

Claire David
Sami Mustapha
Frederi Viens
Nathalie Capron

MATHÉMATIQUES

pour les **SCIENCES** de la **VIE**

TOUT LE COURS EN FICHES

Licence • Prépas • CAPES

140 Fiches de cours
200 exercices corrigés
et exemples
d'application

RESSOURCES



NUMÉRIQUES



DUNOD

Tout le catalogue sur
www.dunod.com



ÉDITEUR DE SAVOIRS

Illustration de couverture : © tr3gi - Fotolia.com

Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements

d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour

les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée. Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du

Centre français d'exploitation du droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).



© Dunod, 2014

5 rue Laromiguière, 75005 Paris

www.dunod.com

ISBN 978-2-10-071800-9

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), 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 les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

Table des matières

<i>Avant-propos</i>	XI
<i>Comment utiliser cet ouvrage ?</i>	XII

Partie 1 Calculus

Nombres réels	2
Fiche 1 Les ensembles de nombres	2
Fiche 2 Intervalles, voisinages, bornes	6
Limites	8
Fiche 3 Limite d'une fonction en un point	8
Fiche 4 Limite d'une fonction en $+\infty$ ou $-\infty$	12
Fiche 5 Propriétés des limites – Opérations sur les limites	14
Fiche 6 Notations de Landau	16
Fonctions numériques	18
Fiche 7 Domaine de définition d'une fonction, graphe	18
Focus <i>La construction de l'ensemble des réels : les coupures de Dedekind</i>	21
Fiche 8 Comment définir une fonction ?	22
Fiche 9 Majorations et minoration	24
Fiche 10 Fonctions monotones	26
Fiche 11 Parité, imparité	28
Fiche 12 Symétries	30
Fiche 13 Fonctions périodiques	32
Fonctions usuelles	33
Fiche 14 Fonctions puissances entières	33
Fiche 15 Fonctions polynômes et fonction valeur absolue	35
Focus <i>John Napier et les tables logarithmiques</i>	38
Fiche 16 La fonction logarithme népérien	39
Fiche 17 La fonction exponentielle	41
Fiche 18 Fonctions puissances « non entières »	43
Focus <i>Leibniz et la fonction exponentielle</i>	44
Fiche 19 Fonctions circulaires	45
Fiche 20 Fonctions hyperboliques	47
Focus <i>L'origine de la trigonométrie</i>	49
Continuité	51
Fiche 21 Continuité d'une fonction en un point	51
Fiche 22 Fonctions continues sur un intervalle	55
Dérivabilité	58
Fiche 23 Dérivabilité en un point	58

Fiche 24	Dérivabilité sur un intervalle	61
Fiche 25	Dérivées successives	65
Fiche 26	Théorème des accroissements finis et théorème de Rolle	67
Fiche 27	Formule de Taylor-Lagrange	71
Fonctions réciproques		72
Fiche 28	Fonctions réciproques	72
Fiche 29	Les fonctions trigonométriques inverses	75
Fiche 30	Les fonctions hyperboliques inverses	79
Développements limités		81
Fiche 31	Développements limités	81
Fiche 32	Formule de Taylor-Young	84
Fiche 33	Développements limités usuels	89
Fiche 34	Opérations algébriques et composition des développements limités	92
Développements asymptotiques		95
Fiche 35	Développements asymptotiques	95
Convexité		96
Fiche 36	Convexité	96
Équations différentielles linéaires du 1^{er} ordre		100
Fiche 37	Équations différentielles linéaires du 1 ^{er} ordre homogènes	100
Fiche 38	Équations différentielles linéaires du 1 ^{er} ordre avec second membre	103
Suites		111
Fiche 39	Qu'est-ce qu'une suite ?	111
Fiche 40	Les différents types de suites	114
Focus	<i>Suites arithmético-géométriques et finance</i>	119
Fiche 41	Étude d'une suite	120
Fiche 42	Majorants, minorants d'une suite réelle Croissance et décroissance	123
Fiche 43	Techniques d'étude des suites réelles	125
Fiche 44	Convergence	127
Fiche 45	Convergence des suites monotones	130
Fiche 46	Opérations sur les limites de suites	132
Fiche 47	Convergence des suites homographiques réelles	135
Fiche 48	Suites extraites	140
Fiche 49	Suites de Cauchy	142
Fiche 50	Comparaison des suites réelles	144
Focus	<i>Suites et systèmes dynamiques – L'attracteur de Hénon</i>	148
Séries		149
Fiche 51	Séries	149
Fiche 52	Quelques séries remarquables	151
Fiche 53	Critères de convergence pour les séries à termes positifs	153

Intégrales	155
Fiche 54 Qu'est-ce qu'une intégrale ?	155
Fiche 55 Intégrale d'une fonction en escaliers	157
Fiche 56 Intégrale d'une fonction continue par morceaux	162
Fiche 57 Calcul intégral	168
Fiche 58 Primitives de fractions rationnelles	174
Fiche 59 Calcul approché d'intégrales	176
Focus <i>Intégrale de Riemann vs intégrale de Lebesgue</i>	183
Fiche 60 Intégrales généralisées	185
Fiche 61 Intégrales doubles sur un pavé du plan \mathbb{R}^2	193
<i>Exercices d'entraînement</i>	195

Partie 2 Algèbre

Le plan complexe – Les nombres complexes	206
Fiche 62 Le corps des nombres complexes	206
Focus <i>Les nombres complexes</i>	209
Fiche 63 Représentation géométrique des nombres complexes	211
Fiche 64 Inversion des nombres complexes	214
Fiche 65 Propriétés fondamentales des nombres complexes	216
Fiche 66 Complément : les polynômes de Tchebychev	218
Fiche 67 Racines $n^{\text{ièmes}}$ de l'unité, racines $n^{\text{ièmes}}$ complexes	221
Fiche 68 Factorisation des polynômes dans le corps \mathbb{C}	224
Fiche 69 Fractions rationnelles et décomposition en éléments simples	229
Fiche 70 Transformations du plan : translations, homothéties	240
Fiche 71 Transformations du plan : rotations	242
Fiche 72 Transformations du plan : similitudes	244
Focus <i>Transformations complexes, fractales, et représentations de la nature</i>	248
Focus <i>L'origine des matrices</i>	250
Matrices	252
Fiche 73 Matrices de taille 2×2	252
Fiche 74 Déterminant de matrices de taille 2×2	254
Fiche 75 Matrices de taille 3×3	256
Fiche 76 Déterminant de matrices de taille 3×3	259
Fiche 77 Matrices de taille $m \times n$	262
Fiche 78 Opérations sur les matrices	264
Fiche 79 Matrices remarquables	266
Fiche 80 Introduction aux déterminants de matrices de taille $n \times n$	270
Fiche 81 Inversion des matrices carrées	272
Focus <i>Les matrices et leurs applications</i>	276
Fiche 82 Systèmes linéaires	278

Focus	<i>Matrices, systèmes linéaires et chimie</i>	281
Fiche 83	Vecteurs	282
Fiche 84	Barycentres	286
Fiche 85	Droites, plans	290
Fiche 86	Produit scalaire	293
Focus	<i>Produit scalaire, espaces fonctionnels et calcul numérique</i>	297
Fiche 87	Produit vectoriel	298
Fiche 88	Aires et volumes	300
Focus	<i>Géométrie euclidienne – ou non ? Encore des matrices !</i>	302
Transformations linéaires du plan		304
Fiche 89	Bases et transformations linéaires du plan	304
Fiche 90	Changement de base en dimension 2, et déterminant d'une application linéaire	308
Fiche 91	Conjugaison – Matrices semblables de taille 2×2	310
Fiche 92	Opérateurs orthogonaux en dimension 2	312
Fiche 93	Rotations vectorielles du plan	314
Transformations linéaires de l'espace		317
Fiche 94	Bases de l'espace \mathbb{R}^3	317
Fiche 95	Transformations linéaires de l'espace \mathbb{R}^3	318
Fiche 96	Changement de base en dimension 3	322
Fiche 97	Conjugaison – Matrices semblables de taille 3×3	324
Fiche 98	Opérateurs orthogonaux de l'espace \mathbb{R}^3	326
Fiche 99	Rotations vectorielles de l'espace \mathbb{R}^3	328
L'espace \mathbb{R}^n		330
Fiche 100	Vecteurs en dimension n , $n \geq 2$	330
Fiche 101	Espace engendré par une famille de vecteurs Sous-espaces vectoriels de \mathbb{R}^n	332
Fiche 102	Transformations linéaires de l'espace \mathbb{R}^n	335
Fiche 103	Changement de base	339
Fiche 104	Conjugaison – Matrices semblables de taille $n \times n$	341
Fiche 105	Réduction des matrices carrées	343
Focus	<i>Groupe spécial orthogonal et cristallographie</i>	347
Focus	<i>Diagonalisation – La toupie de Lagrange (et de Michèle Audin)</i>	349
Espaces vectoriels		350
Fiche 106	Les espaces vectoriels	350
Fiche 107	Sous-espaces vectoriels	354
Fiche 108	Somme de sous-espaces vectoriels	356
Fiche 109	Projecteurs, symétries	357
Exercices d'entraînement		359

Partie 3 Probabilités

Analyse Combinatoire	368
Fiche 110 Factorielle	368
Fiche 111 Arrangements	370
Fiche 112 Combinaisons	372
Espaces probabilisés	376
Fiche 113 Espaces probabilisés	376
Focus <i>Les jeux de hasard</i>	380
Focus <i>Lancer de dés</i>	382
Fiche 114 Conditionnement	384
Fiche 115 Indépendance	388
Probabilités discrètes	389
Fiche 116 Variable aléatoire discrète et loi associée	389
Fiche 117 Fonction de répartition	392
Fiche 118 La loi de Bernoulli, de paramètre $p \in [0, 1]$	394
Fiche 119 La loi uniforme (sur un ensemble de réels)	396
Fiche 120 La loi binomiale	397
Fiche 121 La loi géométrique	400
Fiche 122 La loi binomiale négative	404
Focus <i>Les tortues des Galápagos</i>	406
Fiche 123 La loi hypergéométrique	409
Fiche 124 La loi de Poisson	412
Fiche 125 Somme de variables aléatoires discrètes	416
Fiche 126 Espérance	419
Fiche 127 Moment d'ordre r , $r \in \mathbb{N}^*$, d'une variable aléatoire discrète	425
Focus <i>Application de l'inégalité de Markov</i>	427
Focus <i>Variations de température</i>	429
Fiche 128 Variance, écart-type	431
Focus <i>Inégalité de Tchebychev et concentration de mesure</i>	439
Fiche 129 Covariance	440
Fiche 130 Couples de variables aléatoires discrètes	445
Fiche 131 Un théorème limite pour les variables aléatoires discrètes : la loi des grands nombres	447
Probabilités continues	449
Fiche 132 Du discret au continu : variables aléatoires continues	449
Fiche 133 Variables à densité	450
Fiche 134 Lois uniformes	457
Focus <i>Le coût d'un déminage avec la loi uniforme</i>	460
Fiche 135 Lois exponentielles	461
Fiche 136 Lois normales (ou gaussiennes)	465
Fiche 137 Couples de variables aléatoires réelles	472

Fiche 138	Quelques résultats concernant la somme de variables aléatoires continues	475
Fiche 139	Lois Gamma	480
Focus	<i>Dans l'attente d'un $n^{\text{ième}}$ client rue de la Paix</i>	487
	<i>Exercices d'entraînement</i>	489
Annexes	Formulaire de trigonométrie	493
	Dérivées usuelles	495
	Dérivées des fonctions réciproques usuelles	496
	Primitives usuelles	497
	Limites usuelles des fonctions puissances	498
	Rang d'une matrice	499
	<i>Bibliographie</i>	500
	<i>Index</i>	502

Avant-propos

Cet ouvrage est destiné aux étudiants en licences de Sciences de la Vie et de la Terre, ou en classes préparatoires. Il se base sur nos cours donnés en première année de Licence à l'UPMC (université Pierre-et-Marie-Curie).

Face aux demandes croissantes de nos étudiants, qui recherchaient un ouvrage de référence complet mais abordable, ainsi que des exercices d'application corrigés, nous nous sommes lancés dans la conception de ce livre qui, nous l'espérons, sera un outil utile pour les générations d'étudiants à venir.

Cet ouvrage est donc le fruit d'un compromis : dans un volume condensé, nous avons essayé de donner suffisamment d'éléments recouvrant l'ensemble des mathématiques des premières années d'études supérieures. Il correspond aussi à l'arrivée des nouveaux programmes universitaires et des classes préparatoires. Pour mieux assurer la jonction avec les mathématiques enseignées au lycée, nous avons opté, pour la première partie d'analyse, relative à l'étude des fonctions, à une présentation de type « Calculus », inspirée de l'esprit des « textbooks » anglo-saxons, qui permet d'aborder plus facilement le reste du programme, plus « classique », sur les suites et le calcul intégral. Pour l'algèbre, la présentation reprend celle de l'ouvrage *Calcul Vectoriel* (Collection *Sciences Sup*), en allant un peu plus loin : \mathbb{R}^n , réduction, espaces vectoriels. En ce qui concerne les probabilités, nous allons jusqu'aux probabilités continues, indispensables aux filières SVT, mais qui peuvent aussi intéresser les autres publics.

Malgré tout le soin apporté à la rédaction, nous demandons l'indulgence du lecteur pour les éventuelles imperfections qui pourraient subsister ; qu'il n'hésite pas à nous les signaler.

Claire David et Sami Mustapha
Claire.David@upmc.fr
Sami.Mustapha@imj-prg.fr

Remerciements

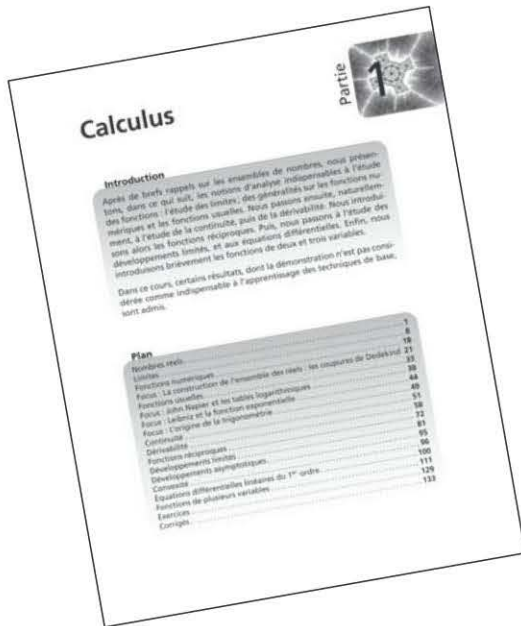
Nous remercions vivement toutes les personnes dont la relecture et les remarques ont contribué à améliorer la version initiale du manuscrit :

les membres du comité de lecture, pour leur relecture extrêmement minutieuse et leurs remarques très pertinentes ;

- Maryse Beguin, laboratoire Jean Kuntzman, Institut Polytechnique de Grenoble.
- Sylvie Benzoni, université Claude Bernard Lyon 1, Institut Camille Jordan.
- Laurent Di Menza, université de Reims, laboratoire de mathématiques de Reims (LMR).
- Jean-Pierre Escofier, université de Rennes, Institut mathématique de Rennes.
- Sandrine Gachet, professeur de mathématiques, lycée Gustave Eiffel, Dijon.
- Chloé Mullaert, professeur de mathématiques, lycée Paul Valéry, Paris.
- Laure Quivy, ENS Cachan et université Paris XIII, Centre de mathématiques et leurs applications (CMLA).
- Olivier Sellès, professeur de mathématiques, lycée Saint-Louis, Paris.
- Lamia Attouche, étudiante à l'UPMC, Paris.
- Alexis Prel, étudiant à l'UPMC, Paris.

mais aussi Albert Cohen, Ramona Anton, Sylvie Delabrière, Patrick Polo, Adnène Benabdesslem, Matthieu Solnon, Eugénie Poulon, Daniel Hoehener, Julien Piera Vest.

Comment utiliser cet ouvrage ?



Un découpage
en trois grandes parties :
Calculus, Algèbre, Probabilités

140 fiches de cours

Elles présentent les notions essentielles du cours

fiche 1 Les ensembles de nombres

Un ensemble E est une collection d'objets, qui constituent les « éléments » de l'ensemble. Le nombre d'éléments de l'ensemble peut être fini, ou infini.

1. Notation

Pour décrire l'ensemble, on utilise des accolades, à l'intérieur desquelles on écrit les éléments de l'ensemble.

Suivant les cas, on peut, simplement, placer, à l'intérieur des accolades, la liste des éléments de l'ensemble ; ainsi, dans le cas d'un ensemble E avec un nombre fini d'éléments e_1, \dots, e_n , où n est un nombre entier positif, on écrit :

$$E = \{e_1, \dots, e_n\}$$

ou bien, dans le cas d'un ensemble d'éléments vérifiant une propriété donnée P , on écrit

$$E = \{x \mid P(x)\} \quad \text{ou encore} \quad \{x, P(x)\}$$

ce qui désigne ainsi l'ensemble des éléments x tels que la propriété P soit vérifiée pour x .

Exemples

- $\{1, 2, 3, 4\}$ est un ensemble. Ses éléments sont les nombres 1, 2, 3 et 4.
- $\{3, 4, 5, 6, \dots\}$ est un ensemble. Ses éléments sont les nombres entiers supérieurs ou égaux à 3.
- $\{x \in \{1, 2, 3, 4, 5, 6\} \mid x \text{ est impair}\} = \{1, 3, 5\}$.

► **Les entiers naturels**
L'ensemble des entiers naturels, c'est-à-dire des entiers positifs ou nuls, est noté \mathbb{N} :

$$\mathbb{N} = \{0, 1, 2, 3, 4, 5, \dots\}$$

► **Les nombres pairs**
L'ensemble des entiers naturels pairs est noté $2\mathbb{N}$:

$$2\mathbb{N} = \{0, 2, 4, 6, \dots\} = \{2n, n \in \mathbb{N}\}$$

► $k\mathbb{N}, k \in \mathbb{N}$
Étant donné un entier naturel non nul k , $k\mathbb{N}$ désigne l'ensemble des entiers naturels multiples de k :

$$k\mathbb{N} = \{kn, n \in \mathbb{N}\}$$

► **Les entiers relatifs**
L'ensemble des entiers relatifs, c'est-à-dire des entiers qui sont soit positifs ou nuls, soit négatifs ou nuls, est noté \mathbb{Z} :

$$\mathbb{Z} = \{\dots, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, \dots\}$$

► $a \in \mathbb{Z}, a \in \mathbb{R}$
Étant donné un réel non nul a , $a \in \mathbb{Z}$ désigne l'ensemble des réels de la forme ak , où k est un entier :

$$a\mathbb{Z} = \{ak, k \in \mathbb{Z}\}$$

Exemple
 $2\pi\mathbb{Z} = \{2k\pi, k \in \mathbb{Z}\}$

► **Les nombres rationnels**
L'ensemble des nombres rationnels, c'est-à-dire de la forme $\frac{p}{q}$, où p et q sont deux entiers relatifs, avec $q \neq 0$, est noté \mathbb{Q} .

► **Les nombres réels**
L'ensemble des nombres réels est noté \mathbb{R} .

► \mathbb{R}
L'ensemble $\mathbb{R} \cup \{-\infty, +\infty\}$ est noté $\overline{\mathbb{R}}$ (c'est ce que l'on appelle la « droite réelle achevée », ou encore, l'adhérence de \mathbb{R}).

► **La notation « ' »**
Lorsque l'on écrit l'un des ensembles précédents avec l'exposant « ' », cela signifie que l'on exclut 0 ; ainsi, \mathbb{N}' désigne l'ensemble des entiers naturels non nuls ; \mathbb{Z}' désigne l'ensemble des entiers relatifs non nuls ; etc.

► **La notation « + »**
Lorsque l'on écrit l'un des ensembles précédents avec l'exposant « + », cela signifie que l'on ne considère que les nombres positifs de cet ensemble ; ainsi, \mathbb{Z}^+ (qui est aussi égal à \mathbb{N}), désigne l'ensemble des entiers positifs ou nuls ; \mathbb{R}^+ désigne l'ensemble des réels positifs ou nuls ; etc.

► **La notation « - »**
Lorsque l'on écrit l'un des ensembles précédents avec l'exposant « - », cela signifie que l'on ne considère que les nombres négatifs de cet ensemble ; ainsi, \mathbb{Z}^- (qui est aussi égal à $-\mathbb{N}$), désigne l'ensemble des entiers strictement négatifs ; \mathbb{R}^- désigne l'ensemble des réels strictement négatifs ; etc.

► **La notation « ± »**
Lorsque l'on écrit l'un des ensembles précédents avec l'exposant « ± », cela signifie que l'on ne considère que les nombres strictement positifs de cet ensemble ; ainsi, \mathbb{Z}^\pm (qui est aussi égal à \mathbb{N}^*), désigne l'ensemble des entiers strictement positifs ; \mathbb{R}^\pm désigne l'ensemble des réels strictement positifs ; etc.

Propriété
On a :

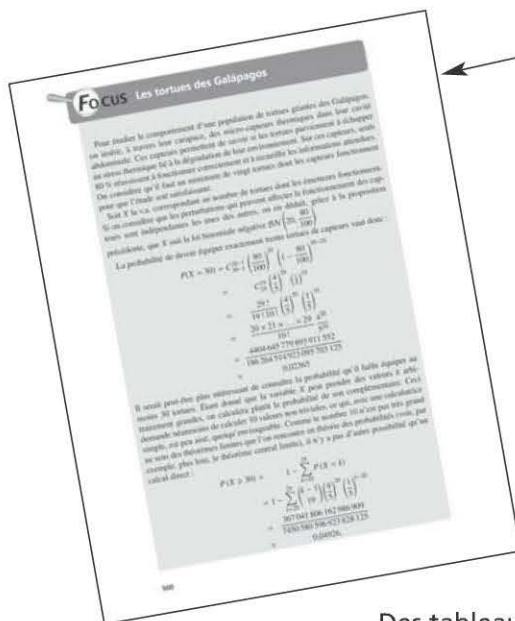
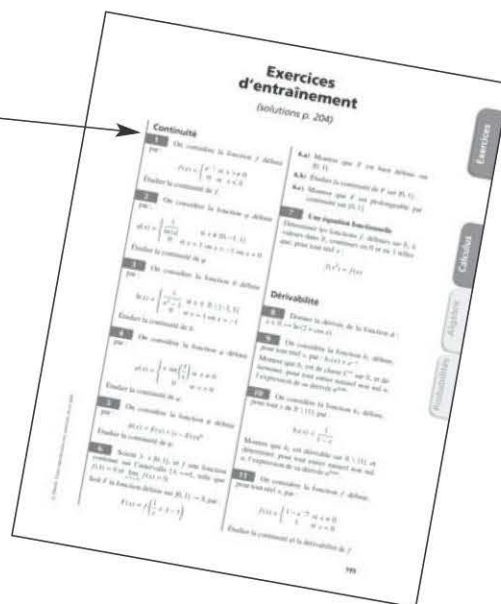
$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$$

Un repérage facile

Les fiches sont regroupées par thème

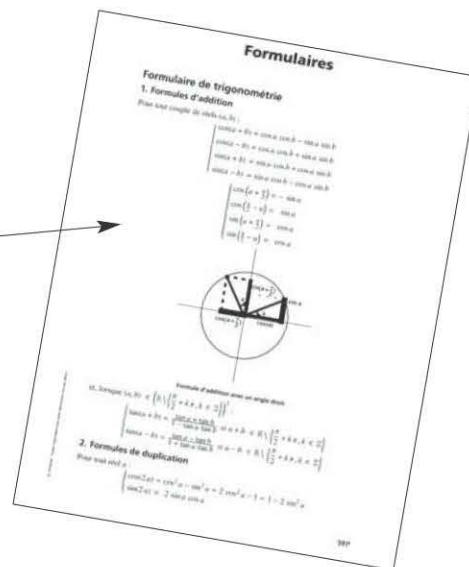
De très nombreux exemples

À la fin de chaque partie, des **exercices corrigés** pour s'entraîner. Les corrigés sont disponibles sur le site dunod.com



Des **focus** pour découvrir des **applications des mathématiques** ou approfondir un point du cours

Des **tableaux et formules** en annexe

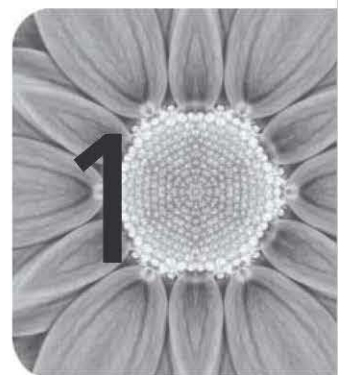


Tous les corrigés sont disponibles en téléchargement sur le site dunod.com à partir de la page de présentation de l'ouvrage.

Calculus

Partie

1



Introduction

Après de brefs rappels sur les ensembles de nombres, nous présentons, dans ce qui suit, les notions d'analyse indispensables à l'étude des fonctions : l'étude des limites ; des généralités sur les fonctions numériques et les fonctions usuelles. Nous passons ensuite, naturellement, à l'étude de la continuité, puis de la dérivabilité. Nous introduisons alors les fonctions réciproques. Puis, nous passons à l'étude des développements limités, et aux équations différentielles. Enfin, nous introduisons brièvement les fonctions de deux et trois variables.

Dans ce cours, certains résultats, dont la démonstration n'est pas considérée comme indispensable à l'apprentissage des techniques de base, sont admis.

Plan

Nombres réels	2
Limites	8
Fonctions numériques	18
<i>Focus : La construction de l'ensemble des réels : les coupures de Dedekind</i>	21
Fonctions usuelles	33
<i>Focus : John Napier et les tables logarithmiques</i>	38
<i>Focus : Leibniz et la fonction exponentielle</i>	44
<i>Focus : L'origine de la trigonométrie</i>	49
Continuité	51
Dérivabilité	58
Fonctions réciproques	72
Développements limités	81
Développements asymptotiques	95
Convexité	96
Équations différentielles linéaires du 1 ^{er} ordre	100
Suites	111
<i>Focus : Suites arithmético-géométriques et finance</i>	119
<i>Focus : Suites et systèmes dynamiques – L'attracteur de Hénon</i>	148
Séries	149
Intégrales	149
<i>Focus : Intégrale de Riemann vs intégrale de Lebesgue</i>	183
Exercices d'entraînement	195

Les bonus web sur Dunod.com

www Les corrigés des exercices sont consultables sur dunod.com sur la page de présentation de l'ouvrage.

Les ensembles de nombres

Un **ensemble** E est une collection d'objets, qui constituent les « éléments » de l'ensemble. Le nombre d'éléments de l'ensemble peut être fini, ou infini.

1. Notation

Pour décrire l'ensemble, on utilise des accolades, à l'intérieur desquelles on écrit les éléments de l'ensemble.

Suivant les cas, on peut, simplement, placer, à l'intérieur des accolades, la liste des éléments de l'ensemble ; ainsi, dans le cas d'un ensemble E avec un nombre fini d'éléments e_1, \dots, e_n , où n est un nombre entier positif, on écrit :

$$E = \{e_1, \dots, e_n\}$$

ou bien, dans le cas d'un ensemble d'éléments vérifiant une propriété donnée \mathcal{P} , on écrit

$$E = \left\{x \mid \mathcal{P}(x)\right\} \quad \text{ou encore} \quad \{x, \mathcal{P}(x)\} \quad \text{ou encore} \quad \{x; \mathcal{P}(x)\}$$

ce qui désigne ainsi l'ensemble des éléments x tels que la propriété \mathcal{P} soit vérifiée pour x .

Exemples

1. $\{1, 2, 3, 4\}$ est un ensemble. Ses éléments sont les nombres 1, 2, 3 et 4.
2. $\{3, 4, 5, 6, \dots\}$ est un ensemble. Ses éléments sont les nombres entiers supérieurs ou égaux à 3.
3. $\left\{x \in \{1, 2, 3, 4, 5, 6\} \mid x \text{ est impair}\right\} = \{1, 3, 5\}$.

► Les entiers naturels

L'ensemble des entiers naturels, c'est-à-dire des entiers positifs ou nuls, est noté \mathbb{N} :

$$\mathbb{N} = \{0, 1, 2, 3, 4, 5, \dots\}$$

► Les nombres pairs

L'ensemble des entiers naturels pairs est noté $2\mathbb{N}$:

$$2\mathbb{N} = \{0, 2, 4, 6, \dots\} = \{2n, n \in \mathbb{N}\}$$

► $k\mathbb{N}, k \in \mathbb{N}$

Étant donné un entier naturel k , $k\mathbb{N}$ désigne l'ensemble des entiers naturels multiples de k :

$$k\mathbb{N} = \{kn, n \in \mathbb{N}\}$$

► Les entiers relatifs

L'ensemble des entiers relatifs, c'est-à-dire des entiers qui sont soit positifs ou nuls, ou négatifs ou nuls, est noté \mathbb{Z} :

$$\mathbb{Z} = \{\dots, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, \dots\}$$

► $\alpha \mathbb{Z}, \alpha \in \mathbb{R}$

Étant donné un réel α , $\alpha \mathbb{Z}$ désigne l'ensemble des réels de la forme αk , où k est un entier :

$$\alpha \mathbb{Z} = \{\alpha k, k \in \mathbb{Z}\}$$

Exemple

$$2\pi \mathbb{Z} = \{2k\pi, k \in \mathbb{Z}\}.$$

► Les nombres rationnels

L'ensemble des nombres rationnels, c'est-à-dire de la forme $\frac{p}{q}$, où p et q sont deux entiers relatifs, avec $q \neq 0$, est noté \mathbb{Q} .

► Les nombres réels

L'ensemble des nombres réels est noté \mathbb{R} .

► $\overline{\mathbb{R}}$

L'ensemble $\mathbb{R} \cup \{-\infty, +\infty\}$ est noté $\overline{\mathbb{R}}$ (c'est ce que l'on appelle la « droite réelle achevée », ou encore, l'adhérence de \mathbb{R})

► La notation « * »

Lorsque l'on écrit l'un des ensembles précédents avec l'exposant « * », cela signifie que l'on exclut 0 ; ainsi, \mathbb{N}^* désigne l'ensemble des entiers naturels non nuls ; \mathbb{Z}^* désigne l'ensemble des entiers relatifs non nuls ; etc.

► La notation « + »

Lorsque l'on écrit l'un des ensembles précédents avec l'exposant « + », cela signifie que l'on ne considère que les nombres positifs de cet ensemble ; ainsi, \mathbb{Z}^+ (qui est aussi égal à \mathbb{N}), désigne l'ensemble des entiers positifs ou nuls ; \mathbb{R}^+ désigne l'ensemble des réels positifs ou nuls ; etc.

► La notation « - »

Lorsque l'on écrit l'un des ensembles précédents avec l'exposant « - », cela signifie que l'on ne considère que les nombres négatifs de cet ensemble ; ainsi, \mathbb{Z}^- (qui est aussi égal à $-\mathbb{N}$), désigne l'ensemble des entiers négatifs ou nuls ; \mathbb{R}^- désigne l'ensemble des réels positifs ou nuls ; etc.

► La notation « *₊ »

Lorsque l'on écrit l'un des ensembles précédents avec l'exposant « *₊ », cela signifie que l'on ne considère que les nombres strictement positifs de cet ensemble ; ainsi, \mathbb{Z}_{+}^* (qui est aussi égal à \mathbb{N}^*), désigne l'ensemble des entiers strictement positifs ; \mathbb{R}_{+}^* désigne l'ensemble des réels strictement positifs ; etc.

► La notation « *₋ »

Lorsque l'on écrit l'un des ensembles précédents avec l'exposant « *₋ », cela signifie que l'on ne considère que les nombres strictement négatifs de cet ensemble ; ainsi, \mathbb{Z}_{-}^* (qui est aussi égal à $-\mathbb{N}^*$), désigne l'ensemble des entiers strictement négatifs ; \mathbb{R}_{-}^* désigne l'ensemble des réels strictement négatifs ; etc.

Propriété

On a : $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$

où le symbole \subset signifie « inclus dans ».

2. Les ensembles

► Ensemble vide

Un ensemble ne contenant aucun élément est appelé **ensemble vide**, et noté \emptyset .

Exemple

$\{n \in 3\mathbb{N}, n \text{ pair}\}$ ne contient aucun nombre : c'est l'ensemble vide.

► Intersection d'ensembles

Étant donnés deux ensembles E_1 et E_2 , leur **intersection**, notée $E_1 \cap E_2$, est l'ensemble des éléments qui appartiennent à la fois à E_1 et à E_2 :

$$E_1 \cap E_2 = \{x, x \in E_1 \text{ et } x \in E_2\}$$

► Union d'ensembles

Étant donnés deux ensembles E_1 et E_2 , leur **union**, notée $E_1 \cup E_2$, est l'ensemble des éléments qui appartiennent à E_1 , ou à E_2 :

$$E_1 \cup E_2 = \{x, x \in E_1 \text{ ou } x \in E_2\}$$

► Différence de deux ensembles

Étant donnés deux ensembles E_1 et E_2 , leur **différence**, notée $E_1 \setminus E_2$, est l'ensemble E_1 privé de E_2 :

$$E_1 \setminus E_2 = \{x, x \in E_1 \text{ et } x \notin E_2\}$$

Exemples

1. $\mathbb{R} \setminus \{1, 2\}$ est l'ensemble des réels différents de 1 et de 2.
2. $\mathbb{R} \setminus \pi\mathbb{Z}$ est l'ensemble des réels qui ne sont pas multiples de π .

► Complémentaire d'un ensemble

Étant donnés deux ensembles E_1 et E_2 tels que E_2 soit inclus dans E_1 (que l'on écrit $E_2 \subset E_1$), l'ensemble $E_1 \setminus E_2$ est le complémentaire de E_2 dans E_1 , noté $\complement_{E_1} E_2$:

$$\complement_{E_1} E_2 = E_1 \setminus E_2$$

Exemple

$$\complement_{\mathbb{R}} \{0\} = \mathbb{R}^*$$

► Produit cartésien de deux ensembles

Étant donnés deux ensembles E_1 et E_2 , leur **produit cartésien**, noté $E_1 \times E_2$, est l'ensemble des couples d'éléments de la forme (x_1, x_2) , où le premier élément x_1 appartient à E_1 , et le second, x_2 , à E_2 :

$$E_1 \times E_2 = \{(x_1, x_2), x_1 \in E_1 \text{ et } x_2 \in E_2\}$$

Exemples

1. $\mathbb{R}^2 = \{(x_1, x_2), x_1 \in \mathbb{R} \text{ et } x_2 \in \mathbb{R}\}$ est l'ensemble des couples de réels.
2. $\mathbb{N}^2 = \{(n_1, n_2), n_1 \in \mathbb{N} \text{ et } n_2 \in \mathbb{N}\}$ est l'ensemble des couples d'entiers naturels.

► Produit cartésien de trois ensembles

Étant donnés trois ensembles E_1 , E_2 et E_3 , leur **produit cartésien**, noté $E_1 \times E_2 \times E_3$, est l'ensemble des triplets d'éléments de la forme (x_1, x_2, x_3) , où le premier élément x_1 appartient à E_1 , le second, x_2 , à E_2 , et le troisième, x_3 , à E_3 :

$$E_1 \times E_2 \times E_3 = \{(x_1, x_2, x_3), x_1 \in E_1, x_2 \in E_2 \text{ et } x_3 \in E_3\}$$

► Produit cartésien de n ensembles, $n \in \mathbb{N}$, $n \geq 2$

Étant donnés un entier naturel $n \geq 2$, et n ensembles E_1, \dots, E_n , leur **produit cartésien**, noté $E_1 \times \dots \times E_n$, est l'ensemble des n -uplets d'éléments de la forme (x_1, \dots, x_n) , où $x_1 \in E_1, \dots, x_n \in E_n$:

$$E_1 \times \dots \times E_n = \{(x_1, \dots, x_n), x_1 \in E_1, \dots, x_n \in E_n\}$$

► Application

Étant donnés deux ensembles E et F , une **application** φ de E dans F associe, à chaque élément de E , un et un seul élément de F . E est l'ensemble de départ, F , celui d'arrivée.

Pour tout élément x de E , l'unique élément de F ainsi mis en relation avec x par l'application φ est noté $\varphi(x)$, et appelé image de x . x est un antécédent de $\varphi(x)$. On écrit :

$$\begin{aligned} \varphi : E &\rightarrow F \\ x &\mapsto \varphi(x) \end{aligned}$$

Exemples

1.

$$\begin{aligned} \varphi : \mathbb{N} &\rightarrow \mathbb{N} \\ x &\mapsto x \end{aligned}$$

est une application de \mathbb{N} dans \mathbb{N} , appelée application identité de \mathbb{N} .

2.

$$\begin{aligned} \psi : \mathbb{Q} &\rightarrow \mathbb{Q} \\ x &\mapsto 2x \end{aligned}$$

est une application de \mathbb{Q} dans \mathbb{Q} .

► Fonction

Étant donnés deux ensembles de nombres E et F , une **fonction** f de E dans F associe, à chaque élément x de E , au plus un élément de F appelé alors image de x par f (ce qui signifie donc que tous les éléments de E n'ont pas nécessairement une image par f). E est l'ensemble de départ, F , celui d'arrivée. L'ensemble des éléments de E possédant une image par f est appelé domaine de définition de f , et noté \mathcal{D}_f . Elle permet de définir une application de \mathcal{D}_f dans F .

Exemple

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto \frac{1}{1-x} \end{aligned}$$

est une fonction de \mathbb{R} dans \mathbb{R} , dont le domaine de définition est $\mathbb{R} \setminus \{1\}$. Elle permet de définir une application de $\mathbb{R} \setminus \{1\}$ dans \mathbb{R} .

Intervalles, voisinages, bornes

L'ensemble des nombres réels est habituellement représenté sous la forme d'une droite graduée, appelée **droite des réels**, où il faut pouvoir se repérer. À cet effet, on introduit les notions d'intervalle et de voisinage d'un point.

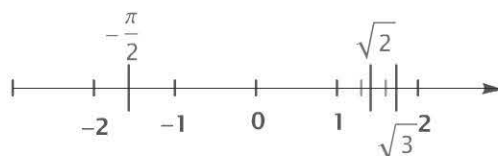


Figure 2.1 – La droite des réels.

1. Intervalles

► Intervalle fermé et borné (ou segment)

On appelle intervalle fermé et borné (ou segment) tout ensemble de la forme

$$[a, b] = \{x \in \mathbb{R}, a \leq x \leq b\} \quad , \quad (a, b) \in \mathbb{R}^2, a \leq b$$

► Intervalle ouvert

On appelle intervalle ouvert tout ensemble de la forme

$$]a, b[= \{x \in \mathbb{R}, a < x < b\} \quad , \quad (a, b) \in \mathbb{R}^2, a < b$$

ou $] -\infty, b[= \{x \in \mathbb{R}, x < b\} \quad , \quad b \in \mathbb{R}$

ou encore $]a, +\infty[= \{x \in \mathbb{R}, a < x\} \quad , \quad a \in \mathbb{R}$

où $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ est l'ensemble des couples de réels.

► Intervalle ouvert et borné

On appelle intervalle ouvert et borné tout ensemble de la forme

$$]a, b[= \{x \in \mathbb{R}, a < x < b\} \quad , \quad (a, b) \in \mathbb{R}^2, a < b$$

► Intervalle semi-ouvert et borné

On appelle intervalle semi-ouvert et borné tout ensemble de la forme

$$[a, b[= \{x \in \mathbb{R}, a \leq x < b\} \quad , \quad (a, b) \in \mathbb{R}^2, a < b$$

ou $]a, b] = \{x \in \mathbb{R}, a < x \leq b\} \quad , \quad (a, b) \in \mathbb{R}^2, a < b$

► Intervalle fermé

Par convention, tout ensemble de la forme

$$[a, +\infty[= \{x \in \mathbb{R}, x \geq a\} \quad , \quad a \in \mathbb{R}$$

ou $] -\infty, b] = \{x \in \mathbb{R}, x \leq b\} \quad , \quad b \in \mathbb{R}$

est considéré comme étant un intervalle fermé.

► Ensemble vide

L'ensemble, noté \emptyset , qui ne contient aucun nombre réel, est aussi un intervalle, appelé ensemble vide.

► Singleton

On appelle singleton un ensemble ne contenant qu'un seul élément, et qui est donc de la forme $\{a\}$, où a est un nombre réel.

► Intervalle

On appelle intervalle de \mathbb{R} l'un des ensembles définis ci-dessus, ou bien \mathbb{R} tout entier.



Un singleton est un intervalle fermé (le singleton $\{a\}$ est donc assimilé à l'intervalle fermé $[a, a]$).

► Adhérence d'un intervalle

Soit I un intervalle de \mathbb{R} . Son adhérence \bar{I} est l'ensemble tel que :

- si I est un segment, alors $\bar{I} = I$;
- si I est de la forme $]a, b[$ ou $]a, b]$ ou $[a, b[$, $(a, b) \in \mathbb{R}^2$, alors $\bar{I} = [a, b]$;
- si I est de la forme $]a, +\infty[$ ou $[a, +\infty[$, $a \in \mathbb{R}$, alors $\bar{I} = [a, +\infty[\cup \{+\infty\}$;
- si I est de la forme $] - \infty, a[$ ou $] - \infty, a]$, $a \in \mathbb{R}$, alors $\bar{I} =] - \infty, a] \cup \{-\infty\}$;
- si I l'ensemble vide \emptyset , alors $\bar{I} = \emptyset$.

2. Voisinage

► Voisinage d'un point

On appelle voisinage d'un point a de \mathbb{R} un sous-ensemble de \mathbb{R} contenant un intervalle ouvert de la forme $]a - \eta, a + \eta[$, où η est un réel strictement positif et tel que $\eta < a$.



On peut étendre la notion de voisinage à $+\infty$ ou $-\infty$; ainsi, un voisinage de $+\infty$ est une partie de \mathbb{R} contenant un intervalle ouvert de la forme $]x_0, +\infty[$, où x_0 est un nombre réel quelconque. De même, un voisinage de $-\infty$ est une partie de \mathbb{R} contenant un intervalle ouvert de la forme $] - \infty, x_0[$, où x_0 est un nombre réel quelconque.

3. Les intervalles de \mathbb{R}

Dans ce qui suit, a, b, x_0 sont des réels tels que $a < b$. Le tableau suivant reprend les différents types d'intervalles de \mathbb{R} .

$[a, b]$	Segment
$]a, b[$	Intervalle ouvert et borné
$]a, b]$	Intervalle semi-ouvert et borné (ouvert à gauche, fermé à droite)
$[a, b[$	Intervalle semi-ouvert et borné (fermé à gauche, ouvert à droite)
\emptyset	Ensemble vide
$\{a\}$	Singleton
$]x_0, +\infty[$	Voisinage de $+\infty$
$[x_0, +\infty[$	
$] - \infty, x_0[$	Voisinage de $-\infty$
$] - \infty, x_0]$	
$] - \infty, +\infty[$	\mathbb{R} tout entier

Limite d'une fonction en un point

1. Limite finie d'une fonction en un point

Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , a un point de I , et ℓ un réel.

On dit que f **admet pour limite (finie) ℓ en a** si, lorsque x devient très proche de a , $f(x)$ devient lui aussi très proche de ℓ , ce qui se traduit mathématiquement par le fait que pour tout réel ε strictement positif, il existe un réel η strictement positif tel que :

$$\forall x \in I, 0 < |x - a| < \eta \Rightarrow |f(x) - \ell| < \varepsilon$$

On écrit : $\lim_{x \rightarrow a} f(x) = \ell$ ou $\lim_a f = \ell$.

Exemple

On considère la fonction qui, à tout x de $] -1, 1[$, associe $\sqrt{1 - x^2}$. Alors :

$$\lim_{x \rightarrow 1} \sqrt{1 - x^2} = 0$$

► Notation 0^+

Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a un point de I .

On dit que f **tend vers 0^+ en a** si, lorsque x devient très proche de a , $f(x)$ tend vers zéro, mais en restant positif, ce qui se traduit mathématiquement par le fait que pour tout réel ε strictement positif, il existe un réel η strictement positif tel que :

$$\forall x \in I, 0 < |x - a| < \eta \Rightarrow 0 \leq f(x) < \varepsilon$$

On écrit : $\lim_{x \rightarrow a} f(x) = 0^+$ ou $\lim_a f = 0^+$.

Lorsque $+\infty$ est une borne de I , on dit que f **tend vers 0^+ en $+\infty$** si, lorsque x devient très grand, $f(x)$ tend vers zéro, mais en restant positif, ce qui se traduit mathématiquement par le fait que pour tout réel ε strictement positif, il existe un réel A strictement positif tel que :

$$\forall x \in I, x > A \Rightarrow 0 \leq f(x) < \varepsilon$$

On écrit : $\lim_{x \rightarrow +\infty} f(x) = 0^+$ ou $\lim_{+\infty} f = 0^+$.

Lorsque $-\infty$ est une borne de I , on dit que f **tend vers 0^+ en $-\infty$** si, lorsque x devient très grand en valeur absolue, mais en restant à valeurs négatives, $f(x)$ tend vers zéro, mais en restant positif, ce qui se traduit mathématiquement par le fait que pour tout réel ε strictement positif, il existe un réel A strictement positif tel que :

$$\forall x \in I, x < -A \Rightarrow 0 \leq f(x) < \varepsilon$$

On écrit : $\lim_{x \rightarrow -\infty} f(x) = 0^+$ ou $\lim_{-\infty} f = 0^+$.

Exemple

$$\lim_{x \rightarrow 0} x^2 = 0^+$$



On utilisera aussi la notation 0^+ pour indiquer que l'on tend vers zéro par valeurs supérieures.

► Notation 0^-

Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a un point de I .

On dit que f **tend vers 0^- en a** si, lorsque x devient très proche de a , $f(x)$ tend vers zéro, mais en restant négatif, ce qui se traduit mathématiquement par le fait que pour tout réel ε strictement positif, il existe un réel η strictement positif tel que :

$$\forall x \in I, 0 < |x - a| < \eta \Rightarrow -\varepsilon < f(x) \leq 0$$

On écrit : $\lim_{x \rightarrow a} f(x) = 0^-$ ou $\lim_a f = 0^-$.

Lorsque $+\infty$ est une borne de I , on dit que f **tend vers 0^- en $+\infty$** si, lorsque x devient très grand, $f(x)$ tend vers zéro, mais en restant négatif, ce qui se traduit mathématiquement par le fait que pour tout réel ε strictement positif, il existe un réel A strictement positif tel que :

$$\forall x \in I, x > A \Rightarrow -\varepsilon < f(x) \leq 0$$

On écrit : $\lim_{x \rightarrow +\infty} f(x) = 0^-$ ou $\lim_{+\infty} f = 0^-$.

Lorsque $-\infty$ est une borne de I , on dit que f **tend vers 0^- en $-\infty$** si, lorsque x devient très grand en valeur absolue, mais en restant à valeurs négatives, $f(x)$ tend vers zéro, mais en restant négatif, ce qui se traduit mathématiquement par le fait que pour tout réel ε strictement positif, il existe un réel A strictement positif tel que :

$$\forall x \in I, x < -A \Rightarrow -\varepsilon < f(x) \leq 0$$

On écrit : $\lim_{x \rightarrow -\infty} f(x) = 0^-$ ou $\lim_{-\infty} f = 0^-$.

Exemple

$$\lim_{x \rightarrow 0} -x^4 = 0^-$$



On utilisera aussi la notation 0^- pour indiquer que l'on tend vers zéro par valeurs inférieures.

Exemple

$$\lim_{x \rightarrow 0^-} x^3 = 0^-$$

► Notation a^+ , $a \in \mathbb{R}$

a étant un réel, la notation a^+ signifie que l'on tend vers a par valeurs supérieures.

► Notation a^- , $a \in \mathbb{R}$

a étant un réel, la notation a^- signifie que l'on tend vers a par valeurs inférieures.

2. Limite infinie d'une fonction en un point

Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a un point de I .

On dit que f **admet pour limite « plus l'infini (on note $+\infty$) » en a** si, lorsque x devient très proche de a , $f(x)$ devient très grand, ce qui se traduit mathématiquement par le fait que pour tout réel A strictement positif, il existe un réel η strictement positif tel que :

$$\forall x \in I, 0 < |x - a| < \eta \Rightarrow f(x) > A$$

On écrit alors : $\lim_{x \rightarrow a} f(x) = +\infty$ ou $\lim_a f(x) = +\infty$.

On dit que f **admet pour limite « moins l'infini (on note $-\infty$) » en a** si, lorsque x devient très proche de a , $f(x)$ devient très grand en valeur absolue, mais en étant à valeurs négatives, ce qui se traduit mathématiquement par le fait que pour tout réel A strictement positif, il existe un réel η strictement positif tel que :

$$\forall x \in I, 0 < |x - a| < \eta \Rightarrow f(x) < -A$$

On écrit : $\lim_{x \rightarrow a} f(x) = -\infty$ ou $\lim_a f = -\infty$.

Exemple

$$\lim_{x \rightarrow 1^+} \frac{1}{\sqrt{x^2 - 1}} = +\infty$$

3. Limite finie à droite (ou par valeurs supérieures)

Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , a un point de I , et ℓ un réel.

On dit que f **admet pour limite (finie) ℓ à droite en a** (ou encore, par valeurs supérieures) si, lorsque x devient très proche de a , en restant plus grand que a , $f(x)$ devient lui aussi très proche de ℓ , ce qui se traduit mathématiquement par le fait que pour tout réel ε strictement positif, il existe un réel η strictement positif tel que :

$$\forall x \in I, 0 < x - a < \eta \Rightarrow |f(x) - \ell| < \varepsilon$$

On écrit : $\lim_{x \rightarrow a^+} f(x) = \ell$ ou $\lim_{a^+} f = \ell$.

Exemple

$$\lim_{x \rightarrow 1^+} (2 + \sqrt{x - 1}) = 2$$

4. Limite finie à gauche (ou par valeurs inférieures)

Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , a un point de I , et ℓ un réel.

On dit que f **admet pour limite (finie) ℓ à gauche en a** (ou encore, par valeurs inférieures) si, lorsque x devient très proche de a , en restant plus petit que a , $f(x)$ devient lui aussi très proche de ℓ , ce qui se traduit mathématiquement par le fait que pour tout réel ε strictement positif, il existe un réel η strictement positif tel que :

$$\forall x \in I, -\eta < x - a < 0 \Rightarrow |f(x) - \ell| < \varepsilon$$

On écrit : $\lim_{x \rightarrow a^-} f(x) = \ell$ ou $\lim_{a^-} f = \ell$.

5. Limite infinie à droite (ou par valeurs supérieures)

Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a un point de I .

On dit que f **admet pour limite $+\infty$ à droite en a** (ou encore, par valeurs supérieures) si, lorsque x devient très proche de a , en restant plus grand que a , $f(x)$ devient très grand, ce qui se traduit mathématiquement par le fait que pour tout réel A strictement positif, il existe un réel η strictement positif tel que :

$$\forall x \in I, 0 < x - a < \eta \Rightarrow f(x) > A$$

On écrit : $\lim_{x \rightarrow a^+} f(x) = +\infty$ ou $\lim_{a^+} f = +\infty$.

On dit que f **admet pour limite $-\infty$ à droite en a** (ou encore, par valeurs supérieures) si, lorsque x devient très proche de a , en restant plus grand que a , $f(x)$ devient très grand en valeur absolue, mais en étant à valeurs négatives, ce qui se traduit mathématiquement par le fait que pour tout réel A strictement positif, il existe un réel η strictement positif tel que :

$$\forall x \in I, 0 < x - a < \eta \Rightarrow f(x) < -A$$

On écrit : $\lim_{x \rightarrow a^+} f(x) = -\infty$ ou $\lim_{a^+} f = -\infty$.

6. Limite infinie à gauche (ou par valeurs inférieures)

Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a un point de I .

On dit que f **admet pour limite $+\infty$ à gauche en a** (ou encore, par valeurs inférieures) si, lorsque x devient très proche de a , en restant plus grand que a , $f(x)$ devient très grand, ce qui se traduit mathématiquement par le fait que pour tout réel A strictement positif, il existe un réel η strictement positif tel que :

$$\forall x \in I, -\eta < x - a < 0 \Rightarrow f(x) > A$$

On écrit : $\lim_{x \rightarrow a^-} f(x) = +\infty$ ou $\lim_{a^-} f = +\infty$.

On dit que f **admet pour limite $-\infty$ à gauche en a** (ou encore, par valeurs inférieures) si, lorsque x devient très proche de a , en restant plus grand que a , $f(x)$ devient très grand en valeur absolue, mais en étant à valeurs négatives, ce qui se traduit mathématiquement par le fait que pour tout réel A strictement positif, il existe un réel η strictement positif tel que :

$$\forall x \in I, \eta < x - a < 0 \Rightarrow f(x) < -A$$

On écrit : $\lim_{x \rightarrow a^-} f(x) = -\infty$ ou $\lim_{a^-} f = -\infty$.

Limite d'une fonction en $+\infty$ ou $-\infty$

1. Limite finie d'une fonction en l'infini

Soient f une fonction définie sur un intervalle de la forme $[a, +\infty[$ de \mathbb{R} , $a \in \mathbb{R}$, et ℓ un réel.

On dit que f **admet pour limite (finie) ℓ en « plus l'infini (on note $+\infty$) »** si, lorsque x devient très grand, $f(x)$ devient très proche de ℓ , ce qui se traduit mathématiquement par le fait que pour tout réel ε strictement positif, il existe un réel « seuil », A , strictement positif tel que :

$$\forall x \in [a, +\infty[, x > A \Rightarrow |f(x) - \ell| < \varepsilon$$

On écrit alors : $\lim_{x \rightarrow +\infty} f(x) = \ell$ ou $\lim_{+\infty} f = \ell$.

Si f est définie sur un intervalle de la forme $] -\infty, a]$ de \mathbb{R} , $a \in \mathbb{R}$, et si ℓ désigne encore un réel, on dit que f **admet pour limite (finie) ℓ en « moins l'infini (on note $-\infty$) »** si, lorsque x devient très grand en valeur absolue, mais en étant à valeurs négatives, $f(x)$ devient très proche de ℓ , ce qui se traduit mathématiquement par le fait que pour tout réel ε strictement positif, il existe un réel A , strictement positif tel que :

$$\forall x \in] -\infty, a], x < -A \Rightarrow |f(x) - \ell| < \varepsilon$$

On écrit alors : $\lim_{x \rightarrow -\infty} f(x) = \ell$ ou $\lim_{-\infty} f = \ell$.

Exemple

$$\lim_{x \rightarrow +\infty} \left(1 - \frac{1}{\sqrt{x^2 - 1}} \right) = 1$$

2. Limite infinie d'une fonction en plus l'infini

Soit f une fonction définie sur un intervalle de la forme $[a, +\infty[$ de \mathbb{R} , $a \in \mathbb{R}$.

