar^2}{2m} \frac{d^2 g}{dx^2} + Vg \rangle$. D'après le théorème de Riesz, ceci est une fonctionnelle linéaire continue de f si et seulement si $-\frac{\hbar^2}{2m} \frac{d^2 g}{dx^2} + Vg$ est un élément de $\mathbb{L}_2(\mathbb{R})$. Cette conclusion est obtenue pour $f \in \mathcal{S}(\mathbb{R})$, mais $\mathcal{S}(\mathbb{R})$ est dense dans $\mathbb{L}_2(\mathbb{R})$ donc la fonctionnelle se prolonge par continuité. Ainsi $\mathcal{D}_{A^*} = \mathcal{D}_A$. Si V n'est pas infiniment dérivable, cette argumentation reste valable à condition d'interpréter correctement $\langle \bar{g}, -\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + Vf \rangle = \langle f, -\frac{\hbar^2}{2m} \frac{d^2 \bar{g}}{dx^2} + V\bar{g} \rangle$: c'est la difficulté classique liée à la multiplication des distributions. $\square$

Exemple XII.8 Plus généralement, sur $\mathbb{L}_2(\mathbb{R}^3)$ au lieu de $\mathbb{L}_2(\mathbb{R})$, l'opérateur $A = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{d^2}{dy^2} + \frac{d^2}{dz^2} + V(x, y, z)$ (même argumentation). $\square$

Ⓡ La relation $A = A^*$ inclut le fait que les domaines sont égaux : $\mathcal{D}_A = \mathcal{D}_{A^*}$. Cela implique que les domaines sont maximaux : un opérateur auto-adjoint est *automatiquement* fermé. Si on restreint A à un domaine plus petit que $\mathcal{D}_A$ (mais toujours dense), l'opérateur ainsi diminué sera préfermé et non plus fermé ; son adjoint sera cependant toujours le même. Beaucoup d'auteurs introduisent pour cela la notion d'opérateur *essentiellement* auto-adjoint : c'est un opérateur préfermé, dont l'extension fermée est auto-adjointe. Pour éviter l'inflation terminologique, on supposera que les opérateurs sont toujours considérés implicitement dans leur extension maximale.

On introduit aussi la notion d'opérateur *symétrique* : c'est un opérateur dont l'adjoint est une extension : A est symétrique si A^* prolonge A ; ainsi un opérateur symétrique est forcément préfermé. Un opérateur symétrique est simplement un opérateur préfermé A pour lequel on a $\forall f \in \mathcal{H}, \forall g \in \mathcal{H}, \langle Af \mid g \rangle = \langle f \mid Ag \rangle$. Un opérateur essentiellement auto-adjoint est symétrique, mais malheureusement il y a des opérateurs symétriques non essentiellement auto-adjoints. On est bien obligé de tenir compte de ces difficultés dans une théorie générale et abstraite des opérateurs, mais pour ce qui nous intéresse on peut s'en passer, car les opérateurs particuliers que nous considérerons seront, eux, auto-adjoints.

Théorème XII.5 Le spectre d'un opérateur auto-adjoint est entièrement contenu dans l'axe réel.

Preuve Si A est auto-adjoint, $\langle Af \mid f \rangle$ est toujours réel (quel que soit $f \in \mathcal{H}$). En effet, le produit scalaire $\langle f \mid g \rangle$ sur un espace de Hilbert complexe a la propriété que $\overline{\langle g \mid f \rangle} = \langle f \mid g \rangle$. D'autre part, si A est auto-adjoint, on a aussi $\langle Af \mid f \rangle = \langle f \mid Af \rangle = \overline{\langle Af \mid f \rangle}$, d'où la conclusion. Maintenant, λ étant un nombre complexe quelconque, on peut écrire :

$$\langle \lambda f - Af \mid f \rangle = \lambda \|f\|^2 - \langle Af \mid f \rangle \quad (\text{XII.41})$$

Ce qui, en ne prenant que la partie imaginaire, donne :

$$\Im \{ \langle \lambda f - Af \mid f \rangle \} = \Im \{ \lambda \} \|f\|^2 \quad (\text{XII.42})$$

(10) En effet :

$$\forall \varphi \in \mathcal{S}(\mathbb{R}), \quad \langle f'', \varphi \rangle = - \int f(x) \varphi''(x) dx$$

est toujours bien défini.

D'autre part, d'après l'inégalité de Schwarz :

$$|\langle \lambda f - Af | f \rangle| \leq \|\lambda f - Af\| \cdot \|f\| \quad (\text{XII.43})$$

En combinant (XII.42) et (XII.43) et en utilisant le fait que le module d'un nombre complexe (ici $\langle \lambda f - Af | f \rangle$) est plus grand que sa partie imaginaire, on obtient l'inégalité :

$$\|\lambda f - Af\| \cdot \|f\| \geq |\Im\{\lambda\}| \|f\|^2 \quad (\text{XII.44})$$

d'où on déduit en simplifiant par $\|f\|$

$$\|\lambda f - Af\| \geq |\Im\{\lambda\}| \|f\| \quad (\text{XII.45})$$

La simplification n'est légitime que si $\|f\| \neq 0$, mais si $\|f\| = 0$, (XII.45) sera vraie pour une autre raison. Cette inégalité prouve déjà que si $\Im\{\lambda\} \neq 0$, l'opérateur $\lambda I - A$ est injectif. Pour montrer qu'il est surjectif (toujours si $\Im\{\lambda\} \neq 0$), montrons que son image $\mathcal{R}$ est fermée et que son supplémentaire orthogonal est nul (un sous-espace de $\mathcal{H}$ qui est fermé et de supplémentaire orthogonal nul, est forcément $\mathcal{H}$ tout entier). Les éléments de $\mathcal{R}$ sont les $\lambda f - Af$ lorsque f parcourt $\mathcal{D}_A$; soit donc g tel que $\langle \lambda f - Af | g \rangle = 0 \forall f \in \mathcal{D}_A$. Si $g \in \mathcal{D}_A$, on aura en particulier (en prenant $f = g$) $\langle \lambda g - Ag | g \rangle = 0$, ce qui d'après (XII.45) implique que $g = 0$. On a ainsi prouvé que tout $g \in \mathcal{D}_A$, orthogonal à $\mathcal{R}$, est nul. Comme $\mathcal{D}_A$ est dense, cela suffit. D'autre part, $\mathcal{R}$ est fermé : si f_n est une suite de $\mathcal{D}_A$ telle que $g_n = \lambda f_n - Af_n$ tend dans $\mathcal{H}$ vers une limite g , alors g_n est une suite de Cauchy (puisque'elle converge). D'après l'inégalité (XII.45) on aura $\|g_n - g_m\| \geq |\Im\{\lambda\}| \|f_n - f_m\|$, qui prouve que f_n est aussi une suite de Cauchy et a donc une limite f dans $\mathcal{H}$. Ainsi, $g_n = \lambda f_n - Af_n$ et f_n ont toutes deux une limite dans $\mathcal{H}$, donc aussi Af_n et puisque l'opérateur A est fermé, on a nécessairement $g = \lambda f - Af$, ce qui prouve bien que $\mathcal{R}$ est fermé et par conséquent égal à $\mathcal{H}$ (On aurait pu énoncer séparément que l'image d'un opérateur fermé est fermée).

On voit ainsi que pour $\Im\{\lambda\} \neq 0$, l'opérateur $\lambda I - A$ est injectif et surjectif. Enfin, le fait que l'inverse est continu résulte directement de l'inégalité (XII.45) : $\|(\lambda I - A)^{-1}g\| \leq \|g\|/|\Im\{\lambda\}|$.

Définition XII.7 On appelle unitaire un opérateur inversible qui laisse invariant le produit scalaire. L'opérateur U est unitaire s'il est surjectif et si pour tous f, g dans $\mathcal{H}$, on a $\langle Uf | Ug \rangle = \langle f | g \rangle$.

Un opérateur unitaire est donc forcément continu, puisqu'il conserve la norme. Pour la même raison, il est forcément injectif (si $\|Uf\| = \|f\|$, il est clair que $\|Uf\| = 0 \Rightarrow \|f\| = 0$). Mais contrairement à ce qui se passe en dimension finie, cela ne garantit pas la surjectivité. L'adjoint U^* d'un opérateur unitaire U vérifie les relations $U^*U = UU^* = I$, autrement dit un opérateur unitaire est inversible et son inverse est unitaire. Les opérateurs unitaires sur les espaces de Hilbert généralisent les matrices unitaires.

Parmi les opérateurs déjà rencontrés, les transformations de Fourier symétriques, les translations, les multiplications par des fonctions de module 1, sont des opérateurs unitaires :

- $\mathcal{F}_\alpha$, qui à la fonction $f \in \mathbb{L}_2(\mathbb{R})$ associe la fonction :

$$g_\alpha(\xi) = \sqrt{\frac{\alpha}{2\pi}} \int_{-\infty}^{+\infty} e^{i\alpha x\xi} f(x) dx \quad (\text{XII.46})$$

est unitaire (pour tout α réel non nul) d'après la relation de Parseval; les cas courants sont $\alpha = 1/\hbar$ (transformation de Fourier de la mécanique quantique) et $\alpha = 1$ (transformation de Fourier symétrique usuelle);

- la translation T_a , qui à toute fonction $f \in \mathbb{L}_2(\mathbb{R})$ associe la fonction $f_a(x) = f(x - a)$, est unitaire, comme le montre un changement de variable évident dans l'intégrale :

$$\langle T_a f | T_a g \rangle = \int_{-\infty}^{+\infty} f(x - a) \overline{g(x - a)} dx \quad (\text{XII.47})$$

- plus généralement, étant donné une fonction $y = \rho(x)$, croissante sur $\mathbb{R}$ et différentiable ainsi que son inverse, la transformation qui à toute fonction $f \in \mathbb{L}_2(\mathbb{R})$ associe la fonction $f_\rho(x) = f(\rho(x)) \cdot \sqrt{\rho'(x)}$, est unitaire ;
- la multiplication par une fonction $V(x) = e^{i\varphi(x)}$ (où la fonction φ est réelle) est aussi un opérateur unitaire, puisque :

$$\langle Vf | Vg \rangle = \int_{-\infty}^{+\infty} e^{i\varphi(x)} f(x) \overline{e^{i\varphi(x)} g(x)} dx = \int_{-\infty}^{+\infty} e^{i\varphi(x)} e^{-i\varphi(x)} f(x) \overline{g(x)} dx \quad (\text{XII.48})$$

Théorème XII.6 Le spectre d'un opérateur unitaire est entièrement contenu dans le cercle $\{|z| = 1\}$.