On dit que f **admet pour limite $+\infty$ en « plus l'infini »** si, lorsque x devient très grand, $f(x)$ devient lui aussi très grand, ce qui se traduit mathématiquement par le fait que pour tout réel B strictement positif, il existe un réel « seuil », A , strictement positif tel que :

$$\forall x \in [a, +\infty[, x > A \Rightarrow f(x) > B$$

On écrit alors : $\lim_{x \rightarrow +\infty} f(x) = +\infty$ ou $\lim_{+\infty} f = +\infty$.

On dit que f **admet pour limite $-\infty$ en « plus l'infini »** si, lorsque x devient très grand, $f(x)$ devient très grand en valeur absolue, mais en étant à valeurs négatives, ce qui se traduit mathématiquement par le fait que pour tout réel B strictement positif, il existe un réel « seuil », A , strictement positif tel que :

$$\forall x \in [a, +\infty[, x > A \Rightarrow f(x) < -B$$

On écrit alors : $\lim_{x \rightarrow +\infty} f(x) = -\infty$ ou $\lim_{+\infty} f = -\infty$.

3. Limite infinie d'une fonction en moins l'infini

Soit f une fonction définie sur un intervalle de la forme $] -\infty, a]$ de \mathbb{R} , $a \in \mathbb{R}$.

On dit que f **admet pour limite $+\infty$ en « moins l'infini »** si, lorsque x devient très grand en valeur absolue, mais en étant à valeurs négatives, $f(x)$ devient lui aussi très grand, ce qui se traduit mathématiquement par le fait que pour tout réel B strictement positif, il existe un réel A , strictement positif tel que :

$$\forall x \in] -\infty, a], x < -A \Rightarrow f(x) > B$$

On écrit alors : $\lim_{x \rightarrow -\infty} f(x) = +\infty$ ou $\lim_{-\infty} f = +\infty$.

On dit que f **admet pour limite $-\infty$ en « moins l'infini »** si, lorsque x devient très grand en valeur absolue, en étant négatif, $f(x)$ devient aussi très grand en valeur absolue, en étant négatif, ce qui se traduit mathématiquement par le fait que pour tout réel B strictement positif, il existe un réel A , strictement positif tel que :

$$\forall x \in] -\infty, a], x < -A \Rightarrow f(x) < -B$$

On écrit alors : $\lim_{x \rightarrow -\infty} f(x) = -\infty$ ou $\lim_{-\infty} f = -\infty$.

4. Forme indéterminée

On appelle **forme indéterminée** une limite que l'on ne sait pas déterminer ; cela correspond donc à des quantités que l'on ne peut pas quantifier **de façon exacte**, comme, par exemple, le quotient de $+\infty$ avec $+\infty$.

Propriétés des limites

Opérations sur les limites

1. Propriétés des limites

► Unicité de la limite

Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} . Si f possède une limite en a , celle-ci est unique.

- Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , a un point de I , et ℓ dans \mathbb{R} .

Alors, si f est définie dans un voisinage à gauche de a , et dans un voisinage à droite de a :

$$\lim_{x \rightarrow a} f(x) = \ell \Leftrightarrow \lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^-} f(x) = \ell$$

- Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} ; m et M sont deux réels. Alors :

- si $\lim_{x \rightarrow a} f(x) < M$, il existe un voisinage de a tel que, pour tout x de ce voisinage :

$$f(x) < M$$

- si $\lim_{x \rightarrow a} f(x) > m$, il existe un voisinage de a tel que, pour tout x de ce voisinage :

$$f(x) > m$$

► Limites et comparaison

Soient f et g deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} ; m et M sont deux réels. Alors, **si f et g ont des limites finies en a** , et s'il existe un voisinage \mathcal{V} de a tel que, pour tout x de ce voisinage,

$$f(x) \leq g(x)$$

on a : $\lim_{x \rightarrow a} f(x) \leq \lim_{x \rightarrow a} g(x)$

► Limites et minoration

Soient f et g deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} . S'il existe un voisinage de a tel que, pour tout x de ce voisinage,

$$f(x) \leq g(x)$$

et si, de plus, $\lim_{x \rightarrow a} g(x) = -\infty$

alors : $\lim_{x \rightarrow a} f(x) = -\infty$

► Limites et majoration

Soient f et g deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} . S'il existe un voisinage de a tel que, pour tout x de ce voisinage, $f(x) \geq g(x)$, et si $\lim_{x \rightarrow a} g(x) = +\infty$, alors :

$$\lim_{x \rightarrow a} f(x) = +\infty$$

► Théorème des gendarmes

Soient f et g et h trois fonction définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} ; ℓ est un réel. S'il existe un voisinage de a tel que, pour tout x de ce voisinage, $f(x) \leq h(x) \leq g(x)$, et si, de plus, $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = \ell$, alors : $\lim_{x \rightarrow a} h(x) = \ell$

2. Opérations sur les limites

► Limite d'une somme de fonctions

Soient f et g deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} ; ℓ et ℓ' sont deux réels finis. Alors :

$\lim_{x \rightarrow a} f(x)$	$\lim_{x \rightarrow a} g(x)$	$\lim_{x \rightarrow a} (f(x) + g(x))$
ℓ	ℓ'	$\ell + \ell'$
ℓ	$+\infty$	$+\infty$
ℓ	$-\infty$	$-\infty$
$+\infty$	$+\infty$	$+\infty$
$-\infty$	$-\infty$	$-\infty$
$+\infty$	$-\infty$	Forme indéterminée

► Limite d'un produit de fonctions

Soient f et g deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} ; ℓ et ℓ' sont deux réels. Alors :

$\lim_{x \rightarrow a} f(x)$	$\lim_{x \rightarrow a} g(x)$	$\lim_{x \rightarrow a} f(x) g(x)$
ℓ	ℓ'	$\ell \ell'$
ℓ , avec $\ell > 0$	$+\infty$	$+\infty$
ℓ , avec $\ell > 0$	$-\infty$	$-\infty$
ℓ , avec $\ell < 0$	$+\infty$	$-\infty$
ℓ , avec $\ell < 0$	$-\infty$	$+\infty$
0	$+\infty$	Forme indéterminée
0	$-\infty$	Forme indéterminée

► Limite d'un quotient de fonctions

Soient f et g deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} ; ℓ et ℓ' sont deux réels. Alors :

$\lim_{x \rightarrow a} f(x)$	$\lim_{x \rightarrow a} g(x)$	$\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$
ℓ	ℓ' , avec $\ell' \neq 0$	$\frac{\ell}{\ell'}$
ℓ	$+\infty$	0
ℓ	$-\infty$	0
ℓ , avec $\ell > 0$	0^+	$+\infty$
ℓ , avec $\ell > 0$	0^-	$-\infty$
ℓ , avec $\ell < 0$	0^+	$-\infty$
ℓ , avec $\ell < 0$	0^-	$+\infty$
$+\infty$	$+\infty$	Forme indéterminée
$+\infty$	$-\infty$	Forme indéterminée
$-\infty$	$+\infty$	Forme indéterminée
$-\infty$	$-\infty$	Forme indéterminée

Notations de Landau

1. Négligeabilité

Définition

Soient f et g deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} .

On suppose que g ne s'annule pas dans un voisinage de a . On dit que f est **négligeable** devant g au voisinage de a si

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0$$

On note alors

$$f(x) \underset{x \rightarrow a}{=} o(g(x)) \quad \text{ou} \quad f \underset{a}{=} o(g)$$

On dit que f est un « petit o » de g au voisinage de a .



La notation « petit o », de même que la notation « grand O », qui sera vue plus loin, est appelée **notation de Landau**, en hommage au mathématicien Edmund Landau¹. Leur paternité est visiblement assez controversée, et reviendrait, a priori, à Paul Bachmann².

Exemple

On considère les fonctions f et g définies, pour tout réel x , par

$$f(x) = x^2, \quad g(x) = x^4$$

Alors, comme $\lim_{x \rightarrow +\infty} \frac{f(x)}{g(x)} = \lim_{x \rightarrow +\infty} \frac{1}{x^2} = 0$, on en déduit : $f \underset{+\infty}{=} o(g)$.



Pour traduire le fait qu'une fonction f possède une limite nulle en a , $a \in \mathbb{R}$, ou, éventuellement, $a = +\infty$ ou $a = -\infty$, on écrit aussi :

$$f(x) \underset{x \rightarrow a}{=} o(1)$$

2. Domination

Définition

Soient f et g deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} . On suppose que g ne s'annule pas dans un voisinage de a , sans, pour autant, que $g(a)$ soit non nul.

1. Edmund Georg Hermann Landau (1877-1938), mathématicien allemand, spécialiste de théorie des nombres.

2. Paul Bachmann (1837-1920), mathématicien allemand lui aussi, et également spécialiste de théorie des nombres.

On dit que f est **dominée** par g au voisinage de a si il existe une constante positive C telle que, pour tout réel x dans un voisinage de a

$$|f(x)| \leq C |g(x)|$$

On note alors :

$$f(x) = O(g(x)) \quad \text{ou} \quad f = O(g)$$

On dit que f est un « grand O » de g au voisinage de a .

Exemple

On considère les fonctions f et g définies, pour tout réel strictement positif x , par :

$$f(x) = \frac{2 - \frac{1}{x}}{2^x}, \quad g(x) = \frac{3}{2^x}$$

Alors, comme, pour tout réel $x > 1$:

$$|f(x)| = \left| \frac{2 - \frac{1}{x}}{2^x} \right| \leq \left| \frac{2}{2^x} \right| \leq \left| \frac{3}{2^x} \right| = |g(x)|$$

on a bien : $f(x) = O(g(x))$.

3. Équivalence

Définition

Soient f et g deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} . On suppose que g ne s'annule pas dans un voisinage de a , sans, pour autant, que $g(a)$ soit non nul.

On dit que f est **équivalente** à g au voisinage de a si :

$$f(x) - g(x) = o(g(x))$$

On note alors :

$$f(x) \sim_{x \rightarrow a} g(x) \quad \text{ou} \quad f \sim_a g$$

Exemple

On considère les fonctions f et g définies, pour tout réel x , par

$$f(x) = x^4 + x^3, \quad g(x) = x^4 + x^2$$

Alors, comme $\lim_{x \rightarrow +\infty} \frac{f(x)}{g(x)} = 1$, on en déduit : $f \sim_{+\infty} g$.

Attention aux manipulations successives et hasardeuses d'équivalents !

Pour cette raison, on ne donnera pas, dans ce cours, de résultats généraux ni de « recettes » pour la manipulation d'équivalents, la meilleure méthode, la plus fiable et la plus sûre, étant de manipuler des « o » ou des « O », suivant les cas et ce qui est le mieux adapté. Mais attention, on ne peut pas utiliser ceux-ci dans des inégalités !



Domaine de définition d'une fonction, graphe

1. Domaine de définition d'une fonction

On s'intéresse ici aux fonctions d'une variable réelle, à valeurs réelles, c'est-à-dire appartenant à \mathbb{R} . Dans ce cadre, une « fonction » est un « procédé » permettant d'associer à un nombre réel x un autre nombre réel, noté $f(x)$.

Comme nous l'avons vu au début de cet ouvrage, il n'est pas nécessaire que ce « procédé » donne un résultat pour tous les nombres réels, mais seulement pour certains d'entre eux. L'ensemble \mathcal{D}_f des nombres réels pour lesquels ce « procédé » donne effectivement un résultat est appelé **domaine de définition de la fonction f** .

Ceci peut être résumé en disant que la fonction f est une application de \mathcal{D}_f dans \mathbb{R} .

Notation

On écrit :

$$\begin{aligned} f : \mathcal{D}_f &\rightarrow \mathbb{R} \\ x &\mapsto f(x) \end{aligned}$$

La première flèche, « \rightarrow », signifie « dans » : f va de \mathcal{D}_f dans \mathbb{R} . \mathcal{D}_f est l'**ensemble de départ** de f , \mathbb{R} , l'**ensemble d'arrivée** de f .

La seconde flèche, « \mapsto », signifie « a pour image » : x a pour image $f(x)$.



Il est essentiel, quand on étudie une fonction, de bien préciser son domaine de définition.

2. Graphe d'une fonction

Le **graphe**, ou **courbe représentative**, C_f d'une fonction $f : \mathcal{D}_f \rightarrow \mathbb{R}$ est l'ensemble des points (x, y) du plan \mathbb{R}^2 tels que x appartienne à \mathcal{D}_f et $y = f(x)$:

$$C_f = \{(x, f(x)); x \in \mathcal{D}_f\}$$

Exemple

Le graphe de la fonction définie sur \mathbb{R} par $x \mapsto x^3$ est :

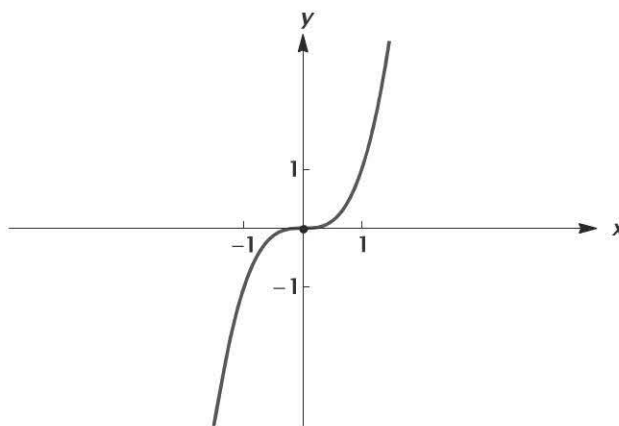


Figure 7.1 – Le graphe de la fonction définie sur \mathbb{R} par $x \mapsto x^3$.



1. Le graphe permet d'associer un aspect géométrique à l'étude d'une fonction.
2. Il convient dès maintenant de bien distinguer les objets suivants : la fonction f (qui est donc une application), le nombre réel $f(x)$, qui désigne la valeur de f en x , et le graphe C_f (qui est une partie du plan \mathbb{R}^2).

► Droite asymptote à une courbe

Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a dans \bar{I} .

Étant donnés deux réels m et p , la droite \mathcal{D} , d'équation $y = mx + p$, $(m, p) \in \mathbb{R}^2$, est dite **asymptote** à la courbe représentative C_f de f lorsque x tend vers a si :

$$\lim_{x \rightarrow a} (f(x) - mx - p) = 0$$

La droite \mathcal{D} , d'équation $x = a$, est dite **asymptote verticale** à la courbe représentative C_f de f lorsque x tend vers a si :

$$\lim_{x \rightarrow a, x < a} f(x) = +\infty \quad \text{ou} \quad \lim_{x \rightarrow a, x > a} f(x) = +\infty$$

ou

$$\lim_{x \rightarrow a, x < a} f(x) = -\infty \quad \text{ou} \quad \lim_{x \rightarrow a, x > a} f(x) = -\infty$$

Exemple

La droite d'équation $y = 2x + 1$ est asymptote à la courbe représentative de la fonction définie sur \mathbb{R} par $x \mapsto 2x + 1 - \frac{1}{x^2 + 1}$ lorsque x tend vers $+\infty$ et x tend vers $-\infty$:

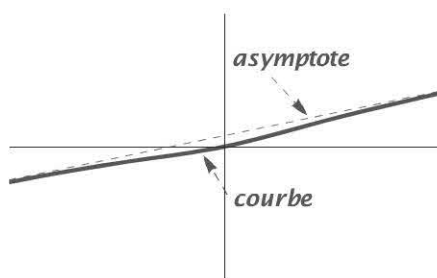


Figure 7.2 – La droite d'équation $y = 2x + 1$ et la courbe représentative de la fonction définie sur \mathbb{R} par $x \mapsto 2x + 1 - \frac{1}{x^2 + 1}$.

► Branche parabolique d'axe vertical

Soit f une fonction définie sur un intervalle I de \mathbb{R} contenant un voisinage de $-\infty$ ou $+\infty$, à valeurs dans \mathbb{R} .

On dit que la courbe représentative C_f de f possède une **branche parabolique d'axe vertical** lorsque x tend vers $+\infty$ si :

$$\lim_{x \rightarrow +\infty} f(x) = +\infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} \frac{f(x)}{x} = +\infty$$

ou

$$\lim_{x \rightarrow +\infty} f(x) = -\infty \quad \text{et} \quad \lim_{x \rightarrow +\infty} \frac{f(x)}{x} = -\infty$$

On dit que la courbe représentative C_f de f possède une **branche parabolique d'axe vertical** lorsque x tend vers $-\infty$ si :

$$\lim_{x \rightarrow -\infty} f(x) = +\infty \quad \text{et} \quad \lim_{x \rightarrow -\infty} \frac{f(x)}{x} = -\infty$$

ou

$$\lim_{x \rightarrow -\infty} f(x) = -\infty \quad \text{et} \quad \lim_{x \rightarrow -\infty} \frac{f(x)}{x} = +\infty$$

Exemple

La fonction définie sur \mathbb{R} par $x \mapsto x^4$ possède une branche parabolique d'axe vertical en $+\infty$ et en $-\infty$:

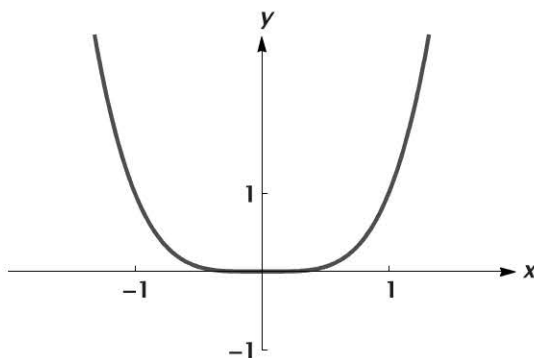


Figure 7.3 – La courbe représentative de la fonction $x \mapsto x^4$.

► Branche parabolique d'axe horizontal

Soit f une fonction définie sur un intervalle I de \mathbb{R} contenant un voisinage de $-\infty$ ou $+\infty$, à valeurs dans \mathbb{R} .

On dit que la courbe représentative C_f de f possède une **branche parabolique d'axe horizontal** lorsque x tend vers $+\infty$ si : $\lim_{x \rightarrow +\infty} |f(x)| = +\infty$ et $\lim_{x \rightarrow +\infty} \frac{f(x)}{x} = 0$

On dit que la courbe représentative C_f de f possède une **branche parabolique d'axe horizontal** lorsque x tend vers $-\infty$ si : $\lim_{x \rightarrow -\infty} |f(x)| = \infty$ et $\lim_{x \rightarrow -\infty} \frac{f(x)}{x} = 0$

► Branche parabolique d'axe d'équation $y = mx$, $m \in \mathbb{R}$

Soit f une fonction définie sur un intervalle I de \mathbb{R} contenant un voisinage de $-\infty$ ou $+\infty$, à valeurs dans \mathbb{R} .

On dit que la courbe représentative C_f de f possède une **branche parabolique d'axe** $y = mx$, $m \in \mathbb{R}$, lorsque x tend vers $+\infty$ si $\lim_{x \rightarrow +\infty} |f(x)| = +\infty$ et si :

$$\lim_{x \rightarrow +\infty} \frac{f(x)}{x} = m \quad \text{et} \quad \lim_{x \rightarrow +\infty} (f(x) - mx) = +\infty \quad \text{ou} \quad \lim_{x \rightarrow +\infty} (f(x) - mx) = -\infty$$

On dit que la courbe représentative C_f de f possède une **branche parabolique d'axe** $y = mx$, $m \in \mathbb{R}$, lorsque x tend vers $-\infty$ si $\lim_{x \rightarrow -\infty} |f(x)| = +\infty$ et si :

$$\lim_{x \rightarrow -\infty} \frac{f(x)}{x} = m \quad \text{et} \quad \lim_{x \rightarrow -\infty} (f(x) - mx) = +\infty \quad \text{ou} \quad \lim_{x \rightarrow -\infty} (f(x) - mx) = -\infty$$

On parle souvent, dans la littérature mathématique, de la « construction de l'ensemble des réels ». Qu'en est-il ?

À l'origine, seuls les nombres entiers et rationnels étaient connus des mathématiciens, même si des irrationnels comme $\sqrt{2}$, longueur de la diagonale d'un carré de côté 1, sont vite apparus comme des nombres « à part », difficilement quantifiables.

La coupure par un rationnel

La démonstration formelle de l'existence – ou construction – d'un ensemble de nombres contenant à la fois les entiers, les rationnels, et les non-rationnels, fut mise en œuvre pour la première fois au XIX^e siècle par le mathématicien allemand Richard Dedekind (1831-1916).

Elle est basée sur l'axiome de la borne supérieure, selon lequel toute partie non vide et majorée de l'ensemble \mathbb{R} des réels possède une borne supérieure.

Richard Dedekind est, tout simplement, parti du fait que tout nombre rationnel $\frac{p}{q}$, $(p, q) \in \mathbb{Z} \times \mathbb{Z}^*$, découpe \mathbb{Q} , ensemble des nombres rationnels, en deux parties, constituées respectivement par les rationnels strictement plus petits que $\frac{p}{q}$, et par les rationnels supérieurs ou égaux à $\frac{p}{q}$.

La coupure par un irrationnel

En étendant ce principe à un découpage par un « irrationnel » comme $\sqrt{2}$, on découpe, de façon analogue, l'ensemble des rationnels en deux parties, constituées respectivement par les rationnels négatifs, ou dont le carré est strictement plus petit que 2, et par les rationnels positifs dont le carré est supérieur ou égal à 2.



Figure 7.4 – Une « coupure ».

Ainsi, $\sqrt{2}$ apparaît comme la « coupure » entre ces deux ensembles, c'est-à-dire le nombre « non rationnel » qui se trouve « entre les deux ».

Une autre construction assez populaire de l'ensemble des nombres réels peut être obtenue par l'intermédiaire de suites de Cauchy.

Comment définir une fonction ?

Dès que l'on connaît quelques fonctions, on peut en construire de (nombreuses) autres en utilisant les procédés suivants :

1. Les opérations algébriques

Si f et g sont deux fonctions définies sur le même intervalle I ,

- la fonction somme $f + g$ est définie, pour tout réel x de l'intervalle I , par :

$$(f + g)(x) = f(x) + g(x)$$

- la fonction produit $f g$ est définie, pour tout réel x de l'intervalle I , par :

$$(f g)(x) = f(x) g(x)$$

- Lorsque la fonction g ne s'annule pas sur l'intervalle I , la fonction quotient $\frac{f}{g}$ est définie, pour tout réel x de l'intervalle I , par :

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$$

2. La composition

Soit f une fonction définie sur un intervalle I de \mathbb{R} , et g une fonction définie un intervalle $J \subset \mathbb{R}$ contenant $f(I)$. La fonction « **composée** » des fonctions f et g est la fonction, que l'on écrit $g \circ f$, définie, pour tout réel x de l'intervalle I , par :

$$(g \circ f)(x) = g(f(x))$$

Exemple

La fonction obtenue en composant la fonction f qui, à tout réel x , associe x^2 avec la fonction g qui, à tout réel x , associe $x + 1$ est définie, pour tout réel x , par : $(g \circ f)(x) = x^2 + 1$.

3. La restriction

Soit f une fonction définie sur l'intervalle I de \mathbb{R} , et $I_0 \subset \mathbb{R}$ un intervalle contenu dans I . On appelle **restriction de f à I_0** , que l'on note $f|_{I_0}$, la fonction définie sur I_0 par :

$$\forall x \in I_0 : f|_{I_0}(x) = f(x)$$

Cela signifie que les fonctions f et $f|_{I_0}$ prennent la même valeur en chaque point de l'intervalle J , mais la fonction $f|_{I_0}$ n'est définie que sur cet intervalle alors que la fonction f est aussi définie aux points de I qui ne sont pas dans I_0 .

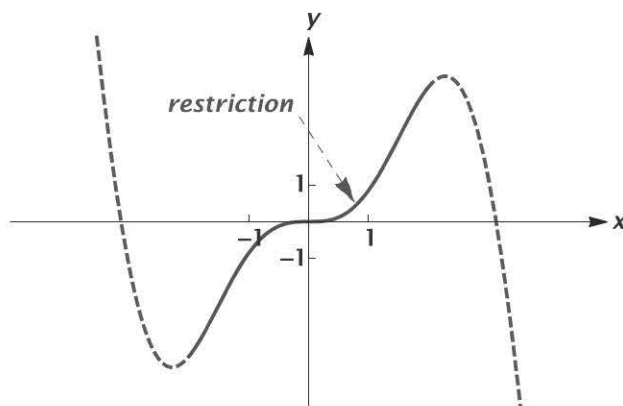


Figure 8.1 – Le graphe d’une fonction et de sa restriction à $[-2, 2]$.

4. Le recollement : les fonctions définies par morceaux

Considérons un intervalle I de \mathbb{R} , divisé en sous-intervalles disjoints I_1, I_2, \dots, I_n , où n est un entier naturel, et envisageons le cas d’une fonction f ayant, sur chacun de ces sous-intervalles, une expression différente :

$$f|_{I_1} = f_1, \quad f|_{I_2} = f_2, \quad \dots, \quad f|_{I_n} = f_n$$

La fonction ainsi obtenue par « recollement », est une fonction « définie par morceaux ».

Exemple

La fonction f définie par :

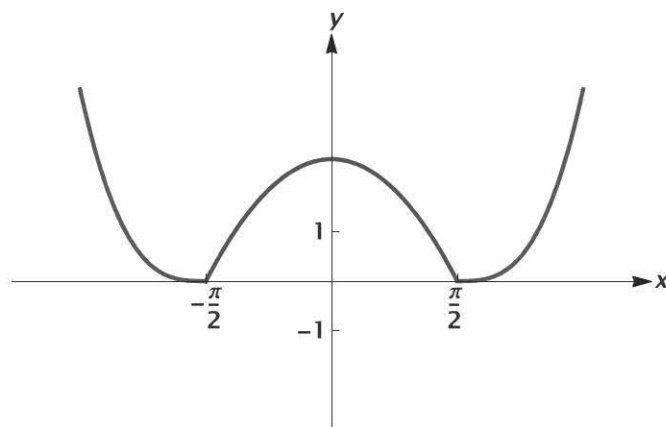


Figure 8.2 – Le graphe d’une fonction définie « par morceaux ».

$$f(x) = \begin{cases} \frac{\pi^2}{4} - x^2 & \text{si } -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ \left(x - \frac{\pi}{2}\right)^2 & \text{si } x > \frac{\pi}{2} \\ \left(x + \frac{\pi}{2}\right)^2 & \text{si } x < -\frac{\pi}{2} \end{cases}$$

est une fonction définie « par morceaux ».

Majorations et minoration

Définition

Soit f une fonction définie sur un intervalle I . Etant donné un réel M , la fonction f est dite **majorée** par M sur I si, pour tout réel $x \in I$:

$$f(x) \leq M$$

Définition

Soit f une fonction définie sur un intervalle I . Etant donné un réel m , la fonction f est dite **minorée** par m sur I si, pour tout réel $x \in I$:

$$f(x) \geq m$$

Définition

Soit f une fonction définie sur un intervalle I . La fonction f est dite **bornée** sur I si elle y est à la fois majorée et minorée.



Cette condition est vérifiée si et seulement si il existe un nombre réel M tel que $|f(x)| \leq M$ pour tout nombre réel x de I .

Exemples

1. La fonction qui, à tout réel x , associe $-1 + \frac{2x^2}{x^2 + 1}$ est bornée sur \mathbb{R} , car majorée par 1 et minorée par -1.

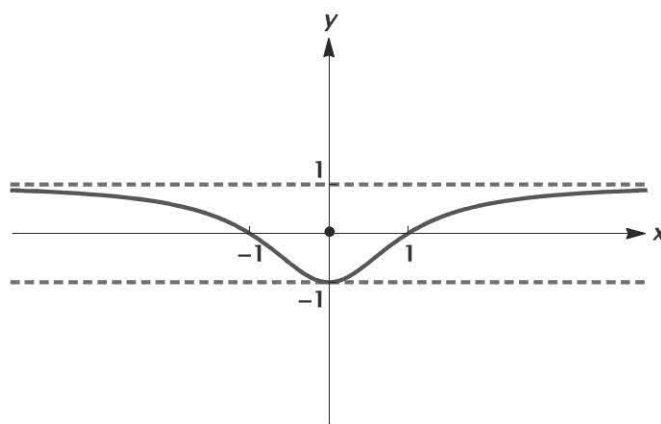


Figure 9.1 – La courbe représentative de la fonction $x \in \mathbb{R} \mapsto -1 + \frac{2x^2}{x^2 + 1}$.

2. La fonction $x \mapsto x^2$ est minorée par 0 mais non majorée sur \mathbb{R} .
3. La fonction $x \mapsto \frac{1}{x} + x$ n'est ni majorée ni minorée sur $]0, \infty[$.

Définition

Soient f et g deux fonctions définies sur un même intervalle I de \mathbb{R} . On dit que f **major**e g , si, pour tout x de I :

$$f(x) \geq g(x)$$

On écrit alors $f \geq g$.

Exemple

Sur l'intervalle $[0, +\infty[$ la fonction $x \mapsto -1 + \frac{2x^2}{x^2+1}$, est majorée par la fonction $x \mapsto x$.

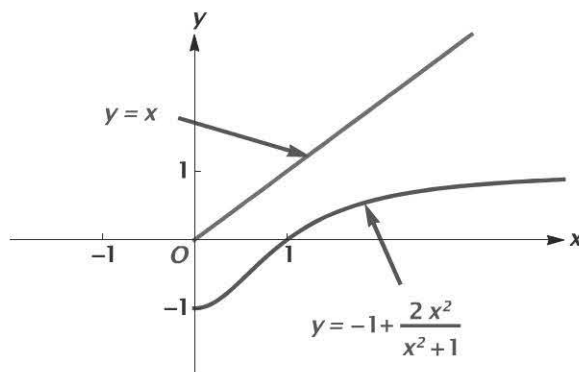


Figure 9.2 – La courbe représentative de la fonction $x \mapsto -1 + \frac{2x^2}{x^2+1}$ majorée sur \mathbb{R}^+ par la fonction $x \mapsto x$.

Définition

Soient f et g deux fonctions définies sur un même intervalle I de \mathbb{R} . On dit que f **min**ore g , si, pour tout x de I :

$$f(x) \leq g(x)$$

On écrit alors $f \leq g$.

1. Définitions

► Croissance

Soit f une fonction définie sur un intervalle $I \subset \mathbb{R}$. Elle est dite **croissante** sur I si :

$$\forall x_1 \in I, \forall x_2 \in I : x_1 \leq x_2 \implies f(x_1) \leq f(x_2)$$

► Décroissance

Soit f une fonction définie sur un intervalle $I \subset \mathbb{R}$. Elle est dite **décroissante** sur I si :

$$\forall x_1 \in I, \forall x_2 \in I : x_1 \leq x_2 \implies f(x_1) \geq f(x_2)$$

► Croissance (au sens strict)

Soit f une fonction définie sur un intervalle $I \subset \mathbb{R}$. Elle est dite **strictement croissante** sur I si :

$$\forall x_1 \in I, \forall x_2 \in I : x_1 < x_2 \implies f(x_1) < f(x_2)$$

► Décroissance (au sens strict)

Soit f une fonction définie sur un intervalle $I \subset \mathbb{R}$. Elle est dite **strictement décroissante** sur I si :

$$\forall x_1 \in I, \forall x_2 \in I : x_1 < x_2 \implies f(x_1) > f(x_2)$$

► Monotonie

Soit f une fonction définie sur un intervalle $I \subset \mathbb{R}$. Elle est dite **monotone** sur I si elle y est croissante ou décroissante.

► Monotonie (au sens strict)

Soit f une fonction définie sur un intervalle $I \subset \mathbb{R}$. Elle est dite **strictement monotone** sur I si elle y est strictement croissante ou strictement décroissante.



Étudier les variations d'une fonction consiste donc à partager son ensemble de définition en intervalles tels que, sur chacun d'eux, la fonction soit monotone.

Exemples

1. La fonction $x \mapsto x + 1$ est croissante sur \mathbb{R} .
2. La fonction $x \mapsto x^2$ est croissante sur $[0, +\infty[$.
3. Les fonctions puissances, de la forme $x \in \mathbb{R} \mapsto x^n, n \in \mathbb{N}^*$, sont :
 - croissantes sur \mathbb{R} si n est impair ;
 - décroissantes sur $] - \infty, 0]$ et croissantes sur $[0, +\infty[$ si n est pair.

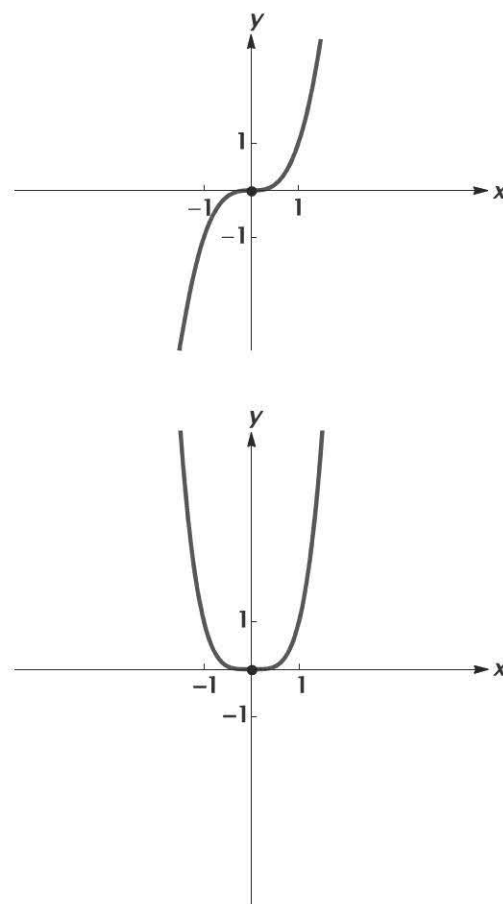


Figure 10.1 – Les courbes représentatives des fonctions $x \mapsto x^3$ et $x \mapsto x^4$.

2. Tableau de variations

Pour rassembler les informations concernant les variations d'une fonction, le plus simple est d'utiliser un **tableau de variations** ; la croissance est représentée par une flèche vers le haut, la décroissance, par une flèche vers le bas. On y indique aussi les valeurs aux bornes (du domaine de définition), qui peuvent n'être que des limites.

Exemple

Le tableau de variations de la fonction définie sur \mathbb{R} par $x \mapsto x^2$ est :

x	$-\infty$	0	$+\infty$
$x \mapsto x^2$	$+\infty$		$+\infty$
		↓	↑
		0	

1. Parité d'une fonction

Définition

Soit f une fonction définie sur un domaine \mathcal{D}_f de \mathbb{R} tel que \mathcal{D}_f soit symétrique, i.e., pour tout x de \mathcal{D}_f :

$$x \in \mathcal{D}_f \Rightarrow -x \in \mathcal{D}_f$$

La fonction f est dite **paire** si, pour tout réel x de son domaine de définition \mathcal{D}_f :

$$f(-x) = f(x)$$



Si la fonction f est paire, sa courbe représentative est symétrique par rapport à l'axe des ordonnées (Oy).

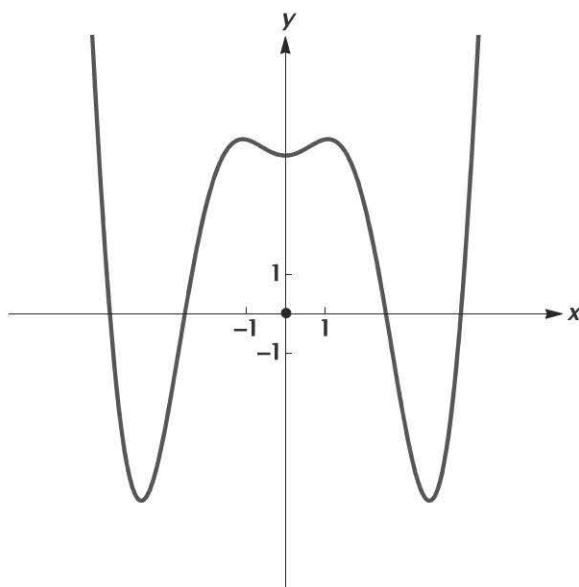


Figure 11.1 – Le graphe d'une fonction paire.

2. Imparité d'une fonction

Définition

Soit f une fonction définie sur un domaine \mathcal{D}_f de \mathbb{R} tel que \mathcal{D}_f soit symétrique, c'est-à-dire, pour tout x de \mathcal{D}_f :

$$x \in \mathcal{D}_f \Rightarrow -x \in \mathcal{D}_f$$

la fonction f est dite **impaire** si, pour tout réel x de son domaine de définition \mathcal{D}_f :

$$f(-x) = -f(x)$$



1. Toute fonction f impaire s'annule en 0 : $f(0) = 0$.
2. Si la fonction f est impaire, sa courbe représentative est symétrique par rapport à l'origine O .

Exemple

La fonction définie sur \mathbb{R} par $x \mapsto x^3$ est impaire.

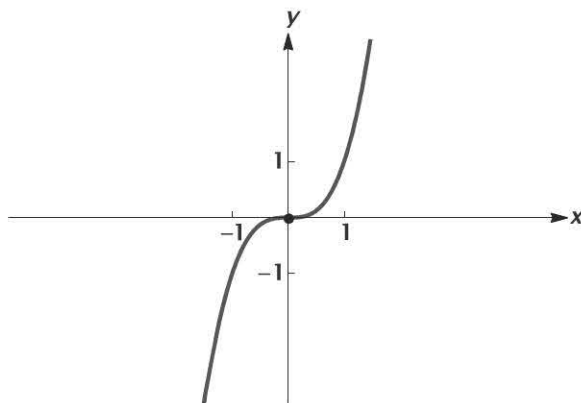


Figure 11.2 – Le graphe d'une fonction impaire.



Les propriétés de parité permettent donc de réduire l'étude de la fonction à l'intervalle $\mathcal{D}_f \cap [0, \infty[$; on trace alors la partie du graphe correspondante, puis on complète par la symétrie convenable.

1. Centre de symétrie de la courbe représentative d'une fonction

Soient f une fonction définie sur un domaine \mathcal{D}_f de \mathbb{R} , et a et b deux réels tels que, pour tout x de \mathcal{D}_f :

$$a + x \in \mathcal{D}_f \quad \text{et} \quad a - x \in \mathcal{D}_f$$

La courbe représentative C_f de f admet le point de coordonnées (a, b) comme centre de symétrie si et seulement si, pour tout réel x de \mathcal{D}_f :

$$f(a + x) + f(a - x) = 2b$$

Démonstration : Considérons un point M d'abscisse $a + x$ appartenant à la courbe représentative C_f de f ; son ordonnée est donc $f(a + x)$.

Le point Ω , de coordonnées (a, b) , est centre de symétrie de C_f si et seulement si le point M' tel que $\overrightarrow{\Omega M'} = \overrightarrow{M\Omega}$ est aussi sur la courbe. Les coordonnées (x', y') sont telles que :

$$x' - a = a - (x + a) = -x \quad , \quad y' - b = b - y = b - f(a + x)$$

Le point M' appartient à la courbe C_f si et seulement si $y' = f(x')$.

La condition précédente devient : $f(a - x) + f(a + x) = 2b$. ■



Le cas $a = 0$ est celui où la fonction est impaire.

Exemple

La courbe représentative de la fonction définie sur \mathbb{R} par $x \mapsto 2 + (x - 2)^3$ admet le point de coordonnées $(2, 2)$ comme centre de symétrie :

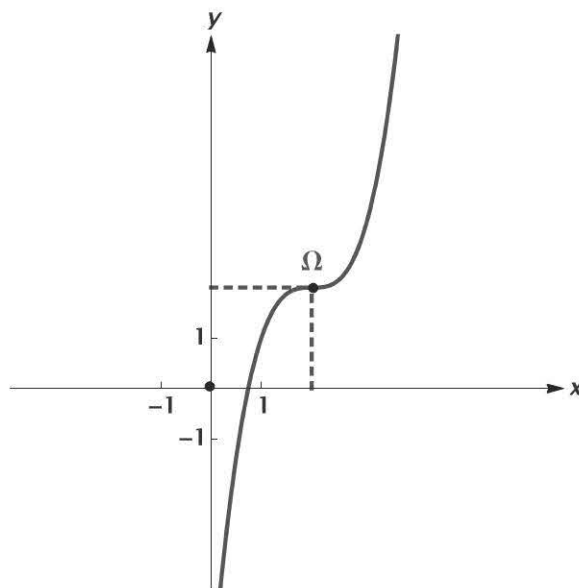


Figure 12.1 – La courbe représentative de la fonction $x \mapsto 2 + (x - 2)^3$.

2. Axe de symétrie vertical de la courbe représentative d'une fonction

Soient f une fonction définie sur un domaine \mathcal{D}_f de \mathbb{R} , et a un réel.

La courbe représentative C_f de f admet la droite d'équation $x = a$ pour axe de symétrie de symétrie si et seulement si, pour tout réel x :

$$f(a + x) = f(a - x)$$

Démonstration : Considérons un point M d'abscisse $a + x$ appartenant à la courbe représentative C_f de f ; son ordonnée est donc $f(a + x)$. Le point M' , symétrique de M par rapport à la droite d'équation $x = a$, a pour coordonnées $(a - x, f(a + x))$. Il appartient à la courbe C_f si et seulement si : $f(a + x) = f(a - x)$. ■

Exemple

La courbe représentative de la fonction définie sur \mathbb{R} par $x \mapsto 2 + (x - 3)^2$ admet la droite d'équation $x = 3$ comme axe de symétrie :

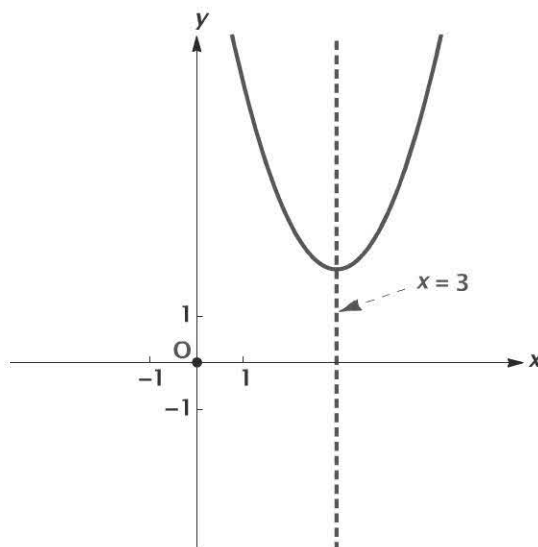


Figure 12.2 – La courbe représentative de la fonction $x \mapsto 2 + (x - 3)^2$.

1. Période

Soit f une fonction définie sur un domaine \mathcal{D}_f de \mathbb{R} , et T un nombre réel non nul tel que, pour tout réel x de \mathcal{D}_f :

$$x + T \in \mathcal{D}_f$$

La fonction f est dite **T -périodique** si, pour tout réel x de son domaine de définition \mathcal{D}_f :

$$f(x + T) = f(x)$$

T est une **période** de f .

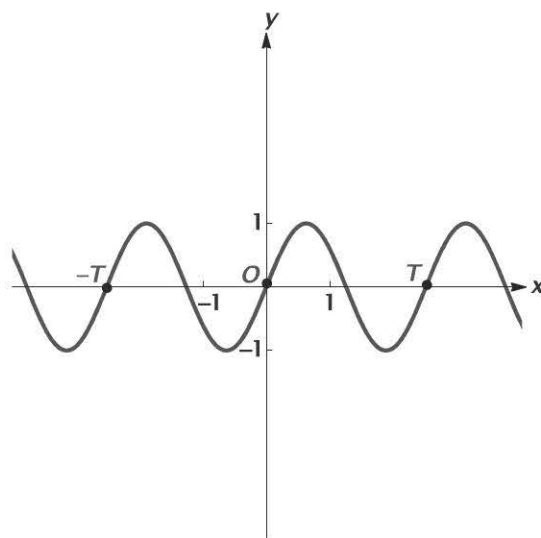


Figure 13.1 – Le graphe d'une fonction périodique.



Si f est une fonction périodique et si T et T' sont des périodes de f telles que

$$T + T' \neq 0$$

alors $-T$ et $T + T'$ sont aussi des périodes de f .

2. Période fondamentale

Soit f une fonction périodique. Si l'ensemble des périodes strictement positives de f a un plus petit élément T_0 , celui-ci est appelé **période fondamentale de f** (la notion de plus petit élément n'ayant pas été définie, on peut supposer que c'est une période qui est plus petite que toutes les autres ; mais il reste à montrer qu'elle existe). Toutes les périodes de f sont alors de la forme nT_0 , $n \in \mathbb{Z}$.



Pour étudier une fonction périodique de période T , il suffit de se placer sur un intervalle I_T de longueur (ou d'amplitude) T . La courbe représentative de la fonction est alors obtenue en « recopiant le motif » obtenu sur I_T !

1. Puissances entières

1. Étant donné un entier naturel non nul n et un réel a , a^n est égal au produit de n facteurs égaux à a :

$$a^n = a \times a \times a \times \dots \times a \quad (n \text{ fois})$$

2. Étant donné un entier relatif strictement négatif $k \in \mathbb{Z}_-^*$, et un réel non nul a , a^k est égal à l'inverse de a^{-k} :

$$a^k = \frac{1}{a^{-k}}$$

3. Pour tout réel non nul a : $a^0 = 1$.

Propriété

Soient a, b des réels et n, p des entiers. On suppose a, b non nuls chaque fois que l'exposant est négatif ou nul. Alors :

$$a^n a^p = a^{n+p}, \quad (a^n)^p = a^{np}, \quad a^n b^n = (ab)^n, \quad \frac{a^n}{b^n} = \left(\frac{a}{b}\right)^n$$

Démonstration : Les propriétés du produit dans \mathbb{R} permettent de justifier simplement les formules précédentes. ■

2. Fonction puissance

Étant donné un entier relatif k non nul, la fonction qui, à tout réel x non nul, associe x^k , est une **fonction puissance**.

► Parité des fonctions puissances

Soit n un entier naturel non nul n . La fonction qui, à tout réel x , associe x^n , a la même parité que l'entier n .

La fonction qui, à tout réel x non nul, associe x^{-n} , a la même parité que l'entier n .

Exemples

1. La fonction qui, à tout réel x associe x^5 , est impaire.
2. La fonction qui, à tout réel x associe x^4 , est paire.

► Sens de variation des fonctions puissances

Soit n un entier naturel non nul. Alors :

- i. La fonction qui, à tout réel x , associe x^{2n} , est décroissante sur $] -\infty, 0]$ et croissante sur $[0, +\infty[$.
- ii. La fonction qui, à tout réel x , associe x^{2n+1} , est croissante sur $\mathbb{R} =] -\infty, +\infty[$.

- iii. La fonction qui, à tout réel x non nul, associe x^{-2n} , est croissante sur $] - \infty, 0[$ et décroissante sur $]0, +\infty[$.
- iv. La fonction qui, à tout réel x non nul, associe x^{-2n-1} , est décroissante sur $] - \infty, 0[$ et sur $]0, +\infty[$.

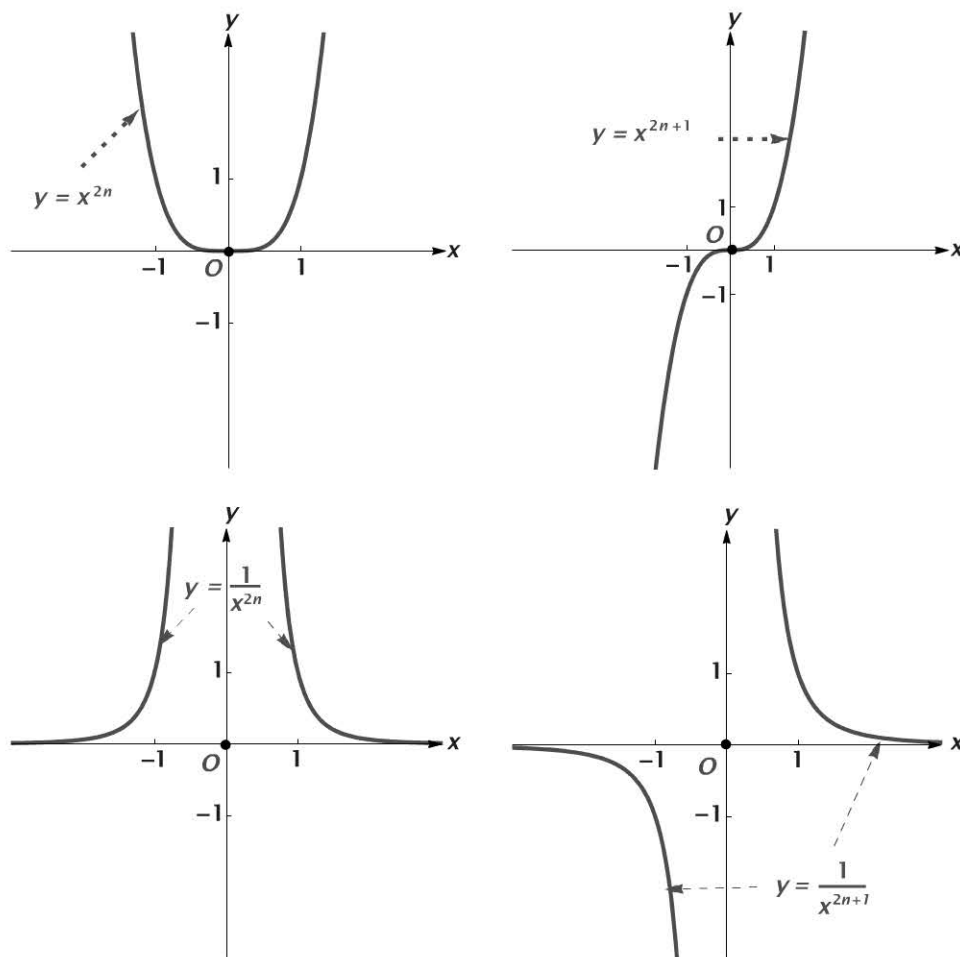


Figure 14.1 – Quelques exemples de graphes de fonctions puissances.

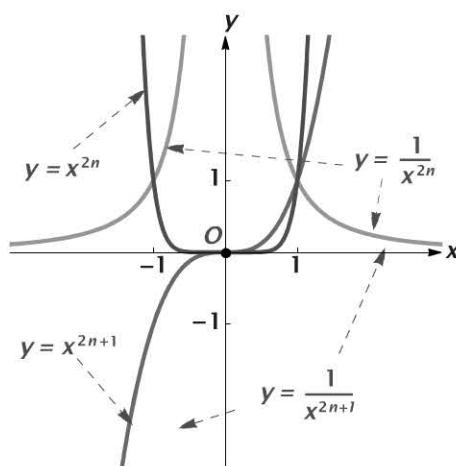


Figure 14.2 – Comparaison des fonctions puissances.

1. Fonction polynômes

► Fonction polynomiale

Étant donné un entier naturel non nul n , toute fonction de la forme

$$x \mapsto a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n = \sum_{i=0}^n a_i x^i$$

où a_0, a_1, \dots, a_n , sont des réels, est une **fonction polynomiale**.

Exemple

La fonction définie sur \mathbb{R} par $x \mapsto \frac{15}{20} x^3 + 2 x^2 - 2 x$ est une fonction polynomiale.

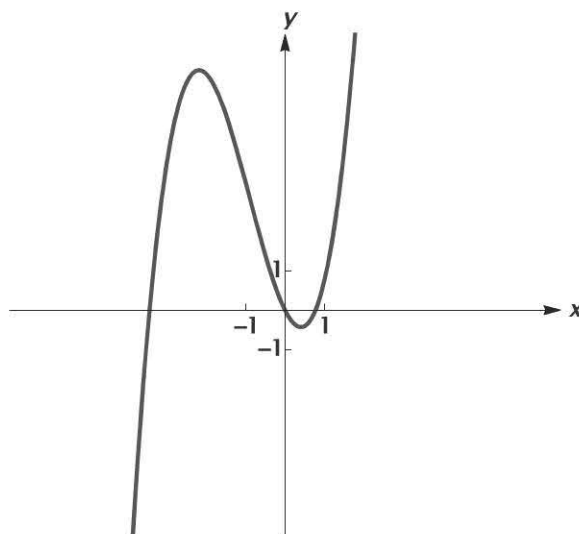


Figure 15.1 – Le graphe de la fonction $x \mapsto \frac{15}{20} x^3 + 2 x^2 - 2 x$.

► Limites d'une fonction polynomiale en $+\infty$ ou $-\infty$

Il suffit de factoriser par le « monôme de plus haut degré », c'est-à-dire par le terme de plus haut degré, pour obtenir facilement le résultat ; étant donné un entier naturel non nul n , et une fonction polynomiale de la forme $x \mapsto a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n$, où a_0, a_1, \dots, a_n , sont des réels :

$$\begin{aligned} \lim_{x \rightarrow +\infty} (a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n) &= \lim_{x \rightarrow +\infty} a_n x^n \left(\frac{a_0}{a_n x^n} + \frac{a_1}{a_n x^{n-1}} + \frac{a_2}{a_n x^{n-2}} + \dots + 1 \right) \\ &= \lim_{x \rightarrow +\infty} a_n x^n \end{aligned}$$

et, de même :

$$\begin{aligned}\lim_{x \rightarrow -\infty} (a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n) &= \lim_{x \rightarrow -\infty} a_n x^n \left(\frac{a_0}{a_n x^n} + \frac{a_1}{a_n x^{n-1}} + \frac{a_2}{a_n x^{n-2}} + \dots + 1 \right) \\ &= \lim_{x \rightarrow -\infty} a_n x^n\end{aligned}$$

Exemple

$$\begin{aligned}\lim_{x \rightarrow +\infty} (-3x^3 + x^2 + 4x - 2) &= \lim_{x \rightarrow +\infty} -3x^3 \left(1 - \frac{1}{x} - \frac{4}{3x} + \frac{2}{3x^2} \right) \\ &= \lim_{x \rightarrow +\infty} -3x^3 \\ &= -\infty\end{aligned}$$

2. Valeur absolue d'un réel

Étant donné un réel x , on appelle valeur absolue de x , que l'on note $|x|$, le réel positif :

$$|x| = \sqrt{x^2} = \begin{cases} x & \text{si } x \geq 0 \\ -x & \text{si } x < 0 \end{cases}$$

► Interprétation géométrique de la valeur absolue

Étant donnés deux réels x et y , $|x - y|$ représente la distance entre les nombres x et y sur la droite réelle.

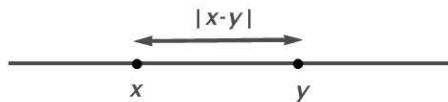


Figure 15.2– Interprétation géométrique de la valeur absolue.

Propriété

Pour tout couple de réels (x, y) : $|x y| = |x| |y|$.