Preuve En tous points semblable à celle du théorème XII.5. On la propose en exercice. •

Exercice XII.7

1. Montrer que $\langle \lambda f - Uf | \lambda f - Uf \rangle = (|\lambda|^2 + 1) \|f\|^2 + 2\Re\{\lambda \langle f | Uf \rangle\}$.
2. Majorer $|\Re\{\lambda \langle f | Uf \rangle\}|$ à l'aide de l'inégalité de Schwarz et en déduire que :

$$\|\lambda f - Uf\|^2 \geq (|\lambda| - 1)^2 \|f\|^2 \quad (\text{XII.49})$$

3. Montrer que si $|\lambda| \neq 1$, $\lambda I - U$ et $\bar{\lambda}I - U^{-1}$ sont injectifs.
4. Pour montrer que si $|\lambda| \neq 1$, $\lambda I - U$ est surjectif, montrer d'abord que $(\lambda I - U)\mathcal{H}$ (image de $\mathcal{H}$ par $\lambda I - U$) est fermé, en considérant des suites f_n telles que $(\lambda I - U)f_n$ converge et en utilisant 2 ; puis montrer en utilisant 3 que si g est orthogonal à tous les $(\lambda I - U)f$, il est forcément nul (en cas de difficulté, lire la démonstration du théorème XII.5).
5. Utiliser à nouveau 2 pour montrer que l'inverse de $\lambda I - U$ est continu.
6. Peut-on étendre ce procédé à d'autres cas ? ■

XII.6 Fonctions d'opérateurs

Dans les espaces de dimension finie, on définit aisément les puissances d'un endomorphisme A : la n^{e} puissance de A , soit A^n , est l'endomorphisme obtenu en itérant n fois A . On peut aussi considérer des polynômes de A : si $P(x) = \sum a_n x^n$, l'endomorphisme $P(A)$ sera $\sum a_n A^n$.

Il est également possible de définir des séries de puissances d'un endomorphisme, mais il faut préciser dans quel sens la série converge. Ainsi, on peut définir l'exponentielle d'un endomorphisme :

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} \quad (\text{XII.50})$$

La convergence s'interprète comme suit. Si A est un endomorphisme sur l'espace euclidien de dimension finie $\mathcal{E}$, il existe une constante M_A telle que $\forall X \in \mathcal{E}, \|AX\| \leq M_A \|X\|$ (rappel : il suffit de prendre pour M_A le maximum de $\|AX\|$ sur la boule $B = \{X \in \mathcal{E} \mid \|X\| \leq 1\}$). L'existence de cette constante M_A correspond à la continuité des opérateurs linéaires sur un espace de dimension finie. Il est clair que si on itère l'endomorphisme A , on aura $\|A^2 X\| \leq M_A \|AX\| \leq M_A^2 \|X\|$, puis $\|A^3 X\| \leq M_A \|A^2 X\| \leq M_A^3 \|X\|$ etc. de sorte que $\|A^n X\| \leq M_A^n \|X\|$. Si on définit $e^A X$ comme la somme de la série $\sum_{n \geq 0} \frac{1}{n!} A^n X$, on voit

que cette série converge car le terme général a pour norme $\frac{1}{n!}\|A^n X\|$, ce qui est majoré par $\frac{1}{n!}M_A^n\|X\|$.

Tout ce qui vient d'être fait se transpose tel quel aux opérateurs *continus* sur les espaces de Hilbert de dimension infinie ; en effet, l'inégalité $\|Af\| \leq M_A\|f\|$ est satisfaite par les opérateurs continus, donc la série $\sum \frac{1}{n!}A^n f$ est normalement convergente. Ce qui implique, puisqu'un espace de Hilbert est complet, que cette série est convergente et définit le vecteur $e^A f$. L'opérateur e^A est alors celui qui transforme f en $e^A f$.

La série $\sum \frac{1}{n!}x^n$ a un rayon de convergence infini, ce qui a évité de se poser des questions sur la valeur de la constante M_A . Si on considère une série $\sum a_n x^n$ ayant un rayon de convergence fini R , il faudra que $M_A < R$ pour définir la somme $\sum a_n A^n$. Voici un exemple :

Théorème XII.7 Si A est un opérateur continu sur un espace de Hilbert $\mathcal{H}$, tel que $\forall f \in \mathcal{H}$, $\|Af\| \leq M_A\|f\|$, avec $M_A < 1$, alors l'opérateur $I - A$ est inversible.

Preuve On a déjà démontré et utilisé ce résultat (voir la démonstration du théorème XII.2). Tout l'argument consiste simplement à remarquer que l'inverse de $I - A$ est la série $\sum_{n \geq 0} A^n$. •

Exercice XII.8 En utilisant ce même argument, montrer que le spectre d'un opérateur continu est borné. ■

On peut généraliser ce résultat :

Théorème XII.8 Soit $\Phi(z)$ une fonction analytique dans le disque $\{|z| < R\}$ et soit A un opérateur continu sur un espace de Hilbert $\mathcal{H}$, tel que $\forall f \in \mathcal{H}$, $\|Af\| \leq M_A\|f\|$, avec $M_A < R$. Alors on peut définir l'opérateur $B = \Phi(A)$. Cet opérateur commute avec A : $AB = BA$.

Preuve La fonction $\Phi(z)$ étant analytique dans le disque $\{|z| < R\}$, est la somme d'une série $\sum a_n z^n$ de rayon de convergence $\geq R$, donc B sera défini par $Bf = \sum a_n A^n f$. Comme toutes les puissances A^n commutent avec A , il en est de même de B . •

Le fait que les opérateurs commutent a une signification bien précise en mécanique quantique : les opérateurs représentent des grandeurs physiques et s'ils commutent ces grandeurs sont compatibles (observables simultanément). Il est logique que si une grandeur b est le carré ou l'exponentielle d'une autre grandeur a , c'est que a et b sont simultanément observables. D'autre part, si on les observe effectivement, on sait, d'après les principes de la mécanique quantique, que les valeurs observées sont nécessairement des valeurs propres de l'opérateur ; la consistance logique de la mécanique quantique exige donc aussi que si $B = \Phi(A)$, on ait la même relation pour les valeurs propres $b = \Phi(a)$. C'est bien le cas :

Théorème XII.9 Soit A un opérateur continu sur un espace de Hilbert $\mathcal{H}$ et $B = \Phi(A)$, Φ étant comme dans le théorème XII.8. Alors les valeurs propres de B sont les images par Φ des valeurs propres de A et les vecteurs propres sont communs. Plus généralement, le spectre de B est l'image par Φ du spectre de A .

Preuve Si λ est une valeur propre, soit φ un vecteur propre associé. On a donc $A\varphi = \lambda\varphi$, d'où $A^n\varphi = \lambda^n\varphi$; en reportant dans la série on aura $\Phi(A)\varphi = \Phi(\lambda)\varphi$. Cela montre bien que $\Phi(\lambda)$ est valeur propre de $\Phi(A)$ avec le même vecteur propre associé.

Pour les autres valeurs $\lambda \in \mathbb{C}$, introduisons la fonction :

$$\chi(z) = \frac{\Phi(\lambda) - \Phi(z)}{\lambda - z}. \quad (\text{XII.51})$$

cette fonction est analytique dans le même domaine que $\Phi(z)$; on peut donc considérer l'opérateur $\chi(A)$ et on a la relation $\Phi(\lambda)I - \Phi(A) = (\lambda I - A)\chi(A)$, qui sera vraie aussi pour les opérateurs :

$$\Phi(\lambda)I - \Phi(A) = (\lambda I - A)\chi(A) = \chi(A)(\lambda I - A) \quad (\text{XII.52})$$

Puisque $\chi(z)$ est analytique dans un disque de rayon $> M_A$, l'opérateur $\chi(A)$ est continu. La relation (XII.52) montre que si l'opérateur $\Phi(\lambda)I - \Phi(A)$ a un inverse continu, il en sera de même de $(\lambda I - A)$, puisqu'il suffira de poser :

$$(\lambda I - A)^{-1} = [\Phi(\lambda)I - \Phi(A)]^{-1} \chi(A) \quad (\text{XII.53})$$

Donc le spectre de $\Phi(A)$ contient l'image par Φ du spectre de A . Pour vérifier que non seulement il la contient, mais qu'il lui est identique, supposons un instant qu'il existe un λ dans le spectre de $\Phi(A)$ qui ne soit pas dans l'image ; cela signifie que λ n'est égal à aucun $\Phi(z)$ tel que z soit dans le spectre de A . Donc la fonction $\psi(z) = 1/(\lambda - \Phi(z))$ est analytique pour tout z appartenant au spectre de A et donc aussi dans un voisinage Ω de ce spectre. Quoique Ω ne soit pas forcément le disque $|z| < R$, on peut quand même définir l'opérateur $\psi(A)$, qui sera forcément l'inverse continu de $\lambda I - \Phi(A)$. •

Jusqu'ici, seuls des opérateurs continus ont été considérés. Toutefois, comme cela a déjà été dit, les opérateurs qui interviennent en mécanique quantique ne sont pas forcément continus. On ne peut pas définir les puissances A^n d'un opérateur fermé discontinu aussi facilement que celles d'un opérateur continu, car l'itération exige que l'image du domaine $\mathcal{D}_A$ soit incluse dans $\mathcal{D}_A$. Si tel n'est pas le cas, on peut poser $\mathcal{D}_{A^2} = \{f \in \mathcal{H} \mid Af \in \mathcal{D}_A\}$, puis $\mathcal{D}_{A^3} = \{f \in \mathcal{H} \mid Af \in \mathcal{D}_{A^2}\}$ et ainsi de suite. Le fait que $\mathcal{D}_A$ soit dense n'entraîne pas automatiquement que $\mathcal{D}_{A^2}$, $\mathcal{D}_{A^3}$ etc. le soient aussi. Même si c'était vrai, cela ne permettrait encore de définir que des polynômes ; pour définir des séries infinies (par exemple l'exponentielle), il faudrait en outre que l'intersection de tous les $\mathcal{D}_{A^n}$ soit dense.

Pour surmonter cette difficulté, il faut d'abord se restreindre aux opérateurs pour lesquels ça marche. Il se trouve que les opérateurs auto-adjoints font partie de ces privilégiés.

Une approche possible consiste à se ramener aux opérateurs continus par une transformation adéquate. Par exemple, J. von Neumann a proposé d'utiliser la *transformation de Cayley*. Si A est un opérateur auto-adjoint non nécessairement continu, son spectre est réel donc $(iI - A)^{-1}$ est continu. Le *transformé de Cayley* de A est alors l'opérateur $U = (iI + A)(iI - A)^{-1}$, qui est unitaire :

Théorème XII.10 Si A est auto-adjoint, $U = (iI + A)(iI - A)^{-1}$ est unitaire (transformation de Cayley). Inversement, si U est un opérateur unitaire tel que $U + I$ soit injectif et $\mathcal{D} = (U + I)\mathcal{H}$ dense, alors l'opérateur $A = -i(U - I)(U + I)^{-1}$, défini sur $\mathcal{D}$, est auto-adjoint et a pour transformé de Cayley U .

Preuve Remarquons d'abord que l'opérateur U est bien défini : puisque A est auto-adjoint, son spectre est inclus dans l'axe réel (théorème XII.5), donc $iI - A$ est inversible et d'inverse continu ; cet inverse transforme n'importe quel $f \in \mathcal{H}$ en un élément de $\mathcal{D}_A$, auquel on peut donc appliquer $iI + A$. Comme $iI + A$ est lui aussi inversible, il appliquera à son tour $\mathcal{D}_A$ *surjectivement* sur $\mathcal{H}$. Il reste encore à vérifier la conservation du produit scalaire.

Pour cela, soient $f, g \in \mathcal{H}$ et posons $p = Uf$ et $q = Ug$; il s'agit de prouver que $\langle p \mid q \rangle = \langle f \mid g \rangle$. D'après la définition de U , on voit que $p = Uf \Leftrightarrow (iI + A)^{-1}p = (iI - A)^{-1}f$ et $q = Ug \Leftrightarrow (iI + A)^{-1}q = (iI - A)^{-1}g$. Posons donc $X = (iI + A)^{-1}p = (iI - A)^{-1}f$ et $Y = (iI + A)^{-1}q = (iI - A)^{-1}g$; X et Y sont ainsi des

éléments du domaine $\mathcal{D}_A$ et on a $f = (iI - A)X$, $g = (iI - A)Y$, $p = (iI + A)X$, $q = (iI + A)Y$. Calculons les produits scalaires $\langle p | q \rangle$ et $\langle f | g \rangle$:

$$\begin{aligned}
 \langle p | q \rangle &= \langle iX + AX | iY + AY \rangle \\
 &= \langle iX | iY \rangle + \langle iX | AY \rangle + \langle AX | iY \rangle + \langle AX | AY \rangle \\
 &= \langle X | Y \rangle + i\langle X | AY \rangle - i\langle AX | Y \rangle + \langle AX | AY \rangle \\
 \langle f | g \rangle &= \langle iX - AX | iY - AY \rangle \\
 &= \langle iX | iY \rangle - \langle iX | AY \rangle - \langle AX | iY \rangle + \langle AX | AY \rangle \\
 &= \langle X | Y \rangle - i\langle X | AY \rangle + i\langle AX | Y \rangle + \langle AX | AY \rangle
 \end{aligned} \tag{XII.54}$$

Comme A est auto-adjoint, le second terme est dans les deux cas annulé par le troisième, d'où :

$$\begin{aligned}
 \langle p | q \rangle &= \langle X | Y \rangle + \langle AX | AY \rangle \\
 \langle f | g \rangle &= \langle X | Y \rangle + \langle AX | AY \rangle
 \end{aligned} \tag{XII.55}$$

ce qui montre bien que $\langle p | q \rangle = \langle f | g \rangle$.

Pour exprimer A en fonction de U , supposons d'abord que si A est donné et $U = (iI + A)(iI - A)^{-1}$. Revenons aux deux relations $f = (iI - A)X$ et $Uf = p = (iI + A)X$. On peut les écrire sous la forme $f = iX - AX$, $Uf = iX + AX$, donc en les ajoutant et soustrayant l'une de l'autre $Uf + f = 2iX$ et $Uf - f = 2AX$, d'où $2X = -i(U + I)f$ et $2AX = -i(U - I)(U + I)^{-1} \cdot 2X$; ceci étant vrai quel que soit $X \in \mathcal{D}_A$, on voit que $U + I$ est un opérateur qui envoie bijectivement $\mathcal{H}$ sur $\mathcal{D}_A$ et on en déduit $A = -i(U - I)(U + I)^{-1}$, ce qui montre aussi que $(U + I)^{-1}$ (qui envoie bijectivement $\mathcal{D}_A$ sur $\mathcal{H}$) est discontinu si A l'est).

Cette formule d'inversion vient d'être obtenue en supposant A donné et $U = (iI + A)(iI - A)^{-1}$, ce qui veut dire que si U est obtenu comme $(iI + A)(iI - A)^{-1}$, A sera égal à $-i(U - I)(U + I)^{-1}$. Si maintenant U est un opérateur unitaire quelconque donné *a priori*, on voit que l'opérateur auto-adjoint A tel que $U = (iI + A)(iI - A)^{-1}$ n'existe pas forcément : il faut pour cela que $U + I$ soit injectif et envoie bijectivement $\mathcal{H}$ sur un sous-espace dense de $\mathcal{H}$ (par exemple si $U = -I$ cela ne marche pas!) Si tel est le cas, soient X et Y deux éléments quelconques de $\mathcal{H}$. Puisque U est unitaire, on aura :

$$\begin{aligned}
 \langle UX - X | UY + Y \rangle &= \langle UX | UY \rangle + \langle UX | Y \rangle - \langle X | UY \rangle - \langle X | Y \rangle \\
 &= \langle UX | Y \rangle - \langle X | UY \rangle \\
 \langle UX + X | UY - Y \rangle &= \langle UX | UY \rangle - \langle UX | Y \rangle + \langle X | UY \rangle - \langle X | Y \rangle \\
 &= -\langle UX | Y \rangle + \langle X | UY \rangle
 \end{aligned} \tag{XII.56}$$

Autrement dit $\langle UX - X | UY + Y \rangle = -\langle UX + X | UY - Y \rangle$, ou encore $\langle -i(UX - X) | UY + Y \rangle = \langle UX + X | i(UY - Y) \rangle$. Si on pose $f = UX + X$, $f \in \mathcal{D}$ d'après l'hypothèse et on aura $Af = -i(UX - X)$, de sorte que pour tout $Y \in \mathcal{H}$:

$$\langle Af | UY + Y \rangle = \langle f | i(UY - Y) \rangle \tag{XII.57}$$

Cela prouve que la fonctionnelle $f \mapsto \langle Af | UY + Y \rangle$ est continue pour tout $Y \in \mathcal{H}$ et donc que la fonctionnelle $f \mapsto \langle Af | g \rangle$ est continue pour tout $g \in \mathcal{D}$; ce qui revient à dire que $\mathcal{D}$ est inclus dans $\mathcal{D}_{A^*}$. Mais en échangeant les rôles de X et Y on arriverait à la conclusion que $\mathcal{D}_{A^*}$ est inclus dans $\mathcal{D}$. •

La transformation de Cayley est aussi une fonction d'opérateur : on peut dire que $U = f(A)$ avec $f(z) = (i + z)/(i - z)$. Contrairement aux transformations envisagées avant, celle-ci s'applique à un opérateur non nécessairement continu. Pour la définir, on a procédé directement, sans passer par l'intermédiaire de séries entières qui doivent converger. Il serait possible de faire la même chose pour d'autres transformations. Par exemple, si on prend $f(z) = 1/(\lambda - z)$, ou $f(z) = (\lambda + z)/(\lambda - z)$, ou encore $f(z) = (\lambda + z)^2/(\lambda - z)^2$ etc. (avec λ non réel) on obtient des opérateurs $f(A)$ qui sont continus, bien que A ne le soit pas. On pourrait prendre aussi $f(z) = z/(\lambda - z)$, qui donnerait un opérateur $f(A)$ non continu, mais cependant bien défini.

On peut vérifier directement pour ces transformations, que les valeurs propres et les spectres sont transformés comme le dit le théorème; de même, les opérateurs A et $f(A)$ commutent.

Si maintenant on prend un opérateur unitaire U et une fonction f analytique dans un disque de centre 0 et de rayon $R > 1$, la fonction $f(U)$ est bien définie par sa série entière. Soit aussi la fonction $g(z) = f\left(\frac{i+z}{i-z}\right)$. Il est équivalent de dire que f est analytique dans un disque de rayon R , ou que g est analytique dans le domaine D_R des points $z = x + iy$ du plan tels que $2(R^2 + 1)y < (R^2 - 1)(x^2 + y^2 + 1)$, qui est l'extérieur d'un cercle comme expliqué sur la figure XII.1.

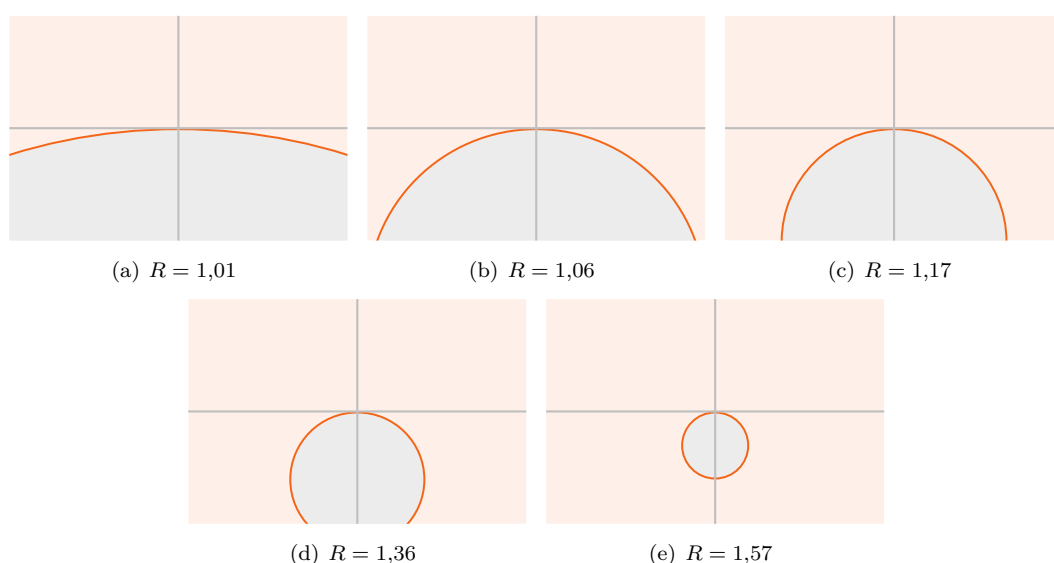


Figure XII.1 – Image du disque de centre 0 et de rayon R : c'est l'extérieur d'un cercle entièrement situé sous l'axe réel

Si $R = 1$, ce cercle dégénère en la droite réelle et D_R est le demi-plan $\{y > 0\}$; si $R > 1$, ce cercle est situé sous la droite réelle et donc le domaine D_R contient le demi-plan $\{y > 0\}$. La transformation homographique $z \mapsto w = (i+z)/(i-z)$ est une transformation conforme du domaine D_R sur le disque de centre 0 et de rayon R , privé du point $w = -1$. Ce point $w = -1$ correspond dans D_R au point à l'infini).

Si $w = (i+z)/(i-z)$, $g(z) = f(w)$. Par conséquent, si A est un opérateur auto-adjoint non nécessairement continu et si g est une fonction analytique dans D_R et sans singularité à l'infini (c'est-à-dire que $g(z)$ doit avoir une limite finie quand z tend dans D_R vers l'infini), on peut définir $g(A)$ sans difficulté en disant que c'est $f(U)$. Si g a une singularité à l'infini, on peut aussi poser $g(A) = f(U)$, si on parvient à résoudre la difficulté par un moyen direct (mais on ne peut se prononcer *a priori*). Par exemple, on peut définir e^{iA} en disant que c'est $f(U)$, où $f(w) = e^{(1-w)/(w+1)}$; mais cette fonction a un point singulier essentiel en $w = -1$, donc il faut prendre des précautions. On verra plus loin ce qui se passe sur des exemples.