► Inégalité triangulaire

Propriété

Pour tout couple de réels (x, y) :

$$|x + y| \leq |x| + |y|$$

Corollaire

Pour tout couple de réels (x, y) :

$$||x| - |y|| \leq |x + y|$$

Propriétés

1. Pour tout couple de réels (x, y) , avec $y \geq 0$:

$$|x| \leq y \Leftrightarrow -y \leq x \leq y$$

2. Pour tout couple de réels (x, y) , avec $y \geq 0$:

$$|x| \geq y \Leftrightarrow (x \geq y \text{ ou } x \leq -y)$$

► Fonction valeur absolue

On appelle **fonction valeur absolue** la fonction définie sur \mathbb{R} par $x \mapsto |x|$.

Propriété

La fonction valeur absolue est paire. Elle est croissante sur \mathbb{R}^+ , et décroissante sur \mathbb{R}^- .

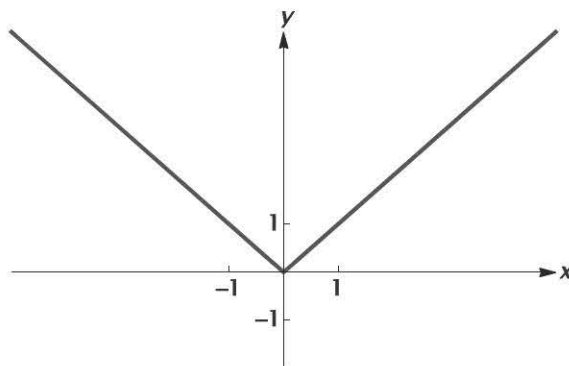


Figure 15.3 – Le graphe de la fonction valeur absolue.



Les premières tables de calcul

À partir du XIV^e siècle, l'astronomie et la navigation deviennent de plus en plus précises. Elles requièrent, de ce fait, des calculs (multiplications, divisions, extraction de racines, ...), qui se révèlent de plus en plus longs et compliqués. Il devient nécessaire de mettre en place des outils permettant de les simplifier. Peu à peu, l'idée de tables permettant de faire, rapidement, multiplications et divisions, émerge.

En 1614, le mathématicien anglais John Napier, ou Neper (1550-1617)¹, se basant sur le lien entre les progressions arithmétiques (des suites arithmétiques) et géométriques (des suites géométriques), publie les premières tables logarithmiques (des logarithmes de sinus, dans *Mirifici logarithmorum canonis descriptio*), qui permettent de transformer des produits en sommes. Ainsi, pour calculer le produit du nombre a par le nombre b , il suffit de chercher sur la table le logarithme de a et celui de b , de faire leur somme, et de retrouver ensuite, par simple lecture sur la table, le nombre dont cette somme est le logarithme !

Après les tables de Neper, Henry Briggs (1556-1630), mathématicien, géomètre et géographe, perfectionna les calculs de John Neper et publia des tables de logarithme de base 10 (aussi dit logarithme décimal), où le logarithme de 1 vaut 0, et celui de 10, 1. Ainsi, le logarithme décimal de $100000000000000 = 10^{14}$ vaut 14.

Des tables à la fonction logarithme

Ces tables de valeur préfiguraient la fonction en elle-même. On l'a appelée « logarithme népérien » en hommage à John Neper. Le logarithme népérien, ou de base e , noté \ln , est celui prenant la valeur 1 en $e \simeq 2,71828$. C'est grâce à ces tables de valeurs que l'on a pu « construire » cette fonction. Le lecteur intéressé pourra trouver plus de précisions sur l'historique des logarithmes dans [1]. Les tables de logarithmes ont été utilisées très longtemps. Avant l'apparition des calculatrices, la règle à calcul était un outil efficace et puissant pour la détermination des logarithmes !

La lecture d'une table de logarithmes.

Nombre	Logarithme
a	$\ln a$
b	$\ln b$
ab	$\ln a + \ln b$

1. Il était aussi astronome, physicien, et théologien.

On admet l'existence de la fonction logarithme népérien comme étant l'unique fonction vérifiant les propriétés suivantes.

1. Propriétés

Les propriétés suivantes sont admises :

► Logarithme népérien du nombre e

Le nombre e tel que :

$$\ln e = 1$$

est appelé base du **logarithme népérien** ($e \approx 2,71828$).

► Logarithme népérien d'un produit

Pour tout couple (a, b) de réels strictement positifs :

$$\ln(ab) = \ln a + \ln b$$

► Logarithme népérien d'un quotient

- Pour tout couple (a, b) de réels strictement positifs :

$$\ln\left(\frac{a}{b}\right) = \ln a - \ln b$$

- Pour tout réel strictement positif a , et tout entier naturel non nul n :

$$\ln(a^n) = n \ln a, \quad \ln(a^{-n}) = -n \ln a$$

- $\lim_{x \rightarrow +\infty} \frac{\ln x}{x} = 0$, $\lim_{x \rightarrow 0^+} \ln x = -\infty$

2. Fonction logarithme népérien

On appelle **fonction logarithme népérien** la fonction, notée \ln , qui, à tout réel x de $]0, +\infty[$, associe son logarithme népérien $\ln x$.

► Tableau de variations de la fonction logarithme népérien

x	0	$+\infty$
$\ln x$	$-\infty$	$+\infty$

- Comme $\lim_{x \rightarrow 0^+} \ln x = -\infty$, l'axe (Oy) est asymptote verticale à la courbe représentative de la fonction logarithme népérien.
- Comme $\lim_{x \rightarrow +\infty} \frac{\ln x}{x} = 0$, la courbe représentative de la fonction logarithme népérien possède une branche parabolique horizontale lorsque x tend vers $+\infty$.

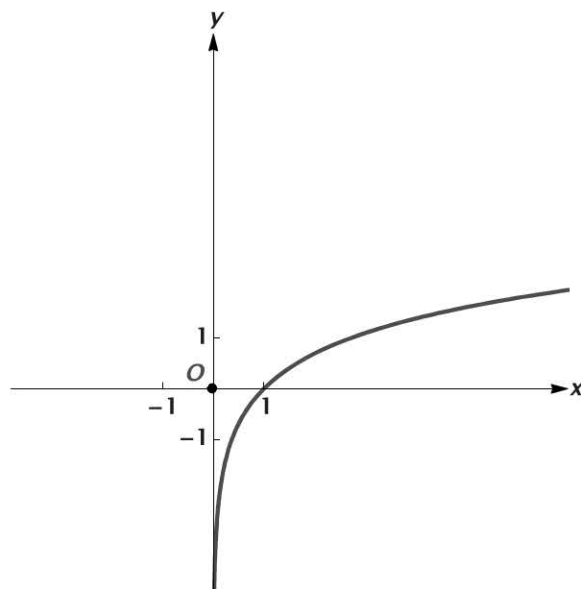


Figure 16.1 – Le graphe de la fonction logarithme népérien.

► **Une inégalité utile**

Pour tout réel strictement positif x : $\ln x \leq x - 1$.

► **Logarithme de base a , $a \in \mathbb{R}_+^*$**

Soit a un réel strictement positif. Pour tout réel strictement positif x , on définit son **logarithme de base a** , noté $\log_a x$, par :

$$\log_a x = \frac{\ln x}{\ln a}$$

On admet l'existence de la fonction exponentielle comme étant l'unique fonction vérifiant les propriétés ci-dessous.

1. Propriétés

Les propriétés suivantes sont admises :

1. Pour tout couple de réels (a, b) :

$$e^{a+b} = e^a e^b, \quad e^{a-b} = \frac{e^a}{e^b}$$

2. Pour tout réel a , et tout entier naturel n :

$$(e^a)^n = e^{na}, \quad (e^a)^{-n} = \frac{1}{e^{na}}$$

3. Pour tout réel strictement positif a : $e^{\ln a} = a$.

4. Pour tout réel a : $\ln e^a = a$.

2. Fonction exponentielle

On appelle **fonction exponentielle** la fonction, notée e , ou \exp , qui, à tout réel x , associe e^x (que l'on peut aussi écrire $\exp(x)$).

► Tableau de variations de la fonction exponentielle

x	$-\infty$	$+\infty$
e^x	0^+	$+\infty$

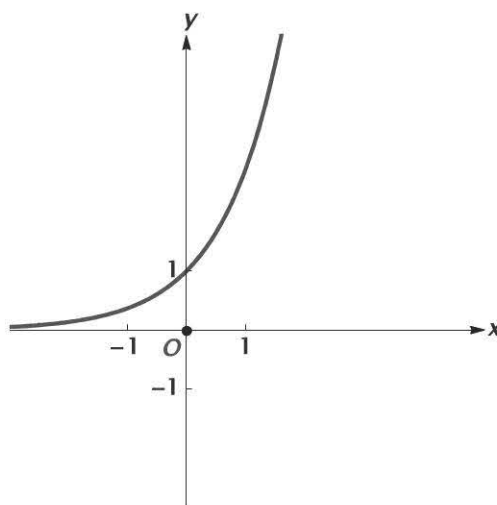


Figure 17.1 – La courbe représentative de la fonction exponentielle.

3. Puissance (quelconque) d'un réel strictement positif

Étant donné un réel a strictement positif, et un réel b , on définit le réel « a puissance b », noté a^b , par :

$$a^b = e^{b \ln a}$$



Cette définition est cohérente avec la définition de la puissance entière, puisque, pour tout réel strictement positif a , et tout entier relatif k :

$$e^{k \ln a} = e^{\ln(a^k)} = a^k$$

Étant donné un réel strictement positif a , et un réel b :

$$a^{-b} = \frac{1}{a^b} = \left(\frac{1}{a}\right)^b$$

Étant donné un réel strictement positif a , et deux réels b_1 et b_2 :

$$a^{b_1+b_2} = a^{b_1} a^{b_2} \quad , \quad a^{b_1 b_2} = (a^{b_1})^{b_2}$$

Étant donné deux réels strictement positif a_1 et a_2 , ainsi qu'un réel b :

$$(a_1 a_2)^b = a_1^b a_2^b$$

4. Racine $n^{\text{ième}}$ d'un réel strictement positif, $n \in \mathbb{N}^*$

Étant donné un réel x strictement positif, et un entier naturel non nul n , on appelle racine $n^{\text{ième}}$ de x , notée $\sqrt[n]{x}$, le réel :

$$\sqrt[n]{x} = x^{\frac{1}{n}}$$

1. Sens de variation des fonctions puissances non entières

Soit α un réel non entier. La fonction $x \mapsto x^\alpha$ est définie sur \mathbb{R}_+^* . De plus :

- Si $\alpha > 0$, la fonction $x \mapsto x^\alpha$ est croissante sur \mathbb{R}_+^* .
- Si $\alpha < 0$, la fonction $x \mapsto x^\alpha$ est décroissante sur \mathbb{R}_+^* .

2. Comparaison des fonctions puissances non entières

Soient α et β deux réels non entiers tels que $\alpha \leq \beta$. Alors :

- pour tout réel x de $]0, 1]$: $x^\alpha \geq x^\beta$;
- pour tout réel $x \geq 1$: $x^\alpha \leq x^\beta$.

Exemple

Pour tout réel x de $]0, 1]$: $0 < x^4 \leq x^3 \leq x^2 \leq x \leq 1 \leq \frac{1}{x^4} \leq \frac{1}{x^3} \leq \frac{1}{x^2} \leq \frac{1}{x}$

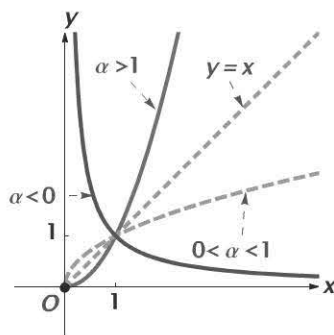


Figure 18.1 – Les graphes des fonctions puissances non entières.

3. Limites usuelles des fonctions puissances non entières

Soit α un réel strictement positif, non entier. Alors :

$$\lim_{x \rightarrow 0^+} x^\alpha = 0^+$$

$$\lim_{x \rightarrow +\infty} x^\alpha = +\infty$$

$$\lim_{x \rightarrow 0^+} x^{-\alpha} = +\infty$$

$$\lim_{x \rightarrow +\infty} x^{-\alpha} = 0^+$$

4. Croissances comparées

Soit α un réel strictement positif, et β un réel quelconque. Alors :

$$\lim_{x \rightarrow +\infty} \frac{e^x}{x^\alpha} = +\infty, \quad \lim_{x \rightarrow +\infty} \frac{e^x}{x^{-\alpha}} = +\infty, \quad \lim_{x \rightarrow +\infty} \frac{x^\alpha}{(\ln x)^\beta} = +\infty, \quad \lim_{x \rightarrow 0^+} x^\alpha (\ln x)^\beta = 0$$



Historiquement, la « naissance » de la fonction exponentielle vient de la nécessité de trouver un moyen de définir les « puissances non entières » d'un réel strictement positif donné. En effet, autant des expressions de la forme a^n , a^{n+1} , ..., où a est un réel strictement positif, et n un entier naturel, ont toujours été « naturelles », autant il n'en a pas toujours été le cas pour des expressions de la forme a^r , où r est, cette fois-ci, un réel (par exemple : $a^{1,38}$, quelque part entre $a^1 = a$ et a^2 ...)

En 1676, Isaac Newton et Gottfried Leibniz sont les premiers à écrire, successivement, un nombre fractionnaire, puis un irrationnel, en exposant. Mais il faut attendre la fin du XVII^e siècle pour que de tels exposants commencent à être perçus comme des logarithmes. G. Leibniz donne la relation explicite en 1679, sous la forme :

$$\ln\left(\frac{1+v}{1-v}\right) = t \Leftrightarrow \frac{1+v}{1-v} = b^t$$

où il désigne par b une « grandeur constante dont le logarithme vaut 1 ». Ultérieurement, le mathématicien suisse Leonhard Euler, introduira, à cet effet, la notation « e ». Jean Bernoulli finalisera l'étude de la fonction « exponentielle » ainsi obtenue, qui apparaît donc, naturellement, comme fonction réciproque du logarithme népérien.

- Isaac Newton (1643-1727) était non seulement mathématicien, mais, aussi, philosophe, physicien, alchimiste, astronome et théologien. C'est lui qui est à l'origine du calcul infinitésimal, c'est-à-dire le calcul différentiel et intégral. Il est, également, l'un des contributeurs majeurs en mécanique classique, avec la théorie de la gravitation universelle (la fameuse « pomme » de Newton, qui tomba d'un arbre sur sa tête).
- Gottfried Leibniz (1646-1716) fut aussi philosophe, diplomate, juriste, et philologue. C'est lui qui, le premier, employa le terme de « fonction », et introduisit le symbole \int utilisé pour désigner une intégrale.
- Leonhard Euler (1707-1783) contribua lui aussi au calcul infinitésimal, introduisit une grande partie des notations mathématiques modernes. Il est aussi l'auteur de nombreux travaux en mécanique, en dynamique des fluides, astronomie, ...
- Jean Bernoulli (1667-1748), frère cadet du mathématicien suisse Jacques Bernoulli (1654-1705), oncle de Daniel (1700-1782) et Nicolas Bernoulli (1695-1726). Il trouva l'équation de la courbe dite « chaînette », correspondant à la fonction *cosinus hyperbolique*, et à la forme prise par un câble suspendu à ses extrémités et soumis à son poids. De façon amusante, il est l'ancêtre des prix Nobel de physique Pierre Curie et Pierre-Gilles de Gennes.

1. Fonction sinus

On appelle **fonction sinus** la fonction, notée \sin , qui, à tout réel x , associe son sinus, $\sin x$.

► Tableau de variations de la fonction sinus sur $[0, \pi]$

La fonction *sinus* étant impaire et périodique de période 2π , il suffit de l'étudier sur une demi-période, par exemple, $[0, \pi]$.

x	0	$\frac{\pi}{2}$	π
$\sin x$	0^+	1	0^+

2. Fonction cosinus

On appelle **fonction cosinus** la fonction, notée \cos , qui, à tout réel x associe son cosinus, $\cos x$.

► Tableau de variations de la fonction cosinus sur $[0, \pi]$

La fonction *cosinus* étant paire et périodique de période 2π , il suffit de l'étudier sur une demi-période, par exemple, $[0, \pi]$.

x	0	$\frac{\pi}{2}$	π
$\cos x$	1	0	-1

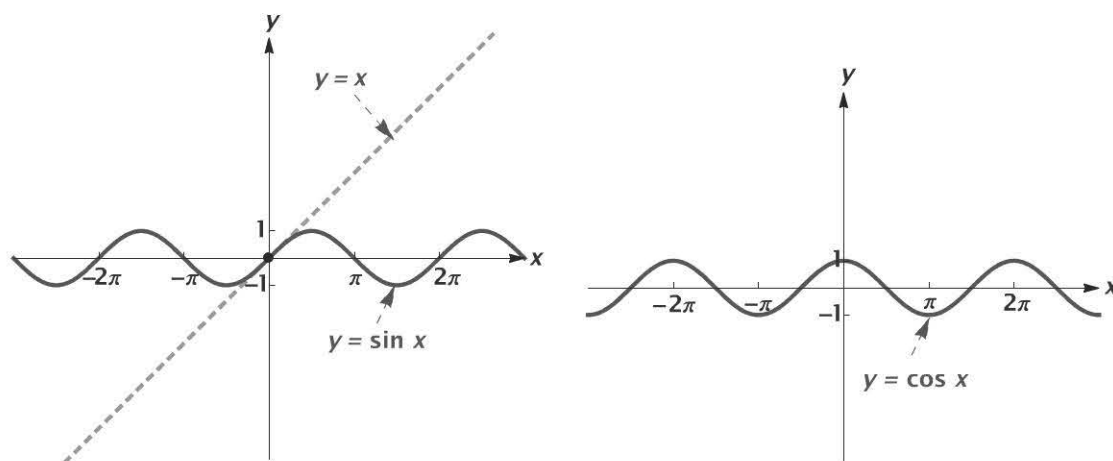


Figure 19.1 – Les courbes représentatives des fonctions sinus et cosinus.

3. Fonction tangente

On appelle **fonction tangente** la fonction, notée \tan , qui, à tout x de $\mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi, k \in \mathbb{Z} \right\}$, associe sa tangente, $\tan x$.

► Tableau de variations de la fonction tangente sur $[0, \pi]$

La fonction *tangente* est définie sur $\mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi, k \in \mathbb{Z} \right\}$ par :

$$\tan x = \frac{\sin x}{\cos x}$$

Elle est impaire et périodique de période π . Il suffit donc de l'étudier sur $\left[0, \frac{\pi}{2}\right[$.

x	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$
$\tan x$	0^+	1	$+\infty$

Comme $\lim_{x \rightarrow \frac{\pi}{2}^-} \tan x = +\infty$, la droite d'équation $x = \frac{\pi}{2}$ est asymptote à la courbe représentative de la fonction tangente lorsque x tend vers $\frac{\pi}{2}$ par valeurs inférieures. Il en est de même, par périodicité, de toute droite d'équation $x = \frac{\pi}{2} + k\pi, k \in \mathbb{Z}$.

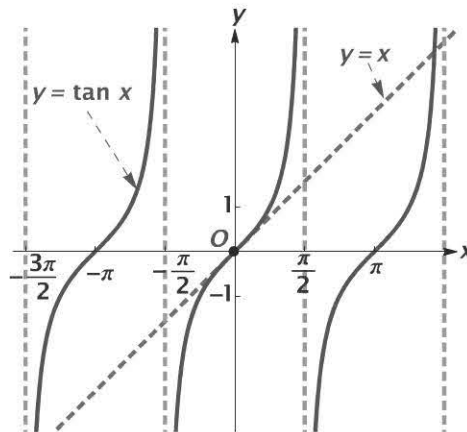


Figure 19.2 – La courbe représentative de la fonction tangente.

4. Valeurs remarquables des fonctions sinus, cosinus et tangente

x	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\sin x$	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1
$\cos x$	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0
$\tan x$	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	non définie

Pour tout réel x :

$$\cos^2 x + \sin^2 x = 1$$

1. Sinus hyperbolique

Définition

Étant donné un réel x , on appelle **sinus hyperbolique** de x le réel, noté $\operatorname{sh} x$, tel que :

$$\operatorname{sh} x = \frac{e^x - e^{-x}}{2}$$

La fonction **sinus hyperbolique**, notée sh , est la fonction qui, à tout réel x , associe son sinus hyperbolique $\operatorname{sh} x$.

► Tableau de variations de la fonction sinus hyperbolique

La fonction *sinus hyperbolique* est définie sur \mathbb{R} , et impaire. Il suffit de l'étudier sur \mathbb{R}_+ .

x	0	$+\infty$
$\operatorname{sh} x$	0^+	$+\infty$

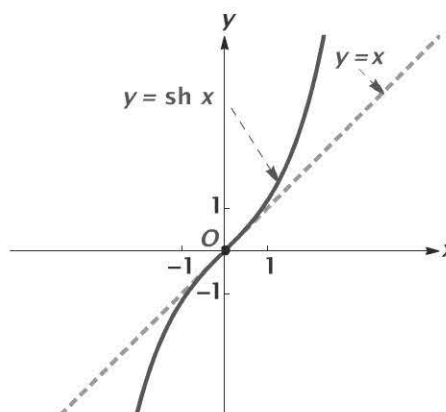


Figure 20.1 – La courbe représentative de la fonction *sinus hyperbolique*.

2. Cosinus hyperbolique

Définition

Étant donné un réel x , on appelle **cosinus hyperbolique** de x le réel, noté $\operatorname{ch} x$, tel que :

$$\operatorname{ch} x = \frac{e^x + e^{-x}}{2}$$

La fonction **cosinus hyperbolique**, notée ch , est la fonction qui, à tout réel x , associe son cosinus hyperbolique $\operatorname{ch} x$.

► Tableau de variations de la fonction cosinus hyperbolique

La fonction *cosinus hyperbolique* est définie sur \mathbb{R} , et paire. Il suffit de l'étudier sur \mathbb{R}_+ .

x	0	$+\infty$
$\text{ch } x$	1	$+\infty$

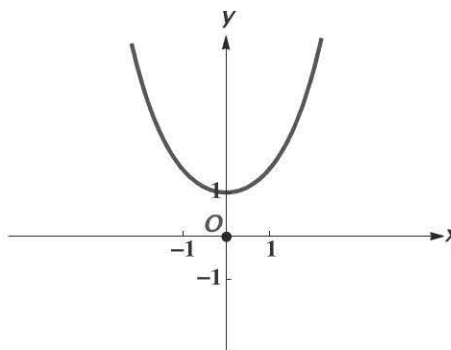


Figure 20.2 – La courbe représentative de la fonction *cosinus hyperbolique*.

3. Tangente hyperbolique

Définition

Étant donné un réel x , on appelle **tangente hyperbolique** de x le réel, noté $\text{th } x$, tel que :

$$\text{th } x = \frac{\text{sh } x}{\text{ch } x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

La fonction **tangente hyperbolique**, notée th , est la fonction qui, à tout réel x , associe sa tangente hyperbolique $\text{th } x$.

► Tableau de variations de la fonction *tangente hyperbolique*

La fonction *tangente hyperbolique* est définie sur \mathbb{R} , et impaire. Il suffit de l'étudier sur \mathbb{R}_+ .

x	0	$+\infty$
$\text{th } x$	0^+	1

Propriétés

- Comme : $\lim_{x \rightarrow +\infty} \text{th } x = 1$, $\lim_{x \rightarrow -\infty} \text{th } x = -1$

les droites d'équations respectives $y = 1$ et $y = -1$ sont asymptotes à la courbe représentative de la fonction tangente hyperbolique lorsque x tend vers $+\infty$ et $-\infty$ respectivement.

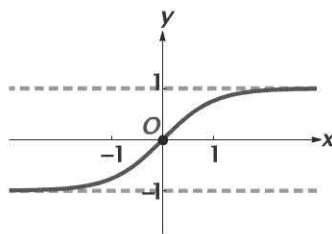
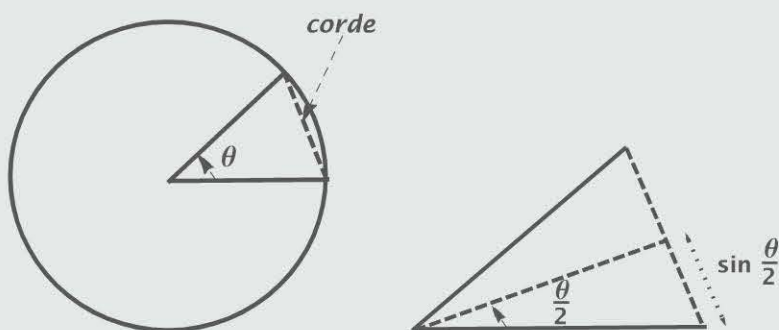


Figure 20.3 – La courbe représentative de la fonction *tangente hyperbolique*.

- Pour tout réel x : $\text{ch}^2 x - \text{sh}^2 x = 1$

Les précurseurs grecs

La chronologie exacte de l'apparition de la trigonométrie, et des fonctions circulaires, demeure incertaine. De tout temps, les astronomes ont eu besoin de tables permettant le passage de la mesure des angles à celle des arcs et des cordes sous-tendues associées (dont la longueur est égale à deux fois le sinus de l'angle moitié). Il semblerait qu'il faille attendre le mathématicien et géomètre Hipparque de Nicée (180 av.J.-C./125 av.J.-C.), qui divise le cercle en 360° , pour qu'apparaissent ces premières tables (dans son ouvrage *De l'étude des droites dans le cercle*), pour lesquelles il passa beaucoup de temps à observer les astres et leur mouvement. C'est lui qui inventa l'« astrolabe », qui permet d'établir la hauteur d'un astre par rapport à l'horizon. Il introduisit aussi la notion de parallèles et de méridiens pour repérer la position d'un point sur la terre. Ultérieurement c'est Ptolémée (environ 90-168, astronome et astrologue grec, qui vécut à Alexandrie.) dans *l'Almageste*, qui expliqua comment calculer la longueur d'une corde, en donnant les tables correspondantes.



Un angle, et la corde sous-tendue.
(le rayon du cercle vaut 1)

Des Indiens aux arabes : les débuts de l'algèbre

La première définition véritable du *sinus*, de même que celle du *cosinus*, est due au mathématicien et astronome indien Aryabhata (476-550, travailla aussi sur l'approximation du nombre π .) Il eut l'idée d'utiliser non pas la corde sous-tendue à un arc, mais la demi-corde, qui correspond donc exactement à la valeur du sinus de l'angle moitié. Le nom de *sinus* en lui-même, qui vient, bien sûr, du latin, semble devoir son origine à une erreur de traduction depuis le sanskrit. Aryabhata établit lui aussi des tables de valeurs, avec quatre décimales, ce qui était, pour l'époque, extrêmement précis.

La formule bien connue qui donne, pour un angle θ , $\cos^2 \theta + \sin^2 \theta = 1$, fut établie par le mathématicien et astronome indien Varahamihira (505-587).

Des calculs plus poussés furent ensuite donnés par le mathématicien perse Al-Khwarizmi, qui est aussi à l'origine de l'introduction de l'algèbre et des chiffres arabes en Europe.

Peu à peu, d'autres mathématiciens apportèrent de nouvelles contributions, et démontrèrent de nouvelles formules. L'Égyptien Habash al-Hasib inventa la tangente, qui permet de mesurer des hauteurs. Abu Al-Wafa compléta les tables de valeurs déjà existantes, et introduisit les notions de « sécante » (l'inverse du *cosinus*) et « co-sécante » (l'inverse du *sinus*). Il démontra les formules d'addition pour la fonction *sinus*.

Nasir Al-Din Al-Tusi perfectionna les tables de valeurs déjà existantes. Il fut suivi au, XIV^e siècle, par Al-Kashi, qui est aussi à l'origine du fameux théorème qui porte son nom (ce théorème est aussi appelé loi des cosinus).

Les notations modernes

En ce qui concerne la notation « sin », elle fut introduite en 1583 par le mathématicien et physicien danois Thomas Fincke (1561-1656) dans son ouvrage « Geometria rotundi » ; la notation « cos » semble pouvoir être attribuée, conjointement, au mathématicien et théologien anglais William Oughtred (1574-1660), et au français Albert Girard. Albert Girard introduisit, aussi, la notation « tan ».

- Abu Abdallah Muhammad ibn Musa al-Khwarizmi (780-850), mathématicien, géographe et astronome perse, sous l'empire de la dynastie des Abbasides. Ce sont ses travaux, où il établit un classement systématique des équations et des méthodes de résolution associées, qui ont permis l'introduction de l'algèbre et des chiffres arabes en Europe. Le mot « algorithme » vient de la latinisation de son nom. Le mot « algèbre » provient, quant à lui, du mot arabe « al-jabr », utilisé pour désigner l'une des deux opérations qu'il utilisait pour résoudre une équation quadratique (c'est-à-dire de degré deux).
- Habash al-Hasib (? -869), était aussi, bien sûr, astronome, et géographe.
- Muhammad ibn Muhammad ibn Yahya ibn Ismail ibn al-Abbas al-Buzjani (940-998), mathématicien et astronome perse, qui apporta aussi de nombreuses contributions à l'arithmétique, il fut le premier à introduire les nombres négatifs.
- Abû Jafar Muhammad ben Muhammad ben al-Hasan Nasîr ad-Dîn at-Tûsî (1201-1274), philosophe, mathématicien, astronome, théologien et médecin perse. Le petit-fils de Genghis Khan, Houlagou Khan, fit construire, à son intention, l'observatoire de Maragha, qui lui permit d'établir des tables très précises permettant de calculer les positions des planètes.
- Ghiyath al-Din Jamshid Masud al-Kashi, 1380-1429, mathématicien et astronome perse.
- Albert Girard (1595-1632), apporta aussi des contributions en arithmétique, où il montra que tout nombre premier congru à 1 modulo 4 est égal à la somme de deux carrés, résultat qui sera ultérieurement démontré par Pierre de Fermat (1601-1665).

1. Continuité en un point (Caractérisation de Weierstrass)

Définition

Soit f une fonction définie sur un intervalle I de \mathbb{R} , non vide, non réduit à un point, et a un point de I . La fonction f est dite **continue en a** si, pour tout réel strictement positif ε , il existe un réel strictement positif η tel que, pour tout réel x de I vérifiant $|x - a| \leq \eta$, on ait $|f(x) - f(a)| \leq \varepsilon$, soit, en langage formalisé¹ :

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in I : |x - a| \leq \eta \Rightarrow |f(x) - f(a)| \leq \varepsilon$$

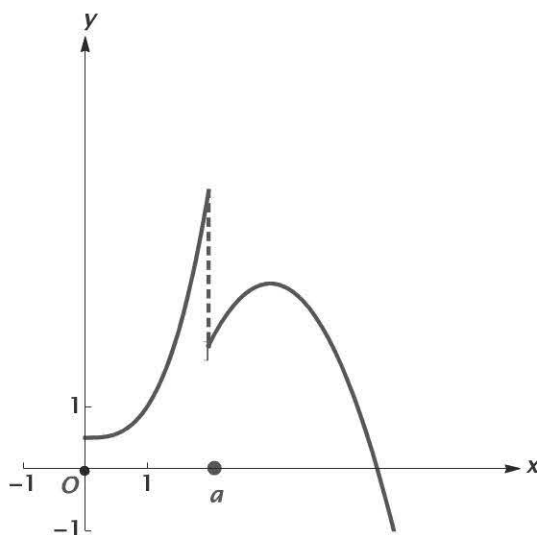


Figure 21.1 – Le graphe d'une fonction discontinue en un point a .



Dire qu'une fonction est continue signifie, tout simplement, que sa courbe représentative ne présente pas de « sauts ». Une fonction continue en un point a admet une limite en a .

► Continuité à gauche

Définition

Soit f une fonction définie sur un intervalle I de \mathbb{R} , non vide, non réduit à un point, et a un point de I . La fonction f est dite **continue à gauche en a** si :

$$\lim_{x \rightarrow a^-} f(x) = f(a)$$

1. Cette caractérisation a été donnée par le mathématicien allemand Karl Weierstrass (1815-1897).

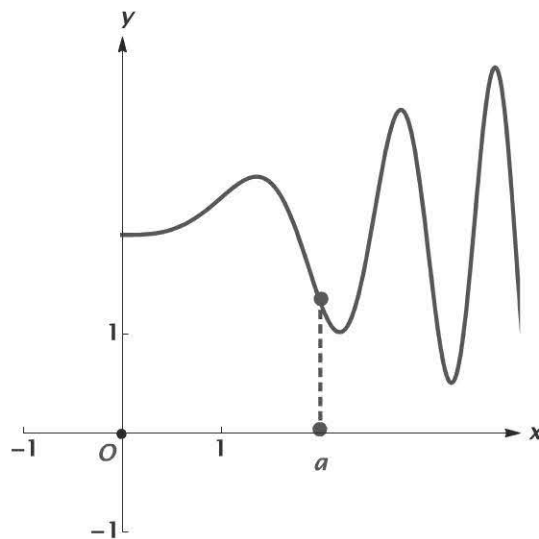


Figure 21.2– Le graphe d'une fonction continue en un point a .

► Continuité à droite

Définition

Soit f une fonction définie sur un intervalle I de \mathbb{R} , non vide, non réduit à un point, et a un point de I . La fonction f est dite **continue à droite en a** si :

$$\lim_{x \rightarrow a^+} f(x) = f(a)$$



Lorsque l'on veut explicitement indiquer sur un graphe qu'une fonction est continue à gauche et non à droite au point a , on place un « point » sur la valeur « à gauche », et un crochet « $] \gg$ » sur la valeur « à droite ».

De même, lorsque l'on veut explicitement indiquer sur un graphe qu'une fonction est continue à droite et non à gauche au point a , on place un « point » sur la valeur « à droite », et un crochet « $[\gg$ » sur la valeur « à gauche ».

Exemple

Voici un exemple de fonction continue à droite en chaque entier, et discontinue à gauche en chaque entier : la fonction « partie entière ».

La fonction « partie entière » d'un réel donné x est le plus grand entier n_x inférieur ou égal à x . Ainsi, la partie entière du nombre 2,3 est 2, celle du nombre 4,7 est 4, etc. :

$$\forall x \geq 0 : E(x) \leq x < E(x) + 1$$

Par définition, la fonction « partie entière », notée E , est la fonction qui, à tout réel positif x , associe sa partie entière.

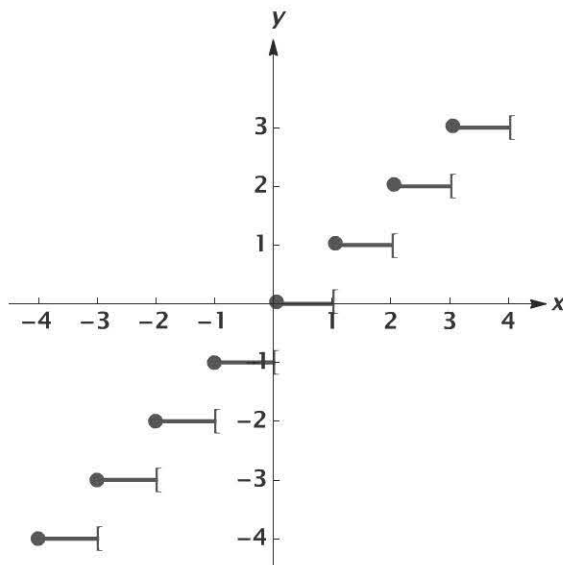


Figure 21.3– La courbe représentative de la fonction « partie entière ».

► Prolongement par continuité en un point

Théorème

Soient f une fonction définie sur un domaine \mathcal{D}_f de \mathbb{R} , à valeurs dans \mathbb{R} , et a un réel donné n'appartenant pas à \mathcal{D}_f .

On suppose que f admet une limite finie ℓ en a . Alors, la fonction \tilde{f} définie pour tout x de $\mathcal{D}_f \cup \{a\}$ par :

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \neq a \\ \ell & \text{si } x = a \end{cases}$$

est continue en a , et constitue le prolongement par continuité de f en a .

Exemple

On considère la fonction qui, à tout réel x non nul, associe $x \cos\left(\frac{1}{x}\right)$. Cette fonction est définie sur \mathbb{R}^* , mais peut être prolongée par continuité sur \mathbb{R} par la fonction :

$$x \mapsto \begin{cases} x \cos\left(\frac{1}{x}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

► Caractérisation séquentielle de la continuité

Théorème

Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , et a un réel donné dans I . Il y a équivalence entre les propriétés suivantes :

- f est continue en a ;
- Pour toute suite $(u_n)_{n \in \mathbb{N}}$, à valeurs dans I , de limite a , la suite $(f(u_n))_{n \in \mathbb{N}}$ converge vers $f(a)$.

Partie 3

► Opérations algébriques sur les fonctions continues

Théorème

Soient f et g des fonctions définies sur un même intervalle I de \mathbb{R} , et a un point de I . Alors :

- Si f et g sont continues en a , les fonctions $f + g$ et fg sont définies sur I et continues en a .
- Si $g(a) \neq 0$, et si g est continue en a , la fonction $\frac{1}{g}$ est définie sur un intervalle de la forme $]a - \eta, a + \eta[\cap I$, $0 < \eta < \alpha$, et est continue en a .

► Continuité des fonctions composées

Théorème

Soient f une fonction définie sur un intervalle I de \mathbb{R} , et g une fonction définie sur un intervalle $J \subset \mathbb{R}$ contenant $f(I)$: $f(I) \subset J$.

Si f est continue en un point a de I et si g est continue au point $f(a) \in J$, la fonction composée $g \circ f$ est définie sur l'intervalle I et continue en a .

► Continuité des fonctions usuelles

Théorème

Les fonctions usuelles, c'est-à-dire :

- la fonction identité $x \mapsto x$;
 - la fonction logarithme népérien $x > 0 \mapsto \ln x$;
 - la fonction exponentielle,
 - les fonctions sinus et cosinus,
 - les fonctions sinus hyperbolique et cosinus hyperbolique,
- sont continues en tout point de leur domaine de définition.

Corollaire

Les fonctions construites à partir des fonctions usuelles par opérations algébriques et composition sont continues en tout point où elles sont définies.



Quelques conséquences simples :

- Les fonctions polynômes, de la forme $x \in \mathbb{R} \mapsto a_n x^n + a_{n-1} x^{n-1} + \dots + a_0$, $n \in \mathbb{N}^*$, $(a_0, \dots, a_n) \in \mathbb{R}^{n+1}$, sont continues en tout point de \mathbb{R} .
- Les fonctions « fractions rationnelles » $x \mapsto \frac{P(x)}{Q(x)}$ (où P et Q sont des polynômes et Q n'est pas identiquement nul) sont continues en tout point où le polynôme Q ne s'annule pas.
- Les fonctions de la forme $x \mapsto x^a$, $a \in \mathbb{R}$, sont continues en tout point où elles sont définies.

1. Continuité sur un intervalle

On étend, ici, la notion de continuité en un point à un intervalle.

Définition

Soit f une fonction définie sur un intervalle I de \mathbb{R} . On dit que la fonction f est **continue sur I** si f est continue en tout point de I , soit, de façon formelle :

$$\forall x_0 \in I, \forall \varepsilon > 0, \exists \eta > 0, \forall x \in I : |x - x_0| \leq \eta \Rightarrow |f(x) - f(x_0)| \leq \varepsilon$$

2. Théorèmes de continuité globale

Les théorèmes de continuité en un point se traduisent immédiatement en des théorèmes de continuité globale :

► Continuité et opérations algébriques

Théorème

Soient f et g des fonctions définies et continues sur un intervalle I de \mathbb{R} . Les fonctions $f + g$ et fg sont définies et continues sur I .

De plus, si la fonction g ne s'annule pas sur I , la fonction $\frac{1}{g}$ est définie et continue sur I .

► Continuité des fonctions composées

Théorème

Soient f une fonction définie et continue sur un intervalle I de \mathbb{R} , et g une fonction définie et continue sur un intervalle $J \subset \mathbb{R}$ contenant $f(I)$.

La fonction composée $g \circ f$ est définie et continue sur l'intervalle I .

► Continuité des fonctions usuelles

Théorème

- Les fonctions polynômes sont continues sur \mathbb{R} .
- La fonction logarithme népérien est continue sur $]0, +\infty[$.
- La fonction exponentielle est continue sur \mathbb{R} .
- Les fonctions sinus et cosinus sont continues sur \mathbb{R} .

Corollaire

Les fonctions construites à partir des fonctions usuelles par opérations algébriques et composition sont continues sur tout intervalle où elles sont définies :

- Les fonctions « fractions rationnelles », de la forme $x \mapsto \frac{P(x)}{Q(x)}$ sont continues sur tout intervalle où le polynôme Q ne s'annule pas.

- La fonction tangente est continue sur tout intervalle de la forme $]-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi[, k \in \mathbb{Z}$.
- Les fonctions $x \mapsto x^a, a \in \mathbb{R}$, sont continues sur $]0, +\infty[$.



Ces théorèmes permettent souvent de conclure à la continuité d'une fonction sur son ensemble de définition, à l'exception éventuelle de quelques points pour lesquels on doit faire une étude directe locale. C'est systématiquement le cas des points de raccordement lorsque la fonction est définie par morceaux, ou quand la fonction a été obtenue en prolongeant par continuité une autre fonction a priori non définie en un point. Ainsi, la fonction :

$$x \mapsto \begin{cases} x \sin\left(\frac{1}{x}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

est continue sur \mathbb{R} .

3. Théorème des valeurs intermédiaires

Théorème

Soient f une fonction définie et continue sur un intervalle I de \mathbb{R} , et a et b deux réels de I . Alors, tout réel compris entre $f(a)$ et $f(b)$ possède au moins un antécédent par la fonction f .

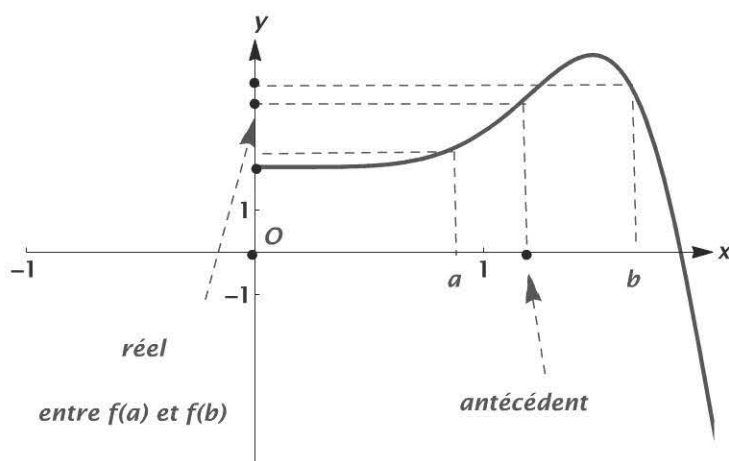


Figure 22.1 – Illustration graphique du théorème des valeurs intermédiaires.



On peut aussi énoncer ce théorème sous la forme suivante : l'image d'un intervalle par une fonction continue est un intervalle.

4. Théorème de Weierstrass, ou des bornes atteintes

Théorème

Soient f une fonction définie et continue sur un segment I de \mathbb{R} , et a et b deux réels de I . Alors, f y est bornée et atteint ses bornes, ce qui signifie qu'il existe deux réels m et M tels que, pour tout réel x de I :

$$m \leq f(x) \leq M$$

et qu'il existe deux réels x_m et x_M de I tels que :

$$f(x_m) = m \quad , \quad f(x_M) = M$$

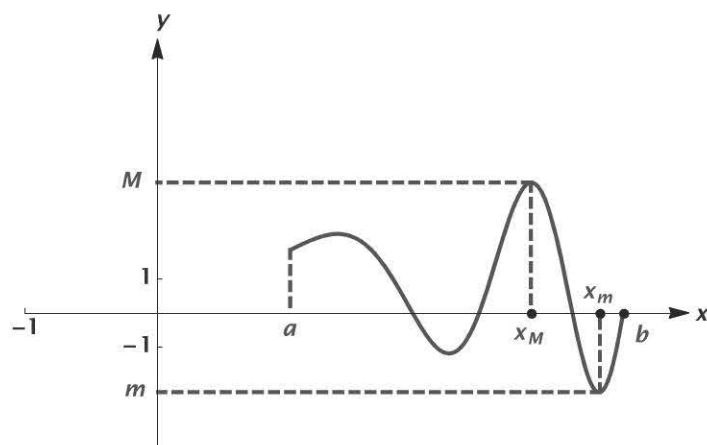


Figure 22.2 – Illustration graphique du théorème de Weierstrass.



Il résulte du théorème précédent que l'image d'un intervalle fermé et borné (c'est-à-dire un segment) par une fonction continue est un intervalle fermé et borné.

Dérivabilité en un point

1. Conditions de dérivabilité

Définition

Soient f une fonction définie sur un intervalle I de \mathbb{R} , non vide, non réduit à un point, et a un point de I . La fonction f est dite **dérivable en a** si la fonction τ , définie, pour tout x de $I \setminus \{a\}$, par :

$$\tau(x) = \frac{f(x) - f(a)}{x - a}$$

admet une limite en a . Dans ces conditions, la limite de la fonction τ en a s'appelle **dérivée de la fonction f en a** , et se note $f'(a)$.

► Interprétation graphique

Étant donnés deux réels distincts x_1 et x_2 de I , le quotient $\frac{f(x_2) - f(x_1)}{x_2 - x_1}$ est, tout simplement, le coefficient directeur de la corde joignant les points M_1 et M_2 , de coordonnées respectives $(x_1, f(x_1))$ et $(x_2, f(x_2))$, si on se place dans un repère orthonormé direct. Étudier ce qui se passe lorsque $x_1 = a$ et x_2 tend vers x_1 revient donc à étudier la position limite de la sécante à la courbe en a :

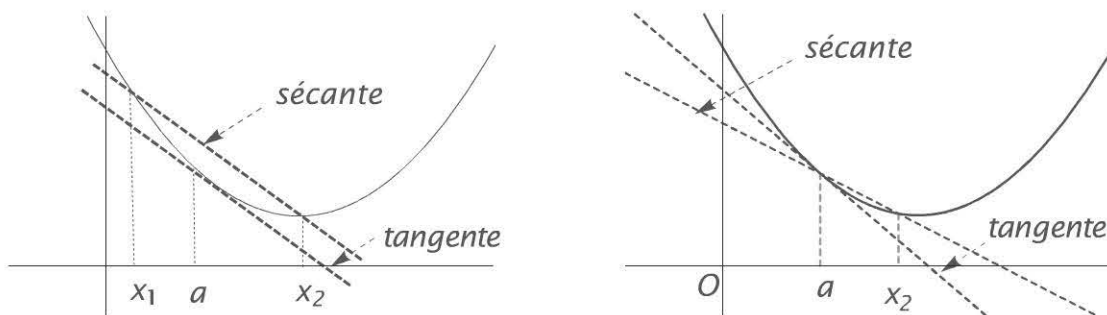


Figure 23.1 – La sécante, la position limite de la sécante, et la tangente.

La tangente à la courbe représentative de f en a ayant pour équation $y = (x - a) f'(a) + f(a)$, la fonction affine définie sur \mathbb{R} par $x \mapsto (x - a) f'(a) + f(a)$ permet donc d'approcher f par une fonction affine au voisinage de a .

On peut, de façon équivalente, donner pour la dérivabilité la définition suivante :

Définition

Soient f une fonction définie sur un intervalle I de \mathbb{R} , non vide, non réduit à un point, et a un point de I .

La fonction f est dite **dérivable en a** s'il existe deux réels A et B tels que :

$$f(x) = A + B(x - a) + (x - a) \varepsilon(x - a)$$

où ε est une fonction de limite nulle en 0. On a alors : $A = f(a)$, $B = f'(a)$



La deuxième définition présente l'avantage de pouvoir être généralisée aux fonctions de plusieurs variables. Elle a aussi la conséquence immédiate suivante :

Propriété

Toute fonction f dérivable en un point a y est continue.



La réciproque de cette proposition est FAUSSE !

Ainsi, la fonction $x \mapsto |x|$ est continue en 0, mais n'est pas dérivable en ce point.

Exemple : Une fonction continue partout, mais nulle part dérivable

Soient $a \in]0, 1[$, et b un réel tel que $ab > 1 + \frac{3\pi}{2}$. La fonction de Weierstrass

$$x \in \mathbb{R} \mapsto \sum_{n=0}^{+\infty} a^n \cos(b^n \pi x)$$

est continue partout, mais nulle part dérivable, ce qui signifie qu'elle n'admet, nulle part, de tangente.

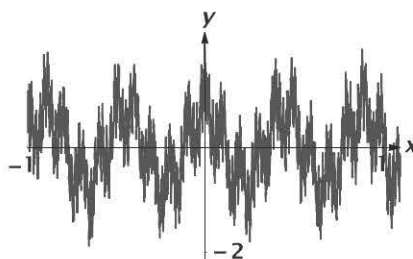


Figure 23.2 – La courbe représentative de la fonction de Weierstrass, pour $a = \frac{7}{10}$, et $b = 10$.

Les sismogrammes constituent, par exemple, des exemples de courbes continues, mais n'admettant nulle part de tangente.

2. Opérations algébriques et composition

► Dérivabilité en un point et opérations algébriques

Théorème

Soient f et g des fonctions définies sur un même intervalle I de \mathbb{R} , et a un point de I . Alors :

- Si f et g sont dérivables en a , les fonctions $f + g$ et fg sont dérivables en a , et :

$$(f + g)'(a) = f'(a) + g'(a) \quad , \quad (fg)'(a) = f'(a)g(a) + f(a)g'(a)$$

- Si $g(a) \neq 0$, et si g est dérivable en a , la fonction $\frac{1}{g}$ est définie sur un intervalle de la forme $]a - \eta, a + \eta[\cap I$, $0 < \eta$, et est dérivable en a . De plus :

$$\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{g(a)^2}$$

► Dérivabilité des fonctions composées

Théorème

Soient f une fonction définie sur un intervalle I de \mathbb{R} , et g une fonction définie sur un intervalle J contenant $f(I)$.

Si f est dérivable en un point a de I , et si g est dérivable au point $f(a)$, alors la fonction composée $g \circ f$ est dérivable en a et : $(g \circ f)'(a) = f'(a) g'(f(a))$

Démonstration : Posons $b = f(a)$. La dérivabilité de la fonction f en a s'écrit :

$$f(x) = f(a) + (x - a)f'(a) + (x - a)\varepsilon_1(x - a)$$

où ε_1 est une fonction de limite nulle en 0.

La dérivabilité de la fonction g en b s'écrit :

$$g(x) = g(b) + (x - b)g'(b) + (x - b)\varepsilon_2(x - b)$$

où ε_2 est une fonction de limite nulle en 0.

En remplaçant x par $f(x)$ dans la deuxième formule, on obtient :

$$\begin{aligned} g(f(x)) &= g(b) + ((x - a)f'(a) + (x - a)\varepsilon_1(x - a))g'(b) + ((x - a)f'(a) \\ &\quad + (x - a)\varepsilon_1(x - a))\varepsilon_2(f(x) - b) \end{aligned}$$

On remarque alors que, f étant continue en a , on peut écrire :

$$f(x) - b = f(x) - f(a) = \varepsilon_3(x - a)$$

où ε_3 est une fonction de limite nulle en 0.

Par suite :

$$\varepsilon_2(f(x) - b) = \varepsilon_2(\varepsilon_3(x - a))$$

où, pour alléger les écritures, et, de façon générique, on décide de désigner par $\varepsilon(.)$: $x \mapsto \varepsilon(x)$ n'importe quelle fonction de limite nulle en 0.

Finalement, en regroupant les termes qui contiennent à la fois un facteur $(x - a)$ et un facteur $\varepsilon(x - a)$, on obtient :

$$(g \circ f)(x) = g(f(x)) = g(b) + (x - a)f'(a)g'(b) + (x - a)\varepsilon(x - a)$$

ce qui montre que la fonction $g \circ f$ est dérivable en a et que :

$$(g \circ f)'(a) = f'(a) g'(f(a))$$

■

1. Conditions de dérivabilité sur un intervalle

Définition

Soit f une fonction définie sur un intervalle I de \mathbb{R} . On dit que la fonction f est dérivable sur I si f est dérivable en tout point de I . On appelle **fonction dérivée** de f la fonction f' définie sur I et qui, à chaque point a de l'intervalle I , associe $f'(a)$.

► Dérivabilité et opérations algébriques

Théorème

Soient f et g des fonctions définies sur un même intervalle I de \mathbb{R} . Alors :

- Si f et g sont dérivables sur I , les fonctions $f + g$ et fg sont dérivables sur I , et, pour tout x de I :

$$(f + g)'(x) = f'(x) + g'(x) \quad , \quad (fg)'(x) = f'(x)g(x) + f(x)g'(x)$$

- Si g ne s'annule pas sur I , et si g est dérivable sur I , la fonction $\frac{1}{g}$ est dérivable sur I .
De plus, pour tout x de I :

$$\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g(x)^2}$$

► Dérivabilité des fonctions composées

Théorème

Soient f une fonction définie sur un intervalle I de \mathbb{R} , et g une fonction définie sur un intervalle J contenant $f(I)$.

Si f est dérivable sur I , et si g est dérivable sur J , alors la fonction composée $g \circ f$ est dérivable sur I et, pour tout x de I :

$$(g \circ f)'(x) = f'(x)g'(f(x))$$

Exemple

On considère la fonction f qui, à tout réel x , associe $\cos(x^2)$. f est dérivable sur \mathbb{R} , et, pour tout réel x :

$$f'(x) = -2x \sin(x^2)$$

Le théorème de dérivabilité des fonctions composées peut être considéré comme l'un des plus importants du calcul des dérivées. Une application est de l'utiliser pour obtenir la dérivée d'un quotient de fonctions, ou, encore, pour obtenir celle d'un produit ; à cet effet, il suffit de



considérer deux fonctions f et g définies sur un même intervalle I de \mathbb{R} . Alors, compte tenu de l'identité :

$$(f + g)^2 = f^2 + 2fg + g^2$$

on obtient, à l'aide de la formule donnant la dérivée d'une fonction composée :

$$2(f + g)(f' + g') = 2ff' + f'g + fg' + 2gg'$$

ce qui conduit donc à :

$$(fg)' = f'g + fg'$$

Théorème

Toute fonction dérivable sur un intervalle I de \mathbb{R} y est continue.

► Primitive

Définition

Soit f une fonction définie sur un intervalle I de \mathbb{R} , continue sur I . On dit que F est une **primitive** de f sur I si F est dérivable sur I et si, pour tout réel x de I :

$$F'(x) = f(x)$$



Une fonction continue sur un intervalle I y admet une infinité de primitives !

Exemple

La fonction qui, à tout réel x , associe x^2 , est une primitive sur \mathbb{R} de la fonction qui, à tout réel x , associe $2x$. Mais la fonction qui, à tout réel x , associe $x^2 + 1$, est aussi une primitive sur \mathbb{R} de la fonction qui, à tout réel x , associe $2x$.