XII.7 Groupes unitaires

Le problème mathématique de la mécanique quantique non relativiste est de résoudre les équations de Schrödinger :

$$i\hbar \frac{\partial \psi}{\partial t} = A\psi \quad (\text{XII.58})$$

de même que le problème mathématique de la mécanique classique est de résoudre les équations de Newton $F = m\gamma$. L'opérateur A est appelé l'hamiltonien et représente l'énergie (ses valeurs propres sont les états d'énergie du système) ; c'est, *par principe*, un opérateur auto-adjoint. Il opère sur l'espace $\mathcal{H} = \mathbb{L}_2(\mathbb{R})$ et est en général discontinu.

Si on considérait (XII.58) sur un espace de dimension finie, c'est-à-dire si A était un endomorphisme sur $\mathcal{E} = \mathbb{C}^n$, les fonctions ψ seraient des vecteurs à n composantes dépendant du temps et l'équation (XII.58) serait un système de n équations différentielles linéaires du premier ordre. Si n était égal à 1, A serait simplement la multiplication par un nombre réel a et (XII.58) serait *une* équation différentielle linéaire du premier ordre ; ψ serait une fonction à valeurs complexes du temps t . La solution serait alors $\psi(t) = \psi(0) e^{-\frac{i}{\hbar}at}$. Pour $n > 1$, (XII.58) serait un système d'équations, mais on pourrait toujours écrire sa solution sous la forme :

$$\psi(t) = e^{-\frac{i}{\hbar}tA} \psi(0) \quad (\text{XII.59})$$

Pour calculer effectivement l'opérateur $e^{-\frac{i}{\hbar}tA}$, il faut diagonaliser la matrice A (on peut toujours diagonaliser une matrice auto-adjointe), mais si on se contente d'une réponse de principe, l'exponentielle $U_t = e^{-\frac{i}{\hbar}tA}$ de la matrice exprime de manière univoque, quoique abstraite, la solution du système d'équations.

On remarque que les matrices U_t forment un groupe : $U_{s+t} = U_s U_t$, $U_0 = I$, $U_t^{-1} = U_{-t}$. La solution d'un système d'équations différentielles linéaires du premier ordre est donc donnée par un groupe d'endomorphismes. Dans le cas du système (XII.58), du fait que A est une matrice auto-adjointe et du coefficient $i = \sqrt{-1}$ de $\frac{\partial \psi}{\partial t}$, les endomorphismes U_t sont unitaires, ce qui ne serait pas le cas pour le système $d\psi/dt = A\psi$.

Le but de la théorie des opérateurs sur les espaces de Hilbert est d'étendre ces considérations aux espaces de dimension infinie tels que les espaces $\mathbb{L}_2(\mathbb{R}^n)$.

Dans le cas de l'équation (XII.58), l'opérateur est en général :

$$A = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + V(x, y, z) \quad (\text{XII.60})$$

et il faut pouvoir préciser le sens de $U_t = e^{-\frac{i}{\hbar}tA}$. Si A était continu, il suffirait de le définir par la série $\sum \frac{1}{n!} \left(-\frac{i}{\hbar}tA \right)^n$, mais justement A n'est pas continu. Il n'existe pas de théorème du type "tout opérateur auto-adjoint possède une exponentielle" mais on peut construire le groupe U_t dans la plupart des cas particuliers utiles. Pour fixer le langage, voici quelques définitions.

Définition XII.8 On appelle groupe unitaire une famille d'opérateurs U_t , paramétrée par $t \in \mathbb{R}$, telle que :

- les U_t sont tous unitaires ;
- $\forall s, t \in \mathbb{R}, U_t U_s = U_{s+t}$;
- pour $t = 0, U_t = I$;
- $\forall f \in \mathcal{H}$, la fonction $t \mapsto U_t f$ est continue sur $\mathbb{R}$, c'est-à-dire que pour tout $t \in \mathbb{R}$, $\|U_{t+h} f - U_t f\|$ tend vers zéro quand h tend vers zéro.

Pour la condition de continuité, il suffit de l'exiger en $t = 0$; la propriété de groupe la garantit alors sur tout $\mathbb{R}$: en effet, $\|U_{t+h} f - U_t f\| = \|(U_{t+h} - U_t)f\| = \|U_t (U_h f - f)\| = \|U_h f - f\|$.

Définition XII.9 Le générateur infinitésimal du groupe est l'opérateur (non nécessairement continu) G défini comme suit :

- le domaine $\mathcal{D}_G$ est l'ensemble des éléments $f \in \mathcal{H}$ pour lesquels $\langle (U_t f - f)/it \mid g \rangle$ a pour tout $g \in \mathcal{H}$ une limite (dans $\mathbb{C}$) lorsque t tend vers zéro ;
- Gf est l'élément de $\mathcal{H}$ tel que, $\forall g \in \mathcal{H}$, $\langle Gf \mid g \rangle$ est cette limite.

On écrit souvent $U_t = e^{itG}$ pour un tel groupe.

Théorème XII.11 Le générateur infinitésimal d'un groupe unitaire est toujours auto-adjoint.

Preuve La principale difficulté est de prouver que $\mathcal{D}_G$ est dense. Voyons d'abord le reste.

La propriété de groupe implique que $U_t^{-1} = U_{-t}$, donc $\forall f, g \in \mathcal{H}$, $\langle U_t f \mid g \rangle = \langle f \mid U_{-t} g \rangle$ et donc aussi :

$$\left\langle \frac{U_t f - f}{it} \mid g \right\rangle = \left\langle f \mid \frac{U_{-t} g - g}{-it} \right\rangle \quad (\text{XII.61})$$

Cette égalité, étant vraie $\forall t$, reste vraie quand on passe à la limite. Si $f \in \mathcal{D}_G$, la limite du premier membre est par définition $\langle Gf \mid g \rangle$ et le second membre tend alors aussi vers cette limite, ce qui d'après la définition signifie que g appartient à $\mathcal{D}_G$ et que $\langle Gf \mid g \rangle = \langle f \mid Gg \rangle$; ce raisonnement restant identique à lui-même lorsqu'on échange f et g , ou lorsqu'on remplace t par $-t$, cela montre que $\mathcal{D}_G = \mathcal{D}_{G^*}$ et $G = G^*$.

Pour montrer que $\mathcal{D}_G$ est dense, on va construire une suite d'opérateurs L_n qui converge vers I et tels que leurs images soient contenues dans $\mathcal{D}_G$. La densité de $\mathcal{D}_G$ résulte alors simplement du raisonnement suivant : pour tout $f \in \mathcal{H}$, $f_n = L_n f \in \mathcal{D}_G$; d'autre part f_n tend vers f puisque $L_n f$ tend vers f . Ce procédé astucieux est emprunté à K. Yosida⁽¹¹⁾ avec quelques modifications.

L'opérateur L_n est défini ainsi : pour tous $f, g \in \mathcal{H}$, on pose :

$$\langle L_n f \mid g \rangle = n \int_0^\infty e^{-ns} \langle U_s f \mid g \rangle ds \quad (\text{XII.62})$$

La convergence de l'intégrale est garantie par l'inégalité de Schwarz $|\langle U_s f \mid g \rangle| \leq \|f\| \|g\|$ et l'opérateur L_n est bien défini par la donnée de tous les produits scalaires $\langle L_n f \mid g \rangle$. L'idée implicite est la suivante : il faut imaginer que $U_t = e^{itG}$ et l'opérateur L_n est la fonction d'opérateur $\Phi(G)$ avec :

$$\Phi(x) = n \int_0^\infty e^{-ns} e^{isx} dx = \frac{n}{n - ix} \quad (\text{XII.63})$$

Cette fonction tend bien vers 1 quand n tend vers l'infini, donc "on peut s'attendre" à ce que L_n tende vers I . Bien entendu, il faut démontrer ce qui n'est ainsi que deviné. Mais prouvons d'abord que pour tout

(11) K YOSIDA, *Functional Analysis* Springer-Verlag, Berlin, 1968, p. 237–238.

$h \in \mathcal{H}$, $f = L_n h \in \mathcal{D}_G$. Il faut montrer que pour tout $g \in \mathcal{H}$, $\langle \frac{1}{i\varepsilon} (U_\varepsilon f - f) | g \rangle$ tend vers une limite quand ε tend vers zéro. Or :

$$\begin{aligned} \langle \frac{1}{i\varepsilon} (U_\varepsilon f - f) | g \rangle &= \frac{n}{i\varepsilon} \int_0^\infty e^{-ns} \langle U_{\varepsilon+s} h - U_s h | g \rangle ds \\ &= \frac{e^{n\varepsilon} - 1}{i\varepsilon} n \int_\varepsilon^\infty e^{-ns} \langle U_s h | g \rangle ds - \frac{n}{i\varepsilon} \int_0^\varepsilon e^{-ns} \langle U_s h | g \rangle ds \\ &= \frac{e^{n\varepsilon} - 1}{i\varepsilon} \left\{ \langle L_n h | g \rangle - n \int_0^\varepsilon e^{-ns} \langle U_s h | g \rangle ds \right\} - \frac{n}{i\varepsilon} \int_0^\varepsilon e^{-ns} \langle U_s h | g \rangle ds \end{aligned} \quad (\text{XII.64})$$

La continuité de la fonction $t \mapsto \langle U_t h | g \rangle$ (qui est exigée par la définition) implique que l'intégrale entre accolades tend vers zéro quand ε tend vers zéro et le terme à l'extérieur des accolades tend vers $-n \langle U_0 h | g \rangle$; enfin, le facteur $(e^{n\varepsilon} - 1)/i\varepsilon$ tend vers $-in$ (quand ε tend vers zéro). Par conséquent, la limite du tout est $-in [\langle L_n h - h | g \rangle]$. On a ainsi prouvé que cette limite existe, donc que $L_n h$ est dans $\mathcal{D}_G$.