2. Dérivabilité des fonctions définies par morceaux

Pour pouvoir étudier la dérivabilité en un point de raccordement des fonctions définies par morceaux, il est nécessaire d'introduire les notions de dérivée à droite et à gauche.

► Dérivée à droite d'une fonction

Soit f une fonction définie sur un intervalle ouvert I de \mathbb{R} , et a un point de I . La fonction f est dite **dérivable à droite** en a si la fonction τ , définie, pour tout x de $I \setminus \{a\}$, par :

$$\tau(x) = \frac{f(x) - f(a)}{x - a}$$

a une limite à droite en a , c'est-à-dire lorsque x tend vers a par valeurs supérieures.

Cette limite est appelée dérivée à droite de la fonction f en a , et notée $f'_d(a)$.

► Demi-tangente à droite

Soit f une fonction définie sur un intervalle ouvert I de \mathbb{R} , et a un point de I . On suppose que la fonction f est dérivable à droite en a . La demi-droite d'équation :

$$y = f(a) + (x - a)f'_d(a) \quad \text{pour } x \geq a$$

est la **demi-tangente à droite en a** de la courbe représentative de f .

► Dérivée à gauche d'une fonction

Soit f une fonction définie sur un intervalle ouvert I de \mathbb{R} , et a un point de I . La fonction f est dite **dérivable à gauche** en a si la fonction τ , définie pour tout x de $I \setminus \{a\}$, par :

$$\tau(x) = \frac{f(x) - f(a)}{x - a}$$

a une limite à gauche en a , c'est-à-dire lorsque x tend vers a par valeurs inférieures.

Cette limite est appelée dérivée à gauche de la fonction f en a , et notée $f'_g(a)$.

► Demi-tangente à gauche

Soient f une fonction définie sur un intervalle ouvert I de \mathbb{R} , et a un point de I . On suppose que la fonction f est dérivable à gauche en a . La demi-droite d'équation

$$y = f(a) + (x - a) f'_g(a) \quad \text{pour } x \leq a$$

est la **demi-tangente à gauche** en a de la courbe représentative de f .

Propriété

Soient f une fonction définie sur un intervalle ouvert I de \mathbb{R} , et a un point de I .

La fonction f est dérivable en a si et seulement si elle est dérivable à gauche et à droite en a , et si ses dérivées à gauche et à droite en a sont égales.

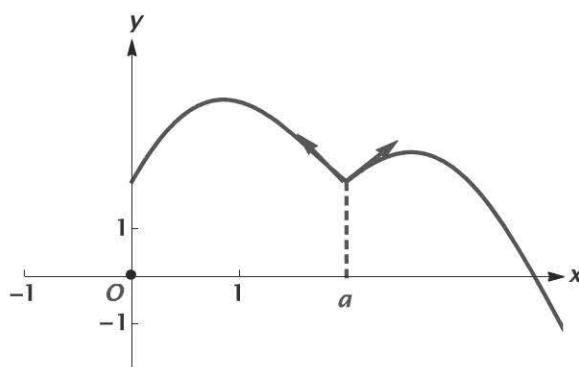


Figure 24.1 – Le graphe d'une fonction non dérivable en a : les deux dérivées à gauche et à droite ne sont pas égales. Il y a une demi-tangente à gauche, et une demi-tangente à droite. La courbe possède un point anguleux.

Exemple

La fonction valeur absolue, qui, à tout réel x , associe sa valeur absolue $|x|$, n'est pas dérivable en zéro, mais possède des dérivées à droite et à gauche en zéro.

3. Dérivabilité et variations

► Caractérisation des fonctions constantes dérivables

Théorème

Soient I un intervalle de \mathbb{R} , et f une fonction dérivable sur I . Alors, f est constante sur I si et seulement si sa dérivée f' est identiquement nulle sur I :

$$\forall x \in I : f'(x) = 0$$

► Caractérisation des fonctions croissantes dérivables

Théorème

Soient I un intervalle de \mathbb{R} , et f une fonction dérivable sur I . Alors, f est croissante sur I si et seulement si sa dérivée f' est positive ou nulle sur I :

$$\forall x \in I : f'(x) \geq 0$$

► Caractérisation des fonctions décroissantes dérivables

Théorème

Soient I un intervalle de \mathbb{R} , et f une fonction dérivable sur I . Alors, f est décroissante sur I si et seulement si sa dérivée f' est négative ou nulle sur I :

$$\forall x \in I : f'(x) \leq 0$$

► Caractérisation des fonctions strictement croissantes dérivables

Théorème

Soient I un intervalle de \mathbb{R} , et f une fonction dérivable sur I . Alors, f est strictement croissante sur I si et seulement si sa dérivée f' est positive ou nulle sur I , et ne s'annule sur aucun intervalle de I non réduit à un point.

► Caractérisation des fonctions strictement décroissantes dérivables

Théorème

Soient I un intervalle de \mathbb{R} , et f une fonction dérivable sur I . Alors, f est strictement décroissante sur I si et seulement si sa dérivée f' est négative ou nulle sur I , et ne s'annule sur aucun intervalle de I non réduit à un point.

1. Dérivée $n^{\text{ième}}$

Étant donné un entier naturel non nul n , on définit, sous réserve d'existence bien sûr (c'est-à-dire lorsque cela est possible), la dérivée $n^{\text{ième}}$ de f , notée $f^{(n)}$, par récurrence :

$$f^{(n)} = (f^{(n-1)})'$$

avec la convention :

$$f^{(0)} = f$$

► Fonction n fois dérivable, $n \in \mathbb{N}$

Si f admet une dérivée $n^{\text{ième}}$, $n \in \mathbb{N}$, on dit que f est n fois dérivable sur I .

► Fonction indéfiniment dérivable

Si, pour tout entier naturel n , f admet une dérivée $n^{\text{ième}}$ on dit que f est indéfiniment dérivable sur I .

Exemples

1. Les fonctions polynômes, les fonctions trigonométriques *sinus* et *cosinus*, la fonction exponentielle $x \mapsto e^x$, sont indéfiniment dérivables sur \mathbb{R} .
2. Les fractions rationnelles sont indéfiniment dérivables sur tout intervalle qui ne contient pas de racine du dénominateur.
3. La fonction *logarithme népérien* $x \mapsto \ln x$ et la fonction racine carrée $x \mapsto \sqrt{x}$ sont indéfiniment dérivables sur $]0, \infty[$.

2. Classes de fonction

► Fonction de classe C^n , $n \in \mathbb{N}$

Soient n un entier naturel non nul, et f une fonction définie sur un intervalle I de \mathbb{R} .

On dit que f est de classe C^n sur I si f est n fois dérivable sur I , et si sa dérivée $n^{\text{ième}}$, $f^{(n)}$, est continue sur I .

► Fonction de classe C^∞

Soit f une fonction définie sur un intervalle I de \mathbb{R} .

On dit que f est de classe C^∞ sur I si, pour tout entier naturel n , f est de classe C^n sur I .

3. Théorèmes

► Dérivée $n^{\text{ième}}$ d'une combinaison linéaire de fonctions, $n \in \mathbb{N}$

Théorème

Soient f et g des fonctions définies sur un même intervalle I de \mathbb{R} , et n un entier naturel. On suppose que f et g sont n fois dérivables sur I . Alors, toute combinaison linéaire de

f et g est n fois dérivable sur I ; pour tout couple (α, β) de réels, $\alpha f + \beta g$ est n fois dérivable sur I et : $(\alpha f + \beta g)^{(n)} = \alpha f^{(n)} + \beta g^{(n)}$

► Formule de Leibniz¹

Théorème

Soient f et g des fonctions définies sur un même intervalle I de \mathbb{R} , et n un entier naturel. On suppose que f et g sont n fois dérivables sur I . Alors, la fonction produit $f g$ est n

fois dérivable sur I , et : $(f g)^{(n)} = \sum_{k=0}^n C_n^k f^{(k)} g^{(n-k)}$

où, pour tout entier k de $\{0, \dots, n\}$, C_n^k désigne le coefficient binomial : $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Par convention, ce coefficient est nul si $k > n$.

Démonstration : On ne donne, ici, que des éléments de preuve. Ce résultat se démontre par récurrence, à l'aide de la formule donnée par le triangle de Pascal ; étant donné un entier naturel non nul n , alors, pour tout entier naturel $k \leq n-1$:

$$C_n^k + C_n^{k+1} = C_{n+1}^{k+1} \quad \text{soit :} \quad \binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1} \quad \blacksquare$$

► Dérivée $n^{\text{ième}}$ d'un quotient de fonctions, $n \in \mathbb{N}$

Théorème

Soient f et g des fonctions définies sur un même intervalle I de \mathbb{R} , et n un entier naturel. On suppose que f et g sont n fois dérivables sur I , et que g ne s'annule pas sur I . Alors, le quotient $\frac{f}{g}$ est n fois dérivable sur I .

► Dérivée $n^{\text{ième}}$ de la composée de deux fonctions, $n \in \mathbb{N}$

Théorème

Soit f une fonction définie sur un intervalle I de \mathbb{R} , et g une fonction définie sur un intervalle J contenant $f(I)$.

Si f est n fois dérivable sur I , et g est n fois dérivable sur J , alors la fonction composée $g \circ f$ est n fois dérivable sur I .



Il existe une formule permettant de calculer la dérivée $n^{\text{ième}}$ d'une fonction composée. Cette formule est due à Faà di Bruno² :

$$(g \circ f)^{(n)} = n! \sum_{1 \leq k \leq n} \frac{g^{(k)} \circ f}{k!} \prod_{n_1 + \dots + n_k = n, n_j \geq 1} \frac{f^{(n_j)}}{n_j!}$$

Le lecteur intéressé pourra trouver plus de précisions dans [4], page 165.

1. Gottfried Wilhelm Leibniz (1646-1716), philosophe, mathématicien, juriste et philologue allemand. Il apporta des contributions fondamentales en calcul différentiel et intégral. Il fut, également, l'inventeur d'une des premières machines à calculer. Ses œuvres philosophiques sont, elles aussi, de première importance.

2. Francesco Faà di Bruno (1825-1888), prêtre et religieux italien, en même temps que mathématicien renommé, et musicien de talent. Il fut l'élève d'Augustin Cauchy à la Sorbonne. Il a été béatifié en 1988.

Une fonction dérivable, présentant un extremum en un point de son domaine de définition, possède, en ce point, une dérivée nulle. En pratique, cette remarque permet de trouver l'image d'un intervalle fermé borné par une fonction dérivable. Plus fondamentalement, elle est le point clef de la démonstration du théorème de Rolle et de ses conséquences : le théorème des accroissements finis et la formule de Taylor-Lagrange.^{1 2}

1. Extremum sur un intervalle

► Maximum d'une fonction sur un intervalle

Soient f une fonction définie sur un intervalle I de \mathbb{R} , et c un point de I . La fonction f admet un **maximum** en c si, pour tout x de I :

$$f(x) \leq f(c)$$

► Minimum d'une fonction sur un intervalle

Soient f une fonction définie sur un intervalle I de \mathbb{R} , et c un point de I . La fonction f admet un **minimum** en c si, pour tout x de I :

$$f(x) \geq f(c)$$

► Extremum d'une fonction sur un intervalle

Soient f une fonction définie sur un intervalle I de \mathbb{R} , et c un point de I . La fonction f admet un **extremum** en c si elle admet en c un maximum ou un minimum.

2. Extremum local

► Maximum local d'une fonction

Soient f une fonction définie sur un intervalle I de \mathbb{R} , et c un point de I . La fonction f admet un **maximum local** en c s'il existe un intervalle ouvert de la forme $]c - \eta, c + \eta[\subset I$, $0 < \eta < c$, tel que la restriction de f à cet intervalle admette un maximum en c :

$$\forall x \in]c - \eta, c + \eta[: f(x) \leq f(c)$$

► Minimum local d'une fonction

Soient f une fonction définie sur un intervalle I de \mathbb{R} , et c un point de I . La fonction f admet un **minimum local** en c s'il existe un intervalle ouvert de la forme $]c - \eta, c + \eta[\subset I$, $0 < \eta < c$, tel que la restriction de f à cet intervalle admette un minimum en c :

$$\forall x \in]c - \eta, c + \eta[: f(x) \geq f(c)$$

1. Brook Taylor (1685-1731), mathématicien, historien des sciences, musicien et peintre anglais. C'est lui qui découvrit l'intégration par parties, et est, bien sûr, à l'origine des « développements de Taylor ».

2. Joseph Louis, comte de Lagrange (1736-1813), mathématicien, mécanicien et astronome italien. Il fut l'initiateur du calcul variationnel. En parallèle, il apporta de nombreuses contributions en algèbre, à la théorie des nombres, au calcul infinitésimal, aux probabilités, mais aussi à la mécanique.

► Extremum local d'une fonction

Soient f une fonction définie sur un intervalle I de \mathbb{R} , et c un point de I . La fonction f admet un **extremum local** en c si elle admet en c un minimum local, ou un maximum local.

► Dérivée et extremum local

Soient f une fonction définie sur un intervalle I de \mathbb{R} , et c un point intérieur à I (c 'est-à-dire qui ne soit pas une borne de I). Si f est dérivable en c et admet en c un extremum local, alors :

$$f'(c) = 0$$

► Interprétation graphique

Une fonction admettant des extremums locaux en certains points intérieurs aura donc, en ces points, des tangentes horizontales.

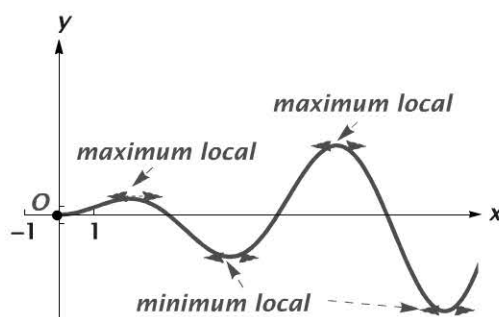


Figure 26.1 – Le graphe d'une fonction avec des extremums locaux.

Démonstration : Le fait que f possède une dérivée nulle au point c est une *propriété locale* (elle ne dépend que des valeurs de f au voisinage du point c).

Quitte à considérer la restriction de f à un sous-intervalle ouvert de I , de la forme $]c - \eta, c + \eta[$, $0 < \eta < c$, on peut supposer que f possède un extremum en c .

Supposons, dans un premier temps, que f possède un maximum en c . Alors, pour tout x de I tel que $x < c$: $f(x) \leq f(c)$. Il en résulte :

$$\frac{f(x) - f(c)}{x - c} \geq 0$$

et donc :

$$f'(c) = \lim_{x \rightarrow c^-} \frac{f(x) - f(c)}{x - c} \geq 0$$

De même, pour tout x de I tel que $x > c$: $f(x) \leq f(c)$. Il en résulte :

$$\frac{f(x) - f(c)}{x - c} \leq 0$$

et donc :

$$f'(c) = \lim_{x \rightarrow c^+} \frac{f(x) - f(c)}{x - c} \leq 0$$

D'où, nécessairement : $f'(c) = 0$.

Le cas où f possède un minimum en c se traite de façon analogue (il suffit de changer f en $-f$). ■



1. On voit bien dans cette démonstration l'importance de supposer que c soit intérieur à I ; ainsi, le point c n'est pas une extrémité de l'intervalle I et les limites à gauche et à droite utilisées dans la démonstration ont un sens.

Si on considère la restriction de l'application identité à l'intervalle $[0, 1]$, cette restriction présente un extremum en 1, alors que la dérivée de la fonction ne s'annule pas en 1.

2. La réciproque de la proposition précédente est fausse : une fonction dérivable sur un intervalle peut avoir une dérivée nulle en un point qui n'est pas un extremum local. C'est le cas, par exemple, de la fonction définie sur \mathbb{R} par $x \mapsto x^3$, au point 0.

► Une application pratique

Soit f une fonction définie et continue sur un intervalle fermé et borné $[a, b]$. Si cette fonction est dérivable sur l'intervalle ouvert $]a, b[$ alors son maximum M et son minimum m sont à rechercher parmi les valeurs de la fonction f en a , ou en b , ou aux points c où sa dérivée f' s'annule.

3. Le théorème de Rolle³

Théorème

Soient a et b deux réels tels que $a < b$, et f une fonction définie et continue sur l'intervalle $[a, b]$, dérivable sur l'intervalle $]a, b[$. Alors, si $f(a) = f(b)$, il existe un nombre c de l'intervalle $]a, b[$ tel que :

$$f'(c) = 0$$

► Interprétation graphique du théorème de Rolle

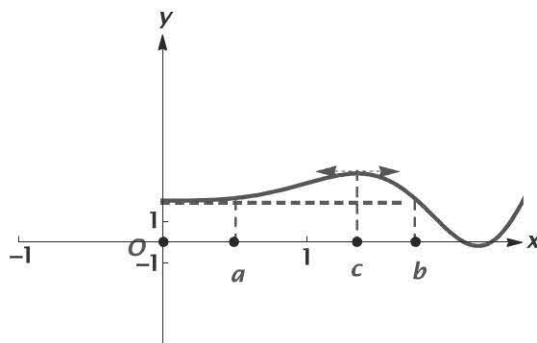


Figure 26.2 – Illustration graphique du théorème de Rolle.

Démonstration : Il faut montrer que la fonction f admet un extremum en un point de l'intervalle $]a, b[$.

Comme f est continue sur l'intervalle fermé borné $[a, b]$, l'image $f([a, b])$ de $[a, b]$, est un intervalle fermé borné, de la forme $[m, M]$, $(m, M) \in \mathbb{R}^2$, $m \leq M$.

On distingue deux cas :

- si la fonction f est constante sur l'intervalle $[a, b]$, alors sa dérivée f' est identiquement nulle sur cet intervalle. N'importe quel point c de l'intervalle $]a, b[$ convient, puisqu'il vérifie $f'(c) = 0$.

3. Ce théorème doit son nom au mathématicien français Michel Rolle (1652-1719), qui fut le premier à établir ce résultat, dans le cas de fonctions polynomiales. Ses contributions portent et sur l'algèbre, et sur l'analyse. Il est à l'origine de la notation $\sqrt[n]{x}$, pour désigner la racine $n^{\text{ème}}$ d'un réel positif x .

- Si la fonction f n'est pas constante, l'intervalle $[m, M]$ n'est pas réduit à un point. Donc l'une, au moins, des deux inégalités strictes : $M > f(a) = f(b)$, ou $m < f(a) = f(b)$ est vérifiée.

On supposera, dans ce qui suit : $M > f(a)$ (le cas $m < f(a)$ se traite de manière analogue).

Comme M est un point de $[m, M]$, il existe un point c de l'intervalle $[a, b]$ tel que $f(c) = M$. Comme $M > f(a)$ et $M > f(b)$, c appartient en fait à l'intervalle $]a, b[$. La fonction f admet en c un maximum, étant dérivable sur $]a, b[$, donc en c , elle vérifie donc : $f'(c) = 0$. ■

4. Le théorème des accroissements finis

Théorème

Soient a et b deux réels tels que $a < b$, et f une fonction définie et continue sur l'intervalle $[a, b]$, dérivable sur l'intervalle $]a, b[$. Alors, il existe un nombre c de l'intervalle $]a, b[$ tel que :

$$f(b) - f(a) = (b - a) f'(c)$$

► Interprétation graphique du théorème des accroissements finis

Le théorème des accroissements finis traduit, tout simplement, le fait que la courbe représentative de f possède, en c , une tangente parallèle à la corde joignant les points $(a, f(a))$ et $(b, f(b))$:

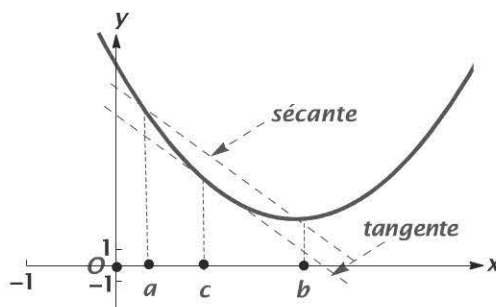


Figure 26.3 – Une illustration graphique du théorème des accroissements finis.

Démonstration : On considère la fonction φ définie sur l'intervalle $[a, b]$ par :

$$\varphi(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a} (x - a)$$

Elle est, comme la fonction f , continue sur $[a, b]$ et dérivable sur $]a, b[$. Par ailleurs elle vérifie : $\varphi(a) = \varphi(b)$. Elle satisfait donc aux hypothèses du théorème de Rolle. Ainsi, il existe c dans l'intervalle $]a, b[$ tel que :

$$0 = \varphi'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}$$

ce qui achève la démonstration. ■



Le théorème des accroissements permet de relier les variations d'une fonction au signe de sa dérivée.

On présente, dans ce qui suit, la formule de Taylor-Lagrange, qui généralise la formule des accroissements finis pour les fonctions plusieurs fois dérivables.

Théorème

Soient a et b deux réels tels que $a < b$, et f une fonction définie sur l'intervalle $[a, b]$. On suppose que :

- la fonction f est n fois dérivable sur l'intervalle $[a, b]$;
- la dérivée $n^{\text{ième}}$ de f , $f^{(n)}$, est continue sur l'intervalle fermé $[a, b]$ et dérivable sur l'intervalle ouvert $]a, b[$.

Alors, il existe un réel c , appartenant à l'intervalle ouvert $]a, b[$, tel que :

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(a)}{2}(b-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(b-a)^n + \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}$$

Démonstration : Soit A le réel défini par :

$$\frac{A}{(n+1)!}(b-a)^{n+1} = f(b) - f(a) - (b-a)f'(a) - \frac{f''(a)}{2}(b-a)^2 - \dots - \frac{f^{(n)}(a)}{n!}(b-a)^n$$

Il s'agit de montrer qu'il existe un réel c , appartenant à l'intervalle ouvert $]a, b[$, tel que :

$$A = f^{(n+1)}(c)$$

On introduit, à cet effet, la fonction φ définie, pour tout x de $[a, b]$, par :

$$\varphi(x) = f(b) - f(x) - f'(x)(b-x) - \frac{f''(x)}{2}(b-x)^2 - \dots - \frac{f^{(n)}(x)}{n!}(b-x)^n - \frac{A}{(n+1)!}(b-x)^{n+1}$$

Il est évident que :

$$\varphi(b) = 0$$

Le choix du nombre A conduit à :

$$\varphi(a) = 0$$

Par ailleurs, la fonction φ est, comme la fonction f et chacune de ses n premières dérivées, continue sur l'intervalle fermé $[a, b]$, et dérivable sur l'intervalle ouvert $]a, b[$.

La fonction φ satisfait donc les hypothèses du théorème de Rolle.

Il existe donc un nombre c , appartenant à l'intervalle ouvert $]a, b[$, tel que :

$$\varphi'(c) = 0$$

Un calcul facile montre que :

$$\varphi'(x) = \frac{A - f^{(n+1)}(x)}{n!}(b-x)^n$$

Comme le nombre c est différent de b , on peut en déduire que :

$$A = f^{(n+1)}(c)$$

On s'intéresse, dans ce qui suit, aux conditions d'existence d'une fonction réciproque pour une fonction définie et continue sur un intervalle, et, sous réserve d'existence, à ses propriétés : continuité, dérivabilité, représentation graphique.

1. Définitions

► Injectivité

Étant données deux parties \mathcal{P}_1 et \mathcal{P}_2 de \mathbb{R} , une application $f : \mathcal{P}_1 \rightarrow \mathcal{P}_2$ est dite **injective** si, pour tout couple $(x, x') \in \mathcal{P}_1 \times \mathcal{P}_1$:

$$f(x) = f(x') \Rightarrow x = x'$$

► Surjectivité

Étant données deux parties \mathcal{P}_1 et \mathcal{P}_2 de \mathbb{R} , une application $f : \mathcal{P}_1 \rightarrow \mathcal{P}_2$ est dite **surjective** si tout élément y de \mathcal{P}_2 admet au moins un antécédent par f :

$$\forall y \in \mathcal{P}_2, \exists x \in \mathcal{P}_1 : y = f(x)$$

► Bijectivité

Étant données deux parties \mathcal{P}_1 et \mathcal{P}_2 de \mathbb{R} , une application $f : \mathcal{P}_1 \rightarrow \mathcal{P}_2$ est dite **bijective** si elle est injective et surjective, c'est-à-dire si tout élément y de \mathcal{P}_2 admet un unique antécédent par f :

$$\forall y \in \mathcal{P}_2, \exists ! x \in \mathcal{P}_1 : y = f(x)$$

► Application identité

Étant donnée une partie \mathcal{P} de \mathbb{R} , l'application

$$\begin{aligned} id_{\mathcal{P}} : \mathcal{P} &\rightarrow \mathcal{P} \\ x &\mapsto x \end{aligned}$$

est appelée **application identité** de \mathcal{P} .

► Fonction réciproque

Soient \mathcal{P}_1 et \mathcal{P}_2 deux parties de \mathbb{R} , et $f : \mathcal{P}_1 \rightarrow \mathcal{P}_2$ une application bijective. Pour chaque élément y de \mathcal{P}_2 , l'unique élément x de \mathcal{P}_1 tel que $f(x) = y$ est noté $f^{-1}(y)$.

L'application $f^{-1} : \mathcal{P}_2 \rightarrow \mathcal{P}_1$ ainsi définie est appelée **application réciproque** de f .

Soient \mathcal{P}_1 et \mathcal{P}_2 deux parties de \mathbb{R} , et $f : \mathcal{P}_1 \rightarrow \mathcal{P}_2$ une application bijective. Alors, pour tout réel x de \mathcal{P}_1 :

$$(f^{-1} \circ f)(x) = f^{-1}(f(x)) = id_{\mathcal{P}_1}(x) = x$$

et, pour tout réel y de \mathcal{P}_2 :

$$(f \circ f^{-1})(y) = f(f^{-1}(y)) = id_{\mathcal{P}_2}(y) = y$$

Il s'agit, ensuite, d'appliquer ces définitions générales au cas particulier des fonctions définies sur un intervalle I , non vide et non réduit à un point.



1. Une fonction f définie sur un intervalle I est, par définition, une surjection de I sur $f(I)$. Par suite, la fonction f admet une application réciproque, définie sur $f(I)$, si et seulement si elle est bijective de I sur $f(I)$, et donc si et seulement si elle est injective sur I .
2. Pour s'assurer que la fonction réciproque d'une fonction donnée f est définie sur un intervalle J , il suffit de supposer que la fonction f est continue sur l'intervalle I car, alors, d'après le théorème des valeurs intermédiaires, on sait que $J = f(I)$ est un intervalle.

2. Injectivité et monotonie

Soit f une fonction définie sur un intervalle I de \mathbb{R} , non vide, et non réduit à un point. Alors :

- Si f est strictement monotone sur I , alors elle est injective sur I .
- Si f est continue et injective sur I , alors elle est strictement monotone sur I .

Démonstration : Ce théorème est admis. ■



On sait déjà que l'image d'un intervalle I par une fonction continue f est un intervalle ; cependant lorsque, de plus, la fonction f est strictement monotone, on peut préciser l'intervalle $f(I)$.

► Image d'un intervalle par une fonction continue strictement monotone

Soit f une fonction continue et strictement monotone sur un intervalle I , de bornes a et b . Alors, l'intervalle $f(I)$ a pour bornes $\lim_{x \rightarrow a} f(x)$ et $\lim_{x \rightarrow b} f(x)$ (ces limites pouvant être elles-mêmes finies ou infinies) et les intervalles I et $f(I)$ sont de même nature : fermés, ouverts, ou semi-ouverts.

3. Théorème des fonctions réciproques

Théorème

Soit f une fonction continue strictement monotone sur un intervalle I . Alors :

- L'ensemble $J = f(I)$ est un intervalle (de même nature que I), et dont les extrémités sont les limites de f aux bornes de I .
- La fonction f admet une fonction réciproque f^{-1} définie sur J .
- La fonction réciproque f^{-1} est continue et strictement monotone sur J , de même sens de monotonie que f .
- Si la fonction f est dérivable en un point a de I et si $f'(a) \neq 0$, la fonction f^{-1} est dérivable au point $b = f(a)$ et :

$$(f^{-1})'(b) = \frac{1}{f'(a)}$$

De façon équivalente :

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}$$

Démonstration : les deux premiers points ont été vus dans les propositions précédentes. Le troisième point, c'est-à-dire la continuité de f^{-1} , est admis. On ne démontrera ici que la dérivabilité de la fonction réciproque.

À cet effet, on suppose que la fonction f est dérivable en un point a de l'intervalle I , et que $f'(a) \neq 0$.

Montrer que la fonction f^{-1} est dérivable au point $b = f(a)$ revient à montrer que le rapport $\frac{f^{-1}(y) - f^{-1}(b)}{y - b}$ a une limite lorsque y tend vers b , en restant bien sûr dans l'intervalle J .

Pour tout y de $J = f(I)$, le nombre $x = f^{-1}(y)$ de I vérifie la condition $y = f(x)$. Il en résulte :

$$\frac{f^{-1}(y) - f^{-1}(b)}{y - b} = \frac{x - a}{f(x) - f(a)}.$$

Comme la fonction f^{-1} est continue en b , le nombre $x = f^{-1}(y)$ tend vers $a = f^{-1}(b)$ lorsque y tend vers b . Le rapport $\frac{x - a}{f(x) - f(a)}$ a donc une limite, puisque la fonction f est dérivable en a et que sa dérivée $f'(a)$ est non nulle. On obtient :

$$(f^{-1})'(b) = \frac{1}{f'(a)}$$

■

4. Représentation graphique

Soit f une fonction continue et strictement monotone sur un intervalle $I \subset \mathbb{R}$.

On note $J = f(I)$. Par suite :

$$(x \in I \text{ et } f(x) = y) \Leftrightarrow (y \in J \text{ et } f^{-1}(y) = x)$$

Ainsi, le graphe de la fonction réciproque f^{-1} , c'est-à-dire l'ensemble des couples $(y, f^{-1}(y))$ lorsque y parcourt l'intervalle J , est donc aussi l'ensemble des couples $(f(x), x)$ lorsque x parcourt l'intervalle I . Or, dans un repère orthonormé, le point de coordonnées (a, b) , $(a, b) \in \mathbb{R}^2$, est le symétrique par rapport à la première bissectrice du point de coordonnées (b, a) .

► Graphe d'une fonction réciproque

Soit f une fonction continue, réalisant une bijection d'un intervalle I de \mathbb{R} sur un intervalle J de \mathbb{R} . Alors, le graphe de f^{-1} est symétrique de celui de f par rapport à la première bissectrice (qui est aussi la droite d'équation $y = x$).

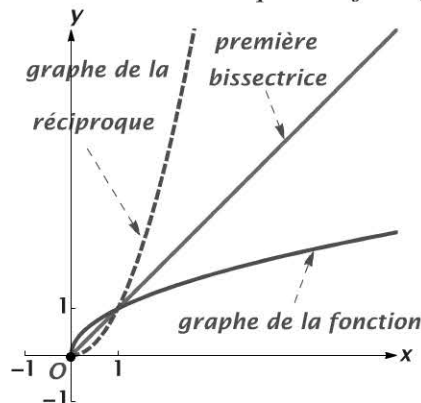


Figure 28.1 – Le graphe d'une fonction bijective, et celui de sa réciproque.

1. La fonction *arcsinus*

La restriction $\sin|_{[-\frac{\pi}{2}, \frac{\pi}{2}]}$ de la fonction sinus à l'intervalle $[-\frac{\pi}{2}, \frac{\pi}{2}]$ est continue et strictement croissante. D'après le théorème des fonctions réciproques :

$$\sin|_{[-\frac{\pi}{2}, \frac{\pi}{2}]} \left(\left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \right) = \left[\sin\left(-\frac{\pi}{2}\right), \sin\left(\frac{\pi}{2}\right) \right] = [-1, 1]$$

Cette restriction établit une bijection de $[-\frac{\pi}{2}, \frac{\pi}{2}]$ sur $[-1, 1]$.

La fonction réciproque de cette restriction est appelée *fonction arcsinus*, et notée \arcsin . C'est une bijection de $[-1, 1]$ sur l'intervalle $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Pour tout réel x de $[-1, 1]$, $\arcsin x$ est donc l'unique élément de l'intervalle $[-\frac{\pi}{2}, \frac{\pi}{2}]$ qui a pour sinus le réel x :

$$\sin(\arcsin x) = x$$

De même, pour tout réel x de $[-\frac{\pi}{2}, \frac{\pi}{2}]$:

$$\arcsin(\sin x) = x$$

Par définition :

$$\arcsin(0) = 0 \quad , \quad \arcsin(1) = \frac{\pi}{2} \quad , \quad \arcsin\left(\frac{1}{2}\right) = \frac{\pi}{6}$$

Les propriétés de la fonction *arcsinus* sont les suivantes :

- La fonction *arcsinus* est continue et strictement croissante sur $[-1, 1]$. C'est une conséquence directe du théorème des fonctions réciproques.
- La fonction *arcsinus* est impaire. Dans le plan rapporté à un repère orthonormé direct, la courbe représentative de la fonction *arcsinus* est la courbe symétrique par rapport à la première bissectrice de la courbe représentative de la restriction de la fonction sinus à l'intervalle $[-\frac{\pi}{2}, \frac{\pi}{2}]$.
- La restriction $\sin|_{[-\frac{\pi}{2}, \frac{\pi}{2}]}$ est dérivable sur $[-\frac{\pi}{2}, \frac{\pi}{2}]$, de dérivée

$$x \mapsto \cos x = \sqrt{1 - \sin^2(x)}.$$

Cette dérivée ne s'annulant pas sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$, la fonction *arcsinus* est donc dérivable sur $] -1, 1[$, sa dérivée étant définie, pour tout x de $] -1, 1[$, par :

$$(\arcsin)'(x) = \frac{1}{\sqrt{1 - \sin^2(\arcsin x)}} = \frac{1}{\sqrt{1 - x^2}}$$

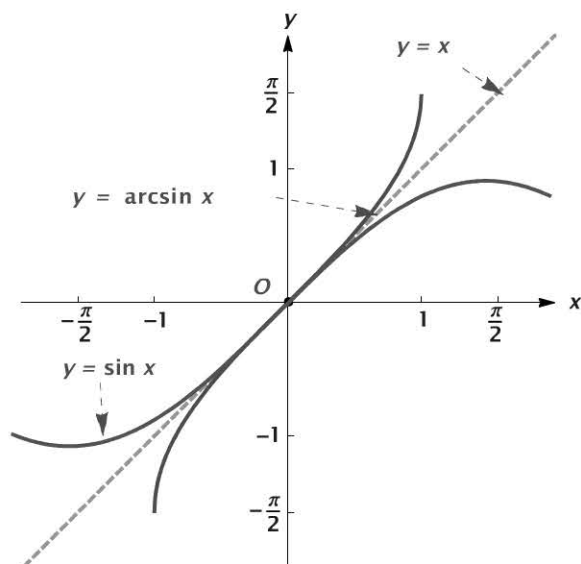


Figure 29.1 – La courbe représentative de la fonction *arcsinus*.

2. La fonction *arccosinus*

La restriction $\cos|_{[0,\pi]}$ de la fonction *cosinus* à l'intervalle $[0, \pi]$ est continue et strictement décroissante. D'après le théorème des fonctions réciproques, elle établit une bijection de $[0, \pi]$ sur $[-1, 1]$. La fonction réciproque de cette restriction est appelée *arccosinus*, et notée « arccos ». C'est une bijection de $[-1, 1]$ sur l'intervalle $[0, \pi]$.

Pour tout réel x de $[-1, 1]$, $\arccos x$ est donc l'unique élément de l'intervalle $[0, \pi]$ qui a pour cosinus le réel x :

$$\cos(\arccos x) = x$$

De même, pour tout réel x de $[0, \pi]$:

$$\arccos(\cos x) = x$$

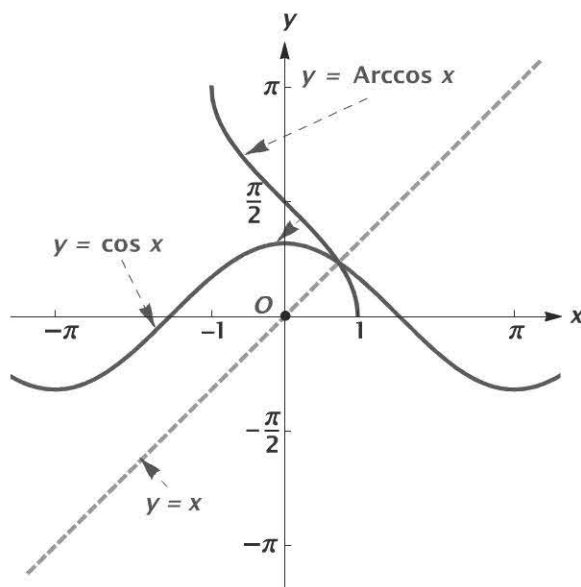
Par définition :

$$\arccos(0) = \frac{\pi}{2}, \quad \arccos(1) = 0, \quad \arccos(-1) = \pi$$

Les propriétés de la fonction *arccosinus* sont les suivantes :

- La fonction *arccosinus* est continue et strictement décroissante sur $[-1, 1]$.
- La restriction $\cos|_{[0,\pi]}$ est dérivable sur $[0, \pi]$, de dérivée $x \mapsto -\sin x = -\sqrt{1 - \cos^2 x}$, qui ne s'annule pas sur $]0, \pi[$. La fonction *arccosinus* est donc dérivable sur $] -1, 1[$, sa dérivée étant définie, pour tout x de $] -1, 1[$, par :

$$(\arccos)'(x) = -\frac{1}{\sqrt{1 - \cos^2(\arccos x)}} = -\frac{1}{\sqrt{1 - x^2}}$$

Figure 29.2 – La courbe représentative de la fonction *arccosinus*.

3. La fonction *arctangente*

La restriction $\tan|_{]-\frac{\pi}{2}, \frac{\pi}{2}[}$ de la fonction *tangente* à l'intervalle $]-\frac{\pi}{2}, \frac{\pi}{2}[$ est continue et strictement croissante sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$.

D'après le théorème « des fonctions réciproques » :

$$\tan|_{]-\frac{\pi}{2}, \frac{\pi}{2}[} \left(]-\frac{\pi}{2}, \frac{\pi}{2}[\right) = \left[\lim_{x \rightarrow -\frac{\pi}{2}^-} \tan x, \lim_{x \rightarrow \frac{\pi}{2}^+} \tan x \right] = \mathbb{R}$$

En outre, cette restriction établit une bijection de $]-\frac{\pi}{2}, \frac{\pi}{2}[$ sur \mathbb{R} . La fonction réciproque de la restriction de la fonction *tangente* à l'intervalle $]-\frac{\pi}{2}, \frac{\pi}{2}[$ est appelée *fonction arctangente*, et notée « arctan ». C'est une bijection de \mathbb{R} sur l'intervalle $]-\frac{\pi}{2}, \frac{\pi}{2}[$.

Pour tout réel x , $\arctan x$ est donc l'unique élément de l'intervalle $]-\frac{\pi}{2}, \frac{\pi}{2}[$ qui a pour tangente le réel x :

$$\tan(\arctan x) = x$$

Par définition : $\arctan(1) = \frac{\pi}{4}$.



Attention ! Il y a une infinité de réels dont la tangente est égale à 1. Mais parmi ces réels, seul $\frac{\pi}{4}$ appartient à l'intervalle $]-\frac{\pi}{2}, \frac{\pi}{2}[$.

Les propriétés de la fonction *arctangente* sont les suivantes :

- La fonction *arctangente* est continue et strictement croissante sur \mathbb{R} . C'est une conséquence directe du théorème des fonctions réciproques. De plus :

$$\lim_{x \rightarrow -\infty} \arctan x = -\frac{\pi}{2} \quad , \quad \lim_{x \rightarrow +\infty} \arctan x = \frac{\pi}{2}$$

- La fonction *arctangente* est impaire. Dans le plan rapporté à un repère orthonormé direct, la courbe représentative de la fonction *arctangente* est la courbe obtenue par symétrie par rapport à la première bissectrice de la courbe représentative de la restriction de la fonction tangente à l'intervalle $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$.
- On rappelle que la fonction tangente est dérivable sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, de dérivée $x \mapsto 1 + \tan^2 x$. Cette dérivée n'est jamais nulle. Par suite, la fonction *arctangente* est donc dérivable sur \mathbb{R} , sa dérivée étant donnée par :

$$(\arctan x)' = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}$$

- Pour tout réel strictement positif x :

$$\arctan x + \arctan\left(\frac{1}{x}\right) = \frac{\pi}{2}$$

et, pour tout réel strictement négatif x :

$$\arctan x + \arctan\left(\frac{1}{x}\right) = -\frac{\pi}{2}$$

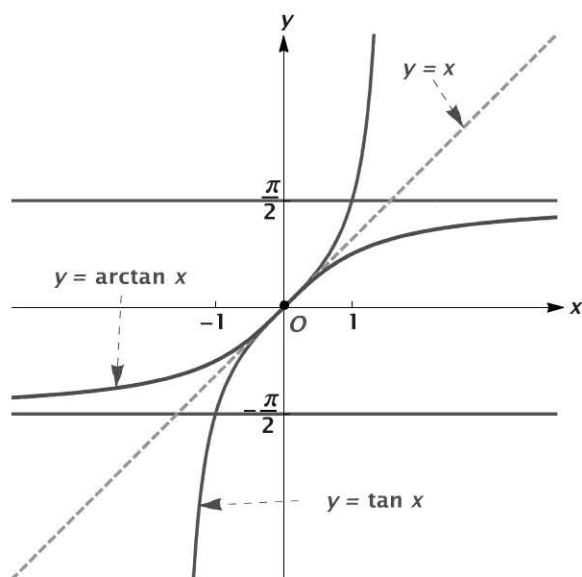


Figure 29.3 – La courbe représentative de la fonction *arctangente*.

1. La fonction argument sinus hyperbolique

La fonction sinus hyperbolique est bijective de \mathbb{R} sur \mathbb{R} . Sa fonction réciproque est appelée **argument sinus hyperbolique**, notée Argsh , et vérifie, pour tout réel x :

$$\text{Argsh } x = \ln \left(x + \sqrt{x^2 + 1} \right)$$

Les propriétés de la fonction *argument sinus hyperbolique* sont les suivantes :

- La fonction *argument sinus hyperbolique* est continue, dérivable et strictement croissante sur \mathbb{R} . Sa dérivée est définie, pour tout réel x , par :

$$\text{Argsh}'(x) = \frac{1}{\sqrt{x^2 + 1}}$$

- La fonction *argument sinus hyperbolique* est impaire.

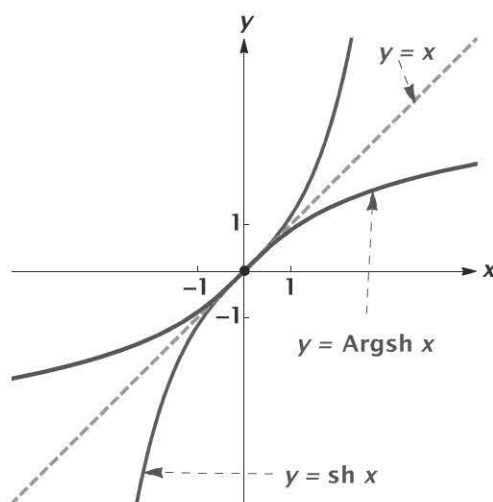


Figure 30.1 – La courbe représentative de la fonction *argument sinus hyperbolique*.

2. La fonction argument cosinus hyperbolique

La fonction cosinus hyperbolique est bijective de \mathbb{R}^+ sur $[1, +\infty[$. Sa fonction réciproque est appelée **argument cosinus hyperbolique**, notée Argch , et telle que, pour tout x de $[1, +\infty[$:

$$\text{Argch } x = \ln \left(x + \sqrt{x^2 - 1} \right)$$

La fonction *argument cosinus hyperbolique* est continue sur $[1, +\infty[$, dérivable sur $]1, +\infty[$, et strictement croissante sur $]1, +\infty[$. Sa dérivée est définie, pour tout x de $]1, +\infty[$, par :

$$\text{Argch}'(x) = \frac{1}{\sqrt{x^2 - 1}}$$

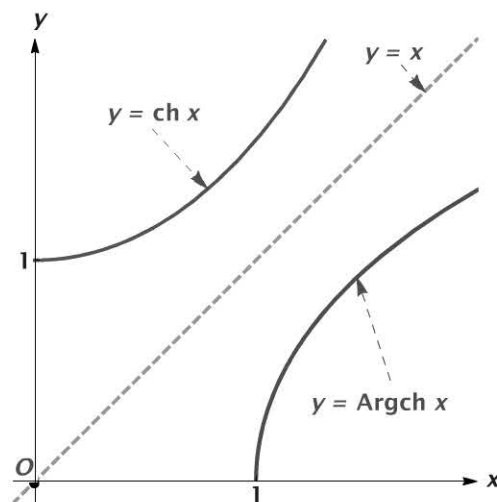


Figure 30.2 – La courbe représentative de la fonction *argument cosinus hyperbolique*.

3. La fonction argument tangente hyperbolique

La fonction tangente hyperbolique est bijective de \mathbb{R} sur $] -1, 1[$. Sa fonction réciproque est appelée **argument tangente hyperbolique**, notée Argth , et telle que, pour tout x de $] -1, 1[$:

$$\text{Argth } x = \frac{1}{2} \ln \left(\frac{1+x}{1-x} \right)$$

Les propriétés de la fonction *argument tangente hyperbolique* sont les suivantes :

- La fonction *argument tangente hyperbolique* est continue, dérivable et strictement croissante sur $] -1, 1[$. Sa dérivée est définie, pour tout x de $] -1, 1[$, par :

$$\text{Argth}'(x) = \frac{1}{1-x^2}$$

- La fonction *argument tangente hyperbolique* est impaire.

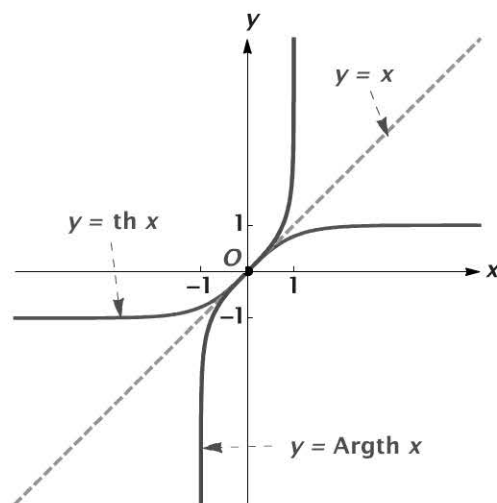


Figure 30.3 – La courbe représentative de la fonction *argument tangente hyperbolique*.

On introduit ici un nouvel outil permettant l'étude locale des fonctions : les développements limités, qui permettent notamment de déterminer la limite en un point d'une fonction donnée sous une forme indéterminée (par exemple, le quotient de deux fonctions ayant toutes les deux une limite nulle). Les développements limités sont aussi d'une grande utilité pour montrer qu'une fonction est dérivable en un point, trouver l'équation de la tangente à son graphe en ce point, et préciser la position relative du graphe et de sa tangente.

1. Développement limité au voisinage d'un point

Définition

Soient f une fonction définie sur un intervalle I de \mathbb{R} , a un point de I , et n un entier naturel non nul. On dit que f admet un **développement limité d'ordre n au voisinage de a** , s'il existe des nombres réels c_0, c_1, \dots, c_n et une fonction ε de limite nulle en 0 tels que, pour tout réel x de I :

$$f(x) = c_0 + c_1(x-a) + \dots + c_n(x-a)^n + (x-a)^n \varepsilon(x-a)$$

ce qui s'écrit aussi, avec la notation « petit o » :

$$f(x) = c_0 + c_1(x-a) + \dots + c_n(x-a)^n + o((x-a)^n)$$



Un développement limité d'une fonction au voisinage d'un point consiste donc à approcher cette fonction par un polynôme.

Exemple

Au voisinage de zéro, la fonction *sinus* peut ainsi être approchée par les fonctions polynomiales $x \mapsto x$, $x \mapsto x - \frac{x^3}{3!}$, et $x \mapsto x - \frac{x^3}{3!} + \frac{x^5}{5!}$:

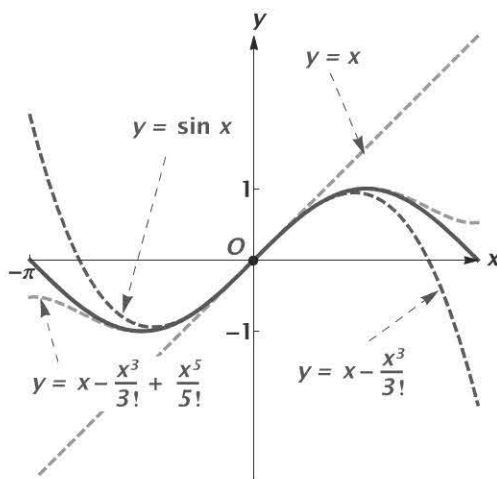


Figure 31.1 – Le graphe de la fonction *sinus* et de trois de ses approximations polynomiales au voisinage de 0.

2. Partie principale d'ordre $n \in \mathbb{N}$ d'un développement limité

Définition

Soit f une fonction définie sur un intervalle I de \mathbb{R} , admettant un développement limité d'ordre n au voisinage de $a \in I$, de la forme :

$$f(x) = c_0 + c_1(x-a) + \dots + c_n(x-a)^n + (x-a)^n \varepsilon(x-a)$$

On appelle **partie principale d'ordre n** du développement limité de f le polynôme :

$$c_0 + c_1(x-a) + \dots + c_n(x-a)^n$$



1. Cette définition respecte la convention introduite dans le paragraphe précédent, où l'on a choisi, de façon générique, de désigner par ε toute fonction de limite nulle en 0.
2. Quand x s'approche de a , c'est-à-dire lorsque $x - a$ devient très petit, chacun des termes du développement limité devient négligeable devant le terme qui le précède. Plus précisément, le terme $c_1(x-a)$ devient petit devant c_0 , qui est une constante ; de même, $c_2(x-a)^2$ devient lui aussi très petit, et négligeable devant $(x-a)$, \dots , $c_n(x-a)^n$ devient très très, très petit, c'est-à-dire négligeable devant $(x-a)^{n-1}$, et, finalement, $(x-a)^n \varepsilon(x-a)$ devient encore plus petit, c'est-à-dire négligeable devant $(x-a)^n$.
3. Même si le développement limité donne apparemment la valeur de la fonction en tout point de l'intervalle I , il ne faut pas oublier que l'on ne connaît aucune propriété de la fonction ε en dehors de sa limite en 0. Autrement dit, à partir du développement limité, on ne peut espérer obtenir aucune information sur le comportement de la fonction f ailleurs qu'au point a . C'est uniquement lorsque l'on recherche la limite en a d'une expression faisant intervenir la fonction f qu'il peut être avantageux de remplacer la fonction f par son développement limité en a .
4. La partie principale d'ordre n , $P(x) = c_0 + c_1(x-a) + \dots + c_n(x-a)^n$, du développement limité est un polynôme de degré n , écrit de manière tout à fait adaptée pour étudier ce qui se passe lorsque x tend vers a . Il serait particulièrement maladroit de le développer suivant les puissances de x car on perdrait ainsi toute l'information contenue dans les coefficients c_i , $i = 0, \dots, n$. Par exemple, le coefficient c_0 donne la valeur du polynôme P au point a , le coefficient c_1 donne la dérivée de P en a , etc.
5. Le reste $(x-a)^n \varepsilon(x-a)$ du développement limité est indispensable : en l'oubliant, on affirme que la fonction f est un polynôme de degré au plus n ce qui n'est pas vrai en général. Il faut le considérer comme un indicateur de l'ordre du développement limité, c'est-à-dire de la précision avec laquelle on veut connaître la fonction f au voisinage du point a . De plus, quand on effectue des opérations sur les développements limités, c'est en suivant les transformations successives de ce terme que l'on connaît l'ordre du développement limité obtenu à la fin. Cet ordre peut être en effet difficile à prévoir avant de faire les calculs explicites.

3. Unicité du développement limité

Théorème

Soient f une fonction définie sur un intervalle I , et a un point de I . Si f admet un développement limité d'ordre $n \in \mathbb{N}^*$ au voisinage de a , alors ce développement est unique.



Le théorème précédent signifie que, si on trouve, par exemple par deux calculs différents, deux développements limités pour la même fonction, ceux-ci sont égaux. Nous pourrions donc parler « du » développement limité à l'ordre n de la fonction f au voisinage de a .

Exemple

La fonction $x \mapsto \frac{1}{1-x}$ est définie sur l'intervalle $]0, \infty[$. Nous voulons en trouver le développement limité en un point $a > 0$. La première étape consiste à établir un développement limité de la fonction $x \mapsto \frac{1}{1-x}$ en 0.

On remarque que, pour tout réel $x \neq 1$:

$$\frac{1}{1-x} = 1 + x + x^2 + \dots + x^n + x^n \frac{x}{1-x}$$

grâce à la formule classique donnant la somme des $n+1$ premiers termes d'une suite géométrique de premier terme 1, et de raison $x \neq 1$:

$$1 + x + x^2 + \dots + x^n = \frac{1 - x^{n+1}}{1 - x}$$

En posant $\varepsilon(x) = \frac{x}{1-x}$, on obtient alors le développement limité à l'ordre n , au voisinage de 0, de la fonction $x \mapsto \frac{1}{1-x}$.

Passons à la deuxième étape. Une méthode générale pour obtenir le développement limité d'une fonction au voisinage d'un point a est de se ramener à un développement limité au voisinage de 0 en posant

$$x = a + h$$

où h tend vers 0 (en pratique, les développements limités « classiques » sont tous des développements limités au voisinage de 0).

Grâce au développement limité obtenu plus haut, on obtient :

$$\begin{aligned} \frac{1}{x} &= \frac{1}{a+h} \\ &= \frac{1}{a} \times \frac{1}{1 - \frac{-h}{a}} \\ &= \frac{1}{a} \left(1 + \left(\frac{-h}{a} \right) + \left(\frac{-h}{a} \right)^2 + \dots + \left(\frac{-h}{a} \right)^n + \left(\frac{-h}{a} \right)^n \varepsilon \left(\frac{-h}{a} \right) \right) \\ &= \frac{1}{a} - \frac{1}{a^2} (x-a) + \frac{1}{a^3} (x-a)^2 + \dots + (-1)^n \frac{1}{a^{n+1}} (x-a)^n + (x-a)^n \varepsilon(x-a) \end{aligned}$$

Formule de Taylor-Young

Quoique faisant intervenir une notion globale (fonction dérivable sur un intervalle) dans son énoncé et un résultat qui nécessite l'étude globale des fonctions (théorème des accroissements finis) dans sa démonstration, le résultat suivant est essentiel dans l'étude locale des fonctions.

1. Intégration des développements limités

Propriété

Soient f une fonction définie sur un intervalle I de \mathbb{R} , et a un point de I .

Si f est dérivable sur I , et si sa dérivée f' admet un développement limité d'ordre $n \in \mathbb{N}^*$ au voisinage de a , de la forme :

$$\begin{aligned} f'(x) &= c_0 + c_1(x-a) + \dots + c_n(x-a)^n + (x-a)^n \varepsilon(x-a) \\ &= c_0 + c_1(x-a) + \dots + c_n(x-a)^n + o((x-a)^n) \end{aligned}$$

alors f admet un développement limité d'ordre $n+1$ au voisinage de a , dont la partie principale d'ordre $n+1$ est obtenue en primitivant terme à terme celle du développement limité de f' , de la façon suivante :

$$\begin{aligned} f(x) &= f(a) + c_0(x-a) + \frac{c_1}{2}(x-a)^2 + \dots + \frac{c_n}{n+1}(x-a)^{n+1} + (x-a)^{n+1} \varepsilon(x-a) \\ &= f(a) + c_0(x-a) + \frac{c_1}{2}(x-a)^2 + \dots + \frac{c_n}{n+1}(x-a)^{n+1} + o((x-a)^{n+1}) \end{aligned}$$

Démonstration : Ce résultat est admis. ■

► Développement limité de la fonction logarithme népérien au voisinage de 1

On intègre le développement limité :

$$\frac{1}{1-x} = 1 + x + x^2 + \dots + x^n + x^n \varepsilon(x)$$

Une primitive de la fonction $x \mapsto \frac{1}{1-x}$ est la fonction $x \mapsto -\ln(1-x)$, fonction qui s'annule pour $x=0$. Il en résulte :

$$-\ln(1-x) = x + \frac{x^2}{2} + \dots + \frac{x^{n+1}}{n+1} + x^{n+1} \varepsilon(x)$$

Cette formule est vraie pour tout ordre n . On préfère écrire un développement limité d'ordre n :

$$\ln(1-x) = -x - \frac{x^2}{2} - \dots - \frac{x^n}{n} + x^n \varepsilon(x) = -\sum_{k=1}^n \frac{x^k}{k} + x^n \varepsilon(x)$$

En changeant x en $-x$, on obtient :

$$\ln(1+x) = x - \frac{x^2}{2} + \dots + (-1)^{n+1} \frac{x^n}{n} + x^n \varepsilon(x) = \sum_{k=1}^n \frac{(-1)^{k+1} x^k}{k} + x^n \varepsilon(x)$$

► Développement limité de la fonction exponentielle au voisinage de zéro

La fonction exponentielle est dérivable. Sa dérivée est la fonction exponentielle, qui vaut 1 en 0. On en déduit :

$$e^x = 1 + x + x\varepsilon(x)$$

En intégrant le développement limité :

$$e^x = e^x = 1 + x + x\varepsilon(x)$$

on trouve :

$$e^x = e^0 + x + \frac{x^2}{2} + x^2\varepsilon(x) = 1 + x + \frac{x^2}{2} + x^2\varepsilon(x)$$

qui est aussi un développement limité de la fonction dérivée $x \mapsto e^x$. En continuant à intégrer les développements de la fonction dérivée $x \mapsto e^x$ ainsi obtenus, on trouve finalement :

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + x^n\varepsilon(x) = \sum_{k=0}^n \frac{x^k}{k!} + x^n\varepsilon(x)$$

À partir de là, on obtient facilement le développement limité de la fonction exponentielle à l'ordre n en un point a :

$$\begin{aligned} e^{a+h} &= e^a e^h \\ &= e^a \left(1 + h + \frac{h^2}{2} + \frac{h^3}{3!} + \dots + \frac{h^n}{n!} + h^n\varepsilon(h) \right) \\ &= e^a \left(1 + (x-a) + \frac{(x-a)^2}{2} + \frac{(x-a)^3}{3!} + \dots + \frac{(x-a)^n}{n!} + (x-a)^n\varepsilon(x-a) \right) \end{aligned}$$

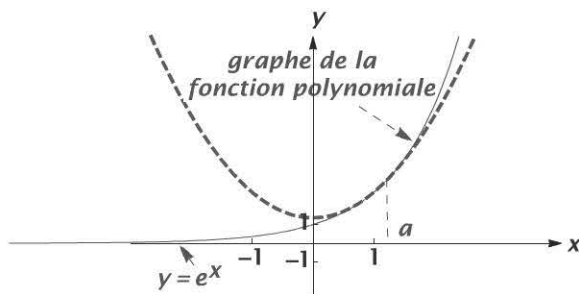


Figure 32.1 – Le graphe d'une approximation polynomiale de la fonction *exponentielle* au voisinage de a .

► Développement limité des fonctions *sinus* et *cosinus* au voisinage de zéro

1. La fonction sinus vaut 0 en 0, elle est dérivable sur \mathbb{R} , sa dérivée est la fonction cosinus. La fonction cosinus vaut 1 en 0, elle est dérivable, sa dérivée est la fonction $-\sinus$. On en déduit :

$$\cos x = 1 + 0 \times x + x\varepsilon(x) = 1 + \varepsilon(x)$$

En intégrant le développement limité :

$$\cos x = 1 + x\varepsilon(x)$$

on obtient :

$$\sin x = \sin(0) + x + x^2 \varepsilon(x) = x + x^2 \varepsilon(x)$$

En intégrant le développement limité :

$$-\sin x = -x + x^2 \varepsilon(x)$$

on obtient :

$$\cos x = \cos(0) - \frac{x^2}{2} + x^3 \varepsilon(x) = 1 - \frac{x^2}{2} + x^3 \varepsilon(x)$$

En continuant à intégrer successivement les développements des dérivées respectives des fonctions *sinus* et *cosinus*, on trouve finalement :

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots + (-1)^n \frac{x^{2n}}{(2n)!} + x^{2n+1} \varepsilon(x) = \sum_{k=0}^n \frac{(-1)^k x^{2k}}{(2k)!} + x^{2n+1} \varepsilon(x)$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + x^{2n+2} \varepsilon(x) = \sum_{k=0}^n \frac{(-1)^k x^{2k+1}}{(2k+1)!} + x^{2n+2} \varepsilon(x)$$

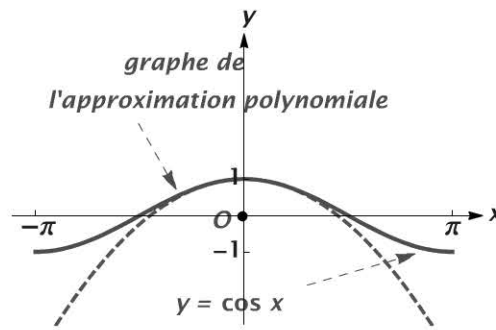


Figure 32.2 – Le graphe d’une approximation polynomiale de la fonction *cosinus* au voisinage de 0.

À partir de là, on trouve facilement le développement limité de la fonction *sinus*, par exemple, à l’ordre $2n$ en un point a :

$$\begin{aligned} \sin x &= \sin(a+h) \\ &= \sin a \cos h + \cos a \sin h \\ &= \left(1 - \frac{h^2}{2!} + \frac{h^4}{4!} - \dots + (-1)^n \frac{h^{2n}}{(2n)!} + h^{2n+1} \varepsilon(h)\right) \\ &\quad + \cos(a) \left(h - \frac{h^3}{3!} + \dots + (-1)^{n-1} \frac{h^{2n-1}}{(2n-1)!} + h^{2n} \varepsilon(h)\right) \\ &= \sin a + (x-a) \cos(a) - \frac{(x-a)^2}{2!} \sin(a) + \dots + (-1)^{n-1} \frac{\cos(a)}{(2n-1)!} (x-a)^{2n-1} \\ &\quad + (-1)^n \frac{\sin(a)}{(2n)!} (x-a)^{2n} + (x-a)^{2n} \varepsilon(x-a) \end{aligned}$$

2. Déterminons la limite, lorsque θ tend vers $\frac{3\pi}{2}$, par valeurs inférieures, de $\frac{\cos(\theta)}{\cos\left(\frac{\theta}{3}\right)}$; on

pose $\theta = \frac{3\pi}{2} - h$, où h tend vers zéro par valeurs positives (ce qui est logique et assez

naturel, car θ tend vers $\frac{3\pi}{2}$ par valeurs inférieures, c'est-à-dire juste avant $\frac{3\pi}{2}$:

$$\begin{aligned}\lim_{\theta \rightarrow \frac{3\pi}{2}^-} \frac{\cos(\theta)}{\cos\left(\frac{\theta}{3}\right)} &= \lim_{h \rightarrow 0} \frac{\cos\left(\frac{3\pi}{2} - h\right)}{\cos\left(\frac{\pi}{2} - \frac{h}{3}\right)} \\ &= \lim_{h \rightarrow 0} -\frac{\sin(h)}{\sin\left(\frac{h}{3}\right)} \\ &= -3\end{aligned}$$

2. Formule de Taylor-Young¹

Théorème

Soit f une fonction définie sur un intervalle I et soit a un point de I . Si f est n fois dérivable en a , alors f admet un développement limité d'ordre n en a donné par :

$$f(x) = f(a) + (x-a)f'(a) + \frac{f''(a)}{2}(x-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(x-a)^n + (x-a)^n \varepsilon(x-a)$$

où, pour tout entier i de $\{1, \dots, n\}$, $f^{(i)}$ désigne la dérivée $i^{\text{ème}}$ de f .

Démonstration : On démontre ce résultat par récurrence sur l'entier n .

- La propriété est bien vraie au rang 1 : par hypothèse, la fonction f est (une fois) dérivable en a . Elle satisfait donc la propriété caractéristique des fonctions dérivables :

$$f(x) = f(a) + (x-a)f'(a) + (x-a)\varepsilon(x-a)$$

Ainsi, elle admet un développement limité d'ordre 1 en a , qui est bien de la forme annoncée.

- Supposons la propriété vraie jusqu'à un rang $n > 0$; si la fonction f est $n+1$ fois dérivable en a , sa dérivée f' est donc n fois dérivable en a . L'hypothèse de récurrence permet d'en déduire qu'elle admet donc un développement limité donné par :

$$f'(x) = f'(a) + (x-a)f''(a) + \frac{f^{(3)}(a)}{2}(x-a)^2 + \dots + \frac{f^{(n+1)}(a)}{(n)!}(x-a)^n + o((x-a)^n)$$

(la $i^{\text{ème}}$ dérivée de f' est la $(i+1)^{\text{ème}}$ dérivée de f).

Il suffit d'intégrer ce développement limité pour obtenir celui de f et la formule cherchée. ■



Une fonction peut admettre un développement limité à un ordre $n \geq 2$ au voisinage d'un point, sans être n fois dérivable en ce point. C'est le cas, par exemple, de la fonction f qui, à tout réel x non nul, associe $x^3 \sin\left(\frac{1}{x}\right)$.

1. William Henry Young (1863-1942), mathématicien anglais, spécialiste de calcul différentiel, théorie de la mesure, analyse spectrale, analyse complexe.

Au voisinage de zéro, f admet le développement limité à l'ordre 2 suivant :

$$x^3 \sin\left(\frac{1}{x}\right) = 0 + x^2 x \sin\left(\frac{1}{x}\right)$$

avec :

$$\lim_{x \rightarrow 0} x \sin\left(\frac{1}{x}\right) = 0$$

La dérivée de f peut être prolongée par continuité en zéro, puisque :

$$\lim_{x \rightarrow 0} \frac{x^3 \sin\left(\frac{1}{x}\right) - 0}{x - 0} = 0$$

mais ce n'est pas le cas de sa dérivée seconde, puisque :

$$\frac{3x^2 \sin\left(\frac{1}{x}\right) - x \cos\left(\frac{1}{x}\right) - 0}{x - 0} = 3x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)$$

n'a pas de limite lorsque x tend vers zéro.

Dans ce qui suit, n désigne un entier naturel.

► Développement limité au voisinage de 0 de la fonction $x \mapsto \frac{1}{1-x}$

Théorème

Pour tout réel x tendant vers 0 :

$$\begin{aligned}\frac{1}{1-x} &= 1 + x + x^2 + \dots + x^n + x^n \varepsilon(x) \\ &= 1 + x + x^2 + \dots + x^n + o(x^n) \\ &= \sum_{k=0}^n x^k + o(x^n)\end{aligned}$$

► Développement limité au voisinage de 0 de la fonction $x \mapsto \frac{1}{1+x}$

Théorème

Pour tout réel x tendant vers 0 :

$$\begin{aligned}\frac{1}{1+x} &= 1 - x + x^2 - \dots + (-1)^n x^n + x^n \varepsilon(x) \\ &= 1 - x + x^2 - \dots + (-1)^n x^n + o(x^n) \\ &= \sum_{k=0}^n (-1)^k x^k + o(x^n)\end{aligned}$$

► Développement limité au voisinage de 0 de la fonction $x \mapsto \ln(1+x)$

Théorème

Pour tout réel x tendant vers 0 :

$$\begin{aligned}\ln(1+x) &= x - \frac{x^2}{2} + \dots + (-1)^{n+1} \frac{x^n}{n} + x^n \varepsilon(x) \\ &= x - \frac{x^2}{2} + \dots + (-1)^{n+1} \frac{x^n}{n} + o(x^n) \\ &= \sum_{k=1}^n \frac{(-1)^{k+1} x^k}{k} + o(x^n)\end{aligned}$$

► Développement limité au voisinage de 0 de la fonction $x \mapsto \ln(1-x)$

Théorème

Pour tout réel x tendant vers 0 :

$$\begin{aligned}\ln(1-x) &= -x - \frac{x^2}{2} - \dots - \frac{x^n}{n} + x^n \varepsilon(x) \\ &= -x - \frac{x^2}{2} - \dots - \frac{x^n}{n} + o(x^n) \\ &= -\sum_{k=1}^n \frac{x^k}{k} + o(x^n)\end{aligned}$$

► Développement limité au voisinage de 0 de la fonction exponentielle

Théorème

Pour tout réel x tendant vers 0 :

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + x^n \varepsilon(x) \\ &= 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + o(x^n) \\ &= \sum_{k=0}^n \frac{x^k}{k!} + o(x^n) \end{aligned}$$

► Développement limité au voisinage de 0 des fonctions trigonométriques

Théorème

Pour tout réel x tendant vers 0 :

$$\begin{aligned} \cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots + (-1)^n \frac{x^{2n}}{(2n)!} + x^{2n+1} \varepsilon(x) \\ &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots + (-1)^n \frac{x^{2n}}{(2n)!} + o(x^{2n+1}) \\ &= \sum_{k=0}^n \frac{(-1)^k x^{2k}}{(2k)!} + o(x^{2n+1}) \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + x^{2n+2} \varepsilon(x) \\ &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2}) \\ &= \sum_{k=0}^n \frac{(-1)^k x^{2k+1}}{(2k+1)!} + o(x^{2n+2}) \\ \tan x &= x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + x^8 \varepsilon(x) \\ &= x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + o(x^8) \end{aligned}$$



On ne donnera pas, ici, de développement limité à l'ordre n , au voisinage de zéro de la fonction tangente ; il faut savoir que ce développement existe, mais fait intervenir des quantités qu'il serait trop compliqué de définir ici (les nombres de Bernoulli en l'occurrence).

$$\begin{aligned} \arctan x &= x - \frac{x^3}{3} + \frac{x^5}{5} - \dots + \frac{(-1)^n x^{2n+1}}{2n+1} + x^{2n+2} \varepsilon(x) \\ &= x - \frac{x^3}{3} + \frac{x^5}{5} - \dots + \frac{(-1)^n x^{2n+1}}{2n+1} + o(x^{2n+2}) \end{aligned}$$

► Développement limité au voisinage de 0 des fonctions trigonométriques hyperboliques en 0

Théorème

Pour tout réel x tendant vers 0 :

$$\begin{aligned}
 \operatorname{ch}(x) &= 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \dots + \frac{x^{2n}}{(2n)!} + x^{2n+1} \varepsilon(x) \\
 &= 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \dots + \frac{x^{2n}}{(2n)!} + o(x^{2n+1}) \\
 &= \sum_{k=0}^n \frac{x^{2k}}{(2k)!} + o(x^{2n+1}) \\
 \operatorname{sh}(x) &= x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \dots + \frac{x^{2n+1}}{(2n+1)!} + x^{2n+2} \varepsilon(x) \\
 &= x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \dots + \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+2}) \\
 &= \sum_{k=0}^n \frac{x^{2k+1}}{(2k+1)!} + o(x^{2n+2}) \\
 \operatorname{th}(x) &= x - \frac{x^3}{3} + \frac{2x^5}{15} - \frac{17x^7}{315} + x^8 \varepsilon(x) \\
 &= x - \frac{x^3}{3} + \frac{2x^5}{15} - \frac{17x^7}{315} + o(x^8)
 \end{aligned}$$



De même que pour la fonction tangente, on ne donne pas, ici, de développement limité à l'ordre n , au voisinage de zéro de la fonction tangente hyperbolique, qui fait aussi intervenir les nombres de Bernoulli.

$$\begin{aligned}
 \operatorname{argth} x &= x + \frac{x^3}{3} + \frac{x^5}{5} + \dots + \frac{x^{2n+1}}{2n+1} + x^{2n+2} \varepsilon(x) \\
 &= x + \frac{x^3}{3} + \frac{x^5}{5} + \dots + \frac{x^{2n+1}}{2n+1} + o(x^{2n+2})
 \end{aligned}$$

► Développement limité au voisinage de 0 de la fonction $x \mapsto (1+x)^\alpha$, $\alpha \in \mathbb{R}^*$

Théorème

$$\begin{aligned}
 (1+x)^\alpha &= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2} x^2 + \dots + \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} x^n + x^n \varepsilon(x) \\
 &= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2} x^2 + \dots + \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} x^n + o(x^n)
 \end{aligned}$$

Opérations algébriques et composition des développements limités

Lorsque l'on veut calculer le développement limité d'une fonction, on commence par se ramener à un développement limité au voisinage de 0 (si cela est nécessaire, on pose donc $x = a + h$, avec h tendant vers 0). Nous nous contenterons donc, dans ce qui suit, de parler des développements limités au voisinage de 0.

1. Premier principe

Dans un développement limité, il est inutile de conserver des termes plus petits que le reste.

Ainsi, pour un développement limité à un ordre n donné ($n \in \mathbb{N}^*$), si p est un entier strictement plus grand que n :

$$x^p + x^n \varepsilon(x) = x^n \varepsilon(x) \quad , \quad x^p \varepsilon(x) + x^n \varepsilon(x) = x^n \varepsilon(x)$$

ou encore, avec la notation « petit o » :

$$x^p + o(x^n) = o(x^n) \quad , \quad o(x^p) + o(x^n) = o(x^n)$$

(ceci vient du fait que $x^p = x^n x^{p-n} = x^n \varepsilon(x)$).



On prendra bien garde au fait que ces égalités se lisent de la gauche vers la droite, et non l'inverse !

À partir de ce principe, il est facile de calculer la somme ou le produit de deux développements limités.

Exemple

Au voisinage de 0 :

$$x^4 = o(x^3) \quad , \quad o(x^5) + o(x^7) = o(x^4) \quad , \quad x^2 o(x^5) = o(x^5)$$

2. Second principe

Pour éviter des calculs trop lourds et pénibles, on n'explicite les « factorielles » qu'à la fin !

Exemples

1. Calcul à éviter

Pour déterminer un développement limité à l'ordre 4 en zéro de la fonction $x \mapsto e^x \sin x$, il est peu judicieux de procéder ainsi :

$$\begin{aligned}
 e^x \sin x &= \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + o(x^4)\right) \left(x - \frac{x^3}{3!} + o(x^4)\right) \\
 &= \left(x - \frac{x^3}{6} + o(x^4)\right) \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + o(x^4)\right) \\
 &= x \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + o(x^3)\right) \\
 &\quad - \frac{x^3}{6} \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + o(x^4)\right) \\
 &\quad + o(x^4) \left(1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + o(x^4)\right) \\
 &= x + x^2 + \frac{x^3}{2} + \frac{x^4}{6} + \frac{x^5}{24} + o(x^4) \\
 &\quad - \frac{x^3}{6} - \frac{x^4}{6} - \frac{x^5}{12} + \frac{x^6}{36} - \frac{x^7}{144} + o(x^6) \\
 &\quad + \dots + \text{horribles fractions}
 \end{aligned}$$

2. Calcul modèle

Pour déterminer un développement limité à l'ordre 4 en zéro de la fonction $x \mapsto e^x \sin x$, il est préférable de procéder ainsi :

$$\begin{aligned}
 e^x \sin x &= \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + o(x^4)\right) \left(x - \frac{x^3}{3!} + o(x^4)\right) \\
 &= x \left(1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + o(x^4)\right) \\
 &\quad - \frac{x^3}{3!} \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + o(x^4)\right) \\
 &\quad + o(x^4) \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + o(x^4)\right) \\
 &= x + x^2 + \frac{x^3}{2} + \frac{x^4}{3!} + o(x^4) - \frac{x^3}{3!} - \frac{x^4}{3!} + o(x^4) \\
 &= x + x^2 + \frac{x^3}{2} + \frac{x^4}{3!} + o(x^4) - \frac{x^3}{3!} - \frac{x^4}{3!} + o(x^4) \\
 &= x + x^2 + \frac{x^3}{3} + o(x^4)
 \end{aligned}$$

3. Troisième principe

Pour déterminer le développement limité au voisinage de zéro, à un ordre $n \in \mathbb{N}^*$, de la composée de deux fonctions f et g , où g tend vers zéro : on commence par poser

$g(x) = X$, et on effectue le développement limité de f en zéro à un ordre suffisamment élevé p afin que le développement de la fonction composée soit bien d'ordre n par rapport à x :

$$(f \circ g)(x) = f(X) = \alpha_0 + \alpha_1 X + \alpha_2 X^2 + \dots + \alpha_n X^n + o(X^n)$$

Il reste ensuite à développer à l'ordre n par rapport à x les termes $X = g(x)$, $X^2 = (g(x))^2$, ..., $X^p = (g(x))^p$, en ne gardant que les termes en x^n .

Exemple

Au voisinage de 0, à l'ordre 4 :

$$\begin{aligned} \cos(\sin x) &= 1 - \frac{\sin^2 x}{2} + \frac{\sin^4 x}{4!} + o(\sin^4 x) \\ &= 1 - \frac{1}{2} \left(x - \frac{x^3}{3!} + o(x^4) \right)^2 + \frac{1}{4!} \left(x - \frac{x^3}{3!} + o(x^4) \right)^2 + o(x^4) \\ &= 1 - \frac{1}{2} \left(x^2 - \frac{2x^4}{3!} + o(x^4) \right) + \frac{x^4}{4!} + o(x^4) \\ &= 1 - \frac{x^2}{2} + \frac{5x^4}{24} + o(x^4) \end{aligned}$$

4. Quatrième principe

Pour déterminer le développement limité, au voisinage de zéro, à un ordre $n \in \mathbb{N}^*$, de la composée de deux fonctions f et g , il faut faire très attention dans les cas où la fonction g ne tend pas vers zéro en zéro !

Exemple

Au voisinage de 0, à l'ordre 2 :

$$\begin{aligned} e^{\cos x} &= e^{1 - \frac{x^2}{2} + o(x^2)} \\ &= e e^{-\frac{x^2}{2} + o(x^2)} \\ &= e \left(1 - \frac{x^2}{2} + o(x^2) \right) \end{aligned}$$

À la fin du XIX^e siècle, le mathématicien Henri Poincaré¹ s'intéresse au problème « des trois corps », qui est un cas particulier du problème dit « des N corps », où on cherche à décrire le système formé par N corps célestes (étoiles, planètes), dont les mouvements sont régis par les lois de l'attraction universelle, et à étudier leur stabilité. Dans le cas « des trois corps », qui sont en fait le Soleil, la Terre, la Lune, de masses respectives m_{soleil} , m_{terre} , m_{lune} , le rapport $\frac{m_{\text{lune}}}{m_{\text{soleil}}}$ est très petit : $\frac{m_{\text{lune}}}{m_{\text{soleil}}} \ll 1$. Pour résoudre ce problème, Poincaré eut l'idée d'effectuer un développement limité par rapport à cette dernière quantité. C'est ainsi qu'est apparue la notion de « développement asymptotique », la grandeur m_{soleil} étant « très grande ».

De façon générale, un **développement asymptotique** est obtenu lorsqu'une des quantités en jeu tend vers une limite infinie. L'inverse de cette quantité tendant vers zéro, il est alors possible de se ramener à un développement limité au voisinage de zéro.

Exemples

1. Lorsque n tend vers $+\infty$:

$$\left(1 + \frac{1}{n}\right)^n = e^{n \ln\left(1 + \frac{1}{n}\right)} = e^{n\left(\frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} + o\left(\frac{1}{n^3}\right)\right)} = e^{1 - \frac{1}{2n} + \frac{1}{3n^2} + o\left(\frac{1}{n^2}\right)}$$

(On a effectué un développement asymptotique à l'ordre 2 en $\frac{1}{n}$.)

Ainsi :

$$\lim_{n \rightarrow +\infty} \left(1 - \frac{1}{n}\right)^n = e$$

2. Lorsque n tend vers $+\infty$:

$$\begin{aligned} \sin(\pi \sqrt{n^2 + n}) &= \sin\left(\pi n \sqrt{1 + \frac{1}{n}}\right) = \sin\left(n\pi \left\{1 + \frac{1}{2n} - \frac{1}{8n^2} + o\left(\frac{1}{n^2}\right)\right\}\right) \\ &= \sin\left(n\pi + \frac{\pi}{2} - \frac{\pi}{8n} + o\left(\frac{1}{n}\right)\right) \\ &= \cos\left(n\pi - \frac{\pi}{8n} + o\left(\frac{1}{n}\right)\right) \\ &= (-1)^n \cos\left(-\frac{\pi}{8n} + o\left(\frac{1}{n}\right)\right) \\ &= (-1)^n \cos\left(\frac{\pi}{8n} + o\left(\frac{1}{n}\right)\right) \\ &= (-1)^n \left\{1 - \frac{1}{2} \frac{\pi^2}{64n^2} + o\left(\frac{1}{n^2}\right)\right\} \\ &= (-1)^n \left\{1 - \frac{\pi^2}{128n^2} + o\left(\frac{1}{n^2}\right)\right\} \end{aligned}$$

(On a effectué un développement asymptotique à l'ordre 2 en $\frac{1}{n}$.)

Ainsi, $\sin(\pi \sqrt{n^2 + n})$ n'a pas de limite lorsque n tend vers l'infini.

1. 1854-1912. Outre ses apports aux mathématiques (il est considéré comme un des fondateurs de la topologie), il apporta aussi de nombreuses contributions à la physique théorique, en optique et en relativité.

La notion de convexité correspond à une réalité physique ; ainsi, en optique, une lentille dite « convexe » est un verre « bombé vers l'extérieur ». Lorsqu'on la pose sur une table, sa forme « bombée » fait que, quelle que soit sa position, la table reste toujours tangente au verre. De façon équivalente, en mathématiques, la première caractérisation de la convexité, qui s'applique à des courbes, est liée au fait que le barycentre d'un système de points situés sur la courbe doit se trouver au-dessus de celle-ci (les fameuses « inégalités de convexité »), ou encore, que la courbe est située au-dessous de toutes ses cordes.

1. Définitions

► Fonction convexe

Une fonction f , définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , est dite **convexe** si, pour tout couple (a, b) d'éléments de I^2 , et tout réel t de l'intervalle $[0, 1]$:

$$f(ta + (1 - t)b) \leq t f(a) + (1 - t) f(b)$$



Cette définition traduit, tout simplement, le fait que tout point situé sur une corde joignant deux points de la courbe, de coordonnées respectives $(a, f(a))$ et $(b, f(b))$, $(a, b) \in I^2$, est au-dessus de la courbe. Ce point étant sur la corde, son abscisse est de la forme $ta + (1 - t)b$, $t \in [0, 1]$, et son ordonnée de la forme $t f(a) + (1 - t) f(b)$, qui est donc plus grande que celle du point de la courbe d'abscisse $ta + (1 - t)b$, c'est-à-dire $f(ta + (1 - t)b)$.

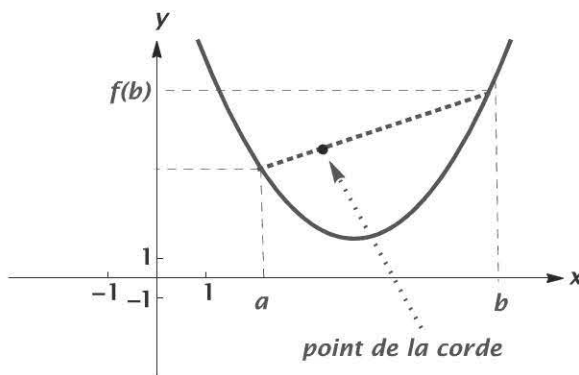


Figure 36.1 – Le graphe d'une fonction convexe.

► Fonction concave

Une fonction f , définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , est dite **concave** si, pour tout couple (a, b) d'éléments de I^2 , et tout réel t de l'intervalle $[0, 1]$:

$$f(ta + (1 - t)b) \geq t f(a) + (1 - t) f(b)$$



Dire qu'une fonction f est concave revient donc à dire que son opposée $-f$ est convexe.

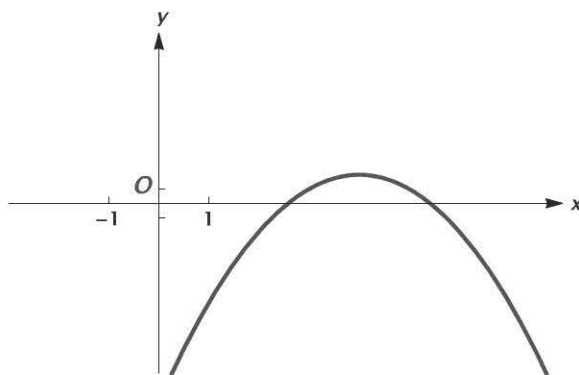


Figure 36.2 – Le graphe d'une fonction concave.

2. Théorèmes

► Position de la courbe représentative d'une fonction convexe par rapport à ses cordes

Une fonction f , définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , est convexe si et seulement si sa courbe représentative est située en-dessous de toutes les cordes joignant deux points de cette courbe.

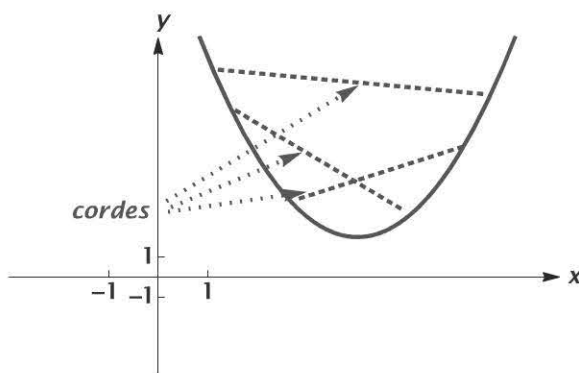


Figure 36.3 – Illustration graphique de la position de la courbe représentative d'une fonction convexe par rapport à ses cordes.

► Inégalité de convexité (ou inégalité de Jensen¹)

Théorème

Soient f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , convexe, et t_1, \dots, t_n , n réels positifs dont la somme vaut 1 : $t_1 + \dots + t_n = \sum_{i=1}^n t_i = 1$. Alors, pour tout n -uplet (x_1, \dots, x_n) de I^n :

$$f\left(\sum_{i=1}^n t_i x_i\right) \leq \sum_{i=1}^n t_i f(x_i)$$

Cette définition traduit, tout simplement, le fait que tout barycentre d'un ensemble de points situés sur la courbe, de coordonnées respectives $(x_1, f(x_1)), \dots, (x_n, f(x_n))$, $(x_1, \dots, x_n) \in I^n$,

1. Mathématicien danois, 1859-1925.



est au-dessus de la courbe. L'abscisse de ce point est donc de la forme $\sum_{i=1}^n t_i x_i$, où t_1, \dots, t_n , sont n réels positifs dont la somme vaut 1, et son ordonnée de la forme $\sum_{i=1}^n t_i f(x_i)$, qui est donc plus grande que celle du point de la courbe d'abscisse $\sum_{i=1}^n t_i x_i$, c'est-à-dire $f\left(\sum_{i=1}^n t_i x_i\right)$.

► Régularité d'une fonction convexe

Théorème

Une fonction f , définie sur un intervalle **ouvert** I de \mathbb{R} , à valeurs dans \mathbb{R} , convexe, est continue, et admet, en tout point de I , une dérivée à droite et une dérivée à gauche.



L'hypothèse selon laquelle l'intervalle I doit être **ouvert** est essentielle !

► Inégalité des pentes croissantes

Théorème

Une fonction f , définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , dérivable sur I , est convexe si et seulement si, pour tout réel a de I , la fonction qui, à tout x de I , distinct de a , associe le rapport $\frac{f(x) - f(a)}{x - a}$ est croissante, ce qui est équivalent à dire que, pour tout triplet de réels (a, b, c) de I tel que $a < b < c$:

$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a} \leq \frac{f(c) - f(b)}{b - c}$$



Ce théorème traduit, tout simplement, le fait que la pente, ou coefficient directeur, de la droite joignant les points de coordonnées $(a, f(a))$ et $(b, f(b))$, est plus petite que la pente de la droite joignant les points de coordonnées $(a, f(a))$ et $(c, f(c))$, qui est elle-même plus petite que la pente de la droite joignant les points de coordonnées $(b, f(b))$ et $(c, f(c))$, comme illustré sur le dessin suivant :

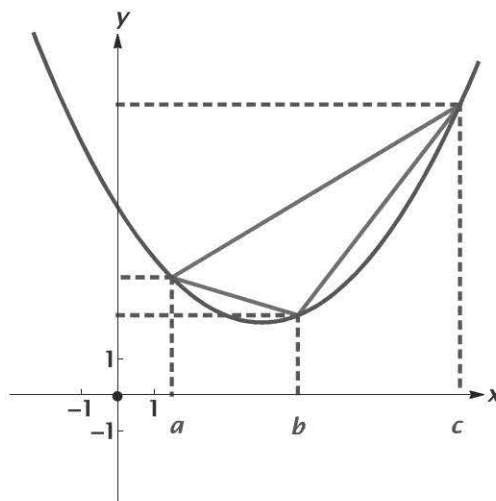


Figure 36.4– Illustration graphique de l'inégalité des pentes croissantes.

► Position de la courbe représentative d'une fonction convexe par rapport à ses tangentes

Théorème

Une fonction f , définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , dérivable sur I , est convexe si et seulement si sa courbe représentative est située au-dessus de toutes ses tangentes.

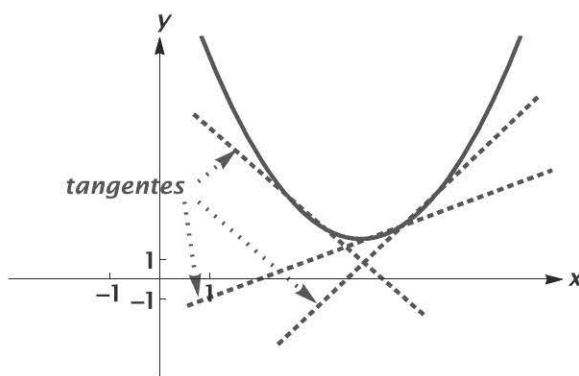


Figure 36.5 – Illustration graphique de la position de la courbe représentative d'une fonction convexe par rapport à ses tangentes.

► Convexité et dérivabilité

Théorème

Une fonction f , définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , dérivable sur I , est convexe si et seulement si sa dérivée f' est croissante.

► Cas des fonctions deux fois dérivables

Corollaire

Une fonction f , définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , deux fois dérivable sur I , est convexe si et seulement si sa dérivée seconde f'' est à valeurs positives, c'est-à-dire si et seulement si, pour tout réel x de I :

$$f''(x) \geq 0$$

Équations différentielles linéaires du 1^{er} ordre homogènes

Une équation différentielle est un type d'équation un peu particulier, dans la mesure où l'inconnue est une fonction, en général désignée par la notation « y ». On parle d'« équation différentielle », dans la mesure où les dérivées de la fonction inconnue figurent aussi dans l'équation. Par exemple, si I est un intervalle de \mathbb{R} ,

$$\begin{aligned}y'(x) &= 2y(x) \quad \forall x \in I \\y(x)y'(x) + x &= 1 \quad \forall x \in I \\(y'(x))^2 + y(x) + x &= 0 \quad \forall x \in I\end{aligned}$$

sont des équations différentielles.

Dans ce qui suit, on ne s'intéressera qu'au cas des équations différentielles linéaires du premier ordre, c'est-à-dire linéaires par rapport à la fonction inconnue y , et ne faisant intervenir que la dérivée première y' de la fonction cherchée.

Définition

Soit a une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , continue.

On appelle **équation différentielle linéaire du premier ordre homogène** une équation de la forme :

$$y'(x) = a(x)y(x) \quad \forall x \in I$$

où y est une fonction dérivable sur I . On peut aussi écrire, pour alléger les notations :

$$y' = a(x)y$$

1. Méthode pratique de résolution

Soit a une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , continue.

On s'intéresse à l'équation différentielle linéaire du premier ordre homogène, que l'on désignera, dans ce qui suit, par (\mathcal{E}_0) :

$$y' = a(x)y$$

La fonction identiquement nulle sur l'intervalle I est solution de (\mathcal{E}_0) .

Considérons une solution y de (\mathcal{E}_0) , définie sur un intervalle $J \subset I$, non identiquement nulle. On suppose que y ne s'annule pas sur J . Quitte à changer y en $-y$, on supposera que y est à valeurs strictement positives.

On peut alors écrire :

$$\frac{y'(x)}{y(x)} = a(x)$$

La fonction qui, à tout x de J , associe $\frac{y'(x)}{y(x)}$, est la dérivée de la fonction qui, à tout x de J , associe $\ln |y(x)| = \ln (y(x))$.

Ainsi, si on arrive à trouver une fonction \mathcal{F}_a dont a soit la dérivée sur J (c'est-à-dire $\mathcal{F}'_a = a$), on aura, à une constante réelle C près :

$$\ln |y(x)| = \mathcal{F}_a(x) + C \quad \forall x \in J$$

(la dérivée de la fonction constante qui, à tout x de J , associe C , est la fonction identiquement nulle).

On en déduit alors, pour tout x de J :

$$y(x) = e^{\mathcal{F}_a(x)+C} = e^C e^{\mathcal{F}_a(x)}$$

On vérifie que, réciproquement, toute fonction de la forme $x \in J \mapsto e^C e^{\mathcal{F}_a(x)}$ est solution de (\mathcal{E}_0) .

On peut écrire que l'ensemble des solutions de l'équation différentielle initiale est :

$$\{x \in J \mapsto e^C e^{\mathcal{F}_a(x)}, C \in \mathbb{R}\}$$

► Solutions d'une équation différentielle linéaire du premier ordre homogène

Théorème

Soit a une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , continue.

Les solutions de l'équation différentielle linéaire du premier ordre homogène :

$$y'(x) = a(x)y(x) \quad \forall x \in I$$

sont les fonctions de la forme :

$$x \in I \mapsto K e^{\mathcal{F}_a(x)}$$

où K est une constante réelle, et \mathcal{F}_a une fonction telle que, pour tout x de I :

$$\mathcal{F}'_a(x) = a(x)$$

De façon équivalente, on peut écrire que l'ensemble des solutions de l'équation différentielle initiale est :

$$\{x \in I \mapsto K e^{\mathcal{F}_a(x)}, K \in \mathbb{R}\}$$

Exemple

Considérons l'équation différentielle :

$$y'(x) = 2xy(x)$$

On recherche les solutions y qui ne s'annulent pas sur \mathbb{R} . On a alors, pour tout réel x :

$$\frac{y'(x)}{y(x)} = 2x$$

La fonction qui, à tout réel x , associe $2x$, est la dérivée de la fonction qui, à tout réel x , associe x^2 .

On en déduit, pour tout réel x :

$$\ln |y(x)| = x^2 + C$$

où C est une constante réelle ; puis :

$$y(x) = e^{x^2+C} = e^C e^{x^2}$$

que l'on peut encore écrire sous la forme :

$$y(x) = K e^{x^2}$$

où K est une constante réelle positive (cela revient juste à appeler autrement e^C , qui est aussi une constante).

On vérifie que, réciproquement, toute fonction de la forme $x \in \mathbb{R} \mapsto K e^{x^2}$, $K \in \mathbb{R}$, est solution de $y'(x) = 2x y(x)$.

De façon équivalente, on peut écrire que l'ensemble des solutions de l'équation différentielle initiale est :

$$\{x \in \mathbb{R} \mapsto K e^{x^2}, K \in \mathbb{R}\}$$

2. Condition initiale (pour une équation différentielle linéaire du premier ordre homogène)

Définition

Soient a une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , continue, x_0 un réel de I , et y_0 un réel.

On appelle **équation différentielle linéaire du premier ordre homogène**, avec la **condition initiale** $y(x_0) = y_0$, la donnée de l'équation différentielle :

$$y'(x) = a(x)y(x) \quad \forall x \in I$$

avec la condition :

$$y(x_0) = y_0$$

► Solutions d'une équation différentielle linéaire du premier ordre homogène avec une condition initiale

Théorème

Soient a une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , continue, x_0 un réel de I , et y_0 un réel.

L'équation différentielle linéaire du premier ordre homogène :

$$y'(x) = a(x)y(x) \quad \forall x \in I$$

avec la condition initiale :

$$y(x_0) = y_0$$

admet pour unique solution la fonction :

$$x \in I \mapsto y_0 e^{-\mathcal{F}_a(x_0)} e^{\mathcal{F}_a(x)}$$

où \mathcal{F}_a est une fonction telle que, pour tout x de I :

$$\mathcal{F}'_a(x) = a(x)$$

(on admet l'existence d'une telle fonction).

Équations différentielles linéaires du 1^{er} ordre avec second membre

1. Définitions et théorèmes

► Équation différentielle linéaire du premier ordre avec second membre

Soient a et b deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , continues.

On appelle **équation différentielle linéaire du premier ordre avec second membre** une équation de la forme :

$$y'(x) = a(x)y(x) + b(x) \quad \forall x \in I$$

que l'on peut aussi écrire, pour alléger les notations :

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

► Équation homogène associée à une équation différentielle linéaire du premier ordre avec second membre

Soient a et b deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , continues.

On appelle **équation homogène** associée à l'équation différentielle linéaire du premier ordre avec second membre $y' = a(x)y + b(x)$ l'équation :

$$y'(x) = a(x)y(x) \quad \forall x \in I$$

► Principe de superposition

Théorème

Soient a , b_1 et b_2 trois fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , continues.

Si y_1 est une solution sur I de l'équation différentielle linéaire $y' = a(x)y + b_1(x)$, et y_2 une solution sur I de l'équation différentielle linéaire $y' = a(x)y + b_2(x)$, alors, pour tout couple de réels (λ_1, λ_2) , $\lambda_1 y_1 + \lambda_2 y_2$ est une solution sur I de l'équation différentielle linéaire $y' = a(x)y + \lambda_1 b_1(x) + \lambda_2 b_2(x)$.

Corollaire

Soient a et b deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , continues.

Si y_1 et y_2 sont deux solutions sur I de l'équation différentielle linéaire avec second membre $y' = a(x)y + b(x)$, alors la fonction différence $y_2 - y_1$ est une solution sur I de l'équation différentielle homogène associée $y' = a(x)y$.

Démonstration : Sachant que y_1 et y_2 sont solutions sur I de l'équation différentielle $y' = a(x)y + b(x)$, on obtient :

$$y_2 - y_1 = a(x)y_2' + b(x) - (a(x)y_1' + b(x)) = a(x)(y_2 - y_1)'$$

D'où le résultat. ■

Exemple

On considère l'équation différentielle :

$$y' = y + 2x + \sin x$$

La fonction qui, à tout réel x , associe $-2x - 2$ est solution sur \mathbb{R} de l'équation différentielle :

$$y' = y + 2x$$

La fonction qui, à tout réel x , associe $-\frac{\sin x + \cos x}{2}$ est solution sur \mathbb{R} de l'équation différentielle :

$$y' = y + \sin x$$

La fonction qui, à tout réel x , associe $-2x - 2 - \frac{\sin x + \cos x}{2}$ est donc solution sur \mathbb{R} de l'équation différentielle :

$$y' = y + 2x + \sin x$$

► Solution générale d'une équation différentielle linéaire du premier ordre avec second membre

Théorème

Soient a et b deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , continues. La solution générale sur I de l'équation différentielle linéaire du premier ordre avec second membre $y' = a(x)y + b(x)$ s'obtient comme somme de la solution générale sur I de l'équation différentielle homogène associée, et d'une solution particulière sur I de l'équation différentielle avec second membre.



Pour mémoriser ce résultat, on peut le retenir sous la forme :

$$\begin{aligned} &\text{Solution générale de l'équation avec second membre} \\ &= \\ &\text{Solution générale de l'équation sans second membre} \\ &+ \\ &\text{Solution particulière de l'équation avec second membre} \end{aligned}$$

ou, sous forme « abrégée » :

$$\begin{aligned} &\text{Solution générale E.A.S.M.} \\ &= \\ &\text{Solution générale E.S.S.M.} \\ &+ \\ &\text{Solution particulière E.A.S.M.} \end{aligned}$$

Démonstration : Ce résultat vient du corollaire précédent ; en effet, si on connaît une solution y_1 de l'équation différentielle avec second membre, définie sur un intervalle J , n'importe quelle solution sera de la forme $y_1 + z$, où z est une solution sur J de l'équation

homogène associée. Ainsi, toute solution sur J de l'équation avec second membre est telle que, pour tout réel x de J :

$$y(x) = y_1(x) + \lambda e^{\mathcal{F}_a(x)}$$

où λ est une constante réelle arbitraire, et \mathcal{F}_a une fonction telle que, pour tout x de J :

$$\mathcal{F}_a'(x) = a(x) \quad \blacksquare$$

► Extension aux fonctions à valeurs complexes

On admet que l'ensemble des résultats précédents est généralisable aux fonctions à valeurs complexes, c'est-à-dire dans \mathbb{C} .

2. Recherche de solutions particulières de l'équation avec second membre, dans le cas où la fonction a est constante

► Cas d'un second membre de type « exponentielle \times polynôme »

Soit :

$$y'(x) = \alpha y(x) + P(x) e^{rx}$$

où P est une fonction polynomiale, et r et α deux réels.

• Premier cas : $r = \alpha$.

On cherche une solution particulière, définie sur \mathbb{R} , de la forme :

$$x \mapsto Q(x) e^{\alpha x}$$

où Q est une fonction polynomiale de degré inférieur ou égal à degré de $P + 1$.

Exemples

1. Pour l'équation différentielle :

$$y' = 3y - e^{3x}$$

le second membre est de type « exponentielle \times polynôme », la fonction polynomiale étant de degré 0.

On cherche donc une solution particulière, définie sur \mathbb{R} , de la forme :

$$x \mapsto y(x) = (\lambda x + \mu) e^{3x}$$

où λ et μ sont deux réels, ce qui conduit, pour tout réel x , à :

$$y'(x) = (\lambda + 3\lambda x + 3\mu) e^{3x}$$

En injectant dans l'équation différentielle, on en déduit, pour tout réel x :

$$(\lambda + 3\lambda x + 3\mu) e^{3x} = 3(\lambda x + \mu) e^{3x} - e^{3x}$$

soit :

$$\lambda e^{3x} = -e^{3x}$$

et donc :

$$\lambda = -1$$

Une solution particulière de l'équation avec second membre est donc :

$$x \mapsto y(x) = -e^{3x}$$

(Comme il n'y a pas de condition sur μ , le cas $\mu = 0$ convient.)

2. Pour l'équation différentielle :

$$y' = 2y + (x+1)e^{2x}$$

le second membre est de type « exponentielle \times polynôme », la fonction polynomiale étant de degré 1.

On cherche donc une solution particulière, définie sur \mathbb{R} , de la forme :

$$x \mapsto y(x) = (\lambda x^2 + \mu x + \nu) e^{2x}$$

où λ , μ et ν sont des réels, ce qui conduit, pour tout réel x , à :

$$y'(x) = (2\lambda x + \mu + 2\lambda x^2 + 2\mu x + 2\nu) e^{2x}$$

En injectant dans l'équation différentielle, on en déduit, pour tout réel x :

$$y'(x) = (2\lambda x + \mu + 2\lambda x^2 + 2\mu x + 2\nu) e^{2x} = 2(\lambda x^2 + \mu x + \nu) e^{2x} + (x+1)e^{2x}$$

soit :

$$y'(x) = (2\lambda x + \mu) e^{2x} = (x+1)e^{2x}$$

et donc : $2\lambda = 1, \mu = 1$.

Une solution particulière de l'équation avec second membre est donc :

$$x \mapsto y(x) = \left(\frac{x^2}{2} + x\right) e^{2x}$$

• **Deuxième cas : $r \neq \alpha$.**

On cherche une solution particulière, définie sur \mathbb{R} , de la forme :

$$x \mapsto Q(x) e^{rx}$$

où Q est une fonction polynomiale de degré inférieur ou égal au degré de P .

Exemples

1. Pour l'équation différentielle :

$$y' = y + e^{3x}$$

le second membre est de type « exponentielle \times polynôme », la fonction polynomiale étant de degré 0.

On cherche donc une solution particulière, définie sur \mathbb{R} , de la forme :

$$x \mapsto y(x) = \lambda e^{3x}$$

où λ est un réel, ce qui conduit, pour tout réel x , à :

$$y'(x) = 3\lambda e^{3x}$$

En injectant dans l'équation différentielle, on en déduit, pour tout réel x :

$$3\lambda e^{3x} = \lambda e^{3x} + e^{3x}$$

soit :

$$2\lambda e^{3x} = e^{3x}$$

et donc :

$$\lambda = \frac{1}{2}$$

Une solution particulière de l'équation avec second membre est donc :

$$x \mapsto y(x) = \frac{1}{2} e^{3x}$$

2. Pour l'équation différentielle :

$$y' = 2y + x^2 + 1$$

le second membre est de type « exponentielle \times polynôme », la fonction polynomiale étant de degré 2.

On cherche donc une solution particulière, définie sur \mathbb{R} , de la forme :

$$x \mapsto y(x) = (\lambda x^2 + \mu x + \nu) e^{0 \times x} = \lambda x^2 + \mu x + \nu$$

où λ , μ et ν sont des réels, ce qui conduit, pour tout réel x , à :

$$y'(x) = 2\lambda x + \mu$$

En injectant dans l'équation différentielle, on en déduit, pour tout réel x :

$$2\lambda x + \mu = 2\lambda x^2 + 2\mu x + 2\nu + x^2 + 1$$

soit :

$$2(\lambda - \mu)x + \mu - 2\nu = (2\lambda + 1)x^2 + 1$$

et donc :

$$\lambda = \mu, \quad \mu - 2\nu = 1, \quad 2\lambda + 1 = 0$$

Il en résulte :

$$\lambda = \mu = -\frac{1}{2}, \quad \nu = \frac{\mu - 1}{2} = -\frac{3}{4}$$

Une solution particulière de l'équation avec second membre est donc :

$$x \mapsto y(x) = -\frac{x^2}{2} - \frac{x}{2} - \frac{3}{4}$$

3. Pour l'équation différentielle :

$$y' = 2y + x e^x$$

le second membre est de type « exponentielle \times polynôme », la fonction polynomiale étant de degré 1.

On cherche donc une solution particulière, définie sur \mathbb{R} , de la forme :

$$x \mapsto y(x) = (\lambda x + \mu) e^x$$

où λ et μ sont des réels, ce qui conduit, pour tout réel x , à :

$$y'(x) = (\lambda x + \lambda + \mu) e^x$$

En injectant dans l'équation différentielle, on en déduit, pour tout réel x :

$$y'(x) = (\lambda x + \lambda + \mu) e^x = 2(\lambda x + \mu) e^x + x e^x$$

soit :

$$(-\lambda x + \lambda - \mu) e^x = x e^x$$

et donc :

$$\lambda = -1, \quad \lambda - \mu = 0$$

Une solution particulière de l'équation avec second membre est donc :

$$x \mapsto y(x) = -(x + 1) e^x$$

► Cas d'un second membre de type « polynôme × cosinus »

Soit :

$$y'(x) = \alpha y(x) + P(x) \cos(rx)$$

où P est une fonction polynomiale, et r un réel, on se ramène au cas précédent en recherchant une solution particulière sur \mathbb{C} de l'équation différentielle :

$$y'(x) = \alpha y(x) + P(x) e^{irx}$$

Il suffit ensuite de prendre la partie réelle de la solution obtenue.

Exemple

Pour l'équation différentielle :

$$y' = y + \cos x$$

on se ramène à l'équation différentielle :

$$y' = y + e^{ix}$$

On cherche donc une solution particulière, définie sur \mathbb{R} , de la forme :

$$x \mapsto y(x) = \lambda e^{ix}$$

où λ est un nombre complexe, ce qui conduit, pour tout réel x , à :

$$y'(x) = i\lambda e^{ix}$$

En injectant dans l'équation différentielle, on en déduit, pour tout réel x :

$$y'(x) = i\lambda e^{ix} = \lambda e^{ix} + e^{ix}$$

ce qui conduit à :

$$\lambda = \frac{1}{i-1} = -\frac{1+i}{2}$$

Une solution particulière de l'équation avec second membre initiale est donc :

$$\begin{aligned} x \mapsto \operatorname{Re} \left(-\left(\frac{1+i}{2}\right) e^{ix} \right) &= \operatorname{Re} \left(-\left(\frac{1+i}{2}\right) (\cos x + i \sin x) \right) \\ &= \operatorname{Re} \left(\frac{-\cos x - i \sin x - i \cos x + \sin x}{2} \right) \\ &= \frac{-\cos x + \sin x}{2} \end{aligned}$$

► Cas d'un second membre de type « polynôme × sinus »

Soit :

$$y'(x) = \alpha y(x) + P(x) \sin(rx)$$

où P est une fonction polynomiale, et r un réel, on se ramène au cas « exponentielle × polynôme » en recherchant une solution particulière sur \mathbb{C} de l'équation différentielle :

$$y'(x) = \alpha y(x) + P(x) e^{irx}$$

Il suffit ensuite de prendre la partie imaginaire de la solution obtenue.