Il reste à vérifier que, quand n tend vers l'infini, $\|L_n f - f\|$ tend vers zéro. L'opérateur L_n a été défini par (XII.62), qui doit se vérifier pour tous f et g ; d'autre part, $\int_0^\infty n e^{-ns} ds = 1$, donc on a aussi :

$$\langle f | g \rangle = \int_0^\infty n e^{-ns} \langle f | g \rangle ds \quad (\text{XII.65})$$

En soustrayant cela de (XII.62), on obtient :

$$\langle L_n f - f | g \rangle = \int_0^\infty n e^{-ns} \langle U_s f - f | g \rangle ds \quad (\text{XII.66})$$

et cela est vrai pour tous f et g dans $\mathcal{H}$; prenons donc $g = L_n f - f$, ce qui donne :

$$\|L_n f - f\|^2 = \int_0^\infty n e^{-ns} \langle U_s f - f | L_n f - f \rangle ds \quad (\text{XII.67})$$

Pour majorer cela, on utilise l'inégalité de Schwarz $|\langle U_s f - f | L_n f - f \rangle| \leq \|U_s f - f\| \cdot \|L_n f - f\|$, puis l'inégalité de la moyenne :

$$\left| \int_0^\infty n e^{-ns} \langle U_s f - f | L_n f - f \rangle ds \right| \leq \int_0^\infty n e^{-ns} \|U_s f - f\| ds \times \|L_n f - f\| \quad (\text{XII.68})$$

Par conséquent :

$$\|L_n f - f\| \leq \int_0^\infty n e^{-ns} \|U_s f - f\| ds = \int_0^\infty n e^{-ns} \varphi(s) ds \quad (\text{XII.69})$$

où $\varphi(s) = \|U_s f - f\|$. Or, d'après la définition des groupes unitaires, la fonction $\varphi(s)$ est continue, majorée uniformément par $2\|f\|$ et s'annule pour $s = 0$. La conclusion résulte donc du fait que pour une telle fonction, l'intégrale $\int_0^\infty n e^{-ns} \varphi(s) ds$ tend vers zéro quand n tend vers l'infini. •

Ce théorème ne se prononce pas sur une réciproque du type "tout opérateur auto-adjoint est le générateur infinitésimal d'un groupe unitaire" car ce serait faux. Cela reste cependant vrai pour les opérateurs qui ont déjà été présentés :

1. si G est l'opérateur $i \frac{d}{dx}$ (ou $i\hbar \frac{d}{dx}$), le groupe est celui des translations $U_t f(x) = f(x+t)$ puisqu'en effet on a bien :

$$\lim \langle \frac{U_t f - f}{it} | g \rangle = \lim \int \overline{g(x)} \frac{f(x+t) - f(x)}{it} dx = \int \overline{g(x)} f'(x) dx \quad (\text{XII.70})$$

La solution de l'équation différentielle $\frac{d\psi}{dt} = \frac{d\psi}{dx}$ est bien $\psi(t, x) = f(x+t)$;

2. si G est l'opérateur de multiplication par une fonction réelle $V(x)$, le groupe unitaire est celui des multiplications par $e^{itV(x)}$;

3. si G est l'opérateur $-\frac{d^2}{dx^2} + x^2$, le groupe est celui des transformations U_t qui, si on représente les éléments de $\mathcal{H}$ en séries de fonctions d'Hermite, s'expriment par :

$$U_t : f = \sum a_n \phi_n \mapsto \sum a_n e^{it(2n+1)} \phi_n \quad (\text{XII.71})$$

On résout donc ici l'équation de Schrödinger $i \frac{d\psi}{dt} = -\frac{d^2\psi}{dx^2} + x^2\psi$ exactement de la même façon que Fourier a résolu l'équation de la chaleur (section VII.1), sauf que l'on utilise des séries de fonctions d'Hermite à la place des séries de fonctions trigonométriques ;

4. cela se généralise en principe aux opérateurs $-\frac{d^2}{dx^2} + V(x)$, avec une fonction potentiel V vérifiant certaines conditions (on ne peut garantir l'existence d'un groupe unitaire pour n'importe quelle fonction V).

On peut donc résumer tout cela comme suit.

Le groupe unitaire associé à un opérateur auto-adjoint est la solution d'une équation de Schrödinger :

$$-i \frac{d\psi}{dt} = A\psi \quad (\text{XII.72})$$

qui se résout formellement par $\psi(t) = e^{itA}\psi(0)$ Pour une condition initiale $\psi(0) = f$, la solution est en effet $\psi(t) = U_t f$. Par exemple l'équation :

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right) + V(x, y, z) \psi \quad (\text{XII.73})$$

est de ce type ; on cherche une solution $\psi(t, x, y, z)$ qui pour chaque t fixé est une fonction de carré intégrable. Le groupe unitaire U_t qui la résout donne la fonction à l'instant t , $\psi_t(x, y, z) = \psi(t, x, y, z)$, à partir de la fonction initiale $f(x, y, z) = \psi(0, x, y, z)$. Le fait que le groupe soit unitaire se traduit par la conservation de la norme $\|\psi_t\|_2 = \|f\|_2$. Cette propriété de l'équation de Schrödinger est essentielle, car elle correspond à la *conservation de la probabilité* ; le carré de la fonction d'onde, $|\psi_t(x, y, z)|^2$, représente la densité de présence de la particule dans l'espace, dont l'intégrale sur tout l'espace doit être égale à 1. Ainsi, si au départ $\int |f(x, y, z)|^2 dx dy dz = \|f\|^2 = 1$, on aura aussi à tout instant ultérieur $\int |\psi(t, x, y, z)|^2 dx dy dz = \|f\|^2 = 1$. On voit que, du point de vue mathématique, cette conservation de la probabilité se traduit par le fait que l'opérateur hamiltonien est auto-adjoint.

XII.8 Espace $\mathbb{L}_2(S)$ et harmoniques sphériques

L'espace de Hilbert $\mathbb{L}_2(S)$ est celui des fonctions de carré intégrable sur la sphère :

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\} \quad (\text{XII.74})$$

Si on représente les points de l'espace en coordonnées sphériques :

$$\begin{aligned} x &= r \sin \theta \cos \varphi \\ y &= r \sin \theta \sin \varphi \\ z &= r \cos \theta \end{aligned} \quad (\text{XII.75})$$

ceux de la sphère S correspondent à $r = 1$ et sont donc représentés par les deux coordonnées sphériques angulaires θ et φ . Un vecteur de $\mathbb{L}_2(S)$ est donc une fonction de θ et de φ . Sur la sphère, l'élément de surface est $d\sigma = \frac{1}{4\pi} \sin \theta d\theta d\varphi$, donc la norme $\|f\|_{\text{sph}}$ d'une fonction $f \in \mathbb{L}_2(S)$ sera donnée par :

$$\|f\|_{\text{sph}}^2 = \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi |f(\theta, \varphi)|^2 \sin \theta d\theta d\varphi \quad (\text{XII.76})$$

Le produit scalaire est :

$$\langle f | g \rangle_{\text{sph}} = \int_0^{2\pi} \int_0^\pi f(\theta, \varphi) \overline{g(\theta, \varphi)} \sin \theta d\theta d\varphi \quad (\text{XII.77})$$

Cette section est essentiellement consacrée à la construction d'une base orthogonale de l'espace $\mathbb{L}_2(S)$.

Définition XII.10 On appelle harmoniques sphériques, les restrictions à la sphère :

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\} \quad (\text{XII.78})$$

des polynômes harmoniques homogènes des trois variables x, y, z . Une harmonique sphérique de degré n sera la restriction à la sphère d'un polynôme harmonique homogène de degré n .

Rappelons qu'un polynôme homogène est un polynôme dont tous les termes ont le même degré : par exemple $x^3 + 5xy^2 - x^2z + 2xyz$ ou $x^6 + y^6 + x^4z^2 - 3xy^3z^2$ sont homogènes, tandis que $x + yz$ ou $x^3 - z^2 + 2xy^4$ ne le sont pas. Par ailleurs, une fonction est dite harmonique si son laplacien est nul.

Les harmoniques sphériques forment un espace vectoriel, qui est un sous-espace dense de l'espace de Hilbert $\mathbb{L}_2(S)$. L'expression la plus générale d'un polynôme homogène de degré n des trois variables x, y, z est :

$$P(x, y, z) = \sum \sum \sum_{i+j+k=n} a_{ijk} x^i y^j z^k \quad (\text{XII.79})$$

La dimension de l'espace vectoriel $\mathcal{W}_n$ des polynômes homogènes de degré n est donc le nombre de coefficients de cette expression, qui est aussi le nombre de multi-indices (i, j, k) tels que $i + j + k = n$, soit $\frac{1}{6}(n+1)(n+2)$. Cette dimension est évidemment plus petite que la dimension $\frac{1}{6}(n+1)(n+2)(n+3)$ de *tous* les polynômes de degré inférieur ou égal à n .

La dimension des polynômes homogènes *harmoniques* de degré n sera encore plus petite, puisque la condition $\Delta P = 0$ implique des relations linéaires entre les coefficients a_{ijk} . Pour la connaître, il faut voir de combien de coefficients primitifs indépendants dépendent effectivement les a_{ijk} , ce qu'on peut faire comme suit. D'après la formule de Taylor, on a :

$$a_{ijk} = \frac{\partial^n P}{\partial x^i \partial y^j \partial z^k} \quad (\text{XII.80})$$

On notera qu'il est inutile de préciser qu'il s'agit des dérivées en $(x, y, z) = (0, 0, 0)$, puisque ces dérivées sont constantes. Or, si le polynôme P est harmonique, on a :

$$\frac{\partial^2 P}{\partial z^2} = -\frac{\partial^2 P}{\partial x^2} - \frac{\partial^2 P}{\partial y^2} \quad (\text{XII.81})$$

ce qui signifie que toutes les dérivées partielles d'ordre pair par rapport à z pourront s'exprimer linéairement en fonction des seules dérivées par rapport à x et y ; quant à celles d'ordre impair par rapport à z , on pourra les ramener à des dérivées par rapport à x , y et z , celle restant par rapport à z étant d'ordre 1. Si on fait cela dans (XII.80), on aura exprimé tous les a_{ijk} uniquement à partir des $n+1$ dérivées $\partial^n P / \partial x^i \partial y^j$ (nombre de couples i, j tels que $i+j=n$), plus les n dérivées $\partial^n P / \partial x^i \partial y^j \partial z$ (nombre de couples i, j tels que $i+j+1=n$), soit $2n+1$.

En conclusion, la dimension de l'espace vectoriel des polynômes homogènes harmoniques de degré n est $2n+1$. Cela implique que pour en trouver une base, il suffit d'exhiber $2n+1$ polynômes homogènes harmoniques linéairement indépendants. Pour cela, on considère les intégrales de la forme :

$$f(x, y, z) = \int_{-\pi}^{\pi} (ix \cos t + iy \sin t + z)^n u(t) dt \quad (\text{XII.82})$$

où $u(t)$ est une fonction quelconque, disons continue sur $[-\pi; \pi]$. Ces fonctions sont visiblement des polynômes homogènes de degré n . Si on dérive, on obtient :

$$\Delta f(x, y, z) = \int_{-\pi}^{\pi} n(n-1)[1 - \cos^2 t - \sin^2 t](ix \cos t + iy \sin t + z)^{n-2} u(t) dt \quad (\text{XII.83})$$

ce qui est nul en vertu de l'identité $\cos^2 t + \sin^2 t = 1$; la dérivation sous le signe d'intégration est justifiée par les théorèmes généraux, puisque les dérivées d'ordre supérieur (par exemple quatre) sont toujours des fonctions continues et bornées de t sur $[-\pi; \pi]$.

Ainsi les fonctions (XII.82) sont toutes des polynômes homogènes harmoniques de degré n et les fonctions de θ, φ qu'on en déduit en prenant $r=1$ sont alors des harmoniques sphériques de degré n :

$$\begin{aligned} f(\theta, \varphi) &= \int_{-\pi}^{\pi} [i \sin \theta \cos \varphi \cos t + i \sin \theta \sin \varphi \sin t + \cos \theta]^n u(t) dt \\ &= \int_{-\pi}^{\pi} [i \sin \theta \cos(t - \varphi) + \cos \theta]^n u(t) dt \end{aligned} \quad (\text{XII.84})$$

Les fonctions de cette forme ne représentent pas toutes les harmoniques sphériques de degré n , mais suffisent pour en construire une base. En effet, les fonctions :

$$Y_{n,\ell}(\theta, \varphi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [i \sin \theta \cos(t - \varphi) + \cos \theta]^n e^{i\ell t} dt \quad (\text{XII.85})$$

pour $\ell = -n, -n+1, -n+2, \dots, n-1, n$, qui correspondent aux cas où $u(t) = \frac{1}{2\pi} e^{i\ell t}$ sont linéairement indépendantes et au nombre de $2n+1$.