Exemple

Pour l'équation différentielle :

$$y' = y + \sin x$$

on se ramène à l'équation différentielle :

$$y' = y + e^{ix}$$

On cherche donc une solution particulière de la forme :

$$x \mapsto y(x) = \lambda e^{ix}$$

où λ est un nombre complexe, ce qui conduit, pour tout réel x , à :

$$y'(x) = i\lambda e^{ix}$$

En injectant dans l'équation différentielle, on en déduit, pour tout réel x :

$$y'(x) = i\lambda e^{ix} = \lambda e^{ix} + e^{ix}$$

ce qui conduit à :

$$\lambda = \frac{1}{i-1} = -\frac{1+i}{2}$$

Une solution particulière de l'équation avec second membre initiale est donc :

$$\begin{aligned} x \mapsto \operatorname{Re}\left(-\left(\frac{1+i}{2}\right) e^{ix}\right) &= \operatorname{Im}\left(-\left(\frac{1+i}{2}\right) (\cos x + i \sin x)\right) \\ &= \operatorname{Im}\left(\frac{-\cos x - i \sin x - i \cos x + \sin x}{2}\right) \\ &= \frac{-\cos x - \sin x}{2} \end{aligned}$$

3. Une méthode générale pour trouver une solution particulière de l'équation avec second membre : la méthode de variation de la constante

Soient a et b deux fonctions définies sur un intervalle I de \mathbb{R} , à valeurs dans \mathbb{R} , continues.

Cherchons une solution particulière de l'équation différentielle linéaire du premier ordre avec second membre :

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

La solution générale de l'équation homogène associée étant donnée par :

$$x \mapsto y(x) = \lambda e^{\mathcal{F}_a(x)}$$

où λ est une constante réelle arbitraire, et \mathcal{F}_a une fonction telle que, pour tout x de I :

$$\mathcal{F}_a'(x) = a(x)$$

on peut se demander s'il n'existerait pas des fonctions de la forme :

$$x \mapsto y(x) = \lambda(x) e^{\mathcal{F}_a(x)}$$

qui soient solutions de l'équation avec second membre.

En injectant cette expression dans l'équation avec second membre, on obtient, pour tout x de I :

$$(a(x) \lambda(x) + \lambda'(x) e^{\mathcal{F}_a(x)} = a(x) \lambda(x) e^{\mathcal{F}_a(x)} + b(x)$$

ce qui conduit à :

$$\lambda'(x) = b(x) e^{-\mathcal{F}_a(x)}$$

Il suffit donc de prendre, pour λ , une fonction dont $x \mapsto b(x) e^{-\mathcal{F}_a(x)}$ soit la dérivée sur I .

Exemple

Réolvons, sur $]0, \infty[$, l'équation différentielle :

$$y' = \frac{y}{x} + 1 \quad (\mathcal{E})$$

L'équation homogène associée est :

$$y' = \frac{y}{x}$$

qui admet pour solutions, sur $]0, +\infty[$, les fonctions de la forme $x \mapsto \lambda x$, où λ est une constante réelle.

À l'aide de la méthode de variation de la constante, on recherche une solution particulière de l'équation différentielle avec second membre (\mathcal{E}) , sous la forme :

$$x \mapsto y(x) = \lambda(x) x$$

On a alors, pour tout réel x strictement positif :

$$y'(x) = \lambda'(x) x + \lambda(x)$$

En injectant dans (\mathcal{E}) , on en déduit, pour tout réel x strictement positif :

$$\lambda'(x) x + \lambda(x) = \frac{1}{x} \lambda(x) x + 1 = \lambda(x) + 1$$

et donc :

$$\lambda'(x) = \frac{1}{x}$$

Par suite, pour tout réel x strictement positif :

$$\lambda(x) = \ln x + C$$

où C est une constante réelle.

L'ensemble des solutions de (\mathcal{E}) est donc l'ensemble des fonctions définies sur \mathbb{R}_+^* par :

$$x \mapsto x \ln x + C x$$

où C est une constante réelle.

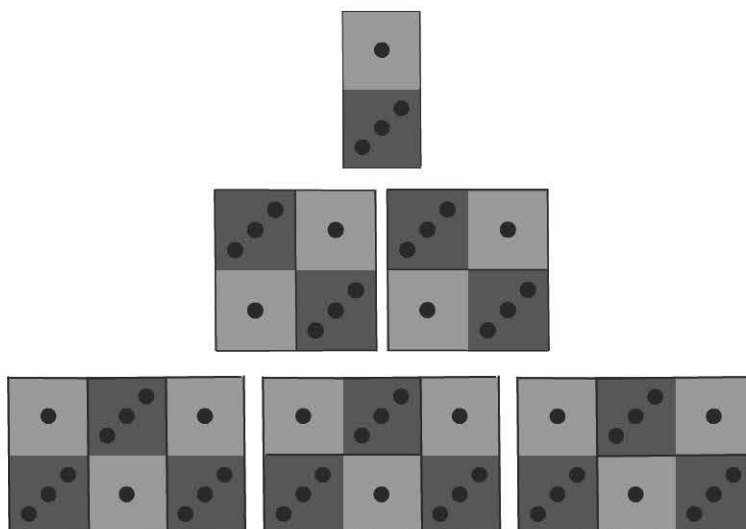


Figure 39.1 – Le pavage des dominos.

Soit n un entier naturel non nul. Considérons le nombre f_n de façons de recouvrir un échiquier ayant une hauteur de 2 carreaux, et une longueur de n carreaux (il comporte donc $2n$ carreaux), par des dominos bicolores. Chaque domino doit recouvrir deux carreaux, et les extrémités de deux dominos adjacents doivent être de couleurs différentes. Seule compte la position du domino (verticale ou horizontale), et non son sens (cela ne change rien que la face portant le numéro 1 soit en haut ou en bas). Il est clair que $f_1 = 1$ (un domino vertical), $f_2 = 2$ (deux dominos verticaux, ou deux dominos horizontaux), $f_3 = 3$ (trois dominos verticaux, ou un domino vertical et deux dominos horizontaux (deux choix possibles)) ; on s'aperçoit ensuite que, pour tout entier naturel $n \geq 2$:

$$f_{n+1} = f_{n-1} + f_n \quad (\mathcal{R})$$

Lorsque n parcourt \mathbb{N} , l'ensemble des f_n constitue donc une suite de nombres réels, appelée **suite réelle**, vérifiant la relation de récurrence (\mathcal{R}) .

Les suites sont un outil mathématique fondamental. Elles ont de nombreuses applications en physique, chimie, biologie.

Elles peuvent servir, notamment, à étudier des systèmes dynamiques, des systèmes de particules (en physique quantique, par exemple, en théorie des orbitales moléculaires, pour un système de N noyaux et n électrons...), ou encore à résoudre numériquement une équation, ou une équation différentielle.

Définition

On appelle **suite**, à valeurs dans un ensemble E , une famille d'éléments de E indexée par les entiers naturels.



Une **suite numérique**, qui est donc une liste de nombres, peut donc être considérée comme une application d'une partie de \mathbb{N} (qui peut être égale à \mathbb{N}), dans \mathbb{R} ou \mathbb{C} . On notera, dans ce qui suit, \mathbb{K} pour désigner \mathbb{R} ou \mathbb{C} . La plupart des résultats qui suivent sont présentés dans le cas des suites réelles.

On désigne, habituellement, par $(u_n)_{n \in \mathbb{N}}$ la suite numérique qui, à un entier naturel n , associe le nombre u_n .

Soit n_0 un entier naturel. On désigne, habituellement, par $(u_n)_{n \geq n_0}$ la suite numérique qui, à un entier naturel $n \geq n_0$, associe le nombre u_n .

1. Somme de suites

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites à valeurs dans le corps \mathbb{K} .

Alors, la suite somme $(w_n)_{n \in \mathbb{N}} = (u_n + v_n)_{n \in \mathbb{N}}$ est définie par :

$$\forall n \in \mathbb{N} : w_n = u_n + v_n$$

2. Multiplication d'une suite par un scalaire

Soit $(u_n)_{n \in \mathbb{N}}$ une suite à valeurs dans le corps \mathbb{K} , et $\lambda \in \mathbb{K}$.

Alors, la suite $(w_n)_{n \in \mathbb{N}} = \lambda (u_n)_{n \in \mathbb{N}}$ est définie par :

$$\forall n \in \mathbb{N} : w_n = \lambda u_n$$

3. L'espace vectoriel des suites

Théorème

L'ensemble, noté $\mathbb{K}^{\mathbb{N}}$ des suites à valeurs dans \mathbb{K} est un **\mathbb{K} -espace vectoriel**, c'est-à-dire un ensemble non vide¹, muni de l'addition et de la multiplication par un scalaire, tel que :

1. étant données deux suites $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$, à valeurs dans \mathbb{K} , la somme de ces suites est aussi une suite à valeurs dans \mathbb{K} :

$$(u_n + v_n)_{n \in \mathbb{N}} = (u_n)_{n \in \mathbb{N}} + (v_n)_{n \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$$

2. étant données deux suites $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$, et un scalaire λ , $(\lambda u_n + v_n)_{n \in \mathbb{N}}$ est aussi une suite à valeurs dans \mathbb{K} :

$$(\lambda u_n + v_n)_{n \in \mathbb{N}} = \lambda (u_n)_{n \in \mathbb{N}} + (v_n)_{n \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$$

3. étant donnée une suite $(u_n)_{n \in \mathbb{N}}$, et deux scalaires λ et μ dans \mathbb{K} , $(\lambda + \mu) (u_n)_{n \in \mathbb{N}}$ est aussi une suite à valeurs dans \mathbb{K} :

$$(\lambda + \mu) (u_n)_{n \in \mathbb{N}} = \lambda (u_n)_{n \in \mathbb{N}} + \mu (u_n)_{n \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$$

4. étant donnée une suite $(u_n)_{n \in \mathbb{N}}$, à valeurs dans \mathbb{K} , et deux scalaires λ et μ dans \mathbb{K} :

$$(\lambda \mu) (u_n)_{n \in \mathbb{N}} = \lambda (\mu u_n)_{n \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$$

1. Il est clair que la suite nulle ($u_n = 0$ pour tout entier naturel n) appartient à cet ensemble.

5. étant donnée une suite $(u_n)_{n \in \mathbb{N}}$, à valeurs dans \mathbb{K} , le réel 1 est l'élément neutre de la multiplication par un scalaire :

$$1 \times (u_n)_{n \in \mathbb{N}} = (u_n)_{n \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$$

Ainsi, $\mathbb{K}^{\mathbb{N}}$ est stable par combinaisons linéaires : **toute combinaison linéaire de suites à valeurs dans \mathbb{K} est une suite à valeurs dans \mathbb{K} .**

Définition

Un sous-ensemble \mathcal{E} de l'espace $\mathbb{K}^{\mathbb{N}}$ des suites est un **sous-espace vectoriel** s'il contient la suite nulle et s'il est stable par combinaisons linéaires, c'est-à-dire, étant données deux suites $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ de \mathcal{E} , ainsi qu'un scalaire $\lambda \in \mathbb{K}$:

$$(u_n)_{n \in \mathbb{N}} + \lambda (v_n)_{n \in \mathbb{N}} \in \mathcal{E}$$

Définition

Un sous-espace \mathcal{E} de l'espace $\mathbb{K}^{\mathbb{N}}$ est dit *de dimension finie* $N \in \mathbb{N}^*$ si il existe une base $((u_n^1)_{n \in \mathbb{N}}, (u_n^2)_{n \in \mathbb{N}}, \dots, (u_n^N)_{n \in \mathbb{N}})$ de $\mathbb{K}^{\mathbb{N}}$, permettant d'exprimer, de façon unique, toute suite $(u_n)_{n \in \mathbb{N}}$ de \mathcal{E} , sous la forme :

$$(u_n)_{n \in \mathbb{N}} = \sum_{k=1}^N s_n^k \times (u_n^k)_{n \in \mathbb{N}} \quad , \quad (s_n^1, \dots, s_n^N) \in \mathbb{K}^N$$

4. Produit de suites

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites à valeurs dans \mathbb{K} .

Alors, la suite produit $(w_n)_{n \in \mathbb{N}} = (u_n v_n)_{n \in \mathbb{N}}$ est définie par :

$$\forall n \in \mathbb{N} : w_n = u_n v_n$$

Les différents types de suites

1. Suites définies explicitement

Définition

Une suite $(u_n)_{n \in \mathbb{N}}$ est *définie explicitement* si, pour tout $n \in \mathbb{N}$, l'expression de u_n en fonction de l'entier n est connue, ce qui est le cas pour des expressions de la forme $u_n = f(n)$, où f est une fonction donnée.

Exemple

C'est le cas de la suite $(u_n)_{n \in \mathbb{N}}$ telle que, pour tout n de \mathbb{N} :

$$u_n = 1 + \frac{1}{2^n}$$

2. Suites définies par une relation de récurrence d'ordre 1

Définition

Une suite $(u_n)_{n \in \mathbb{N}}$ est *définie par une relation de récurrence d'ordre 1* si son premier terme u_0 est donné et si, pour tout entier n , u_{n+1} s'exprime en fonction de u_n , ce qui est le cas pour des expressions de la forme $u_{n+1} = f(u_n)$, où f est une fonction donnée.

Exemple

C'est le cas de la suite $(u_n)_{n \in \mathbb{N}}$ telle que, pour tout n de \mathbb{N} :

$$u_0 = 1, \quad \forall n \in \mathbb{N} : u_{n+1} = \sqrt{1 + u_n}$$

► Suite arithmétique

Définition

Une suite $(u_n)_{n \in \mathbb{N}}$ est dite **arithmétique de raison** $r \in \mathbb{K}$ si, pour tout entier naturel n :

$$u_{n+1} = u_n + r$$

► Somme des termes d'une suite arithmétique

$$\frac{\text{nombre de termes} \times (\text{premier terme} + \text{dernier terme})}{2}$$

Mathématiquement, cela se traduit de la façon suivante : pour une suite arithmétique de premier terme u_p , $p \in \mathbb{N}$, on obtient, pour tout entier naturel $n \geq p$:

$$\sum_{k=p}^n u_k = \frac{(n - p + 1) \times (u_p + u_n)}{2}$$

(de u_p à u_n , il y a exactement $n - p + 1$ termes).

► Suite géométrique

Définition

Une suite $(u_n)_{n \in \mathbb{N}}$ est dite **géométrique de raison** $q \in \mathbb{K}$ si, pour tout entier naturel n :

$$u_{n+1} = q u_n$$

► Somme des termes d'une suite géométrique de raison $\neq 1$

$$\text{premier terme} \times \frac{1 - \text{raison}^{\text{nombre de termes}}}{1 - \text{raison}}$$

Mathématiquement, cela se traduit de la façon suivante : pour une suite géométrique de raison $q \neq 1$, de premier terme u_p , $p \in \mathbb{N}$, on obtient, pour tout entier naturel $n \geq p$,

$$\sum_{k=p}^n u_k = u_p \times \frac{1 - q^{n-p+1}}{1 - q}$$

(de u_p à u_n , il y a exactement $n - p + 1$ termes).

► Suite arithmético-géométrique

Définition

Soient a et b dans \mathbb{K} tels que $a \neq 1$. Une suite $(u_n)_{n \in \mathbb{N}}$ telle que, pour tout entier naturel n :

$$u_{n+1} = a u_n + b$$

est dite **arithmético-géométrique**.

► Étude d'une suite arithmético-géométrique

On commence par déterminer s'il existe un nombre $r \in \mathbb{K}$ tel que la suite $(v_n)_{n \in \mathbb{N}}$ définie, pour tout entier naturel n , par :

$$v_n = u_n + r$$

soit géométrique de raison a . Si un tel scalaire r existe, alors, pour tout entier naturel n :

$$v_{n+1} = u_{n+1} + r = a u_n + b + r = a v_n = a u_n + a r$$

ce qui conduit donc à :

$$r = \frac{b}{a - 1}$$

Réciproquement, on vérifie que la suite $(v_n)_{n \in \mathbb{N}}$ définie, pour tout entier naturel n , par :

$$v_n = u_n + \frac{b}{a - 1}$$

est géométrique de raison a .

On peut alors en déduire, pour tout entier naturel n :

$$v_n = a^n v_0 = a^n \left(u_0 + \frac{b}{a - 1} \right)$$

ce qui conduit, pour tout entier naturel n , à :

$$u_n = v_n - \frac{b}{a-1} = a^n \left(u_0 + \frac{b}{a-1} \right) - \frac{b}{a-1} = a^n u_0 + \frac{b(a^n - 1)}{a-1}$$

3. Suites définies par une relation de récurrence d'ordre 2

Définition

Une suite $(u_n)_{n \in \mathbb{N}}$ est *définie par une relation de récurrence d'ordre 2* si ses deux premiers termes, u_0 et u_1 , sont donnés, et si, pour tout entier $n \geq 1$, u_{n+1} s'exprime en fonction de u_n et u_{n-1} , ce qui est le cas pour des expressions de la forme $u_{n+1} = g(u_n, u_{n-1})$, où g est une fonction donnée.

Exemple

C'est le cas de la suite $(u_n)_{n \in \mathbb{N}}$ telle que :

$$u_0 = 1, \quad u_1 = 1, \quad \forall n \in \mathbb{N}^* : u_{n+1} = u_n + u_{n-1}^2$$

Définition

Une suite $(u_n)_{n \in \mathbb{N}}$, à valeurs dans \mathbb{K} , est *définie par une relation de récurrence linéaire d'ordre 2* si ses deux premiers termes, u_0 et u_1 , sont donnés, et si, pour tout entier $n \geq 1$, u_{n+1} s'exprime **linéairement** en fonction de u_n et u_{n-1} , c'est-à-dire s'il existe trois scalaires a, b et c dans \mathbb{K} , $(a, c) \in \mathbb{K}^* \times \mathbb{K}^*$, tels que :

$$\forall n \in \mathbb{N}^* : a u_{n+1} + b u_n + c u_{n-1} = 0$$

Théorème

a, b et c , étant trois scalaires tels que $(a, c) \in \mathbb{K}^* \times \mathbb{K}^*$, l'ensemble $\mathcal{E}_{n;a,b,c}$ des suites $(u_n)_{n \in \mathbb{N}}$ vérifiant la relation de récurrence linéaire d'ordre 2

$$\forall n \in \mathbb{N}^* : a u_{n+1} + b u_n + c u_{n-1} = 0$$

est un \mathbb{K} -espace vectoriel de dimension 2.

Démonstration :

- Pour montrer que $\mathcal{E}_{n;a,b,c}$ est un \mathbb{K} -espace vectoriel de dimension 2, il suffit de montrer que c'est un sous-espace vectoriel de l'espace $\mathbb{K}^{\mathbb{N}}$ des suites réelles.

Il est non vide, car la suite identiquement nulle appartient à $\mathcal{E}_{n;a,b,c}$.

Il ne reste donc plus qu'à montrer la stabilité de $\mathcal{E}_{n;a,b,c}$ par combinaisons linéaires ; à cet effet, on considère deux suites $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ de $\mathcal{E}_{n;a,b,c}$, ainsi qu'un scalaire λ .

On a, clairement, pour tout entier naturel non nul n :

$$\begin{aligned} a(u_{n+1} + \lambda v_{n+1}) + b(u_n + \lambda v_n) + c(u_{n-1} + \lambda v_{n-1}) &= a u_{n+1} + b u_n + c u_{n-1} \\ &\quad + \lambda (a v_{n+1} + b v_n + c v_{n-1}) = 0 \end{aligned}$$

Ainsi, la suite $(u_n)_{n \in \mathbb{N}} + \lambda (v_n)_{n \in \mathbb{N}}$ appartient bien à $\mathcal{E}_{n;a,b,c}$.

- Pour montrer que $\mathcal{E}_{n;a,b,c}$ est de dimension 2, on va montrer que l'application linéaire, qui, à toute suite de $\mathcal{E}_{n;a,b,c}$, associe ses deux premiers termes, est bijective (c'est un isomorphisme, c'est-à-dire une application linéaire injective et surjective) ; les deux espaces seront donc isomorphes, ce qui garantit l'égalité de leurs dimensions respectives.

Soit donc l'application :
$$\varphi : \mathcal{E}_{n;a,b,c} \rightarrow \mathbb{K}^2$$
$$(u_n)_{n \in \mathbb{N}} \mapsto (u_0, u_1)$$

Il est clair que l'application φ est linéaire.

Le noyau de l'application φ , c'est-à-dire l'ensemble des suites $(u_n)_{n \in \mathbb{N}}$ dont les deux premiers termes sont nuls, est réduit à la suite nulle : φ est donc injective.

L'application φ est surjective, car, si on se donne deux réels u_0 et u_1 , on peut définir une suite de $\mathcal{E}_{n;a,b,c}$ dont les deux premiers termes soient u_0 et u_1 .

L'application φ est donc bien un isomorphisme : \mathbb{K}^2 étant de dimension 2, il en est donc de même de $\mathcal{E}_{n;a,b,c}$. ■

Afin de pouvoir exprimer, le plus facilement possible, les termes d'une suite vérifiant une relation de récurrence linéaire d'ordre 2 de la forme (40.1), il peut être intéressant de déterminer une base de l'espace $\mathcal{E}_{n;a,b,c}$. À cet effet, on commence par rechercher les suites géométriques (distinctes de la suite identiquement nulle) qui satisfont ce type de relations ; on cherche donc à déterminer les scalaires r , supposés non nuls, tels que :

$$\forall n \in \mathbb{N}^* : ar^{n+1} + br^n + cr^{n-1} = 0$$

r étant non nul, il en résulte : $ar^2 + br + c = 0$. Par suite :

- si le discriminant $\Delta = b^2 - 4ac$ est non nul, le trinôme $ar^2 + br + c$ admet deux racines distinctes r_1 et r_2 dans \mathbb{C} .

Les suites géométriques $(r_1^n)_{n \in \mathbb{N}}$ et $(r_2^n)_{n \in \mathbb{N}}$ vérifient bien la relation de récurrence linéaire d'ordre 2 donnée, et sont linéairement indépendantes. Elles forment donc une base de $\mathcal{E}_{n;a,b,c}$.

Ainsi, toute suite $(u_n)_{n \in \mathbb{N}}$ vérifiant cette relation de récurrence est telle que :

$$\forall n \in \mathbb{N} : u_n = \lambda r_1^n + \mu r_2^n, \quad (\lambda, \mu) \in \mathbb{C}^2$$

Si la suite est réelle, et si les deux racines sont complexes, de la forme $r_1 = \rho e^{i\theta}$ et $r_2 = \rho e^{-i\theta}$, toute suite $(u_n)_{n \in \mathbb{N}}$ vérifiant cette relation de récurrence est aussi telle que :

$$\forall n \in \mathbb{N} : u_n = \alpha \rho^n \cos(n\theta) + \beta \rho^n \sin(n\theta), \quad (\alpha, \beta) \in \mathbb{R}^2$$

- si le discriminant $\Delta = b^2 - 4ac$ est nul, le trinôme $ar^2 + br + c$ admet une racine double $r_0 \in \mathbb{K}$.

Les suites $(r_0^n)_{n \in \mathbb{N}}$ et $(nr_0^n)_{n \in \mathbb{N}}$ vérifient bien la relation de récurrence linéaire d'ordre 2 donnée, et sont linéairement indépendantes. Elles forment donc une base de $\mathcal{E}_{n;a,b,c}$.

Ainsi, toute suite $(u_n)_{n \in \mathbb{N}}$ vérifiant cette relation de récurrence est telle que :

$$\forall n \in \mathbb{N} : u_n = (\lambda + \mu n) r_0^n, \quad (\lambda, \mu) \in \mathbb{K}^2$$

Dans tous les cas, pour ce type de suite, il y a donc deux constantes a priori inconnues, λ et μ . La connaissance de deux termes de la suite peut permettre de lever l'indétermination. Lorsque ce sont les deux premiers termes de la suite qui sont connus, on parle de **conditions initiales**.



► Équation caractéristique

Définition

Étant donnée une suite $(u_n)_{n \in \mathbb{N}}$, à valeurs dans \mathbb{K} , vérifiant la relation de récurrence linéaire d'ordre 2 :

$$\forall n \in \mathbb{N}^* : a u_{n+1} + b u_n + c u_{n-1} = 0 \quad , \quad (a, c) \in \mathbb{K}^* \times \mathbb{K}^* \quad , \quad b \in \mathbb{K}$$

l'équation $a r^2 + b r + c = 0$ est appelée **équation caractéristique**.

Exemples

1. Considérons la suite $(u_n)_{n \in \mathbb{N}}$ définie par :

$$u_0 = 1 \quad , \quad u_1 = 2 \quad , \quad \forall n \in \mathbb{N}^* : u_{n+1} = 2 u_n + 3 u_{n-1}$$

L'équation caractéristique est : $r^2 - 2r - 3 = 0$

-1 et 3 sont racines évidentes.

Ainsi, il existe deux réels λ et μ tels que : $\forall n \in \mathbb{N} : u_n = \lambda(-1)^n + \mu 3^n$

Les conditions sur u_0 et u_1 conduisent alors à : $u_0 = \lambda + \mu = 1 \quad , \quad u_1 = -\lambda + 3\mu = 2$

puis : $\lambda = \frac{1}{4} \quad , \quad \mu = \frac{3}{4}$

Par suite : $\forall n \in \mathbb{N} : u_n = \frac{(-1)^n + 3^{n+1}}{4}$

2. Considérons la suite $(v_n)_{n \in \mathbb{N}}$ définie par :

$$v_0 = 1 \quad , \quad v_1 = -1 \quad , \quad \forall n \in \mathbb{N} : v_{n+2} + v_{n+1} + v_n = 0$$

L'équation caractéristique est : $r^2 + r + 1 = 0$, de racines complexes $j = e^{\frac{2i\pi}{3}}$ et $j^2 = e^{\frac{4i\pi}{3}} = \bar{j}$.

Comme $|j| = |j^2| = 1$, il existe deux réels λ et μ tels que :

$$\forall n \in \mathbb{N} : v_n = \lambda \cos\left(\frac{2n\pi}{3}\right) + \mu \sin\left(\frac{2n\pi}{3}\right)$$

Les conditions sur v_0 et v_1 conduisent alors à :

$$v_0 = \lambda = 1 \quad , \quad v_1 = \lambda \cos\left(\frac{2\pi}{3}\right) + \mu \sin\left(\frac{2\pi}{3}\right) = -\lambda \frac{1}{2} + \mu \frac{\sqrt{3}}{2} = -1$$

puis : $\mu = -\frac{1}{\sqrt{3}}$. Par suite : $\forall n \in \mathbb{N} : v_n = \cos\left(\frac{2n\pi}{3}\right) - \frac{1}{\sqrt{3}} \sin\left(\frac{2n\pi}{3}\right)$

4. Suites définies par une relation de récurrence d'ordre p , $p \in \mathbb{N}^*$

Définition

Une suite $(u_n)_{n \in \mathbb{N}}$ est définie par une relation de récurrence d'ordre p , $p \in \mathbb{N}^*$, si ses p premiers termes, u_0, u_1, \dots, u_{p-1} , sont donnés, et si, pour tout entier $n \geq p$, u_{n+1} s'exprime en fonction de $u_n, u_{n-1}, \dots, u_{n-p+1}$, ce qui est le cas pour des expressions de la forme $u_{n+1} = h(u_n, u_{n-1}, \dots, u_{n-p+1})$, où h est une fonction donnée.

Une « start-up » emprunte un capital C , $C \in \mathbb{R}_+^*$.

On désigne par u_n le capital dû à la fin de l'année n ($n \in \mathbb{N}^*$), par i le taux annuel ($i \in [0, 1]$), et par M le montant d'une mensualité ($M \in \mathbb{R}_+^*$).

L'amortissement de l'emprunt est supposé constant.

On a alors :

$$\forall n \in \mathbb{N} : u_{n+1} = (1 + i)u_n - 12M$$

En reprenant la technique d'étude explicitée précédemment, on obtient, pour tout entier naturel n :

$$u_n = (1 + i)^n C - \frac{12M((1 + i)^n - 1)}{i}$$

Soit N le nombre d'années nécessaires pour rembourser l'emprunt ($N \in \mathbb{N}^*$).

Au bout des N années (on est donc à l'année $N + 1$), le capital dû est donc nul :

$$u_{N+1} = 0$$

soit

$$(1 + i)^N C - \frac{12M((1 + i)^N - 1)}{i} = 0$$

On obtient alors le montant d'une mensualité :

$$M = \frac{i(1 + i)^N C}{12((1 + i)^N - 1)} = \frac{iC}{12(1 - (1 + i)^{-N})}$$

Cette dernière formule est bien connue des financiers, qui l'utilisent très fréquemment.

Pour étudier une suite, il est important de pouvoir étudier – et démontrer – ses propriétés, mais aussi de pouvoir comparer ses termes les uns avec les autres, et en particulier savoir s'ils augmentent ou diminuent...

1. Récurrence simple

Considérons une suite $(u_n)_{n \in \mathbb{N}}$. Si jamais on souhaite démontrer qu'elle vérifie une propriété, que nous noterons \mathcal{P} , démontrer celle-ci par récurrence peut s'avérer extrêmement intéressant.

En quoi consiste cette méthode ? Tout simplement, on va commencer par regarder si la propriété \mathcal{P} est vérifiée à un rang initial $n_0 \in \mathbb{N}$ donné (n_0 peut, bien sûr, être égal à zéro) ; si c'est le cas, on suppose ensuite qu'elle est vraie à un rang $n \geq n_0$ quelconque, et on cherche à déterminer si elle est encore vraie au rang $n + 1$: le fait qu'elle soit vraie au rang n_0 permettra d'en déduire qu'elle est vérifiée pour tout entier $n \geq n_0$.

Exemple

Considérons la suite $(v_n)_{n \in \mathbb{N}}$ définie par :

$$\forall n \in \mathbb{N} : v_n = \sum_{k=0}^n k^2$$

On remarque que :

$$v_0 = 0 \quad , \quad v_1 = 0 + 1 = 1 \quad , \quad v_2 = 0 + 1 + 4 = 5 \quad , \quad v_3 = 0 + 1 + 4 + 9 = 14$$

Démontrons alors, par récurrence, que :

$$\forall n \in \mathbb{N} : \sum_{k=0}^n k^2 = \frac{n(n+1)(2n+1)}{6}$$

- La propriété est bien vraie au rang 0.
- Supposons la propriété vraie à un rang $n > 0$; on a alors :

$$\begin{aligned} \sum_{k=0}^{n+1} k^2 &= \sum_{k=0}^n k^2 + (n+1)^2 \\ &= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \\ &= \frac{(n+1)(n(2n+1) + 6(n+1))}{6} \\ &= \frac{(n+1)(n+2)(2n+3)}{6} \\ &= \frac{(n+1)(n+2)(2(n+1)+1)}{6} \end{aligned}$$

(On factorise par $n + 2$, car -2 est racine évidente.)

La propriété est donc vraie au rang $n + 1$.

Comme elle est vraie au rang 0, elle est donc vraie pour tout entier naturel n .



Le raisonnement par récurrence permet de démontrer une propriété vérifiée, suivant les cas, pour tous les entiers naturels, dans d'autres, pour une infinité d'entiers naturels, à partir d'un certain rang.

2. Récurrence forte

Considérons une suite $(u_n)_{n \in \mathbb{N}}$. On cherche toujours à démontrer une propriété \mathcal{P} pour une suite $(u_n)_{n \in \mathbb{N}}$.

Comme pour la récurrence simple, on commence par regarder si la propriété \mathcal{P} est vérifiée à un rang initial $n_0 \in \mathbb{N}$ donné (n_0 peut, bien sûr, être égal à zéro) ; si c'est le cas, on suppose ensuite qu'elle est vraie **jusqu'à un rang** $n \geq n_0$ quelconque, et on cherche à déterminer si elle est encore vraie au rang $n + 1$: si oui, le fait qu'elle soit vraie au rang n_0 permettra d'en déduire qu'elle est vérifiée pour tout entier $n \geq n_0$.

Cette récurrence qui semble, en apparence, plus forte que la récurrence simple, lui est en fait équivalente, dans la mesure où elle revient à démontrer par récurrence simple la propriété à tout rang $k \leq n$.

3. Récurrence double

Il s'agit, cette fois-ci, d'un autre type de démonstration par récurrence : on cherche toujours à démontrer une propriété \mathcal{P} pour une suite $(u_n)_{n \in \mathbb{N}}$, mais, pour commencer, au lieu de regarder si \mathcal{P} est vérifiée à un rang initial $n_0 \in \mathbb{N}$ donné, on regarde si elle est vraie aux rangs n_0 et $n_0 + 1$ (n_0 peut toujours être égal à zéro) ; si c'est le cas, on suppose ensuite que \mathcal{P} est vraie aux rangs $n \geq n_0$ et $n + 1$, pour n quelconque, et on cherche à déterminer si elle est encore vraie aux rangs $n + 1$ et $n + 2$: si oui, le fait qu'elle soit vraie aux rangs n_0 et $n_0 + 1$ permettra d'en déduire qu'elle est vérifiée pour tout entier $n \geq n_0$.

Exemple

Soit θ un réel.

Démontrons par récurrence que, pour tout entier naturel non nul n , $\cos(n\theta)$ est un polynôme, noté T_n , de degré n en $\cos \theta$, de coefficient dominant 2^{n-1} .

- **La propriété est vraie au rang 1 :**

$\cos \theta$ est bien un polynôme de degré 1 en $\cos \theta$, de coefficient dominant $2^{1-1} = 1$.

- **La propriété est vraie au rang 2 :**

$\cos(2\theta) = 2 \cos^2 \theta - 1$ est bien un polynôme de degré 2 en $\cos \theta$, de coefficient dominant $2^{2-1} = 2$.

- **Supposons la vraie jusqu'à un rang $n > 1$.**

Les formules d'addition permettent alors d'écrire :

$$\cos((n+1)\theta) = \cos(n\theta) \cos \theta - \sin(n\theta) \sin \theta$$

et :

$$\cos((n-1)\theta) = \cos(n\theta) \cos \theta + \sin(n\theta) \sin \theta$$

Par suite, en additionnant membre à membre ces deux relations, on en déduit :

$$\cos((n+1)\theta) + \cos((n-1)\theta) = 2 \cos(n\theta) \cos \theta$$

puis :

$$\cos((n+1)\theta) = 2 \cos(n\theta) \cos \theta - \cos((n-1)\theta) \quad (\star)$$

Par hypothèse de récurrence, $\cos(n\theta)$ est un polynôme de degré n en $\cos \theta$. $\cos(n\theta) \cos \theta$ est donc un polynôme de degré $n+1$ en $\cos \theta$. Comme $\cos((n-1)\theta)$ est un polynôme de degré $n-1$ en $\cos \theta$, $\cos((n+1)\theta)$ est bien un polynôme de degré $n+1$ en $\cos \theta$.

Le coefficient du terme de plus haut degré est :

$$2 \times 2^{n-1} = 2^n = 2^{n+1-1}$$

La propriété est donc vraie au rang $n+1$.

- Comme elle est vraie au rang 1, elle est donc vraie pour tout entier naturel $n \geq 1$.

T_n est le $n^{\text{ième}}$ polynôme de Tchebychev¹. Il est à noter que ce polynôme ne dépend pas de θ .
L'unicité de ce polynôme est admise.

1. Majorant et minorant d'une partie non vide de \mathbb{R}

► Majorant

Soit \mathcal{P} une partie non vide de \mathbb{R} . Un réel $M_{\mathcal{P}}$ est un **majorant** de \mathcal{P} si, pour tout x de \mathcal{P} : $x \leq M_{\mathcal{P}}$.

Exemple

Considérons l'ensemble $\left\{2 + \frac{1}{n}, n \in \mathbb{N}^*\right\}$: 3 est un majorant de cet ensemble.

► Minorant

Soit \mathcal{P} une partie non vide de \mathbb{R} . Un réel $m_{\mathcal{P}}$ est un **minorant** de \mathcal{P} si, pour tout x de \mathcal{P} : $x \geq m_{\mathcal{P}}$.

Exemple

2 est un minorant de l'ensemble $\left\{2 + \frac{1}{n}, n \in \mathbb{N}^*\right\}$.

► Deux propriétés fondamentales de l'ensemble des réels

Toute partie \mathcal{P}_M non vide et majorée de \mathbb{R} admet une borne supérieure finie. Cette borne supérieure est le plus petit des majorants de \mathcal{P}_M .

De même, toute partie \mathcal{P}_m non vide et minorée de \mathbb{R} admet une borne inférieure finie. Cette borne inférieure est le plus grand des minorants de \mathcal{P}_m .

Exemple

Considérons l'ensemble $\left\{2 + \frac{1}{2^n}, n \in \mathbb{N}\right\}$. Il est bien non vide. $2 + \frac{1}{2^0} = 3$ est sa borne supérieure, qui est atteinte, et 2 sa borne inférieure, qui n'est pas atteinte.

2. Le cas des suites

► Suite majorée

Une suite réelle $(u_n)_{n \in \mathbb{N}}$ est dite **majorée** s'il existe un réel M tel que :

$$\forall n \in \mathbb{N} : u_n \leq M$$

► Suite minorée

Une suite réelle $(u_n)_{n \in \mathbb{N}}$ est dite **minorée** s'il existe un réel m tel que :

$$\forall n \in \mathbb{N} : u_n \geq m$$



Le majorant positif d'une suite réelle n'est **jamais unique** : si la suite $(u_n)_{n \in \mathbb{N}}$ est majorée par le réel positif M , elle l'est aussi par le réel $2M$, ou encore par le réel $3M$, ou encore, par $M+1$, $M+2$, ...

$$\forall n \in \mathbb{N} : u_n \leq M \leq 2M \leq 3M \leq \dots$$

De même, le minorant positif d'une suite n'est pas unique ! si la suite $(u_n)_{n \in \mathbb{N}}$ est minorée par le réel positif m , elle l'est aussi par le réel $\frac{m}{2}$, ou encore par le réel $\frac{m}{2^2}$, ...

$$\forall n \in \mathbb{N} : u_n \geq m \geq \frac{m}{2} \geq \frac{m}{2^2} \geq \dots$$

► Suite bornée

Une suite réelle $(u_n)_{n \in \mathbb{N}}$ est dite **bornée** si et seulement si elle est à la fois majorée et minorée, c'est-à-dire s'il existe deux réels m et M tels que :

$$\forall n \in \mathbb{N} : m \leq u_n \leq M$$

ou, de façon équivalente, s'il existe une constante positive C telle que :

$$\forall n \in \mathbb{N} : |u_n| \leq C$$

Une suite complexe $(u_n)_{n \in \mathbb{N}}$ est dite **bornée** si et seulement si les suites $(\operatorname{Re}(u_n))_{n \in \mathbb{N}}$ et $(\operatorname{Im}(u_n))_{n \in \mathbb{N}}$ sont bornées.

► Suite positive

Une suite réelle $(u_n)_{n \in \mathbb{N}}$ est dite **positive** si : $\forall n \in \mathbb{N} : u_n \geq 0$.

► Suite négative

Une suite réelle $(u_n)_{n \in \mathbb{N}}$ est dite **négative** si : $\forall n \in \mathbb{N} : u_n \leq 0$.

3. Croissance et décroissance

Une suite réelle $(u_n)_{n \in \mathbb{N}}$ est dite :

- croissante à partir du rang $n_0 \in \mathbb{N}$ si : $\forall n \geq n_0 : u_n \leq u_{n+1}$.
- décroissante à partir du rang $n_0 \in \mathbb{N}$ si : $\forall n \geq n_0 : u_n \geq u_{n+1}$.
- stationnaire à partir du rang $n_0 \in \mathbb{N}$ si : $\forall n \geq n_0 : u_n = u_{n+1}$.

Une suite réelle $(u_n)_{n \in \mathbb{N}}$ est dite :

- strictement croissante à partir du rang $n_0 \in \mathbb{N}$ si :

$$\forall n \geq n_0 : u_n < u_{n+1}$$

- strictement décroissante à partir du rang $n_0 \in \mathbb{N}$ si :

$$\forall n \geq n_0 : u_n > u_{n+1}$$

Une suite réelle ou complexe $(u_n)_{n \in \mathbb{N}}$ est dite **constante** si :

$$\forall n \in \mathbb{N} : u_n = u_{n+1}$$

Une suite réelle $(u_n)_{n \in \mathbb{N}}$ est dite **monotone à partir du rang** $n_0 \in \mathbb{N}$ si elle est croissante ou décroissante à partir du rang n_0 .

Une suite réelle $(u_n)_{n \in \mathbb{N}}$ est dite **strictement monotone à partir du rang** $n_0 \in \mathbb{N}$ si elle est strictement croissante ou strictement décroissante à partir du rang n_0 .

1. Étude de la différence $u_{n+1} - u_n$

Exemple

On considère la suite $(u_n)_{n \in \mathbb{N}}$ définie par :

$$\forall n \in \mathbb{N} : u_n = 1 + \frac{1}{2^n}$$

Alors, pour tout entier naturel n :

$$u_{n+1} - u_n = 1 + \frac{1}{2^{n+1}} - 1 - \frac{1}{2^n} = \frac{1}{2^n} \left(\frac{1}{2} - 1 \right) = -\frac{1}{2^{n+1}} < 0$$

La suite $(u_n)_{n \in \mathbb{N}}$ est donc strictement décroissante.

2. Étude du quotient $\frac{u_{n+1}}{u_n}$ dans le cas d'une suite à termes strictement positifs

Exemple

On considère la suite $(u_n)_{n \in \mathbb{N}}$ définie par :

$$\forall n \in \mathbb{N} : u_n = \frac{1}{2^n}$$

Alors, pour tout entier naturel n :

$$\frac{u_{n+1}}{u_n} = \frac{1}{2} < 1$$

La suite $(u_n)_{n \in \mathbb{N}}$ est donc strictement décroissante.

3. Étude de suites réelles définies par une relation de récurrence de la forme $u_{n+1} = f(u_n)$

La connaissance des variations de la fonction $x \mapsto f(x) - x$ peut être très utile pour l'étude de la suite $(u_n)_{n \in \mathbb{N}}$.

Exemple

On considère la suite $(u_n)_{n \in \mathbb{N}}$ définie par :

$$u_0 = 1, \quad \forall n \in \mathbb{N} : u_{n+1} = \sqrt{u_n + 1}$$

Alors, pour tout entier naturel n :

$$u_{n+1} - u_n = f(u_n) - u_n$$

Par récurrence immédiate, on constate donc que la suite $(u_n)_{n \in \mathbb{N}}$ est à termes positifs. Étudions alors la fonction $g : \mathbb{R}^+ \rightarrow \mathbb{R}, x \mapsto \sqrt{x+1} - x$.

La fonction g est définie, continue et dérivable sur \mathbb{R}^+ :

$$\forall x \in \mathbb{R}^+ : g'(x) = \frac{1}{2\sqrt{x+1}} - 1 = \frac{1 - 2\sqrt{x+1}}{2\sqrt{x+1}}$$

Comme, pour tout réel positif x : $\sqrt{x+1} \geq 1$, $g'(x)$ ne s'annule jamais sur \mathbb{R}^+ , où elle prend des valeurs négatives.

La fonction g est donc strictement décroissante sur \mathbb{R}^+ . Soit φ la racine positive de l'équation $g(x) = x$.

Si $0 < u_0 < \varphi$, la suite $(u_n)_{n \in \mathbb{N}}$ croît vers φ ; si $u_0 > \varphi$, la suite $(u_n)_{n \in \mathbb{N}}$ décroît vers φ .

4. Un cas particulier : les suites homographiques

Définition

Étant donnés quatre scalaires a, b, c et d tels que :

$$ad - bc \neq 0, \quad c \neq 0$$

on appelle **suite homographique** une suite $(u_n)_{n \in \mathbb{N}}$ vérifiant une relation de récurrence de la forme :

$$\forall n \in \mathbb{N} : u_{n+1} = \frac{au_n + b}{cu_n + d}$$



Le nom vient du fait qu'une application de la forme :

$$x \in \mathbb{K} \setminus \left\{ -\frac{d}{c} \right\} \mapsto \frac{ax + b}{cx + d}$$

avec

$$ad - bc \neq 0, \quad c \neq 0$$

est une **homographie**, c'est-à-dire une application du plan complexe dans lui-même, qui laisse invariant l'ensemble des droites et des cercles de celui-ci (on parle de transformation projective bijective). Une homographie s'obtient comme la composée de translations, de rotations, d'homothéties et éventuellement d'une inversion. La condition précédente assure, tout simplement, que la fonction n'est pas constante. En effet, dans le cas où $ad - bc = 0$, on obtient, pour tout x de $\mathbb{K} \setminus \left\{ -\frac{d}{c} \right\}$:

$$\frac{ax + b}{cx + d} = \frac{acx + bc}{c(cx + d)} = \frac{acx + ad}{c(cx + d)} = \frac{a(cx + d)}{c(cx + d)} = \frac{a}{c}$$

Proposition

On considère quatre scalaires a, b, c et d tels que :

$$ad - bc \neq 0, \quad c \neq 0$$

La **suite homographique** $(u_n)_{n \in \mathbb{N}}$ telle que :

$$\forall n \in \mathbb{N} : u_{n+1} = \frac{au_n + b}{cu_n + d}$$

est définie si et seulement si :

$$u_0 \neq -\frac{d}{c}$$

1. Définitions

► Suite convergente

Une suite réelle $(u_n)_{n \in \mathbb{N}}$ converge vers une limite finie $\ell \in \mathbb{R}$ si, pour tout réel $\varepsilon > 0$, il existe un rang n_0 tel que :

$$\forall n \geq n_0 : |u_n - \ell| \leq \varepsilon$$

La suite $(u_n)_{n \in \mathbb{N}}$ est dite **convergente** (ou converge), de limite

$$\ell = \lim_{n \rightarrow +\infty} u_n$$



1. Le fait d'écrire « pour tout $\varepsilon > 0$ » signifie que la quantité ε , qui est positive, peut être choisie aussi petite que l'on veut, et permet donc de comprendre de façon « intuitive » la notion de limite.
2. Le choix de la quantité ε conditionne celle de l'entier n_0 : une fois ε fixé, n_0 l'est aussi.

Une suite complexe $(u_n)_{n \in \mathbb{N}}$ converge vers une limite finie $\ell \in \mathbb{C}$ si et seulement si les suites $(\operatorname{Re}(u_n))_{n \in \mathbb{N}}$ et $(\operatorname{Im}(u_n))_{n \in \mathbb{N}}$ convergent respectivement vers $\operatorname{Re}(\ell)$ et $\operatorname{Im}(\ell)$.

► Suite divergente

Une suite réelle $(u_n)_{n \in \mathbb{N}}$ diverge vers $+\infty$ si, pour tout réel strictement positif A , il existe un rang n_0 tel que :

$$\forall n \geq n_0 : u_n \geq A$$

On écrit alors :

$$\lim_{n \rightarrow +\infty} u_n = +\infty$$

Une suite réelle $(u_n)_{n \in \mathbb{N}}$ diverge vers $-\infty$ si, pour tout réel strictement positif A , il existe un rang n_0 tel que :

$$\forall n \geq n_0 : u_n \leq -A$$

On écrit alors :

$$\lim_{n \rightarrow +\infty} u_n = -\infty$$

Une suite $(u_n)_{n \in \mathbb{N}}$ **qui n'a pas de limite finie** est dite **divergente**.



Une suite réelle qui a pour limite $+\infty$ est ainsi une suite divergente. Mais ce n'est pas parce qu'une suite diverge qu'elle tend vers $+$ ou $-\infty$!

Exemple

La suite $(v_n)_{n \in \mathbb{N}}$ définie par :

$$\forall n \in \mathbb{N} : v_n = 1 + (-1)^n$$

diverge, mais ne tend pas vers $+\infty$ ou $-\infty$; elle est en effet bornée

$$\forall n \in \mathbb{N} : |v_n| \leq 2$$

2. Propriétés fondamentales

Propriété

Si la limite d'une suite existe, alors elle est unique.

Démonstration : On démontre ce résultat par l'absurde, dans le cas d'une suite réelle convergente $(u_n)_{n \in \mathbb{N}}$.

On suppose qu'il existe deux réels ℓ_1 et ℓ_2 tels que la suite $(u_n)_{n \in \mathbb{N}}$ converge vers ℓ_1 et ℓ_2 avec $\ell_1 \neq \ell_2$.

Soit $\varepsilon > 0$. Il existe alors deux entiers n_1 et n_2 tels que :

$$\forall n \geq n_1 : |u_n - \ell_1| \leq \frac{\varepsilon}{2}$$

$$\forall n \geq n_2 : |u_n - \ell_2| \leq \frac{\varepsilon}{2}$$

On a alors, pour tout $n \geq \max(n_1, n_2)$:

$$|\ell_1 - \ell_2| = |\ell_1 - u_n + u_n - \ell_2| \leq |\ell_1 - u_n| + |u_n - \ell_2| \leq 2 \frac{\varepsilon}{2} = \varepsilon$$

ε étant quelconque, on obtient, pour $\varepsilon = \frac{|\ell_1 - \ell_2|}{2}$:

$$|\ell_1 - \ell_2| \leq \frac{|\ell_1 - \ell_2|}{2}$$

Si $\ell_1 \neq \ell_2$, c'est impossible. Il en résulte :

$$\ell_1 = \ell_2$$

■

Propriété

Une suite convergente est bornée.

Démonstration : On démontre ce résultat dans le cas d'une suite réelle.

Soit ainsi $(u_n)_{n \in \mathbb{N}}$ une suite réelle convergente, de limite $\ell \in \mathbb{R}$.

Soit $\varepsilon > 0$. Il existe un entier n_0 tel que :

$$\forall n \geq n_0 : |u_n - \ell| \leq \varepsilon$$

Par suite, pour tout entier $n \geq n_0$:

$$|u_n| = |u_n - \ell + \ell| \leq |u_n - \ell| + |\ell| \leq \varepsilon + |\ell|$$

Ainsi, en considérant

$$M = \max\{|u_0|, |u_1|, \dots, |u_{n_0-1}|, \varepsilon + |\ell|\}$$

on en déduit, pour tout entier naturel n :

$$|u_n| \leq M$$

qui est le résultat cherché !

■



La réciproque est fausse ! Une suite bornée n'est pas nécessairement convergente. C'est le cas, par exemple, de la suite $(u_n)_{n \in \mathbb{N}}$, définie par :

$$\forall n \in \mathbb{N} : u_n = 1 + (-1)^n$$

La suite $(u_n)_{n \in \mathbb{N}}$ est bien bornée, puisque, pour tout entier naturel n :

$$|u_n| \leq 2$$

Elle ne converge pas, car $(-1)^n$ n'est pas le terme d'une suite convergente (les valeurs prises étant, alternativement, -1 et 1).

3. De l'utilité des suites pour l'étude des fonctions : caractérisation séquentielle de la limite

Théorème

Soient f une fonction définie sur un intervalle I de \mathbb{R} , et x_0 un réel donné dans I . Il y a équivalence entre les propriétés suivantes :

- $\lim_{x \rightarrow x_0} f(x) = \ell$
- Pour toute suite réelle $(u_n)_{n \in \mathbb{N}}$, à valeurs dans I , de limite x_0 , la suite $(f(u_n))_{n \in \mathbb{N}}$ converge vers ℓ .



Ce résultat est extrêmement puissant, dans la mesure où, pour une application continue $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$, et une suite $(u_n)_{n \in \mathbb{N}}$, à valeurs dans I , convergeant vers une limite $\ell_u \in I$, on peut en déduire le résultat suivant :

$$\lim_{n \rightarrow +\infty} f(u_n) = f\left(\lim_{n \rightarrow +\infty} u_n\right) = f(\ell_u)$$

Convergence des suites monotones

Propriété

Une suite réelle croissante et majorée converge.

Démonstration : Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle croissante et majorée. Désignons par \mathcal{B}_s la borne supérieure (finie) de l'ensemble $\{u_n, n \in \mathbb{N}\}$.

Comme \mathcal{B}_s est le plus petit des majorants de cet ensemble, alors, pour tout $\varepsilon > 0$, $\mathcal{B}_s - \varepsilon$, qui est plus petit que \mathcal{B}_s , n'est pas un majorant. Ainsi, il existe un entier n_0 tel que :

$$\mathcal{B}_s - \varepsilon \leq u_{n_0} \leq \mathcal{B}_s$$

Or, la suite $(u_n)_{n \in \mathbb{N}}$ étant croissante, on aura donc, pour tout entier $n \geq n_0$:

$$\mathcal{B}_s - \varepsilon \leq u_{n_0} \leq u_n \leq \mathcal{B}_s$$

On retrouve ainsi la définition de la limite d'une suite : $(u_n)_{n \in \mathbb{N}}$ converge donc vers \mathcal{B}_s . ■

Propriété

Une suite réelle décroissante et minorée converge.

Démonstration : Exercice. ■

1. Convergence des suites réelles définies par une relation de récurrence de la forme $u_{n+1} = f(u_n)$

Théorème

Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle définie par une relation de récurrence de la forme $u_{n+1} = f(u_n)$, où f est une fonction donnée.

Si $(u_n)_{n \in \mathbb{N}}$ converge vers une limite ℓ , et si la fonction f est continue en ℓ , alors :

$$\ell = f(\ell)$$

Démonstration : On suppose donc que la suite $(u_n)_{n \in \mathbb{N}}$ converge vers le réel ℓ ; ainsi, pour tout $\varepsilon > 0$, il existe un rang n_0 tel que :

$$\forall n \geq n_0 : |u_{n+1} - \ell| \leq \frac{\varepsilon}{2}$$

soit :

$$\forall n \geq n_0 : |f(u_n) - \ell| \leq \frac{\varepsilon}{2}$$

Comme la suite $(f(u_n))_{n \in \mathbb{N}}$ converge vers $f(\ell)$, il existe un rang n_1 à partir duquel :

$$\forall n \geq n_1 : |f(u_n) - f(\ell)| \leq \frac{\varepsilon}{2}$$

soit :

$$\forall n \geq n_1 : |u_{n+1} - f(\ell)| \leq \frac{\varepsilon}{2}$$

On a alors, pour tout entier n tel que $n \geq \max(n_0, n_1)$:

$$|\ell - f(\ell)| = |\ell - u_{n+1} + u_{n+1} - f(\ell)| \leq |\ell - u_{n+1}| + |u_{n+1} - f(\ell)| \leq 2 \frac{\varepsilon}{2} = \varepsilon$$

Le réel ε ayant été choisi aussi petit que l'on veut, on a donc, nécessairement :

$$\ell = f(\ell)$$

■

2. Théorème du point fixe (de Banach)

Théorème

Soit f une fonction définie sur un intervalle I de \mathbb{R} , à valeurs dans I , **contractante**¹.

Alors, la suite $(u_n)_{n \in \mathbb{N}}$ définie par :

$$\forall n \in \mathbb{N} : u_{n+1} = f(u_n)$$

converge vers l'unique point fixe ℓ de f appartenant à l'intervalle I .

De plus, pour tout entier naturel n :

$$|u_n - \ell| \leq k^n |u_0 - \ell|$$

Démonstration : On admettra l'existence et l'unicité du point fixe ℓ de f (cela est dû au fait que f est contractante).