Exercice XII.9 Prouver que pour tout n fixé, les $2n+1$ fonctions $Y_{n,\ell}$, pour $\ell = -n, -n+1, -n+2, \dots, n-1, n$, sont linéairement indépendantes en effectuant, dans l'intégrale (XII.85), le changement de variable $s = t - \varphi$ et tenant compte du fait que tout est périodique. En déduire qu'elles forment une base de l'espace des harmoniques sphériques de degré n .

Appelons $\mathcal{Y}_n$, l'espace vectoriel (de dimension $2n+1$) des harmoniques sphériques de degré n . La somme directe des $\mathcal{Y}_n$ pour n allant de zéro à l'infini sera un espace de dimension infinie, qu'on désignera par $\mathcal{Y}$.

On peut démontrer un théorème similaire à ceux de Weierstrass en section XI.5 :

Théorème XII.12 Toute fonction continue sur la sphère et à valeurs réelles ou complexes, est la limite uniforme d'une suite d'harmoniques sphériques.

Nous ne donnons pas de démonstration détaillée mais il n'y aurait rien de vraiment nouveau par rapport aux théorèmes de la section XI.5. Le théorème de Weierstrass XI.2 concernant les polynômes algébriques, s'appliquait à l'intervalle $[0; 1]$, ou, moyennant un changement de variable linéaire, à n'importe quel intervalle borné, par exemple $[-1; 1]$. On peut étendre la démonstration du théorème aux polynômes à deux ou trois variables ⁽¹²⁾ et prouver ainsi que les fonctions continues sur le cube $[-1; 1] \times [-1; 1] \times [-1; 1]$ sont des limites uniformes de polynômes. Or la sphère est contenue dans ce cube; en outre, il se trouve que la restriction d'un polynôme quelconque à la sphère est automatiquement aussi la restriction d'un polynôme homogène harmonique (ce point, au demeurant élémentaire, serait en fait la seule nouveauté).

On en déduit alors aisément, comme cela avait été fait pour le théorème XI.3, que les harmoniques sphériques sont denses dans l'espace $\mathbb{L}_2(S)$. On peut énoncer :

Théorème XII.13 La famille *infinie* des fonctions $Y_{n,\ell}$ pour $n = 0, 1, 2, \dots$ et $-n \leq \ell \leq +n$, forme une base orthogonale de l'espace de Hilbert $\mathbb{L}_2(S)$.

Preuve Pour montrer qu'une famille de fonctions forme une base hilbertienne, il faut vérifier qu'elle est orthogonale et que le sous-espace des combinaisons linéaires finies de ces fonctions est dense. La seconde partie de ce programme est déjà réglée par l'extension du théorème de Weierstrass mentionnée ci-dessus. Il ne reste donc plus qu'à calculer les intégrales :

$$\int_0^{2\pi} \int_0^\pi \sin \theta Y_{n\ell}(\theta, \varphi) Y_{n'\ell'}(\theta, \varphi) d\theta d\varphi \quad (\text{XII.86})$$

Exercice XII.10 On rappelle la définition des polynômes de Legendre (sous-section XI.4.1) :

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n ([x^2 - 1]^n)}{dx^n} \quad (\text{XII.87})$$

Montrer que l'on peut écrire aussi :

$$P_n(x) = \frac{1}{2i\pi} \int_\gamma \frac{(z-1)^n (z+1)^n}{2^n (z-x)^{n+1}} dz \quad (\text{XII.88})$$

où γ est un lacet simple du plan complexe qui entoure le point $z = x$. ■

Exercice XII.11 On suppose ici que $-1 < x < +1$. Prendre pour γ un cercle de centre x et de rayon $R = \sqrt{1-x^2}$. Paramétrer ce lacet avec un paramètre angulaire t qui varie de $-\pi$ à $+\pi$ et montrer que l'on a :

$$P_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [x + \sqrt{1-x^2} \cos t]^n dt \quad (\text{XII.89})$$

En déduire que $Y_{n,0}(\theta, \varphi) = P_n(\cos \theta)$. ■

(12) Weierstrass a lui-même donné cette extension du théorème à la fin de son article publié en 1885, déjà cité (cf. note 3 du chapitre XI).

Exercice XII.12 Montrer que pour tout $n \geq 0$ on a :

$$\int_0^{2\pi} \int_0^\pi Y_{n\ell}(\theta, \varphi) Y_{nk}(\theta, \varphi) \sin \theta \, d\theta \, d\varphi = 0 \quad (\text{XII.90})$$

si $k \neq \ell$. Conclure que pour tout n fixé, les $2n + 1$ fonctions $Y_{n\ell}$ ($-n \leq \ell \leq +n$) forment une base orthogonale de $\mathcal{Y}_n$. Remarquer que l'orthogonalité provient uniquement de l'intégration par rapport à φ et que la coordonnée θ n'y joue aucun rôle. ■

Exercice XII.13 Montrer que les fonctions $f(x, y, z) = r^n Y(\theta, \varphi)$, où r, θ, φ sont les coordonnées sphériques de x, y, z , sont des polynômes homogènes harmoniques. ■

Exercice XII.14 On rappelle l'expression du laplacien en coordonnées sphériques :

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \Delta_{\text{sph}} \quad (\text{XII.91})$$

où Δ_{sph} est la partie sphérique du laplacien, qui n'opère que sur les coordonnées θ et φ :

$$\Delta_{\text{sph}} = \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \quad (\text{XII.92})$$

Montrer que $\Delta_{\text{sph}} Y_{n\ell} = -n(n+1) Y_{n\ell}$. ■

Exercice XII.15 Vérifier que $\langle \Delta_{\text{sph}} Y_{n\ell} | Y_{n\ell} \rangle_{\text{sph}} = \langle Y_{n\ell} | \Delta_{\text{sph}} Y_{n\ell} \rangle_{\text{sph}}$ à l'aide d'intégrations par parties. En déduire, en utilisant le résultat obtenu en section XII.6, que si $n \neq m$, $Y_{n\ell}$ est orthogonale à Y_{mk} (indépendamment de k et ℓ). Autrement dit, les sous-espaces $\mathcal{Y}_n$ sont orthogonaux dans $\mathbb{L}_2(S)$. ■

L'exercice XII.15 consiste à prouver que la restriction à $\mathcal{Y}_n$ de l'opérateur Δ_{sph} est égale à $-n(n+1)I$; cela montre que les nombres $\lambda_n = -n(n+1)$ sont des valeurs propres de Δ_{sph} et les $Y_{n\ell}$ des vecteurs propres. Comme par ailleurs les $Y_{n\ell}$ forment (pour l'ensemble des indices $n = 0, 1, 2, \dots$ et $-n \leq \ell \leq +n$) une base de $\mathbb{L}_2(S)$, ils fournissent forcément la totalité des vecteurs propres possibles (s'il existait une valeur propre autre que les $\lambda_n = -n(n+1)$, le sous-espace propre correspondant serait orthogonal à tous les $Y_{n\ell}$, ce qui ne peut pas se produire puisque justement les $Y_{n\ell}$ forment une base. Par le même argument que dans la démonstration du théorème XII.3, on voit aussi que le spectre est réduit à ces valeurs propres. On peut donc énoncer :

Théorème XII.14 L'opérateur Δ_{sph} sur l'espace $\mathbb{L}_2(S)$ a un spectre formé exclusivement des valeurs propres $\lambda_n = -n(n+1)$, de multiplicité $2n+1$; les vecteurs propres associés sont les fonctions $Y_{n\ell}$ données en (XII.85).

L'opérateur Δ_{sph} est discontinu ; son domaine est celui des fonctions $\sum \sum a_{n\ell} Y_{n\ell}$ telles que la série $\sum \sum (2n+1) |a_{n\ell}|^2$ converge. On peut aussi caractériser ce domaine comme suit : les dérivations partielles de l'opérateur étant entendues au sens des distributions, le domaine est l'ensemble des fonctions $f \in \mathbb{L}_2(S)$ telles que $\Delta_{\text{sph}} f \in \mathbb{L}_2(S)$.

XII.9 Théorie de l'atome d'hydrogène

En dehors de l'oscillateur quantique, un autre exemple traité par E. Schrödinger dans les mémoires de 1926⁽¹³⁾ est celui d'un électron dans un champ coulombien. Il s'agit du problème

(13) Pages 3 à 12, dans la traduction française de A. PROCA rééditée récemment par Jacques Gabay.

avec *un seul* électron dans le champ coulombien, car s'il y en a deux ou plus on ne sait pas résoudre l'équation. Le système physique qui correspond à ce problème est celui d'un atome avec un seul électron, donc l'atome d'hydrogène, ou un atome d'hélium ionisé positivement. Le champ coulombien est alors celui du noyau et son potentiel est $V(x, y, z) = -\epsilon^2/r$ où ϵ est la charge élémentaire notée en gothique pour ne pas confondre avec l'exponentielle $e = 2,71828\dots$ et $r = \sqrt{x^2 + y^2 + z^2}$.

Jusqu'ici, nous avons presque toujours considéré des problèmes à une seule variable ; ici il y en a trois et l'espace de Hilbert est $\mathcal{H} = \mathbb{L}_2(\mathbb{R}^3)$. Cela ne change rien au principe.

Dans le problème de l'oscillateur quantique de la section XII.4, la base de vecteurs propres, les fonctions d'Hermite, a été fournie gratuitement. Ici, on ne pourra pas fournir une *base* de vecteurs propres car il n'en existe pas, mais on peut fournir un système orthogonal *non complet* de fonctions propres. Pour l'oscillateur quantique, on a pu démontrer que le spectre ne contenait que les valeurs propres $\lambda_n = \hbar\omega(n + \frac{1}{2})$ parce que les vecteurs propres associés formaient une base complète. Ici, on ne pourra pas arriver à cette conclusion car il n'y a pas de base complète de vecteurs propres et il se trouve que le spectre n'est pas formé que de valeurs propres : il contient en outre la demi-droite $[0; +\infty[$, qui constitue le spectre essentiel et il n'y a aucune valeur propre > 0 .

Le système orthogonal non complet de fonctions propres s'obtient comme suit. On introduit d'abord la famille $f_{nk}(r)$ des *fonctions de Laguerre*, définies pour $t > 0$ par :

$$f_{nk}(t) = Q_{n-k-1}^{2k+1}(t) t^k e^{-t/2} \quad (\text{XII.93})$$

qui sont représentées sur la figure XI.5. Les Q_n^j sont les polynômes de Laguerre déjà introduits en sous-section XI.4.3. Pour tout j , Q_n^j est de degré n .

Le système cherché est alors celui des fonctions :

$$\psi_{nk\ell}(x, y, z) = f_{nk}\left(\frac{2m\epsilon^2}{\hbar^2 n} r\right) Y_{k\ell}(\theta, \varphi) \quad (\text{XII.94})$$

pour les indices entiers $n = 1, 2, 3, \dots$, $0 \leq k \leq n-1$ et $-n \leq \ell \leq +n$, qui seront les vecteurs propres associés aux valeurs propres :

$$E_n = -\frac{m\epsilon^4}{2\hbar^2 n^2} \quad (\text{XII.95})$$

L'importance historique de ce résultat provient de ce qu'il correspond à la loi de Bohr, qui au lieu d'être devinée à partir des observations, est déduite mathématiquement du *principe de correspondance*, c'est-à-dire de l'hypothèse que la mécanique classique serait l'optique géométrique d'un phénomène ondulatoire.

Voici ce qu'écrivait E. Schrödinger dans une présentation d'ensemble de ses mémoires :

« L'analogie entre la mécanique et l'optique signalée par Hamilton ne concerne que l'optique géométrique ; en effet, d'après cette analogie, à toute trajectoire du point représentatif dans l'espace de configuration correspond un rayon de lumière et cette dernière notion ne peut être définie sans ambiguïté qu'en optique géométrique. L'analyse de cette image optique, du point de vue ondulatoire, conduit à l'abandon de la notion de trajectoire du système, dès que les dimensions sont petites par rapport à la longueur d'onde. Cette notion et avec elle toute la mécanique classique, subsistent comme approximations, valables seulement dans le

cas où les dimensions de la trajectoire sont grandes par rapport à la longueur d'onde. [...] Par analogie avec ce qui se passe en optique, les équations fondamentales doivent être remplacées par une seule équation d'ondes dans l'espace de configuration. (...) Cette équation contient un paramètre de valeurs propres E , qui correspond à l'énergie mécanique. [...] En général, cette équation des ondes ne possède de solutions finies, continues, à détermination unique et dont les dérivées satisfassent aux mêmes conditions⁽¹⁴⁾, que pour certaines valeurs particulières du paramètre E , qu'on appelle les valeurs propres. Les valeurs propres sont, ou bien identiques aux niveaux d'énergie de la théorie quantique admise jusqu'à présent, ou bien elles s'en écartent d'une manière qui est en parfait accord avec les résultats de l'expérience. »

On peut donc conclure que cet accord, constaté dans de nombreux cas dès 1926 et jamais démenti depuis, est une très forte confirmation du principe de correspondance.

On va maintenant vérifier que les fonctions $\psi_{nkl}(x, y, z)$ de (XII.94) sont bien les fonctions propres de l'opérateur :

$$H = -\frac{\hbar^2}{2m}\Delta - \frac{e^2}{r} \quad (\text{XII.96})$$

Prouvons d'abord que les polynômes de Laguerre $Q_m^j(t)$ qui interviennent dans l'expression de la fonction ψ_{nkl} satisfont la relation :

$$t Q''(t) + (j+1-t) Q'(t) + m Q(t) = 0 \quad (\text{XII.97})$$

Rappelons que les polynômes $Q_m^j(t)$ sont définis par la dérivée d'ordre m de la fonction $t^{m+j}e^{-t}$ (cf. sous-section XI.4.3) :

$$\frac{d^m(t^{m+j}e^{-t})}{dt^m} = Q_m^j(t) t^j e^{-t} \quad (\text{XII.98})$$

On déduit immédiatement de cette définition que :

$$\begin{aligned} Q_{m+1}^j(t) t^j e^{-t} &= \frac{d^{m+1}(t^{m+1+j}e^{-t})}{dt^{m+1}} = \\ &= \frac{d}{dt} \frac{d^m(t^{m+j+1}e^{-t})}{dt^m} = \frac{d}{dt} \left(Q_m^{j+1}(t) t^{j+1} e^{-t} \right) = \\ &= \left(t \frac{d}{dt} Q_m^{j+1}(t) + (j+1-t) Q_m^{j+1}(t) \right) t^j e^{-t} \end{aligned} \quad (\text{XII.99})$$

d'où :

$$Q_{m+1}^j(t) = t \frac{d}{dt} Q_m^{j+1}(t) + (j+1-t) Q_m^{j+1}(t) \quad (\text{XII.100})$$

En dérivant ce résultat il vient :

$$\frac{d}{dt} Q_{m+1}^j(t) = t \frac{d^2}{dt^2} Q_m^{j+1}(t) + (j+1-t) \frac{d}{dt} Q_m^{j+1}(t) - Q_m^{j+1}(t) \quad (\text{XII.101})$$

Si on fait une récurrence sur m , on suppose comme hypothèse de récurrence que (XII.97) est vraie $\forall j \in \mathbb{R}$ (j n'a aucune raison d'être entier) ; combinée avec (XII.101) cette hypothèse donne :

$$\frac{d}{dt} Q_{m+1}^j(t) = -(m+1) Q_m^{j+1}(t) \quad (\text{XII.102})$$

⁽¹⁴⁾ La vraie nature mathématique de cette condition a été comprise plus tard par J. von Neumann : la solution doit être dans l'espace $\mathbb{L}_2$.

d'où on déduit :

$$\begin{aligned} t \frac{d^2}{dt^2} Q_{m+1}^j(t) + (j+1-t) \frac{d}{dt} Q_{m+1}^j(t) &= \\ &= -(m+1) \left(t \frac{d^2}{dt^2} Q_m^{j+1}(t) + (j+1-t) \frac{d}{dt} Q_m^{j+1}(t) \right) \end{aligned} \quad (\text{XII.103})$$

Revenant à (XII.100), on voit que ceci est égal à $-(m+1) Q_{m+1}^j(t)$, ce qui établit la récurrence. Pour prouver complètement la récurrence, il faut encore s'assurer que (XII.97) est vraie pour $m=0$, ce qui est évident puisque les polynômes Q_m^j sont alors constants.

Revenons maintenant à l'équation de Schrödinger. En écrivant le laplacien en coordonnées sphériques :

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \Delta_{\text{sph}} \quad (\text{XII.104})$$

on obtient :

$$\Delta \psi_{nkl} = \frac{\partial^2 \psi_{nkl}}{\partial r^2} + \frac{2}{r} \frac{\partial \psi_{nkl}}{\partial r} + \frac{1}{r^2} \Delta_{\text{sph}} \psi_{nkl} \quad (\text{XII.105})$$

La partie radiale n'opère que sur la fonction $f_{nk}(\frac{2m\epsilon^2}{\hbar^2 n} r)$ et la partie sphérique Δ_{sph} n'opère que sur les fonctions $Y_{k\ell}(\theta, \varphi)$. La première donne donc :

$$\left(\frac{2m\epsilon^2}{\hbar^2 n} \right)^2 \left(f_{nk}''(t) + \frac{2}{t} f_{nk}'(t) \right) Y_{k\ell}(\theta, \varphi) \quad (\text{XII.106})$$

où on a posé $t = \frac{2m\epsilon^2}{\hbar^2 n} r$ et la seconde, d'après le théorème XII.14, donne :

$$-\frac{1}{r^2} k(k+1) f_{nk}(t) Y_{k\ell}(\theta, \varphi) = -\left(\frac{2m\epsilon^2}{\hbar^2 n} \right)^2 \frac{1}{t^2} k(k+1) f_{nk}(t) Y_{k\ell}(\theta, \varphi) \quad (\text{XII.107})$$

En regroupant tout, on obtient :

$$\Delta \psi_{nkl}(x, y, z) = \left(\frac{2m\epsilon^2}{\hbar^2 n} \right)^2 \left(f_{nk}''(t) + \frac{2}{t} f_{nk}'(t) - \frac{1}{t^2} k(k+1) f_{nk}(t) \right) Y_{k\ell}(\theta, \varphi) \quad (\text{XII.108})$$

Par conséquent $H\psi_{nkl}$ devient, en convertissant tout à la variable t :

$$H\psi_{nkl} = -\frac{2m\epsilon^4}{\hbar^2 n^2} \left(f_{nk}''(t) + \frac{2}{t} f_{nk}'(t) - \frac{1}{t^2} k(k+1) f_{nk}(t) - \frac{n}{t} f_{nk}(t) \right) Y_{k\ell}(\theta, \varphi) \quad (\text{XII.109})$$

Or d'après (XII.97), les polynômes $Q_{n-k-1}^{2k+1}(t)$ qui interviennent dans l'expression des fonctions $f_{nk}(t)$ vérifient la relation :

$$t Q''(t) + (2k+2-t) Q'(t) + (n-k-1) Q(t) = 0 \quad (\text{XII.110})$$

Des calculs élémentaires conduisent alors à la relation correspondante pour les fonctions $f_{nk}(t) = Q_{n-k-1}^{2k+1}(t) t^k e^{-t/2}$:

$$f'' + \frac{2}{t} f' + \left(-\frac{k(k+1)}{t^2} + \frac{n}{t} \right) f = -\frac{1}{4} f \quad (\text{XII.111})$$

Ainsi on voit que :

$$H\psi_{nkl} = \frac{m\epsilon^4}{2\hbar^2 n^2} \psi_{nkl} \quad (\text{XII.112})$$

ce qui montre que les ψ_{nkl} sont des vecteurs propres, associés aux valeurs propres $E_n = m\epsilon^4/2\hbar^2 n^2$.

Cela ne prouve cependant en rien qu'il n'existe pas d'autres fonctions propres, car les ψ_{nkl} ne forment pas une base.

On voit que la valeur propre E_n ne dépend que de n ; par conséquent le nombre de ψ_{nkl} associées à la valeur propre E_n est le nombre des couples d'indices k, ℓ . Or on a vu que ℓ prend les $2k + 1$ valeurs comprises entre $-k$ et $+k$, tandis que k prend les valeurs de 0 à $n - 1$. Le nombre de couples k, ℓ est donc la somme des nombres $2k + 1$ pour k allant de 0 à $n - 1$, soit n^2 . Comme les ψ_{nkl} sont linéairement indépendantes, la multiplicité de la valeur propre E_n est donc n^2 .

Pour prouver qu'il n'existe pas d'autres fonctions propres que les ψ_{nkl} , on remarque que s'il en existait, elles se décomposeraient aussi sous la forme $f(t) Y(\theta, \varphi)$, puisque l'opérateur H n'agit en fait que sur la variable radiale. La fonction $f(t)$ vérifierait donc toujours l'équation (XII.111), mais cette fois pour des valeurs non nécessairement entières ≥ 1 de n . La forme de l'équation (du type de Fuchs) garantit que les solutions sont des séries de Laurent en t . En cherchant ces séries, on obtient une relation de récurrence pour leurs coefficients et on se rend compte que le seul cas où elles peuvent donner des polynômes est celui où n est entier ≥ 1 ; dans les autres cas, la récurrence entre les coefficients montre que leur signe reste constant à partir d'un certain rang et on peut en déduire que ces séries donnent des fonctions qui croissent exponentiellement quand t tend vers $\pm\infty$, ou (puisque'il s'agit de séries de Laurent) qu'elles sont singulières pour $t = 0$. Dans tous ces cas, il est impossible que l'intégrale :

$$\int_0^\infty t^2 |f(t)|^2 dt \quad (\text{XII.113})$$

converge, donc les fonctions ψ correspondantes ne seront pas dans $\mathbb{L}_2(\mathbb{R}^3)$.