Pour tout entier naturel n :

$$|u_n - \ell| = |f(u_{n-1}) - f(\ell)| \leq k |u_{n-1} - \ell| \leq k^2 |u_{n-2} - \ell| \leq \dots \leq k^n |u_0 - \ell|$$

Comme $0 < k < 1$:

$$\lim_{n \rightarrow +\infty} k^n = 0$$

D'où le résultat !

Il est à noter que ce théorème peut aussi se démontrer à l'aide de suites de Cauchy, comme cela sera fait plus loin. ■

1. c'est-à-dire lipschitzienne de rapport $k < 1$; pour tout couple de réels $(x, y) \in I^2$:

$$|f(x) - f(y)| \leq k|x - y|$$

1. Somme de suites

Propriété

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites, à valeurs dans \mathbb{R} ou \mathbb{C} , convergentes, de limites respectives ℓ_u et ℓ_v .

Alors, la suite $(u_n + v_n)_{n \in \mathbb{N}}$ est elle aussi convergente, de limite $\ell_u + \ell_v$.

Propriété

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles, de limites respectives ℓ_u et $+\infty$.

Alors, la suite $(u_n + v_n)_{n \in \mathbb{N}}$ est divergente, de limite $+\infty$.

Propriété

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles, de limites respectives ℓ_u et $-\infty$.

Alors, la suite $(u_n + v_n)_{n \in \mathbb{N}}$ est divergente, de limite $-\infty$.



Dans le cas de deux suites réelles divergentes $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$, de limites respectives $-\infty$ et $+\infty$, il n'est pas possible de savoir, sans une étude supplémentaire, quelle est la limite (et si elle existe), de la suite $(u_n + v_n)_{n \in \mathbb{N}}$.

2. Produit de suites

Propriété

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites, à valeurs dans \mathbb{R} ou \mathbb{C} , convergentes, de limites respectives ℓ_u et ℓ_v .

Alors, la suite $(u_n v_n)_{n \in \mathbb{N}}$ est elle aussi convergente, de limite $\ell_u \ell_v$.

Propriété

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles, de limites respectives $\ell_u > 0$ et $+\infty$.

Alors, la suite $(u_n v_n)_{n \in \mathbb{N}}$ est divergente, de limite $+\infty$.



Dans le cas où $\ell_u = 0$, il n'est pas possible de conclure sans une étude plus poussée (qui n'aboutira pas nécessairement).

Propriété

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles, de limites respectives $\ell_u > 0$ et $-\infty$.

Alors, la suite $(u_n v_n)_{n \in \mathbb{N}}$ est divergente, de limite $-\infty$.



1. Dans le cas où $\ell_u = 0$, il n'est pas possible de conclure sans une étude plus poussée (qui n'aboutira pas nécessairement).
2. On dispose bien sûr de propriétés analogues pour le cas où $\ell_u < 0$.

Propriété

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles divergentes, de limites respectives $-\infty$ et $-\infty$.

Alors, la suite $(u_n v_n)_{n \in \mathbb{N}}$ est divergente, de limite $+\infty$.

Propriété

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles divergentes, de limites respectives $-\infty$ et $+\infty$.

Alors, la suite $(u_n v_n)_{n \in \mathbb{N}}$ est divergente, de limite $-\infty$.

Propriété

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles divergentes, de limites respectives $+\infty$ et $+\infty$.

Alors, la suite $(u_n v_n)_{n \in \mathbb{N}}$ est divergente, de limite $+\infty$.

3. Inverse d'une suite**Propriété**

Soit $(u_n)_{n \in \mathbb{N}}$ une suite, complexe ou réelle, à termes non nuls, convergente, de limite $\ell_u \neq 0$.

Alors, la suite $(\frac{1}{u_n})_{n \in \mathbb{N}}$ est elle aussi convergente, de limite $\frac{1}{\ell_u}$.

Propriété

Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle, à termes non nuls, divergente, de limite $\pm\infty$.

Alors, la suite $(\frac{1}{u_n})_{n \in \mathbb{N}}$ est convergente, de limite nulle.

Propriété

Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle, à termes non nuls et strictement positifs à partir d'un rang $n_0 \in \mathbb{N}$, convergente, de limite nulle.

Alors, la suite $(\frac{1}{u_n})_{n \geq n_0}$ est divergente, de limite $+\infty$.

Propriété

Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle, à termes non nuls et strictement négatifs à partir d'un rang $n_0 \in \mathbb{N}$, convergente, de limite nulle.

Alors, la suite $(\frac{1}{u_n})_{n \geq n_0}$ est divergente, de limite $-\infty$.

4. Multiplication d'une suite par un scalaire**Propriété**

Soient $(u_n)_{n \in \mathbb{N}}$ une suite réelle, à valeurs dans \mathbb{K} , convergente, de limite ℓ_u , et λ dans \mathbb{K} .

Alors, la suite $(\lambda u_n)_{n \in \mathbb{N}}$ est elle aussi convergente, de limite $\lambda \ell_u$.

Propriété

Soient $(u_n)_{n \in \mathbb{N}}$ une suite réelle, divergente, de limite $+\infty$, et λ un réel.

Si le réel λ est strictement positif, la suite $(\lambda u_n)_{n \in \mathbb{N}}$ est elle aussi divergente, de limite $+\infty$.

Si le réel λ est strictement négatif, la suite $(\lambda u_n)_{n \in \mathbb{N}}$ est elle aussi divergente, de limite $-\infty$.

Propriété

Soient $(u_n)_{n \in \mathbb{N}}$ une suite réelle, divergente, de limite $-\infty$, et λ un réel.

Si le réel λ est strictement positif, la suite $(\lambda u_n)_{n \in \mathbb{N}}$ est elle aussi divergente, de limite $-\infty$.

Si le réel λ est strictement négatif, la suite $(\lambda u_n)_{n \in \mathbb{N}}$ est elle aussi divergente, de limite $+\infty$.



Dans le cas de deux suites divergentes $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$, de limites respectives $-\infty$ et $+\infty$, il n'est pas possible de savoir, sans une étude supplémentaire, quelle est la limite (et si elle existe), de la suite $(u_n + v_n)_{n \in \mathbb{N}}$.

5. Complément : le théorème de Cesàro¹

Théorème

Soit $(u_n)_{n \in \mathbb{N}^*}$ une suite réelle, de limite finie $\ell \in \mathbb{R}$.

La suite $(v_n)_{n \in \mathbb{N}^*}$, telle que : $\forall n \in \mathbb{N}^* : v_n = \frac{u_1 + u_2 + \dots + u_n}{n} = \frac{1}{n} \sum_{k=1}^n u_k$, a même limite ℓ que la suite $(u_n)_{n \in \mathbb{N}^*}$.

Démonstration :

- Par définition de la limite, pour tout $\varepsilon > 0$, il existe un rang n_0 tel que, pour tout entier naturel $n \geq n_0$:

$$|u_n - \ell| \leq \frac{\varepsilon}{2}$$

- Pour tout entier naturel $n \geq n_0$:

$$|v_n - \ell| = \frac{1}{n} \left| \sum_{k=1}^n (u_k - \ell) \right| = \frac{1}{n} \left| \sum_{k=1}^{n_0-1} (u_k - \ell) + \sum_{k=n_0}^n (u_k - \ell) \right|$$

$$\left| \sum_{k=1}^{n_0-1} (u_k - \ell) \right| \text{ étant une quantité finie : } \lim_{n \rightarrow +\infty} \frac{1}{n} \left| \sum_{k=1}^{n_0-1} (u_k - \ell) \right| = 0.$$

Il existe donc un rang n_1 tel que, pour tout entier naturel $n \geq n_1$:

$$\frac{1}{n} \left| \sum_{k=1}^{n_0-1} (u_k - \ell) \right| \leq \frac{\varepsilon}{2}$$

Par suite, pour tout entier naturel $n \geq \max \{n_0, n_1\}$:

$$|v_n - \ell| \leq \frac{1}{n} \left| \sum_{k=1}^{n_0-1} (u_k - \ell) \right| + \frac{1}{n} \left| \sum_{k=n_0}^n (u_k - \ell) \right| \leq \frac{\varepsilon}{2} + \frac{n - n_0 + 1}{n} \frac{\varepsilon}{2} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Posons alors $n_2 = \max \{n_0, n_1\}$. On a alors, pour tout entier naturel $n \geq n_2$:

$$|v_n - \ell| \leq \varepsilon$$

La suite $(v_n)_{n \in \mathbb{N}^*}$ a donc même limite que la suite $(u_n)_{n \in \mathbb{N}^*}$. ■

1. Ernesto Cesàro (1859-1906), mathématicien italien, qui apporta de nombreuses contributions à l'étude des séries numériques. C'était aussi un spécialiste de géométrie différentielle.

Convergence des suites homographiques réelles

Proposition

On considère quatre réels a, b, c et d tels que :

$$ad - bc \neq 0, \quad c \neq 0$$

et la **suite homographique** réelle $(u_n)_{n \in \mathbb{N}}$ vérifiant une relation de récurrence de la forme :

$$\forall n \in \mathbb{N} : u_{n+1} = \frac{au_n + b}{cu_n + d}$$

Pour que la suite soit définie pour tout entier naturel n , il faut que, pour tout entier naturel n :

$$cu_n + d \neq 0$$

Si la suite $(u_n)_{n \in \mathbb{N}}$ converge, sa limite ℓ est solution de l'équation :

$$\ell = \frac{a\ell + b}{c\ell + d}$$

soit : $c\ell^2 + (d - a)\ell - b = 0$.

On supposera, dans ce qui suit : $u_0 \neq \frac{a\ell + b}{c\ell + d}$. Alors :

- Si $(d - a)^2 + 4bc < 0$, l'équation précédente n'a pas de solution réelle, et la suite diverge.
- Si $(d - a)^2 + 4bc = 0$, l'équation précédente admet une racine double, $\frac{a - d}{2c}$. La suite converge vers $\frac{a - d}{2c}$.
- Si $(d - a)^2 + 4bc > 0$, l'équation précédente admet deux racines réelles distinctes ℓ_1 et ℓ_2 , avec $\ell_1 < \ell_2$:
 - si $\left| \frac{c\ell_2 + d}{c\ell_1 + d} \right| < 1$, la suite $(u_n)_{n \in \mathbb{N}}$ a pour limite ℓ_1 lorsque n tend vers $+\infty$.
 - Si $\left| \frac{c\ell_2 + d}{c\ell_1 + d} \right| > 1$, la suite $(u_n)_{n \in \mathbb{N}}$ a pour limite ℓ_2 lorsque n tend vers $+\infty$.
 - Si $\frac{c\ell_2 + d}{c\ell_1 + d} = -1$, la suite $(u_n)_{n \in \mathbb{N}}$ diverge.

Démonstration : Si la suite converge, sa limite est un point fixe de la fonction :

$$x \in \mathbb{R} \setminus \left\{ -\frac{d}{c} \right\} \mapsto \frac{ax + b}{cx + d}$$

Un tel point fixe x est donc solution de :

$$x = \frac{ax + b}{cx + d}$$

ce qui conduit à :

$$c x^2 + (d - a) x - b = 0$$

Le discriminant est :

$$\Delta = (d - a)^2 + 4 b c$$

Il faut donc envisager les cas suivants :

1. Si $(d - a)^2 + 4 b c = 0$:

$\ell = \frac{a - d}{2 c}$ est alors racine double du trinôme $c x^2 + (d - a) x - b = 0$.

Considérons alors la suite $(v_n)_{n \in \mathbb{N}}$ définie, pour tout entier naturel n , par :

$$v_n = \frac{1}{u_n - \ell}$$

On obtient :

$$\begin{aligned} v_{n+1} &= \frac{1}{u_{n+1} - \ell} \\ &= \frac{1}{\frac{a u_n + b}{c u_n + d} - \frac{a \ell + b}{c \ell + d}} \\ &= \frac{(c \ell + d)(c u_n + d)}{(a d - b c)(u_n - \ell)} \\ &= \frac{(c \ell + d)(c(u_n - \ell) + c \ell + d)}{(a d - b c)(u_n - \ell)} \\ &= c \frac{c \ell + d}{a d - b c} + \frac{(c \ell + d)^2}{(a d - b c)(u_n - \ell)} \end{aligned}$$

Or :

$$c \ell + d = \frac{a + d}{2}$$

et, compte tenu de $(d - a)^2 + 4 b c = 0$:

$$-b c = \frac{(d - a)^2}{4}$$

ce qui entraîne :

$$\begin{aligned} a d - b c &= \frac{(d - a)^2}{4} + a d \\ &= \frac{4 a d + (d - a)^2}{4} \\ &= \frac{4 a d + d^2 + a^2 - 2 a d}{4} \\ &= \frac{2 a d + d^2 + a^2}{4} \\ &= \frac{(a + d)^2}{4} \end{aligned}$$

Ainsi :

$$\frac{(c\ell + d)^2}{(ad - bc)} = \frac{(a + d)^2}{4} \frac{4}{(a + d)^2} = 1$$

Il en résulte :

$$v_{n+1} = c \frac{c\ell + d}{ad - bc} + \frac{1}{u_n - \ell} = c \frac{c\ell + d}{ad - bc} + v_n$$

La suite $(v_n)_{n \in \mathbb{N}}$ est donc une suite arithmétique, de raison $c \frac{c\ell + d}{ad - bc}$.

Par suite, pour tout entier naturel n :

$$v_n = v_0 + cn \frac{c\ell + d}{ad - bc}$$

La suite $(v_n)_{n \in \mathbb{N}}$ diverge. La convergence de la suite $(u_n)_{n \in \mathbb{N}}$ vers $\frac{a - d}{2c}$ en résulte.

2. Si $(d - a)^2 + 4bc > 0$:

Le discriminant Δ est strictement positif, l'équation « aux points fixes » admet deux solutions distinctes ℓ_1 et ℓ_2 :

$$\ell_1 = \frac{a - d - \sqrt{\Delta}}{2c}, \quad \ell_2 = \frac{a - d + \sqrt{\Delta}}{2c}$$

Ainsi :

$$|\ell_1| < |\ell_2|$$

Considérons alors la suite $(w_n)_{n \in \mathbb{N}}$ définie, pour tout entier naturel n , par :

$$w_n = \frac{u_n - \ell_1}{u_n - \ell_2}$$

On obtient :

$$\begin{aligned} w_{n+1} &= \frac{u_{n+1} - \ell_1}{u_{n+1} - \ell_2} \\ &= \frac{\frac{au_n + b}{cu_n + d} - \ell_1}{\frac{au_n + b}{cu_n + d} - \ell_2} \\ &= \frac{\frac{au_n + b}{cu_n + d} - \frac{a\ell_1 + b}{c\ell_1 + d}}{\frac{au_n + b}{cu_n + d} - \frac{a\ell_2 + b}{c\ell_2 + d}} \\ &= \frac{c\ell_2 + d}{c\ell_1 + d} \frac{(ad - bc)(u_n - \ell_1)}{(ad - bc)(u_n - \ell_2)} \\ &= \frac{c\ell_2 + d}{c\ell_1 + d} \frac{u_n - \ell_1}{u_n - \ell_2} \\ &= \frac{c\ell_2 + d}{c\ell_1 + d} w_n \end{aligned}$$

puisque ℓ_1 et ℓ_2 vérifient respectivement :

$$\ell_1 = \frac{a\ell_1 + b}{c\ell_1 + d}, \quad \ell_2 = \frac{a\ell_2 + b}{c\ell_2 + d}$$

La suite $(w_n)_{n \in \mathbb{N}}$ est donc une suite géométrique, de raison $\frac{c\ell_2 + d}{c\ell_1 + d}$.

Par suite, pour tout entier naturel n :

$$w_n = w_0 \left(\frac{c\ell_2 + d}{c\ell_1 + d} \right)^n$$

Ainsi :

- si $\left| \frac{c\ell_2 + d}{c\ell_1 + d} \right| < 1$, la suite $(w_n)_{n \in \mathbb{N}}$ a pour limite 0 lorsque n tend vers $+\infty$;
comme, pour tout entier naturel n , $w_n = \frac{u_n - \ell_1}{u_n - \ell_2}$, la suite $(u_n)_{n \in \mathbb{N}}$ a donc pour limite ℓ_1 lorsque n tend vers $+\infty$.
- Si $\left| \frac{c\ell_2 + d}{c\ell_1 + d} \right| > 1$, la suite $(w_n)_{n \in \mathbb{N}}$ diverge vers $+\infty$; comme, pour tout entier naturel n , $|w_n| = \left| \frac{u_n - \ell_1}{u_n - \ell_2} \right|$, la suite $(u_n)_{n \in \mathbb{N}}$ a donc pour limite ℓ_2 lorsque n tend vers $+\infty$.
- Si $\frac{c\ell_2 + d}{c\ell_1 + d} = -1$, la suite $(w_n)_{n \in \mathbb{N}}$ diverge ; comme, pour tout entier naturel n , $w_n = \frac{u_n - \ell_1}{u_n - \ell_2}$, la suite $(u_n)_{n \in \mathbb{N}}$ diverge également.

3. Si $(d - a)^2 + 4bc < 0$:

Le discriminant Δ est strictement négatif, l'équation « aux points fixes » admet deux solutions distinctes complexes ℓ et $\bar{\ell}$:

$$\ell = \frac{a - d - i\delta}{2c}, \quad \bar{\ell} = \frac{a - d + i\delta}{2c}$$

où $\delta \in \mathbb{R}^+$ est solution de :

$$\delta^2 = -\Delta$$

Considérons alors la suite $(z_n)_{n \in \mathbb{N}}$ définie, pour tout entier naturel n , par :

$$z_n = \frac{u_n - \ell}{u_n - \bar{\ell}}$$

On obtient :

$$\begin{aligned}
 z_{n+1} &= \frac{u_{n+1} - \ell}{u_{n+1} - \bar{\ell}} \\
 &= \frac{\frac{a u_n + b}{c u_n + d} - \ell}{\frac{a u_n + b}{c u_n + d} - \bar{\ell}} \\
 &= \frac{\frac{a u_n + b}{c u_n + d} - \frac{a \ell + b}{c \ell + d}}{\frac{a u_n + b}{c u_n + d} - \frac{a \bar{\ell} + b}{c \bar{\ell} + d}} \\
 &= \frac{c \bar{\ell} + d}{c \ell + d} \frac{(a d - b c)(u_n - \ell)}{(a d - b c)(u_n - \bar{\ell})} \\
 &= \frac{c \bar{\ell} + d}{c \ell + d} \frac{u_n - \ell}{u_n - \bar{\ell}} \\
 &= \frac{c \bar{\ell} + d}{c \ell + d} z_n
 \end{aligned}$$

puisque ℓ et $\bar{\ell}$ vérifient respectivement

$$\ell = \frac{a \ell + b}{c \ell + d}, \quad \bar{\ell} = \frac{a \bar{\ell} + b}{c \bar{\ell} + d}$$

La suite $(z_n)_{n \in \mathbb{N}}$ est donc une suite géométrique, de raison $\frac{c \bar{\ell} + d}{c \ell + d}$.

Par suite, pour tout entier naturel n :

$$z_n = z_0 \left(\frac{c \bar{\ell} + d}{c \ell + d} \right)^n$$

Or :

$$\left| \frac{c \bar{\ell} + d}{c \ell + d} \right| = 1 \quad \text{et} \quad \frac{c \bar{\ell} + d}{c \ell + d} \neq 1$$

La suite $(z_n)_{n \in \mathbb{N}}$ n'a pas de limite lorsque n tend vers $+\infty$. Il en est de même de la suite $(u_n)_{n \in \mathbb{N}}$.

Il est à noter que la divergence de la suite peut s'obtenir directement, dans la mesure où l'équation aux points fixes n'admet pas de racine réelle ! La suite étant à valeurs réelles, sa limite est en effet nécessairement réelle. ■

On exclut le cas où le premier terme vaut $\frac{a \ell + b}{c \ell + d}$, puisque l'on a alors, pour tout entier naturel n :

$$u_n = \frac{a \ell + b}{c \ell + d}$$



Définition

Soit $(u_n)_{n \in \mathbb{N}}$ une suite, à valeurs réelles ou complexes.

On appelle **suite extraite** de la suite $(u_n)_{n \in \mathbb{N}}$ (ou **sous-suite**) une suite obtenue en sélectionnant, dans l'ordre, un sous-ensemble infini de termes de $(u_n)_{n \in \mathbb{N}}$.

Comme les termes sont retenus « dans l'ordre », il existe ainsi une application strictement croissante $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ permettant d'indexer les termes de la suite extraite, sous la forme $(u_{\varphi(n)})_{n \in \mathbb{N}}$.

Exemple

Étant donnée une suite réelle $(u_n)_{n \in \mathbb{N}}$, la suite $(u_{2n})_{n \in \mathbb{N}}$, obtenue en ne retenant que les termes d'indices pairs, est une suite extraite de $(u_n)_{n \in \mathbb{N}}$.

L'application $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ permettant d'indexer les termes de la suite extraite est donc ici : $n \mapsto 2n$.

Lemme

Toute application $\varphi : \mathbb{N} \rightarrow \mathbb{N}$, strictement croissante, n'est pas majorée.

Démonstration : Dans un premier temps, on démontre par récurrence que, pour tout entier naturel n :

$$\varphi(n) \geq n$$

- Cette propriété est bien vraie au rang 0, φ étant à valeurs dans \mathbb{N} :

$$\varphi(0) \geq 0$$

- Supposons la propriété vraie à un rang $n \geq 0$: $\varphi(n) \geq n$.

L'application φ étant strictement croissante, il en résulte :

$$\varphi(n+1) > \varphi(n) \geq n$$

Ainsi, φ étant à valeurs dans \mathbb{N} :

$$\varphi(n+1) > n$$

ce qui conduit à :

$$\varphi(n+1) \geq n+1$$

La propriété est donc vraie au rang $n+1$.

- Comme elle est vraie au rang 0, elle est donc vraie pour tout entier naturel n .

Comme $\lim_{n \rightarrow +\infty} n = +\infty$, il en résulte :

$$\lim_{n \rightarrow +\infty} \varphi(n) = +\infty$$

L'application φ n'est donc pas majorée. ■

Théorème

Toute sous-suite d'une suite convergente converge vers la même limite.

Démonstration : On démontre le résultat dans le cas d'une suite réelle.

Soit ainsi $(u_{\varphi(n)})_{n \in \mathbb{N}}$ une suite extraite d'une suite convergente $(u_n)_{n \in \mathbb{N}}$.

On note ℓ la limite de $(u_n)_{n \in \mathbb{N}}$.

Alors, pour tout $\varepsilon > 0$, il existe un rang n_0 tel que, pour tout entier $n \geq n_0$:

$$|u_n - \ell| \leq \varepsilon$$

L'application φ étant strictement croissante de \mathbb{N} dans \mathbb{N} , et non majorée, il existe donc un rang n_1 tel que, pour tout entier $n \geq n_1$:

$$\varphi(n) \geq n_0$$

Ainsi, pour tout entier $n \geq n_1$:

$$|u_{\varphi(n)} - \ell| \leq \varepsilon$$

D'où le résultat. ■

► Théorème de Bolzano-Weierstrass

Théorème

De toute suite réelle $(u_n)_{n \in \mathbb{N}}$ bornée, on peut extraire une sous-suite convergente.

Démonstration : Ce théorème est admis. ■

Pour étudier le comportement d'une suite à l'infini, il n'est pas toujours nécessaire de connaître sa limite. Intuitivement, on comprend bien que, si une suite converge, ses termes seront, à partir d'un certain rang, assez proches les uns des autres.

Définition

La suite $(u_n)_{n \in \mathbb{N}}$ est une **suite de Cauchy** si, pour tout réel strictement positif ε , il existe un entier n_0 tel que :

$$\forall p \geq n_0, \forall q \geq n_0 : |u_p - u_q| \leq \varepsilon$$

ce qui signifie que, à partir d'un certain rang, l'écart entre deux termes de la suite est toujours, en valeur absolue, aussi petit que l'on veut.

► Critère de Cauchy

Propriété

Une suite réelle converge si et seulement si c'est une suite de Cauchy¹.

Démonstration : Soit $(u_n)_{n \in \mathbb{N}}$ une suite convergente, de limite finie $\ell \in \mathbb{R}$. Alors, pour tout $\varepsilon > 0$, il existe un rang n_0 tel que :

$$\forall n \geq n_0 : |u_n - \ell| \leq \frac{\varepsilon}{2}$$

Par suite, pour tout couple d'entiers (p, q) tels que $p \geq n_0$ et $q \geq n_0$:

$$|u_p - u_q| = |u_p - \ell + \ell - u_q| \leq |u_p - \ell| + |u_q - \ell| \leq 2 \frac{\varepsilon}{2} = \varepsilon$$

La suite $(u_n)_{n \in \mathbb{N}}$ est donc une suite de Cauchy.

Réciproquement, considérons une suite réelle $(u_n)_{n \in \mathbb{N}}$ telle que, pour tout réel strictement positif ε , il existe un rang n_0 tel que, pour tout couple d'entiers (p, q) tels que $p \geq n_0$ et $q \geq n_0$:

$$|u_p - u_q| \leq \frac{\varepsilon}{2}$$

Il en résulte, en particulier, pour tout entier $n \geq n_0$:

$$|u_n - u_{n_0}| \leq \frac{\varepsilon}{2}$$

puis, par inégalité triangulaire :

$$|u_n| = |u_n - u_{n_0} + u_{n_0}| \leq \frac{\varepsilon}{2} + |u_{n_0}|$$

La suite $(u_n)_{n \in \mathbb{N}}$ est donc bornée par $\max \{u_0, \dots, |u_{n_0-1}|, \frac{\varepsilon}{2} + |u_{n_0}|\}$.

1. \mathbb{R} est un espace complet, c'est-à-dire un espace où toute suite de Cauchy converge.

D'après le théorème de Bolzano-Weierstrass, on peut donc en extraire une sous-suite convergente $(u_{\varphi(n)})_{n \in \mathbb{N}}$, dont on notera $\ell_{u_{\varphi(n)}}$ la limite.

Il existe donc un rang n_1 tel que, pour tout entier $n \geq n_1$:

$$|u_{\varphi(n)} - \ell_{u_{\varphi(n)}}| \leq \frac{\varepsilon}{2}$$

Par suite, pour tout entier $n \geq \max(n_0, n_1)$, comme $\varphi(n) \geq n$:

$$|u_n - \ell_{u_{\varphi(n)}}| = |u_n - u_{\varphi(n)} + u_{\varphi(n)} - \ell_{u_{\varphi(n)}}| \leq 2 \frac{\varepsilon}{2} = \varepsilon$$

La suite $(u_n)_{n \in \mathbb{N}}$ converge donc aussi vers $\ell_{u_{\varphi(n)}}$. ■

Comparaison des suites réelles

1. Négligeabilité

Définition

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles, non identiquement nulles.

La suite $(u_n)_{n \in \mathbb{N}}$ est dite **négligeable** devant la suite $(v_n)_{n \in \mathbb{N}}$ s'il existe une suite $(\varepsilon_n)_{n \in \mathbb{N}}$, de limite nulle, et un rang n_0 à partir duquel :

$$u_n = \varepsilon_n v_n$$

On note alors :

$$u_n \underset{n \rightarrow +\infty}{=} o(v_n)$$

On dit que u_n est un « petit o » de v_n .



La notation « petit o », de même que la notation « grand O », qui sera vue plus loin, est appelée **notation de Landau**, en hommage au mathématicien Edmund Landau¹. Leur paternité est visiblement assez controversée, et reviendrait, a priori, à Paul Bachmann².

Exemple

On considère les suites $(u_n)_{n \in \mathbb{N}^*}$ et $(v_n)_{n \in \mathbb{N}^*}$ définies par :

$$\forall n \in \mathbb{N}^* : u_n = \frac{1}{n^4} \quad , \quad v_n = \frac{1}{n^3}$$

Alors, comme :

$$\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = \lim_{n \rightarrow +\infty} \frac{1}{n} = 0$$

on a bien :

$$u_n \underset{n \rightarrow +\infty}{=} o(v_n)$$

2. Domination

Définition

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles.

La suite $(u_n)_{n \in \mathbb{N}}$ est dite **dominée** par la suite $(v_n)_{n \in \mathbb{N}}$ s'il existe un réel positif M et un rang n_0 à partir duquel :

$$|u_n| \leq M |v_n|$$

On note alors :

$$u_n \underset{n \rightarrow +\infty}{=} O(v_n)$$

On dit que u_n est un « grand O » de v_n .

1. Edmund Georg Hermann Landau (1877-1938), mathématicien allemand, spécialiste de théorie des nombres.

2. Paul Bachmann (1837-1920), mathématicien allemand lui aussi, et également spécialiste de théorie des nombres.

Exemple

On considère les suites $(u_n)_{n \in \mathbb{N}^*}$ et $(v_n)_{n \in \mathbb{N}^*}$ définies par :

$$\forall n \in \mathbb{N}^* : u_n = \frac{5 + \frac{1}{n}}{2^n} \quad , \quad v_n = \frac{3}{2^n}$$

Alors, comme, pour tout entier naturel non nul n :

$$|u_n| = \left| \frac{5 + \frac{1}{n}}{2^n} \right| \leq \left| \frac{6}{2^n} \right| = 2 |v_n|$$

on a bien :

$$u_n \underset{n \rightarrow +\infty}{=} O(v_n)$$

3. Équivalence**Définition**

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles, non identiquement nulles.

Les suites $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ sont dites **équivalentes** s'il existe une suite $(\alpha_n)_{n \in \mathbb{N}}$, de limite égale à 1, et un rang n_0 à partir duquel :

$$u_n = \alpha_n v_n$$

On note alors :

$$u_n \underset{n \rightarrow +\infty}{\sim} v_n$$

Exemple

On considère les suites $(u_n)_{n \in \mathbb{N}^*}$ et $(v_n)_{n \in \mathbb{N}^*}$ définies par :

$$\forall n \in \mathbb{N}^* : u_n = 1 + \frac{1}{n^4} \quad , \quad v_n = 1 + \frac{1}{n^3}$$

Alors, comme :

$$\lim_{n \rightarrow +\infty} \frac{u_n}{v_n} = \lim_{n \rightarrow +\infty} \frac{1 + \frac{1}{n^3}}{1 + \frac{1}{n^4}} = 1$$

les suites $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ sont bien équivalentes lorsque n tend vers $+\infty$.



Attention aux manipulations successives et hasardeuses d'équivalents !

Pour cette raison, on ne donnera pas, dans ce cours, de résultats généraux ni de « recettes » pour la manipulation d'équivalents, la meilleure méthode, la plus fiable et la plus sûre, étant de manipuler, suivant les cas et ce qui est le mieux adapté, des « o » ou des « O ».

Théorème

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles, équivalentes lorsque n tend vers $+\infty$.

Alors :

$$u_n \underset{n \rightarrow +\infty}{=} v_n + o(v_n)$$

Réciproquement, si $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ sont deux suites réelles telles que, lorsque n tend vers $+\infty$

$$u_n = v_n + o(v_n)$$

elles sont aussi équivalentes lorsque n tend vers $+\infty$.



Pouvoir déterminer, pour des suites, des relations de négligeabilité, équivalence, domination, est donc extrêmement utile pour étudier leur convergence !

4. Développement asymptotique

Définition

On appelle **développement asymptotique** d'une suite réelle $(u_n)_{n \in \mathbb{N}}$ une décomposition de la forme :

$$u_n = u_{n,1} + u_{n,2} + \dots + u_{n,p} + o(u_{n,p})$$

où

$$|u_{n,p}| \text{ est très petit devant } |u_{n,p-1}|$$

$$|u_{n,p-1}| \text{ est très petit devant } |u_{n,p-2}|$$

.....

$$|u_{n,2}| \text{ est très petit devant } |u_{n,1}|$$



Les développements asymptotiques peuvent être extrêmement utiles pour déterminer des limites.

On parle de *développement asymptotique*, car on étudie ce qui se passe lorsque n tend vers $+\infty$.

Exemple

On considère la suite $(u_n)_{n \in \mathbb{N}^*}$ définie par :

$$\forall n \in \mathbb{N}^* : u_n = \left(1 - \frac{1}{n}\right)^n$$

Lorsque n tend vers $+\infty$:

$$\left(1 - \frac{1}{n}\right)^n = e^{n \ln(1 - \frac{1}{n})} = e^{n(-\frac{1}{n} + o(\frac{1}{n}))} = e^{-1 + o(1)}$$

(On a effectué un développement asymptotique à l'ordre 1 en $\frac{1}{n}$.)

Ainsi :

$$\lim_{n \rightarrow +\infty} \left(1 - \frac{1}{n}\right)^n = e^{-1} = \frac{1}{e}$$

5. Des propriétés intéressantes : suites adjacentes, théorème des gendarmes

► Suites adjacentes

Définition

Deux suites réelles $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ sont dites **adjacentes** si :

- l'une des suites est croissante ;
- l'autre suite est décroissante ;
- $\lim_{n \rightarrow +\infty} (u_n - v_n) = 0$.

Théorème

Deux suites adjacentes convergent et ont même limite.

► Théorème des gendarmes

Théorème

Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites réelles convergeant vers une même limite ℓ .

Alors, pour toute suite réelle $(w_n)_{n \in \mathbb{N}}$ telle que, pour tout entier naturel n :

$$u_n \leq w_n \leq v_n$$

la suite $(w_n)_{n \in \mathbb{N}}$ est elle aussi convergente, de limite ℓ .

Démonstration : Comme les suites $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ convergent vers ℓ , il existe un rang n_0 tel que, pour tout entier $n \geq n_0$:

$$\ell - \varepsilon \leq u_n \leq \ell + \varepsilon$$

et :

$$\ell - \varepsilon \leq v_n \leq \ell + \varepsilon$$

ce qui conduit à :

$$\ell - \varepsilon \leq u_n \leq w_n \leq v_n \leq \ell + \varepsilon$$

puis :

$$|w_n - \ell| \leq \varepsilon$$

■



En physique, ou en mécanique, un système dynamique est un système évoluant, au cours du temps :

- de façon causale, c'est-à-dire dont l'avenir est directement conditionné soit par son passé, soit par son état actuel ;
- de façon déterministe, c'est-à-dire que la donnée d'une condition initiale correspond à une unique évolution.

Historiquement, l'un des premiers exemples de système dynamique est le système solaire, auquel s'intéressa le mathématicien Joseph-Louis Lagrange¹. De nos jours, la théorie des systèmes dynamiques a de nombreuses applications, notamment, lorsqu'il s'agit d'étudier la stabilité d'un système, ou d'un ensemble de particules.

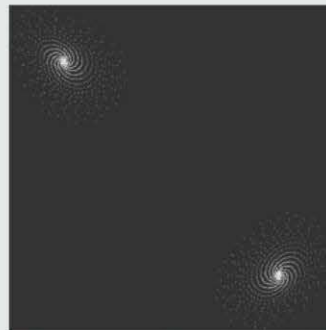
Modéliser le chaos

Un exemple intéressant est l'attracteur (c'est-à-dire l'ensemble des limites des solutions du système) de Hénon. Il est construit par la donnée initiale d'un couple de réels (x_0, y_0) , et de deux réels a et b , à partir desquels on génère la suite de points du plan de coordonnées (x_n, y_n) tels que, pour tout entier naturel n :

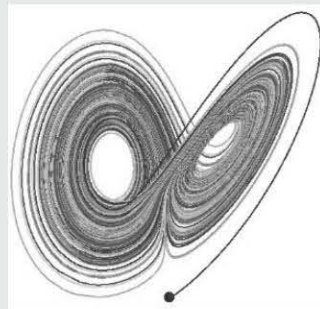
$$\begin{cases} x_{n+1} = y_n + 1 - a x_n^2 \\ y_{n+1} = b x_n \end{cases}$$

Lorsque $a = 1,4$ et $b = 0,3$, l'attracteur est chaotique, ce qui signifie qu'il devient fortement instable ; son comportement à long terme n'est pas prédictible. Le fait que le système dynamique associé soit déterministe permet donc de contrôler, dans certains cas, le chaos.

L'attracteur de Hénon peut apparaître comme une simplification de l'attracteur de Lorenz, qui a de nombreuses applications en météorologie, plus précisément, pour modéliser le comportement du fluide turbulent qu'est l'atmosphère.



L'attracteur de Hénon pour
 $a = 0,0941$ et $b = 0,99681$.



L'attracteur de Lorenz.

1. Joseph Louis, comte de Lagrange (1736-1813), mathématicien, mécanicien et astronome italien. Il fut l'initiateur du calcul variationnel. En parallèle, il apporta de nombreuses contributions en algèbre, à la théorie des nombres, au calcul infinitésimal, aux probabilités, mais aussi à la mécanique.

1. Des sommes partielles aux séries

► Sommes partielles

Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle. On appelle **suite des sommes partielles de $(u_n)_{n \in \mathbb{N}}$** la suite $(S_n)_{n \in \mathbb{N}}$ telle que, pour tout entier naturel n :

$$S_n = u_0 + u_1 + \dots + u_n = \sum_{k=0}^n u_k$$

► Série

Soit $(u_n)_{n \in \mathbb{N}}$ une suite réelle. On appelle **série de terme général u_n** la suite des sommes partielles $\left(\sum_{k=0}^n u_k \right)_{n \in \mathbb{N}}$, notée $\sum u_n$.

2. Convergence et divergence

► Convergence d'une série numérique

Soit $\sum u_n$ une série réelle. On dit que $\sum u_n$ **converge** si la suite des sommes partielles $\left(\sum_{k=0}^n u_k \right)_{n \in \mathbb{N}}$ converge (i.e. admet une limite finie).

► Divergence d'une série numérique

Soit $\sum u_n$ une série réelle. On dit que $\sum u_n$ **diverge** si la suite des sommes partielles $\left(\sum_{k=0}^n u_k \right)_{n \in \mathbb{N}}$ diverge (i.e. n'admet pas de limite finie).

► Somme d'une série convergente

Soit $\sum u_n$ une série réelle convergente. On appelle **somme de la série $\sum u_n$** la quantité

$$\sum_{n=0}^{+\infty} u_n = \lim_{n \rightarrow +\infty} \sum_{k=0}^n u_k$$

qui est donc la limite de la suite des sommes partielles $\left(\sum_{k=0}^n u_k \right)_{n \in \mathbb{N}}$.

► **Reste d'ordre n , $n \in \mathbb{N}$, d'une série convergente**

Soit $\sum u_n$ une série réelle convergente. n étant un entier naturel non nul, on appelle **Reste d'ordre n de la série $\sum u_n$** la quantité

$$R_n = \sum_{k=n+1}^{+\infty} u_k$$

► **Série absolument convergente**

Soit $\sum u_n$ une série réelle. La série $\sum u_n$ est dite **absolument convergente** si la série $\sum |u_n|$ converge.

Théorème

Soit $\sum u_n$ une série absolument convergente. Alors, la série $\sum u_n$ converge.

► **Série semi-convergente**

Soit $\sum u_n$ une série réelle. On dit que $\sum u_n$ est **semi-convergente** si elle est convergente, mais non absolument convergente.

► **Condition nécessaire de convergence d'une série**

Théorème

Soit $\sum u_n$ une série réelle. Une condition nécessaire de convergence de $\sum u_n$ est :

$$\lim_{n \rightarrow +\infty} u_n = 0$$



Attention ! C'est une condition nécessaire, et non suffisante ! Ainsi, ce n'est pas parce que son terme général tend vers zéro qu'une série converge. Ainsi, la série de terme général $\frac{1}{\sqrt{n}}$ diverge.

► **Somme de deux séries convergentes**

Théorème

Soient $\sum u_n$ et $\sum v_n$ deux séries convergentes. Alors, la série $\sum (u_n + v_n)$ converge, et a pour somme :

$$\sum_{n=0}^{+\infty} (u_n + v_n) = \sum_{n=0}^{+\infty} u_n + \sum_{n=0}^{+\infty} v_n$$

1. Les séries géométriques

Définition

On appelle **série géométrique** toute série de la forme $\sum r^n$, où r est un réel.
Le réel r est la **raison** de la série géométrique $\sum r^n$.



Attention ! Toutes les séries géométriques ne sont pas convergentes !

► Condition nécessaire et suffisante de convergence d'une série géométrique

Théorème

La série géométrique $\sum r^n$ converge si et seulement si $|r| < 1$. Dans ce cas, sa somme vaut :

$$\sum_{n=0}^{+\infty} r^n = \frac{1}{1-r}$$

Démonstration : Pour tout entier naturel n :

- Si $r \neq 1$, on peut utiliser la formule donnant la somme des $n+1$ premiers termes d'une suite géométrique :

$$\sum_{k=0}^n r^k = \frac{1-r^{n+1}}{1-r}$$

Lorsque n tend vers l'infini, cette dernière quantité n'admet de limite que si $|r| < 1$.

- Si $r = 1$:

$$\sum_{k=0}^n r^k = \sum_{k=0}^n 1 = n+1$$

Lorsque n tend vers l'infini, cette dernière quantité tend vers l'infini.

Si $|r| < 1$, on a alors :

$$\sum_{n=0}^{+\infty} r^n = \lim_{n \rightarrow +\infty} \sum_{k=0}^n r^k = \lim_{n \rightarrow +\infty} \frac{1-r^{n+1}}{1-r} = \frac{1}{1-r}$$

puisque $\lim_{n \rightarrow +\infty} r^{n+1} = 0$. ■

Exemple

$$\sum_{n=0}^{+\infty} \frac{1}{2^n} = \frac{1}{1-\frac{1}{2}} = 2$$

Proposition

La fonction qui, à tout réel x de $] - 1, 1[$, associe $\sum_{n=0}^{+\infty} x^n$ est dérivable, et a pour dérivée la fonction qui, à tout réel x de $] - 1, 1[$, associe :

$$x \mapsto \sum_{n=0}^{+\infty} n x^{n-1} = \sum_{n=1}^{+\infty} n x^{n-1} = \sum_{n=0}^{+\infty} (n+1) x^n = \left(\frac{1}{1-x} \right)' = \frac{1}{(1-x)^2}$$



Grâce à ce résultat, on peut dériver terme à terme.

Ce résultat est admis. $\sum_{n \geq 0} x^n$ est ce que l'on appelle une **série entière**. L'intervalle $] - 1, 1[$ est son domaine de convergence, et 1, son rayon de convergence. Pour tout réel x de ce domaine, la série converge, et on peut dériver ou intégrer terme à terme.

2. Les séries de Riemann

Définition

On appelle **série de Riemann** toute série de la forme $\sum_{n \geq 1} \frac{1}{n^\alpha}$, où α est un réel.

► Condition nécessaire et suffisante de convergence d'une série de Riemann

Théorème

La série de Riemann $\sum_{n \geq 1} \frac{1}{n^\alpha}$ converge si et seulement si $\alpha > 1$.

Exemple

1. La série de Riemann $\sum_{n \geq 1} \frac{1}{n}$, aussi appelée série harmonique, est une série divergente.
2. La série de Riemann $\sum_{n \geq 1} \frac{1}{n^2}$ est une série convergente, dont la somme vaut :

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

(Ce dernier résultat est admis. On peut le retrouver, notamment, à l'aide de séries un peu particulières : les séries de Fourier.)

3. La série de Riemann $\sum_{n \geq 1} \frac{1}{n^4}$ est une série convergente, dont la somme vaut :

$$\sum_{n=1}^{+\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$$

(Ce dernier résultat est admis.)

Critères de convergence pour les séries à termes positifs

1. Outils d'étude

► Comparaison avec une série convergente

Théorème

Soient $\sum u_n$ et $\sum v_n$ deux séries à termes positifs, telles que, pour tout entier naturel n : $u_n \leq v_n$.

Alors, si la série $\sum v_n$ converge, il en est de même de la série $\sum u_n$.

Exemple

Pour étudier la convergence de la série $\sum_{n \geq 1} \frac{e^{-n}}{n^3}$, il suffit de remarquer que, pour tout entier

naturel non nul n : $\frac{e^{-n}}{n^3} \leq \frac{1}{n^3}$.

Comme $\sum_{n \geq 1} \frac{1}{n^3}$ est une série de Riemann convergente, on en déduit la convergence de la série

$$\sum_{n \geq 1} \frac{e^{-n}}{n^3}.$$

► Comparaison avec une série divergente

Théorème

Soient $\sum u_n$ et $\sum v_n$ deux séries à termes positifs, telles que, pour tout entier naturel n : $u_n \leq v_n$.

Alors, si la série $\sum u_n$ diverge, il en est de même de la série $\sum v_n$.

► Équivalence

Théorème

Soient $\sum u_n$ et $\sum v_n$ deux séries à termes positifs, telles que, pour tout entier naturel n : $u_n \sim v_n$.

Alors, si la série $\sum v_n$ converge, il en est de même de la série $\sum u_n$.

Si la série $\sum v_n$ diverge, il en est de même de la série $\sum u_n$.

Exemple

Pour étudier la convergence de la série $\sum_{n \geq 1} \frac{\sin(\frac{1}{n})}{n^4}$, il suffit de remarquer que, lorsque n tend vers l'infini :

$$\frac{\sin(\frac{1}{n})}{n^4} \sim \frac{\frac{1}{n}}{n^4}, \quad \text{soit :} \quad \frac{\sin(\frac{1}{n})}{n^4} \sim \frac{1}{n^5}$$

Comme $\sum_{n \geq 1} \frac{1}{n^5}$ est une série de Riemann convergente, on en déduit la convergence de la série $\sum_{n \geq 1} \frac{\sin\left(\frac{1}{n}\right)}{n^4}$.

2. Critère de d'Alembert

Théorème

Soit $\sum u_n$ une série à termes positifs, non nuls à partir d'un certain rang. Alors :

- si $\lim_{n \rightarrow +\infty} \frac{u_{n+1}}{u_n} < 1$: $\sum u_n$ converge ;
- si $\lim_{n \rightarrow +\infty} \frac{u_{n+1}}{u_n} > 1$: $\sum u_n$ diverge ;
- si $\lim_{n \rightarrow +\infty} \frac{u_{n+1}}{u_n} = 1$: on ne peut pas conclure sur la convergence de $\sum u_n$.

Exemple

On s'intéresse à la convergence de la série $\sum_{n \geq 1} \frac{e^{-2n}}{n^4}$. Comme :

$$\lim_{n \rightarrow +\infty} \frac{\frac{e^{-2(n+1)}}{(n+1)^4}}{\frac{e^{-2n}}{n^4}} = \lim_{n \rightarrow +\infty} \frac{e^{-2(n+1)}}{(n+1)^4} \frac{n^4}{e^{-2n}} = \lim_{n \rightarrow +\infty} \frac{n^4 e^{-2}}{(n+1)^4} = e^{-2} < 1$$

on peut en déduire que la série $\sum_{n \geq 1} \frac{e^{-2n}}{n^4}$ converge.

Définition

Considérons une fonction $f : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$, continue sur $[a, b]$.

On appelle **intégrale de a à b de f** la valeur de l'aire comprise entre la courbe représentative de f sur $[a, b]$ et l'axe des abscisses.

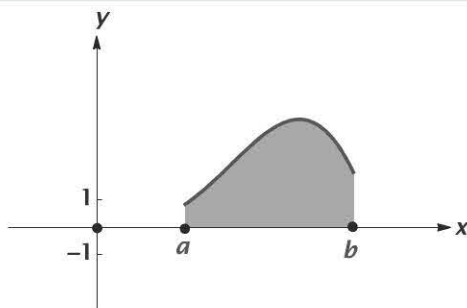


Figure 54.1 – L'aire sous la courbe représentative de f entre a et b .

Le cas présenté ci-dessus ne reflète, hélas, pas la réalité des fonctions que l'on rencontre dans la vie de tous les jours du physicien, du chimiste, du biologiste, ...

En effet, la plupart des fonctions ne possèdent pas toutes des propriétés de régularité, en particulier, elles ne sont pas nécessairement continues (ce qui est le cas, par exemple, d'un effort appliqué localement sur une poutre).

Afin de pouvoir calculer l'intégrale d'une fonction qui n'est pas toujours continue, les mathématiciens ont commencé par construire l'intégrale de fonctions « échelonnées », ou « en escaliers », qui correspondent à des réalités physiques, comme, par exemple, une répartition de chaleur sur un barreau métallique de longueur infinie, en fonction de l'abscisse x :

$$T(x) = \begin{cases} 1 & \text{sur } [-1, 1] \\ 0 & \text{ailleurs} \end{cases}$$

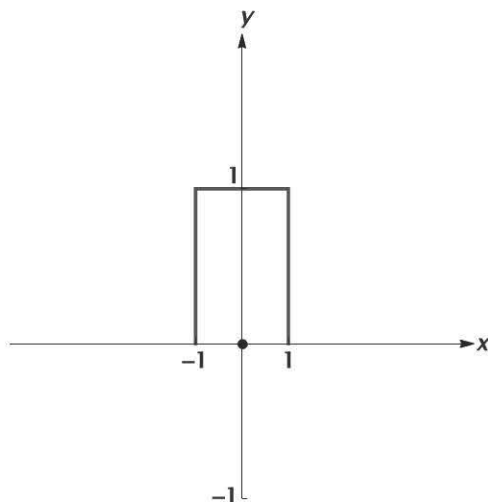


Figure 54.2 – Une fonction « créneau ».

Si on cherche à représenter, par exemple, l'évolution du prix du blé en fonction du temps, en considérant qu'il garde une valeur constante chaque semaine, tout en pouvant augmenter ou diminuer la semaine suivante, on obtient aussi une fonction en escaliers :

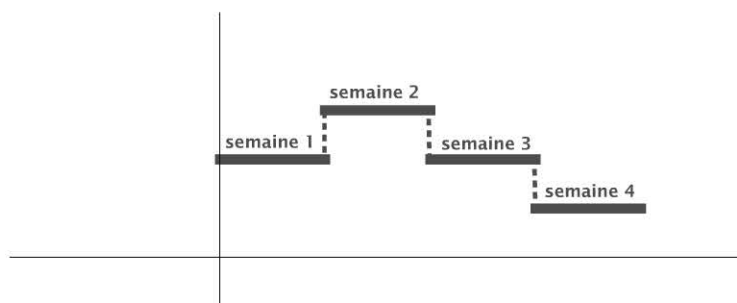


Figure 54.3 – Un exemple de fonction en escalier : l'évolution du prix du blé en fonction du temps.

Ainsi, Bernhard Riemann¹ choisit, pour calculer l'intégrale d'une fonction quelconque, d'approcher celle-ci par des fonctions « en escaliers ».

Il est ensuite très facile d'étendre les résultats obtenus pour ces fonctions « basiques » à des fonctions beaucoup plus régulières !

1. Georg Friedrich Bernhard Riemann (1826-1866), mathématicien allemand. Outre ses travaux sur l'intégration, il a créé la théorie des fonctions algébriques, et développé les travaux de Cauchy sur les fonctions de variables complexes ; en géométrie différentielle, il introduisit le concept de *variété*, qui conduira, ultérieurement, à la géométrie riemannienne.

1. Définitions

► Subdivision d'un intervalle

Étant donné un intervalle $[a, b]$ de \mathbb{R} , et n un entier naturel non nul, on appelle **subdivision de $[a, b]$** tout ensemble σ de réels $\{x_0, x_1, \dots, x_n\}$ tel que :

$$x_0 = a < x_1 < x_2 < \dots < x_{n-1} < x_n = b$$

Figure 55.1 – Une subdivision de l'intervalle $[a, b]$.

Le **pas** de la subdivision est :

$$|\sigma| = \sup \{x_i - x_{i-1}, 1 \leq i \leq n\}$$

► Subdivision régulière

Une subdivision σ est dite **régulière** si elle est de pas constant.

Figure 55.2 – Une subdivision régulière de l'intervalle $[a, b]$.

► Finesse

Étant donnés deux entiers naturels non nuls n et p , et deux subdivisions $\sigma_x = \{x_0, x_1, \dots, x_n\}$ et $\sigma_y = \{y_0, y_1, \dots, y_p\}$ d'un intervalle $[a, b] \subset \mathbb{R}$, la subdivision σ_x est dite **plus fine** que σ_y si :

$$\{y_0, y_1, \dots, y_p\} \subseteq \{x_0, x_1, \dots, x_p\}$$

► Fonction en escalier

Une fonction f , définie sur un intervalle $[a, b]$ de \mathbb{R} , est dite **en escaliers** sur $[a, b]$ s'il existe une subdivision $\sigma = \{x_0, x_1, \dots, x_n\}$ de $[a, b]$ telle que, pour tout i de $\{0, 1, 2, \dots, n-1\}$, f soit constante sur $]x_i, x_{i+1}[$.

► Subdivision adaptée à une fonction en escaliers

Étant donnée une fonction f en escaliers sur un intervalle $[a, b]$ de \mathbb{R} , une subdivision $\sigma = \{x_0, x_1, \dots, x_n\}$ de $[a, b]$ est dite **adaptée à f** si, pour tout i de $\{0, 1, 2, \dots, n-1\}$, f est constante sur $]x_i, x_{i+1}[$.

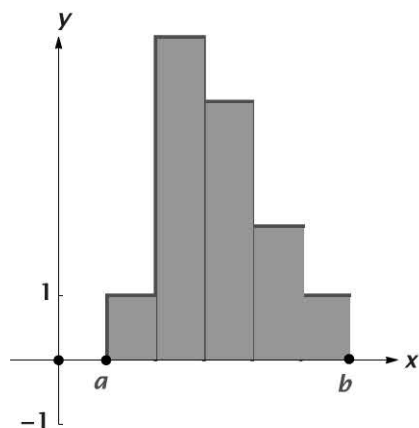


Figure 55.3 – Le graphe d’une fonction en escaliers sur l’intervalle $[a, b]$.

Propriété

Étant donnés deux fonctions f et g en escaliers sur un intervalle $[a, b]$ de \mathbb{R} , et un réel λ , la fonction $f + \lambda g$ est aussi en escaliers sur $[a, b]$ (l’existence d’une subdivision adaptée à f et g est admise).

► Intégrale d’une fonction en escaliers

Soient f une fonction en escaliers sur un intervalle $[a, b]$ de \mathbb{R} , n un entier naturel non nul, et $\sigma = \{x_0, x_1, \dots, x_n\}$ une subdivision de $[a, b]$. Pour tout entier i de $\{0, 1, 2, \dots, n-1\}$, on désigne par y_i la valeur de f sur l’intervalle $]x_i, x_{i+1}[$.

On appelle **intégrale (de Riemann) de f sur $[a, b]$** le réel

$$\int_a^b f(t) dt = \sum_{i=0}^{n-1} (x_{i+1} - x_i) y_i$$

Étant donnée une fonction en escaliers f sur un intervalle $[a, b]$ de \mathbb{R} , la valeur de l’intégrale de f sur $[a, b]$ ne dépend pas de la subdivision (adaptée à f) choisie.