C'est ainsi qu'on montre que les niveaux de Bohr E_n sont les seules valeurs propres possibles, de multiplicité n^2 . Avec les indications ci-dessus, cela est laissé en exercice mais les calculs sont longs.



Bibliographie

- Banach S., *Théorie des opérations linéaires*, 1932. 175
- Cartan H., *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*, Paris, Hermann, 1961. 100
- Cauchy A.-L., De l'influence que peut avoir, sur la valeur d'une intégrale double, l'ordre dans lequel on effectue les intégrations, *Œuvres complètes d'A. Cauchy*, vol. 6, 1882. 42
- Cauchy A.-L., Mémoire sur les intégrales définies, *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 1882. 41, 63
- Hadamard J., *Le problème de Cauchy et les équations aux dérivées partielles linéaires hyperboliques*, Paris, Hermann et C^{ie}, 1932. 150
- Kato T., *Perturbation Theory for Linear Operators*, Berlin, Springer-Verlag, 1966. 234
- Kurzweil J., Generalized ordinary differential equations and continuous dependence on a parameter, *Czechoslovak Mathematical Journal*, vol. 7, 1957, pp. 418–446. 17
- Lavrentiev M. et Chabat B., *Méthodes de la théorie des fonctions d'une variable complexe*, Moscou, MIR, 1977. 49
- Lebesgue H., Sur une généralisation de l'intégrale définie, *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, vol. 132, 1901, pp. 1025–1028. 2, 17
- Lelong-Ferrand J. et Arnaudière J. M., *Algèbre*, Paris, Dunod, 1971. 232
- Schrödinger E., Der stetige Übergang von der mikro zur makromechanik, *Die Naturwissenschaften*, vol. 28, 1926, pp. 664–666. 215

- Schwartz L., *Théorie des distributions*, Paris, Hermann, 1966 vol. 1–2. 175, 193
- Schwartz L., *Un mathématicien aux prises avec le siècle*, Paris, Odile Jacob, 1997. 157
- Smirnov V., *Cours de mathématiques supérieures*, Moscou, MIR, 1975 vol. III. 49
- Weierstrass K., Über die analytische darstellbarkeit sogenannter willkürlicher functionen einer reellen veränderlichen, *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, 1885. 210, 257
- Whittaker E.T. et Watson G.N., *A Course of Modern Analysis*, Cambridge, Cambridge University Press, 1927. 49
- Yosida K., *Functional Analysis*, Berlin, Springer-Verlag, 1968, 237–238p. 252



Index

A

- ABEL, Niels
 propriété de 37
 accroissements finis 6–9
 adhérence 243
 affixe 97, 98
 AMPÈRE, André-Marie 42
 anallagmatique 99
 analyse fonctionnelle 138
 analytique
 fonction voir fonction
 prolongement 70, 136
 ARCHIMÈDE
 spirale d' 53
 atome d'hydrogène 258

B

- BANACH, Stefan 142
 équi-continuité voir équi-continuité
 théorème de 175
 base
 algébrique 214
 hilbertienne 212, 215, 217
 orthogonale 236

- orthonormée 212
 BERNOULLI, Daniel 111
 BERNSTEIN, Sergei
 polynôme de 222–226
 BOHR, Niels 262
 BORN, Max 239

C

- calcul infinitésimal 31
 CARTAN, Henri 28
 CAUCHY, Augustin-Louis
 critère de 204
 formule de 42, 43, 57
 inégalité de 8, 44
 intégrale de 40
 suite de 130, 132, 231, 245
 théorie de 41, 42
 valeur principale de 138, 139, 183
 CAYLEY, Arthur
 transformation de 248, 249
 CESÀRO, Ernesto
 convergence au sens de 113
 concaténation 12
 conforme voir transformation
 connexité 13

continuité
 des distributions . . . voir distribution(s)
 prolongement par 22, 72, 130
 simple 16
 uniforme 16
 convergence
 absolue 1
 au sens de CESÀRO voir CESÀRO
 au sens de SCHWARTZ . voir SCHWARTZ
 au sens des distributions 113
 disque de 38
 dominée 5–9
 en moyenne 124, 160
 quadratique, 113, 127, 160
 euclidienne 211
 intrinsèque 204
 normale 40
 point par point 5
 rayon de 38, 39, 137
 simple 5, 112
 uniforme 2, 5, 124, 160, 211
 convolution 177, 188, 192
 courbe équipotentielle 107
 croissance polynomiale 184, 187

D

décomposition 12
 dérivation algébrique 34, 40
 densité
 de charge 153
 de probabilité 115
 dipolaire 154
 domaine dense 230
 espace dense 129, 130, 211, 230
 dipôle 153, 154, 175
 DIRICHLET, Johann 112
 discontinuité
 de première espèce 16, 181
 de seconde espèce 181, 182
 distance
 euclidienne 210, 211
 uniforme 211
 distribution(s) 117, 135, 140, 153, 167, 176
 constante 170
 continuité 160
 convergence voir convergence
 convolution 186, 189

définition 154, 156
 dérivation 113, 168, 179, 182, 185
 de Dirac 111, 113, 169, 175, 181
 limite 174
 régulière 164, 181
 singulière 164, 181
 transformation de Fourier voir FOURIER

E

EGOROFF, Dimitri
 théorème de 5
 électromagnétisme 42
 électrostatique 103, 105, 153, 164
 endomorphisme 227, 246
 équation
 corde vibrante 111
 de la chaleur 111
 équi-continuité 175
 espace
 complet 210
 de Hilbert voir HILBERT
 euclidien 203, 246
 fermé 210
 hermitien 241
 EULER, Leonhard
 constante d' 75, 89
 fonction d' 68, 70, 136
 formule d' 85, 86
 intégrale d' 70

F

FÉJER, Lipót
 théorème de 211
 fonction
 analytique 30, 33, 177
 Beta β 8, 81–87
 Chi χ 2
 d'Euler voir EULER
 d'onde 227
 de Heaviside 169, 179
 généralisée 164, 167
 Gamma Γ 8, 81–96
 harmonique 103
 holomorphe voir analytique
 puissance 78
 fonctionnelle linéaire 155, 157, 160

forme différentielle exacte.....24
 forme linéaire... voir fonctionnelle linéaire
 formule
 de CAUCHY–HADAMARD.....60
 des compléments 82
 du binôme..... voir NEWTON
 FOURIER, Joseph
 série de 112
 transformation de 115
 distribution(s), 171, 185, 195
 inversion, 123, 171
 polynôme, 172
 fractale 12, 181



générateur infinitésimal.....252
 GREEN, George 42
 fonction de 141, 167, 196
 formule de 19, 42
 groupe unitaire.....251



HADAMARD, Jacques 39, 140
 hamiltonien 114, 229, 240, 243
 HANKEL, Hermann
 chemin de 93, 96
 formule de.....96
 intégrale de.....10, 96
 HERMITE, Charles
 fonction de 243
 polynôme de 204, 215, 217, 220, 224, 240
 série de 217
 HILBERT, David
 base hilbertienne.....243
 espace de ... 129, 206, 227–229, 239–251
 métrique hilbertienne.....229
 homologie.....26–28
 hydrodynamique 105



inégalité de la moyenne 2, 89
 intégrale
 curviligne 12, 18, 26, 42
 double.....14, 42

eulérienne
 de première espèce, 214
 deuxième espèce, 81
 première espèce, 81
 généralisée 164
 partielle..... 14
 semi-convergente 135, 138
 simple.....15
 intégration algébrique 34



KURZWEIL, Jaroslav.....17
 intégrale de 18



LAGUERRE, Edmon
 polynôme de 217
 LAPLACE, Pierre-Simon de
 transformation de 132, 133
 laplacien 104
 LAURENT, Pierre Alphonse
 développement de 60
 série de.....57, 59–80, 262
 théorème de.....59
 LEBESGUE, Henri
 intégrale de.....18, 129
 théorie de 130, 131
 LEGENDRE, Adrien-Marie
 polynôme de 203, 212
 LEIBNIZ, Gottfried
 formule de 125, 161, 194, 220
 limite
 distribution(s) voir distribution(s)
 faible 174, 176, 182, 194
 LIOUVILLE, Joseph
 théorème de.....44
 logarithme complexe.....48, 55
 détermination 77



mécanique quantique.. 116, 138, 215, 220,
 227, 239
 métrique.....206, 207, 229
 euclidienne 207, 224

hilbertienne voir HILBERT
 uniforme 207, 212
 mesure
 ensemble négligeable 16, 17, 131
 MINKOWSKI, Hermann
 inégalité de 127

N

NEWTON, Isaac
 deuxième principe 251
 formule du binôme 80, 223
 norme 207

O

opérateur 227
 auto-adjoint 241, 243
 borné 228
 continu 228
 continuité d'un 228, 246
 de projection 228
 différentiel 172
 fermé 229
 norme 228
 préfermé 242
 symétrique 244
 oscillateur quantique 114, 259

P

PARSEVAL, Marc-Antoine
 formule de 128
 relation de 116
 partie
 entière 59
 finie 150, 181
 singulière 59
 PLANCHEREL, Michel
 formule de 228
 POINCARÉ, Henri 28, 42
 point singulier 61, 68
 POISSON, Siméon Denis
 équation de 103, 105
 pôle 61, 68
 polynôme
 algébrique 212, 225

harmonique 255
 trigonométrique 207, 211, 225
 pontage 19
 principe de correspondance 241, 259
 produit scalaire 203
 propagateur 141
 pseudo-densité 153
 pseudo-fonction 140, 150, 177, 181

R

régularisation 193, 194
 résolvante 232
 règle de l'HOSPITAL 69
 rebroussement 103
 relation de CAUCHY-RIEMANN 34–53, 104
 coordonnées polaires 36
 RIEMANN, Georg
 intégrale de 16
 somme de 2, 115
 RIESZ, Frigyes 142, 167
 RODRIGUES, Benjamin
 formule de 213
 rosette 103, 108

S

série 205
 binomiale 70
 de Laurent voir LAURENT
 entière 33, 37, 69
 géométrique 70
 trigonométrique 111
 saut de discontinuité 73
 SCHRÖDINGER, Erwin 220, 240
 équation de 251, 254
 SCHWARTZ, Laurent 140
 convergence au sens de 124
 espace de 122, 140
 inégalité de 127
 semi-norme 157–165, 184, 194, 195
 spline 222
 suite 205, 206
 de carré intégrable 127
 de Cauchy voir CAUCHY, 210
 numérique 125



taux d'accroissement 31
TAYLOR, Brook
 coefficient de 69
théorème
 d'inversion locale 102
 des résidus 63, 134
topologie algébrique 28
traitement du signal 194
transformation
 anti-analytique 102
 conforme 100, 105
 de Cayley voir CAYLEY
 de Fourier voir FOURIER
 homographie 98
 homothétie 97, 98
 intégrale 115
 inversion 98
 rotation 97
 translation 97
travail d'une force 11



valeur principale voir CAUCHY
VON NEUMANN, John 239, 248



WEIERSTRASS, Karl
 théorème de 88, 211