Si f est une fonction en escaliers sur un intervalle $[a, b]$ de \mathbb{R} , alors :

$$\int_b^a f(t) dt = - \int_a^b f(t) dt$$

2. Propriétés

► Linéarité

Étant donnés deux fonctions f et g en escaliers sur un intervalle $[a, b]$ de \mathbb{R} , et un réel λ :

$$\int_a^b (f(t) + \lambda g(t)) dt = \int_a^b f(t) dt + \lambda \int_a^b g(t) dt$$

► Positivité

Étant donnée une fonction f en escaliers sur un intervalle $[a, b]$ de \mathbb{R} , à **valeurs positives** :

$$\int_a^b f(t) dt \geq 0$$

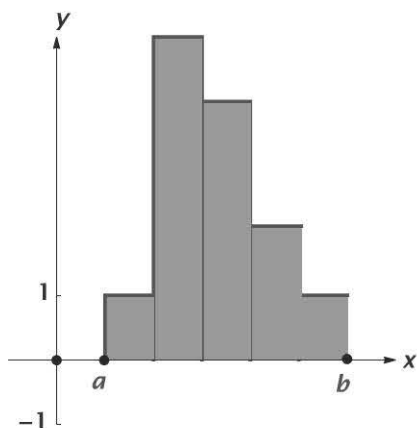


Figure 55.4 – L'intégrale d'une fonction en escaliers sur un intervalle $[a, b]$, à valeurs positives.

► Croissance

Étant données deux fonctions f et g en escaliers sur un intervalle $[a, b]$ de \mathbb{R} telles que :

$$\forall t \in [a, b] : f(t) \leq g(t)$$

alors :

$$\int_a^b f(t) dt \leq \int_a^b g(t) dt$$

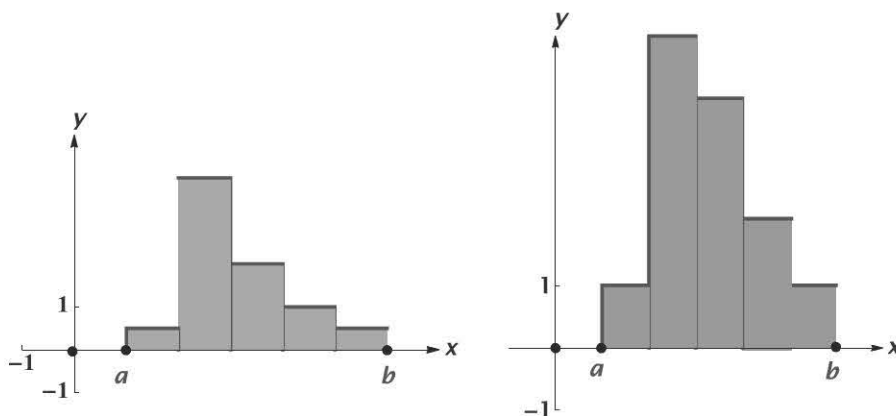


Figure 55.5 – Illustration graphique de la croissance pour une intégrale, dans le cas de fonctions en escaliers (les deux subdivisions ne sont pas nécessairement égales).

► Valeur absolue d'une intégrale

Étant donnée une fonction f en escaliers sur un intervalle $[a, b]$ de \mathbb{R} :

$$\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt$$

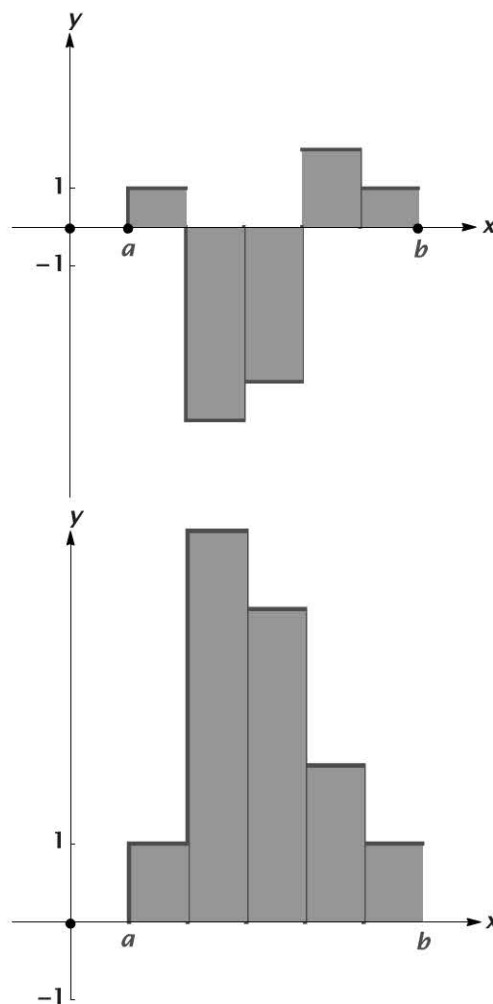


Figure 55.6 – Comparaison entre la valeur absolue d’une intégrale et l’intégrale de la valeur absolue, dans le cas de fonctions en escalier.

► Relation de Chasles

Soit f une fonction f en escaliers sur un intervalle $[a, b]$ de \mathbb{R} . Alors, pour tout c de $[a, b]$:

$$\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$$

► Fonctions « égales presque partout »

Étant données deux fonctions f et g en escaliers sur un intervalle $[a, b]$ de \mathbb{R} , égales sauf en un nombre fini de points¹ :

$$\int_a^b f(t) dt = \int_a^b g(t) dt$$

1. Des fonctions égales, sauf en des points isolés, sont dites « égales presque partout ». L’ensemble de ces points peut être infini. On a ici un cas particulier de « égales presque partout » sur $[a, b]$, puisqu’elles ne diffèrent qu’en un nombre fini de points.

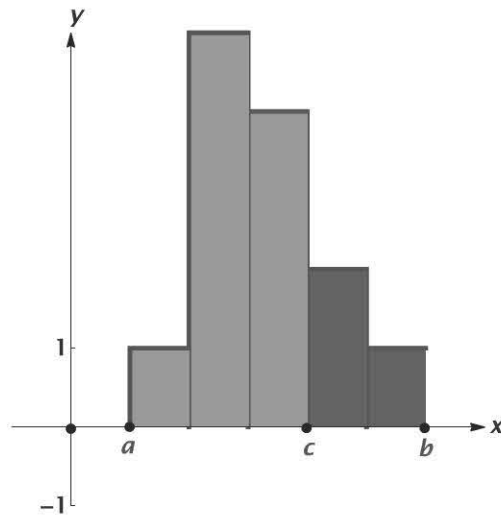


Figure 55.7 – Illustration graphique de la relation de Chasles pour une fonction en escaliers sur un intervalle $[a, b]$.

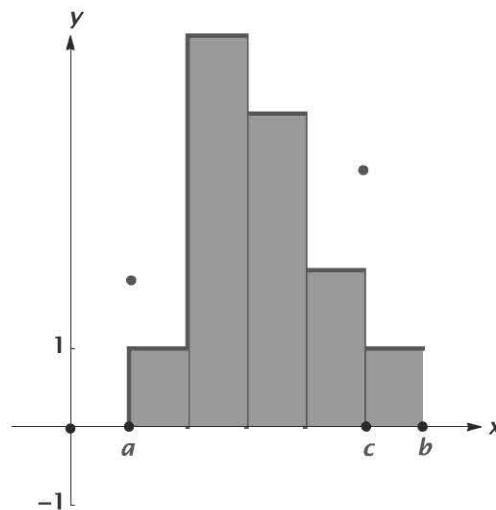


Figure 55.8 – Intégrale de deux fonctions en escaliers sur un intervalle $[a, b]$, égales sauf en a et c .

Intégrale d'une fonction continue par morceaux

1. Définitions

► Fonction continue par morceaux

Une fonction f , définie sur un intervalle $[a, b]$ de \mathbb{R} , est dite **continue par morceaux** sur $[a, b]$ s'il existe une subdivision $\sigma = \{x_0, x_1, \dots, x_n\}$ de $[a, b]$ telle que, pour tout i de $\{0, 1, 2, \dots, n-1\}$, la restriction $f|_{]x_i, x_{i+1}[}$ à l'intervalle $]x_i, x_{i+1}[$ soit continue sur $]x_i, x_{i+1}[$, et prolongeable par continuité en x_i et x_{i+1} .

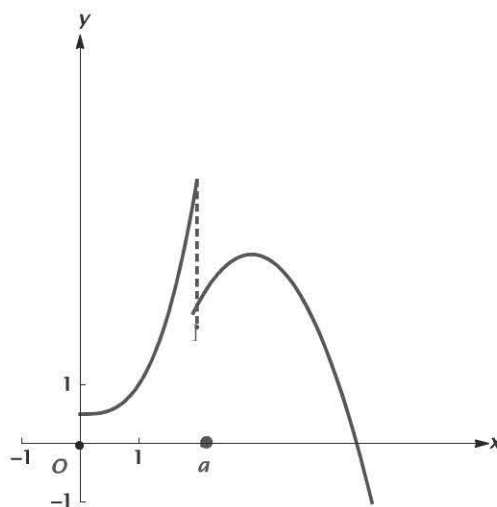


Figure 56.1 – Le graphe d'une fonction continue par morceaux.

► Subdivision adaptée à une fonction continue par morceaux

Étant donnée une fonction f continue par morceaux sur un intervalle $[a, b]$ de \mathbb{R} , une subdivision $\sigma = \{x_0, x_1, \dots, x_n\}$ de $[a, b]$ est dite **adaptée à f** si, pour tout i de $\{0, 1, 2, \dots, n-1\}$, la restriction $f|_{]x_i, x_{i+1}[}$ à l'intervalle $]x_i, x_{i+1}[$ est continue sur $]x_i, x_{i+1}[$, et prolongeable par continuité en x_i et x_{i+1} .

2. Approximation d'une fonction continue par morceaux par des fonctions en escaliers

Théorème

Soit f une fonction continue par morceaux sur un intervalle $[a, b]$ de \mathbb{R} . Alors, il existe deux suites, respectivement croissante et décroissante, de fonctions en escaliers $(\varphi_n)_{n \in \mathbb{N}}$ et $(\psi_n)_{n \in \mathbb{N}}$, et un entier naturel n_0 tels que, pour tout entier naturel $n \geq n_0$, et pour tout t de $[a, b]$:

$$\varphi_n(t) \leq f(t) \leq \psi_n(t) \quad \text{et} \quad 0 \leq \varphi_n(t) - \psi_n(t) < \frac{1}{n}$$

La croissance de la suite $(\varphi_n)_{n \in \mathbb{N}}$ et la décroissance de la suite $(\psi_n)_{n \in \mathbb{N}}$ se traduisent par le fait que, pour tout entier naturel n , et tout t de $[a, b]$:

$$\varphi_n(t) \leq \varphi_{n+1}(t) \quad , \quad \psi_{n+1}(t) \leq \psi_n(t)$$

On dit alors que f est **limite uniforme** des suites $(\varphi_n)_{n \in \mathbb{N}}$ et $(\psi_n)_{n \in \mathbb{N}}$.

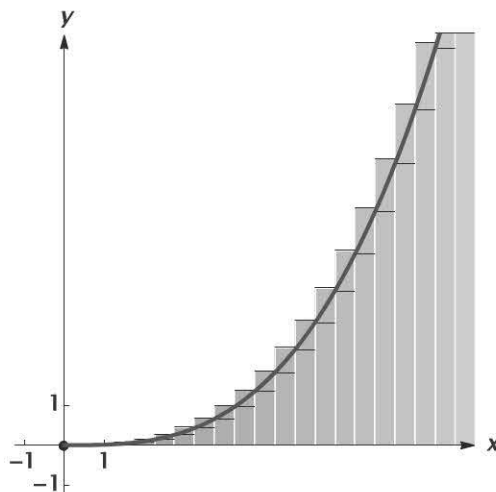


Figure 56.2 – Approximation d'une fonction par une fonction en escaliers.



L'intérêt du théorème précédent est d'étendre la notion d'intégrale vue pour les fonctions en escaliers aux fonctions continues par morceaux.

► Un cas particulier intéressant de limite uniforme : le théorème de Weierstrass

Théorème

Toute fonction continue sur un segment est limite uniforme, sur ce segment, d'une suite de polynômes.

3. Intégrale d'une fonction continue par morceaux

Soit f une fonction continue par morceaux sur un intervalle $[a, b]$ de \mathbb{R} . Alors, pour toute suite de fonctions en escaliers $(\varphi_n)_{n \in \mathbb{N}}$ dont f est limite uniforme, la suite $\left(\int_a^b \varphi_n(t) dt \right)_{n \in \mathbb{N}}$ est convergente, et ne dépend pas du choix de la suite $(\varphi_n)_{n \in \mathbb{N}}$. La limite de la suite $\left(\int_a^b \varphi_n(t) dt \right)_{n \in \mathbb{N}}$ est appelée **intégrale de a à b de f** , et notée $\int_a^b f(t) dt$.

Soit f une fonction continue par morceaux sur un intervalle $[a, b]$ de \mathbb{R} . Alors :

$$\int_b^a f(t) dt = - \int_a^b f(t) dt$$



Cela vient simplement du fait que l'intégrale $\int_b^a f(t) dt$ correspond à une aire comptée négativement.

► Sommes de Riemann, dans le cas d'une subdivision uniforme

Théorème

Soit f une fonction continue sur un intervalle $[a, b]$ de \mathbb{R} .

Alors :

$$\lim_{n \rightarrow +\infty} \frac{b-a}{n} \sum_{k=1}^n f\left(a + k \frac{b-a}{n}\right) = \int_a^b f(t) dt$$

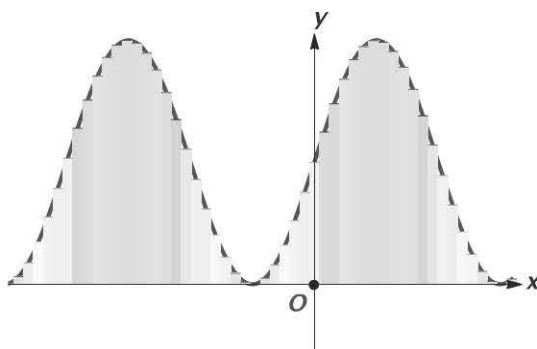


Figure 56.3 – Illustration graphique de la limite d'une somme de Riemann.



1. Ce résultat, que l'on ne démontrera pas dans ce cours, se comprend assez facilement de façon intuitive, dans la mesure où la somme $\frac{b-a}{n} \sum_{k=1}^n f\left(a + k \frac{b-a}{n}\right)$ représente tout simplement la somme des aires des rectangles de longueur $\frac{b-a}{n}$, et de hauteur $f\left(a + k \frac{b-a}{n}\right)$.
2. On a présenté, dans ce qui précède, le cas d'une subdivision uniforme de l'intervalle $[a, b]$. Les sommes de Riemann ne se limitent pas à ce seul cas, et restent encore valables pour des subdivisions non uniformes, sous la condition bien sûr que le pas de celles-ci tende vers zéro lorsque n tend vers $+\infty$.

4. Propriétés

► Linéarité

Étant donnés deux fonctions f et g continues par morceaux sur un intervalle $[a, b]$ de \mathbb{R} , et un réel λ :

$$\int_a^b (f(t) + \lambda g(t)) dt = \int_a^b f(t) dt + \lambda \int_a^b g(t) dt$$

► Positivité

Étant donnée une fonction f continue par morceaux sur un intervalle $[a, b]$ de \mathbb{R} , à **valeurs positives** :

$$\int_a^b f(t) dt \geq 0$$

► Stricte positivité

Étant donnée une fonction f continue par morceaux sur un intervalle $[a, b]$ de \mathbb{R} , à valeurs strictement positives :

$$\int_a^b f(t) dt > 0$$

► Croissance

Étant données deux fonctions f et g continues par morceaux sur un intervalle $[a, b]$ de \mathbb{R} telles que :

$$\forall t \in [a, b] : f(t) \leq g(t)$$

alors :

$$\int_a^b f(t) dt \leq \int_a^b g(t) dt$$

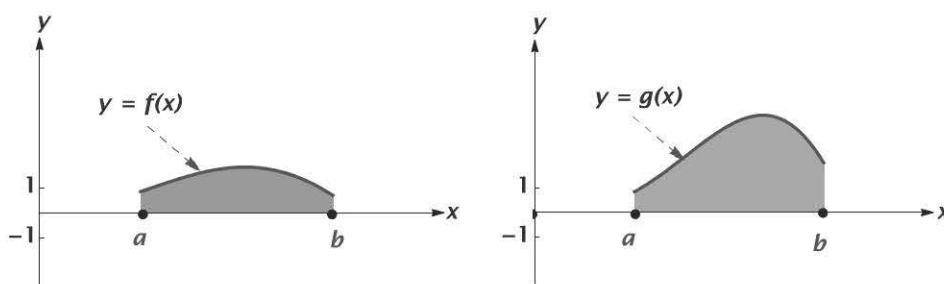


Figure 56.4 – Illustration graphique de la croissance pour une intégrale, dans le cas de fonctions continues par morceaux.

► Valeur absolue d'une intégrale

Étant donnée une fonction f continue par morceaux sur un intervalle $[a, b]$ de \mathbb{R} :

$$\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt$$

► Relation de Chasles

Soit f une fonction f continue par morceaux sur un intervalle $[a, b]$ de \mathbb{R} . Alors, pour tout c de $[a, b]$:

$$\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt$$

► Fonctions « égales presque partout »

Étant données deux fonctions f et g continues par morceaux sur un intervalle $[a, b]$ de \mathbb{R} , égales sauf en un nombre fini de points :

$$\int_a^b f(t) dt = \int_a^b g(t) dt$$

Théorème

Soit f une fonction continue sur un intervalle $[a, b]$ de \mathbb{R} , à valeurs positives ou nulles.

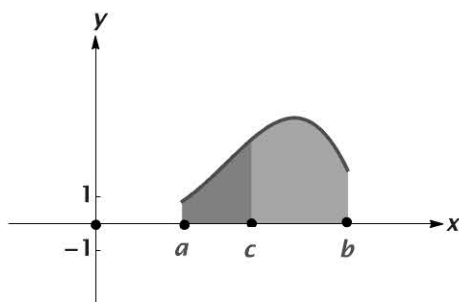


Figure 56.5– Illustration graphique de la relation de Chasles pour l'intégrale d'une fonction continue par morceaux.

Alors, si $\int_a^b f(t) dt = 0$, la fonction f est nulle sur $[a, b]$:

$$\forall t \in [a, b] : f(t) = 0$$

► Inégalité de la moyenne

Théorème

Soit f une fonction continue sur un intervalle $[a, b]$ de \mathbb{R} .

On note :

$$m = \inf_{t \in [a, b]} f(t) \quad , \quad M = \sup_{t \in [a, b]} f(t)$$

Alors :

$$m(b-a) \leq \int_a^b f(t) dt \leq M(b-a)$$

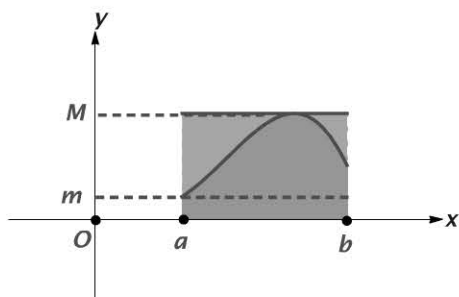


Figure 56.6– Illustration graphique de l'inégalité de la moyenne pour l'intégrale d'une fonction continue sur un intervalle $[a, b]$.

Démonstration : La démonstration ne fait que formaliser les explications de la figure ci-dessus.

Par hypothèse, pour tout t de $[a, b]$:

$$m \leq f(t) \leq M$$

La croissance de l'intégrale permet d'en déduire :

$$\int_a^b m dt \leq \int_a^b f(t) dt \leq \int_a^b M dt$$

soit :

$$m(b-a) \leq \int_a^b f(t) dt \leq M(b-a) \quad \blacksquare$$

► Formule de la moyenne

Théorème

Soient f et g deux fonctions continues sur un intervalle $[a, b]$ de \mathbb{R} . On suppose que la fonction g est à valeurs positives :

$$\forall t \in [a, b] : g(t) \geq 0$$

Alors, il existe un réel c dans l'intervalle $[a, b]$ tel que :

$$\int_a^b f(t) g(t) dt = f(c) \int_a^b g(t) dt$$

Démonstration : f étant continue sur $[a, b]$, il existe deux réels m et M tels que, pour tout t de $[a, b]$:

$$m \leq f(t) \leq M$$

la fonction g étant à valeurs positives, il en résulte, pour tout t de $[a, b]$:

$$m g(t) \leq f(t) g(t) \leq M g(t)$$

Par croissance de l'intégrale, il en résulte :

$$\int_a^b m g(t) dt \leq \int_a^b f(t) g(t) dt \leq \int_a^b M g(t) dt$$

soit :

$$m \int_a^b g(t) dt \leq \int_a^b f(t) g(t) dt \leq M \int_a^b g(t) dt$$

Si $\int_a^b g(t) dt = 0$, n'importe quel c de $[a, b]$ convient.

Si $\int_a^b g(t) dt \neq 0$, on peut diviser membre à membre par $\int_a^b g(t) dt$:

$$m \leq \frac{\int_a^b f(t) g(t) dt}{\int_a^b g(t) dt} \leq M$$

Le théorème des valeurs intermédiaires permet alors d'en déduire l'existence d'un réel c de $[a, b]$ tel que :

$$f(c) = \frac{\int_a^b f(t) g(t) dt}{\int_a^b g(t) dt}$$

ce qui conduit au résultat cherché. ■

1. Primitives

Définition

Soit f une fonction continue sur un intervalle I de \mathbb{R} . Une fonction F , dérivable sur I , est une **primitive** de f sur I si et seulement si :

$$F' = f \text{ sur } I$$

Propriété

Soit f une fonction continue sur un intervalle I de \mathbb{R} . Alors, les primitives de f sur I diffèrent toutes d'une constante.

Démonstration : Soient F_1 et F_2 deux primitives de f sur I :

$$F'_1 = F'_2$$

ce qui conduit à :

$$F'_1 - F'_2 = 0$$

soit :

$$(F_1 - F_2)' = 0$$

Il en résulte :

$$F_1 - F_2 = \text{Fonction à valeur constante sur } I$$

qui est le résultat cherché. ■



Ainsi, si F est une primitive sur I de la fonction f , alors, pour tout réel K , la fonction $t \in I \mapsto F(t) + K$ est aussi une primitive de f sur I , ce qui est assez logique, dans la mesure où la dérivée d'une fonction constante est nulle !

2. Le théorème fondamental de l'analyse

Théorème

Soit f une fonction continue sur un intervalle I de \mathbb{R} , et a un réel appartenant à I . L'application de I dans \mathbb{R} , qui, au réel x de I , associe $\int_a^x f(t) dt$, est une primitive de f sur I .

Démonstration : Désignons par F la fonction qui, à tout x de I , associe $\int_a^x f(t) dt$. Il s'agit donc de montrer que, pour x_0 donné dans I , F est dérivable en x_0 , avec $F'(x_0) = f(x_0)$.

Soit $\varepsilon > 0$. f étant continue en x_0 , il existe un réel η_x tel que, pour tout x vérifiant $|x - x_0| \leq \eta_x$:

$$|f(x) - f(x_0)| \leq \varepsilon$$

Par suite, pour un tel x vérifiant $x > x_0$:

$$\begin{aligned} \left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| &= \left| \frac{\int_a^x f(t) dt - \int_a^{x_0} f(t) dt}{x - x_0} - \frac{(x - x_0) f(x_0)}{x - x_0} \right| \\ &= \left| \frac{\int_a^x f(t) dt + \int_{x_0}^a f(t) dt}{x - x_0} - \frac{f(x_0) \int_{x_0}^x dt}{x - x_0} \right| \\ &= \left| \frac{\int_{x_0}^x f(t) dt}{x - x_0} - \frac{\int_{x_0}^x f(x_0) dt}{x - x_0} \right| \\ &= \left| \frac{\int_{x_0}^x f(t) dt}{x - x_0} - \frac{f(x_0) \int_{x_0}^x dt}{x - x_0} \right| \\ &= \left| \frac{\int_{x_0}^x (f(t) - f(x_0)) dt}{x - x_0} \right| \\ &\leq \frac{\int_{x_0}^x |f(t) - f(x_0)| dt}{|x - x_0|} \\ &\leq \frac{\int_{x_0}^x \varepsilon dt}{|x - x_0|} = \varepsilon \end{aligned}$$

De même, pour un tel x vérifiant $x < x_0$: $\left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| \leq \varepsilon$.

Le réel ε étant quelconque, il peut être choisi aussi petit que possible. Il en résulte :

$$\lim_{x \rightarrow x_0} \frac{F(x) - F(x_0)}{x - x_0} = f(x_0)$$

soit $F'(x_0) = f(x_0)$, qui est le résultat cherché. ■

Théorème

Soit f une fonction continue sur un intervalle I de \mathbb{R} , et a un réel appartenant à I . Pour tout réel C , il existe une unique primitive de f prenant en a la valeur C . C'est l'application :

$$\begin{aligned} I &\rightarrow \mathbb{R} \\ x &\mapsto C + \int_a^x f(t) dt \end{aligned}$$

Théorème

Soit f une fonction continue sur un intervalle $[a, b]$ de \mathbb{R} , et F une primitive de f . Alors :

$$\int_a^b f(t) dt = F(b) - F(a)$$



On notera : $[F(t)]_a^b = F(b) - F(a)$.

3. Propriétés

► Intégration par parties

Théorème

Soient f et g deux fonctions de classe C^1 sur un intervalle $[a, b]$ de \mathbb{R} .

Alors :

$$\int_a^b f(t) g'(t) dt = [f(t) g(t)]_a^b - \int_a^b f'(t) g(t) dt$$

Démonstration : Les fonctions f et g étant de classe C^1 sur $[a, b]$, il en est de même de leur produit $f g$. De plus, pour tout réel t de $[a, b]$:

$$(f(t) g(t))' = f(t) g'(t) + f'(t) g(t)$$

f et g étant de classe C^1 sur $[a, b]$, les fonctions $f g'$ et $f' g$ sont continues sur $[a, b]$, et donc intégrables :

$$\int_a^b (f(t) g(t))' dt = \int_a^b f(t) g'(t) dt + \int_a^b f'(t) g(t) dt$$

soit :

$$[f(t) g(t)]_a^b = \int_a^b f(t) g'(t) dt + \int_a^b f'(t) g(t) dt$$

D'où le résultat. ■

Exemple

Soit x un réel strictement positif. Alors :

$$\begin{aligned} \int_1^x \ln(t) dt &= \int_1^x \ln(t) 1 dt \\ &= [\ln(t) t]_1^x - \int_1^x \frac{1}{t} t dt \\ &= [\ln(t) t]_1^x - \int_1^x 1 dt \\ &= x \ln(x) - \int_1^x 1 dt \\ &= x \ln(x) - (x - 1) \\ &= x \ln(x) - x + 1 \end{aligned}$$

► Changement de variable

Théorème

Soit f une fonction continue sur un intervalle I de \mathbb{R} , et φ une fonction de classe C^1 sur un intervalle J de \mathbb{R} telle que :

$$\varphi(J) \subseteq I$$

Alors, pour tout couple de réels $(\alpha, \beta) \in J^2$:

$$\int_{\alpha}^{\beta} f(\varphi(t)) \varphi'(t) dt = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(t) dt$$

Démonstration : f étant continue sur I , elle possède au moins une primitive F sur I . Par dérivation composée :

$$(F \circ \varphi)' = f \circ \varphi \cdot \varphi'$$

$f \circ \varphi \cdot \varphi'$ est continue sur J , comme produit de fonctions continues sur J . En intégrant de α à β , on en déduit :

$$\int_{\alpha}^{\beta} (F \circ \varphi)'(t) dt = \int_{\alpha}^{\beta} (f \circ \varphi)(t) \cdot \varphi'(t) dt$$

soit :

$$[(F \circ \varphi)(t)]_{\alpha}^{\beta} = \int_{\alpha}^{\beta} (f \circ \varphi)(t) \cdot \varphi'(t) dt$$

ou encore :

$$F(\varphi(\beta)) - F(\varphi(\alpha)) = \int_{\alpha}^{\beta} (f \circ \varphi)(t) \cdot \varphi'(t) dt$$

qui s'écrit aussi :

$$[F(t)]_{\varphi(\alpha)}^{\varphi(\beta)} = \int_{\alpha}^{\beta} (f \circ \varphi)(t) \cdot \varphi'(t) dt$$

mais également :

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(t) dt = \int_{\alpha}^{\beta} (f \circ \varphi)(t) \cdot \varphi'(t) dt$$

D'où le résultat ! ■

Exemple

Considérons l'intégrale $I = \int_0^1 e^t \sqrt{e^t + 2} dt$. Il est judicieux d'effectuer le changement de variable $e^t = u$. L'application $t \mapsto e^t$ est de classe C^1 .

On peut donc appliquer le théorème du changement de variable :

$$I = \int_1^e \sqrt{u+2} du = \int_1^e (u+2)^{\frac{1}{2}} du = \frac{2}{3} [(u+2)^{\frac{3}{2}}]_1^e = \frac{2}{3} (e+2)^{\frac{3}{2}} - 2\sqrt{27} = \frac{2}{3} (e+2)^{\frac{3}{2}} - 2\sqrt{3}$$

► Intégrale et parité

Propriété

Soit a un réel positif, et f une fonction continue et paire sur l'intervalle $[-a, +a]$.

Alors :

$$\int_{-a}^a f(t) dt = 2 \int_0^a f(t) dt$$

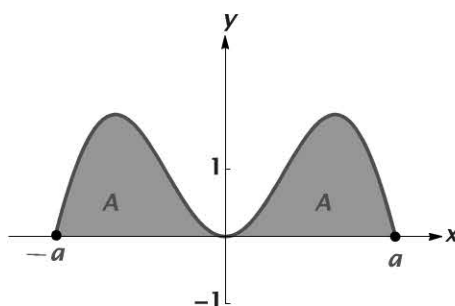


Figure 57.1 – L'intégrale d'une fonction paire.

Démonstration : On décompose tout simplement l'intégrale à l'aide de la relation de Chasles, puis on effectue un changement de variable :

$$\begin{aligned} \int_{-a}^a f(t) dt &= \int_{-a}^0 f(t) dt + \int_0^a f(t) dt \\ &= \int_a^0 f(-t) (-dt) + \int_0^a f(t) dt \\ &= - \int_a^0 f(t) dt + \int_0^a f(t) dt \quad (f \text{ étant paire : } f(-t) = f(t)) \\ &= \int_0^a f(t) dt + \int_0^a f(t) dt \\ &= 2 \int_0^a f(t) dt \end{aligned}$$

■

► Intégrale et imparité

Propriété

Soit a un réel positif, et f une fonction continue et impaire sur l'intervalle $[-a, +a]$.

Alors :

$$\int_{-a}^a f(t) dt = 0$$

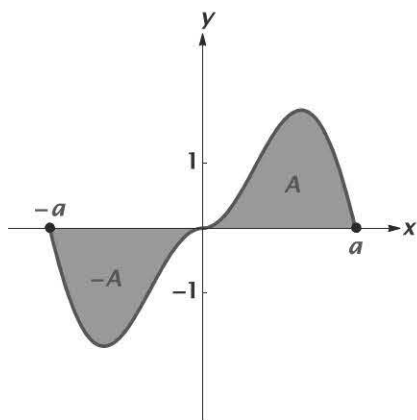


Figure 57.2 – L'intégrale d'une fonction impaire.

Démonstration : On décompose à nouveau l'intégrale à l'aide de la relation de Chasles, puis on effectue un changement de variable :

$$\begin{aligned}
 \int_{-a}^a f(t) dt &= \int_{-a}^0 f(t) dt + \int_0^a f(t) dt \\
 &= \int_a^0 f(-t) (-dt) + \int_0^a f(t) dt \\
 &= - \int_a^0 -f(t) dt + \int_0^a f(t) dt \quad (f \text{ étant impaire : } f(-t) = -f(t)) \\
 &= \int_a^0 f(t) dt + \int_0^a f(t) dt \\
 &= - \int_0^a f(t) dt + \int_0^a f(t) dt \\
 &= 0
 \end{aligned}$$

■

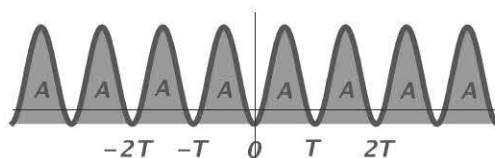
► Intégrale et périodicité

Propriété

Soit T un réel non nul, et f une fonction continue sur \mathbb{R} , de période T .

Alors, pour tout réel a :

$$\int_a^{a+T} f(t) dt = \int_0^T f(t) dt$$

Figure 57.3 – L'intégrale d'une fonction continue de période T .

Primitives de fractions rationnelles

1. Détermination d'une primitive

Connaître la décomposition en éléments simples d'une fonction (réelle) « rationnelle » permet d'en déterminer facilement une primitive :

- lorsque la fonction est de la forme $x \in \mathbb{R} \setminus \{x_0\} \mapsto \frac{\alpha}{x - x_0}$, $\alpha \in \mathbb{R}$, $x_0 \in \mathbb{R}$, une primitive est :

$$x \in \mathbb{R} \setminus \{x_0\} \mapsto \alpha \ln |x - x_0|$$

- lorsque la fonction est de la forme $x > x_0 \mapsto \frac{\alpha}{(x - x_0)^s} = \alpha (x - x_0)^{-s}$, $\alpha \in \mathbb{R}$, $x_0 \in \mathbb{R}$, $s \in \mathbb{R}$, $s \neq 1$, une primitive est :

$$x > x_0 \mapsto \frac{\alpha}{1 - s} (x - x_0)^{1-s} = \frac{\alpha}{(1 - s)(x - x_0)^{s-1}}$$

- lorsque la fonction est de la forme $x \in \mathbb{R} \mapsto \frac{\alpha x + \beta}{x^2 + \lambda x + \mu}$, $(\alpha, \beta, \lambda, \mu) \in \mathbb{R}^4$, $\lambda^2 - 4\mu < 0$, il est intéressant de remarquer que :

$$\begin{aligned} \frac{\alpha}{2} \frac{2x + \lambda}{x^2 + \lambda x + \mu} + \frac{\beta - \frac{\lambda\alpha}{2}}{x^2 + \lambda x + \mu} &= \frac{\alpha}{2} \frac{2x + \lambda}{x^2 + \lambda x + \mu} + \frac{\beta - \frac{\lambda\alpha}{2}}{\left(x + \frac{\lambda}{2}\right)^2 + \mu - \frac{\lambda^2}{4}} \\ &= \frac{\alpha}{2} \frac{2x + \lambda}{x^2 + \lambda x + \mu} + \frac{\frac{\beta - \frac{\lambda\alpha}{2}}{\mu - \frac{\lambda^2}{4}}}{\frac{\left(x + \frac{\lambda}{2}\right)^2}{\mu - \frac{\lambda^2}{4}} + 1} \end{aligned}$$

Une primitive de $x \in \mathbb{R} \mapsto \frac{\alpha x + \beta}{x^2 + \lambda x + \mu}$ est alors de la forme :

$$x \in \mathbb{R} \mapsto \frac{\alpha}{2} \ln |x^2 + \lambda x + \mu| + \frac{\beta - \frac{\lambda\alpha}{2}}{\mu - \frac{\lambda^2}{4}} \sqrt{\mu - \frac{\lambda^2}{4}} \arctan \left(\frac{x + \frac{\lambda}{2}}{\sqrt{\mu - \frac{\lambda^2}{4}}} \right)$$

c'est-à-dire

$$x \in \mathbb{R} \mapsto \frac{\alpha}{2} \ln |x^2 + \lambda x + \mu| + \frac{\beta - \frac{\lambda\alpha}{2}}{\sqrt{\mu - \frac{\lambda^2}{4}}} \arctan \left(\frac{x + \frac{\lambda}{2}}{\sqrt{\mu - \frac{\lambda^2}{4}}} \right)$$

2. Forme canonique

Définition

La factorisation de $x^2 + \lambda x + \mu$ en $\left(x + \frac{\lambda}{2}\right)^2 + \mu - \frac{\lambda^2}{4}$ est appelée « **forme canonique** ».

Exemple

Déterminons une primitive de la fonction $x \in \mathbb{R} \mapsto \frac{x-1}{x^2+x+1}$.

On a :

$$\begin{aligned}\frac{x-1}{x^2+x+1} &= \frac{1}{2} \frac{2x+1}{x^2+x+1} - \frac{3}{2} \frac{1}{x^2+x+1} \\ &= \frac{1}{2} \frac{2x+1}{x^2+x+1} - \frac{3}{2} \frac{1}{\left(x+\frac{1}{2}\right)^2 - \frac{1}{4} + 1} \\ &= \frac{1}{2} \frac{2x+1}{x^2+x+1} - \frac{3}{2} \frac{1}{\left(x+\frac{1}{2}\right)^2 + \frac{3}{4}} \\ &= \frac{1}{2} \frac{2x+1}{x^2+x+1} - \frac{3}{2} \frac{4}{\frac{4}{3} \left(x+\frac{1}{2}\right)^2 + 1} \\ &= \frac{1}{2} \frac{2x+1}{x^2+x+1} - 2 \frac{1}{\frac{4}{3} \left(x+\frac{1}{2}\right)^2 + 1}\end{aligned}$$

Ainsi, une primitive de $x \in \mathbb{R} \mapsto \frac{x-1}{x^2+x+1}$ est :

$$x \mapsto \frac{1}{2} \ln(x^2+x+1) - 2 \frac{\sqrt{3}}{2} \arctan\left(\frac{2}{\sqrt{3}} \left(x + \frac{1}{2}\right)\right)$$

c'est-à-dire :

$$x \mapsto \frac{1}{2} \ln(x^2+x+1) - \sqrt{3} \arctan\left(\frac{2}{\sqrt{3}} \left(x + \frac{1}{2}\right)\right)$$

1. La méthode des rectangles

► Le calcul approché en lui-même

La méthode des rectangles, qui consiste à approcher la valeur d'une intégrale par la somme des aires de rectangles situés sous la courbe, utilise tout simplement les sommes de Riemann vues précédemment ; pour l'intégrale d'une fonction f continue par morceaux sur un intervalle $[a, b]$ de \mathbb{R} , on commence donc par diviser $[a, b]$ en n sous-intervalles, de longueur $\frac{b-a}{n}$.

On choisit ensuite de considérer :

- soit les rectangles de hauteur $f\left(a + k \frac{b-a}{n}\right)$, $k = 0, 1, \dots, n-1$, et de largeur $\frac{b-a}{n}$.

L'intégrale $\int_a^b f(t) dt$ est alors approchée par :

$$\frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right)$$

- soit les rectangles de hauteur $f\left(a + k \frac{b-a}{n}\right)$, $k = 1, \dots, n$, de largeur $\frac{b-a}{n}$. L'intégrale

$\int_a^b f(t) dt$ est alors approchée par :

$$\frac{b-a}{n} \sum_{k=1}^n f\left(a + k \frac{b-a}{n}\right)$$

► Étude de l'erreur commise

Au-delà du calcul approché, il peut être intéressant de quantifier l'erreur ainsi commise, en valeur absolue. Pour la première configuration, elle est donnée par :

$$\begin{aligned} & \left| \int_a^b f(t) dt - \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) \right| \\ &= \left| \sum_{k=0}^{n-1} \int_{a+k \frac{b-a}{n}}^{a+(k+1) \frac{b-a}{n}} f(t) dt - \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) \right| \\ &= \left| \sum_{k=0}^{n-1} \int_{a+k \frac{b-a}{n}}^{a+(k+1) \frac{b-a}{n}} \left(f(t) - f\left(a + k \frac{b-a}{n}\right) \right) dt \right| \\ &\leq \sum_{k=0}^{n-1} \int_{a+k \frac{b-a}{n}}^{a+(k+1) \frac{b-a}{n}} \left| f(t) - f\left(a + k \frac{b-a}{n}\right) \right| dt \end{aligned}$$

Considérons le cas d'une fonction de classe C^1 ; d'après l'inégalité des accroissements finis, pour tout k de $\{0, 1, \dots, n-1\}$, et tout t de $\left[a + k \frac{b-a}{n}, a + (k+1) \frac{b-a}{n}\right]$:

$$\left| f(t) - f\left(a + k \frac{b-a}{n}\right) \right| \leq \left(t - a - k \frac{b-a}{n} \right) \max_{t \in [a, b]} |f'|$$

on a alors :

$$\begin{aligned} \left| \int_a^b f(t) dt - \frac{b-a}{n} \sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) \right| &\leq \max_{t \in [a, b]} |f'| \sum_{k=0}^{n-1} \int_{a+k \frac{b-a}{n}}^{a+(k+1) \frac{b-a}{n}} \left(t - a - k \frac{b-a}{n} \right) dt \\ &= \max_{t \in [a, b]} |f'| \sum_{k=0}^{n-1} \left[\frac{1}{2} \left(t - a - k \frac{b-a}{n} \right)^2 \right]_{a+k \frac{b-a}{n}}^{a+(k+1) \frac{b-a}{n}} \\ &= \max_{t \in [a, b]} |f'| \sum_{k=0}^{n-1} \frac{1}{2} \frac{(b-a)^2}{n^2} \\ &= \max_{t \in [a, b]} |f'| \frac{1}{2} \frac{(b-a)^2}{n^2} \sum_{k=0}^{n-1} 1 \\ &= \max_{t \in [a, b]} |f'| \frac{(b-a)^2}{2n} \end{aligned}$$

(Pour simplifier le calcul, on a choisi de primitiver $t \mapsto t - a - k \frac{b-a}{n}$ par $t \mapsto \frac{1}{2} \left(t - a - k \frac{b-a}{n} \right)^2$ au lieu de $t \mapsto \frac{1}{2} t^2 - at - k \frac{b-a}{n} t$, afin d'obtenir la valeur 0 en $a + k \frac{b-a}{n}$, et de simplifier le calcul en $a + (k+1) \frac{b-a}{n}$!)

L'erreur commise est donc un $O\left(\frac{1}{n}\right)$.



Il est également possible d'obtenir cette majoration en appliquant, tout simplement, la formule de la moyenne.

Exemple

Déterminons la limite, lorsque n tend vers $+\infty$, de $\left(\frac{(2n)!}{n^n n!} \right)^{\frac{1}{n}}$.

On a :

$$\begin{aligned} \left(\frac{(2n)!}{n^n n!} \right)^{\frac{1}{n}} &= \exp\left(\frac{1}{n} \ln \frac{(2n)!}{n^n n!} \right) \\ &= \exp\left(\frac{1}{n} \ln \frac{1 \times 2 \times \dots \times (2n)}{n^n 1 \times 2 \times \dots \times n} \right) \\ &= \exp\left(\frac{1}{n} \ln \frac{(n+1) \times (n+2) \times \dots \times (2n)}{n^n} \right) \\ &= \exp\left(\frac{1}{n} \ln \left(1 + \frac{1}{n} \right) \times \left(1 + \frac{2}{n} \right) \times \dots \times \left(1 + \frac{n}{n} \right) \right) \\ &= \exp\left(\frac{1}{n} \ln \prod_{k=0}^{n-1} \left(1 + \frac{k}{n} \right) \right) \\ &= \exp\left(\frac{1}{n} \sum_{k=0}^{n-1} \ln \left(1 + \frac{k}{n} \right) \right) \end{aligned}$$

On reconnaît alors :

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \ln\left(1 + \frac{k}{n}\right) = \int_0^1 \ln(1+t) dt = [(t+1) \ln(1+t) - t]_0^1 = 2 \ln 2 - 1$$

Par continuité de la fonction *exponentielle*, il en résulte :

$$\lim_{n \rightarrow +\infty} \left(\frac{(2n)!}{n^n n!} \right)^{\frac{1}{n}} = e^{2 \ln 2 - 1} = \frac{4}{e}$$

2. La méthode des trapèzes

Une alternative à la méthode des rectangles consiste à approcher la valeur d'une intégrale par la somme d'aires de trapèzes, qui, graphiquement déjà, sont plus proches de la courbe que les rectangles précédents.

On rappelle la formule donnant l'aire d'un trapèze :

$$\frac{(\text{grande base} + \text{petite base}) \times \text{hauteur}}{2}$$

Ainsi, en considérant la suite des trapèzes de bases respectives $f\left(a + k \frac{b-a}{n}\right)$ et $f\left(a + (k+1) \frac{b-a}{n}\right)$, de hauteur $\frac{b-a}{n}$, $k = 1, 2, \dots, n-1$:

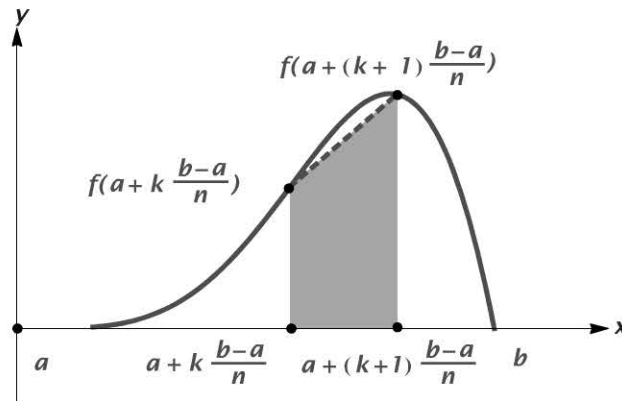


Figure 59.1 – La méthode des trapèzes.

L'intégrale $\int_a^b f(t) dt$ est alors approchée par :

$$(b-a) \sum_{k=0}^{n-1} \frac{f\left(a + k \frac{b-a}{n}\right) + f\left(a + (k+1) \frac{b-a}{n}\right)}{2n}$$

qui s'écrit aussi :

$$\begin{aligned}
 & (b-a) \frac{1}{2n} \left[\sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) + \sum_{k=0}^{n-1} f\left(a + (k+1) \frac{b-a}{n}\right) \right] \\
 = & (b-a) \frac{1}{2n} \left[\sum_{k=0}^{n-1} f\left(a + k \frac{b-a}{n}\right) + \sum_{k=1}^n f\left(a + k \frac{b-a}{n}\right) \right] \\
 = & (b-a) \frac{1}{2n} \left[f(a) + 2 \sum_{k=1}^{n-1} f\left(a + k \frac{b-a}{n}\right) + f(b) \right] \\
 = & \frac{b-a}{n} \left[\frac{f(a) + f(b)}{2} + \sum_{k=1}^{n-1} f\left(a + k \frac{b-a}{n}\right) \right]
 \end{aligned}$$

(On a effectué un changement d'indices dans la deuxième somme, puis regroupé les termes communs aux deux sommes, c'est-à-dire entre 1 et $n-1$).

Graphiquement, on voit bien qu'approcher la fonction par les trapèzes ci-dessus permet une meilleure approximation. L'erreur commise (que nous n'étudierons pas ici) est alors un $O\left(\frac{1}{n^2}\right)$, ce qui est donc plus précis que pour la méthode des rectangles.

Il existe beaucoup de nombreuses méthodes permettant le calcul approché d'une intégrale. Après la méthode des trapèzes, on trouve, en général, la méthode de Simpson¹, qui consiste à approcher la fonction sur l'intervalle d'intégration par un polynôme quadratique, c'est-à-dire de degré 2, prenant les mêmes valeurs que celle-ci aux bornes de l'intervalle. Par rapport à la méthode des trapèzes, l'erreur commise est plus faible, dans la mesure où, graphiquement, on approche la fonction par une courbe et non une droite. En outre, il est clair qu'une expression polynomiale s'intègre facilement.

3. Une application de la méthode des trapèzes : une variante de la formule de Stirling

Considérons la fonction *logarithme népérien* : celle-ci étant concave, sa courbe représentative est située au-dessus des cordes. Ainsi, pour tout entier naturel k , l'aire sous la courbe entre k et $k+1$ est plus grande que l'aire du trapèze ayant pour sommets les points de coordonnées respectives $(k, 0)$, $(k+1, 0)$, $(k+1, f(k+1))$, $(k, f(k))$:

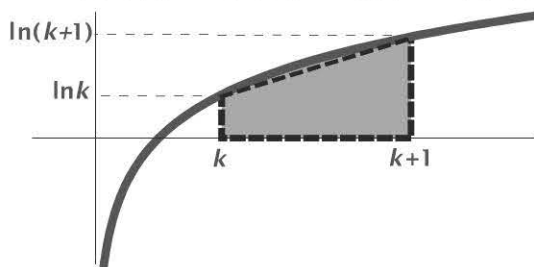


Figure 59.2 – La fonction logarithme népérien, et son intégrale entre k et $k+1$.

$$\int_k^{k+1} \ln t \, dt \geq \frac{\ln k + \ln(k+1)}{2}$$

1. Thomas Simpson (1710-1761), mathématicien anglais, essentiellement connu pour ses travaux sur le calcul infinitésimal (la méthode de Simpson, qui permet un calcul approché de l'aire sous une courbe), mais qui fut aussi l'auteur d'un important traité de trigonométrie, et à qui l'on doit les formules permettant de transformer un produit de *cosinus* ou *sinus* en somme, et vice versa.

On en déduit, en additionnant membre à membre les inégalités ainsi obtenues pour tout entier k de $\{1, \dots, n-1\}$:

$$\sum_{k=1}^{n-1} \int_k^{k+1} \ln t \, dt \geq \sum_{k=1}^{n-1} \frac{\ln k + \ln(k+1)}{2}$$

soit, grâce à la relation de Chasles, et en effectuant un changement d'indices dans

$$\sum_{k=1}^{n-1} \ln(k+1) :$$

$$\int_1^n \ln t \, dt \geq \frac{1}{2} \sum_{k=1}^{n-1} \ln k + \frac{1}{2} \sum_{k=2}^n \ln k$$

c'est-à-dire $[t \ln t - t]_1^n \geq \sum_{k=2}^{n-1} \ln k + \frac{\ln n}{2}$, ou encore :

$$n \ln n - n + 1 \geq \sum_{k=2}^n \ln k - \frac{\ln n}{2}$$

On reconnaît alors : $\ln \left(\prod_{k=2}^n k \right) = \ln(n!) - \frac{\ln n}{2}$. Il en résulte :

$$n \ln n - n + 1 \geq \ln(n!) - \frac{\ln n}{2}$$

puis : $\left(n + \frac{1}{2}\right) \ln n - n + 1 \geq \ln(n!)$.

Par croissance de la fonction exponentielle, on en déduit alors :

$$\exp \left(\left(n + \frac{1}{2} \right) \ln n - n + 1 \right) \geq n!$$

soit :

$$e n^n \sqrt{n} e^{-n} \geq n!$$

La concavité de la fonction logarithme népérien fait aussi que sa courbe représentative est située en-dessous de ses tangentes : ainsi, pour tout entier naturel k , la courbe est située en-dessous de la tangente en k , ce qui se traduit par :

$$f(t) \leq \ln k + \frac{t-k}{k}$$

On intègre alors cette dernière relation entre k et $k + \frac{1}{2}$:

$$\int_k^{k+\frac{1}{2}} \ln t \, dt \leq \frac{\ln k}{2} + \frac{1}{k} \left[\frac{1}{2} (t-k)^2 \right]_k^{k+\frac{1}{2}}$$

(Afin de simplifier les calculs, on primitive $t \mapsto t - k$ par $t \mapsto \frac{1}{2} (t - k)^2$, qui présente l'avantage de s'annuler en $t = k$, et se calcule aisément en $t = k + \frac{1}{2}$.)

Il en résulte :

$$\int_k^{k+\frac{1}{2}} \ln t \, dt \leq \frac{\ln k}{2} + \frac{1}{8k}$$

De même, la courbe est située au-dessus de sa tangente en $k+1$, pour tout entier k de $\{1, \dots, n-1\}$.

Ainsi, pour tout t de $\left[k + \frac{1}{2}, k+1\right]$:

$$\ln t \leq \ln(k+1) + \frac{t-k-1}{k+1}$$

On intègre alors cette dernière relation entre $k + \frac{1}{2}$ et $k+1$:

$$\int_{k+\frac{1}{2}}^{k+1} \ln t \, dt \leq \frac{\ln(k+1)}{2} + \frac{1}{k+1} \left[\frac{1}{2} (t-k-1)^2 \right]_{k+\frac{1}{2}}^{k+1}$$

(Afin de simplifier les calculs, on primitive $t \mapsto t-k-1$ par $t \mapsto \frac{1}{2} (t-k-1)^2$, qui présente l'avantage de s'annuler en $t = k+1$, et se calcule aisément en $t = k + \frac{1}{2}$.) Il en résulte :

$$\int_{k+\frac{1}{2}}^{k+1} \ln t \, dt \leq \frac{\ln(k+1)}{2} - \frac{1}{8(k+1)}$$

Il ne reste plus qu'à sommer les inégalités précédentes pour $k = 1, \dots, n-1$:

$$\int_1^n \ln t \, dt \leq \frac{\ln(1)}{2} + \ln 2 + \ln 3 + \dots + \ln(n-2) + \ln(n-1) + \frac{\ln n}{2} + \frac{1}{8} \sum_{k=1}^{n-1} \left(\frac{1}{k} - \frac{1}{k+1} \right)$$

soit :

$$[t \ln t - t]_1^n \leq \frac{\ln 1}{2} + \ln 2 + \ln 3 + \dots + \ln(n-2) + \ln(n-1) + \frac{\ln n}{2} + \frac{1}{8} \sum_{k=1}^{n-1} \left(\frac{1}{k} - \frac{1}{k+1} \right)$$

$$\text{soit : } n \ln n - n + 1 \leq \sum_{k=1}^{n-1} \ln k + \frac{\ln(n)}{2} + \frac{1}{8} \sum_{k=1}^{n-1} \left(\frac{1}{k} - \frac{1}{k+1} \right).$$

$\sum_{k=1}^{n-1} \left(\frac{1}{k} - \frac{1}{k+1} \right)$ est une somme télescopique :

$$\begin{aligned} \sum_{k=1}^{n-1} \left(\frac{1}{k} - \frac{1}{k+1} \right) &= \left(1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \dots + \frac{1}{n-1} - \frac{1}{n} \right) \\ &= 1 - \frac{1}{n} \end{aligned}$$

On peut aussi retrouver ce résultat en séparant les sommes, et en effectuant un *changement d'indices* :

$$\begin{aligned}\sum_{k=1}^{n-1} \left(\frac{1}{k} - \frac{1}{k+1} \right) &= \sum_{k=1}^n \frac{1}{k} - \sum_{k=1}^{n-1} \frac{1}{k+1} \\ &= \sum_{k=1}^n \frac{1}{k} - \sum_{k=2}^n \frac{1}{k} \\ &= 1 + \sum_{k=2}^{n-1} \frac{1}{k} - \sum_{k=2}^{n-1} \frac{1}{k} - \frac{1}{n} \\ &= 1 - \frac{1}{n}\end{aligned}$$

Ainsi : $n \ln n - n + 1 \leq \sum_{k=1}^{n-1} \ln k + \frac{\ln n}{2} + \frac{1}{8} \left(1 - \frac{1}{n} \right)$, ou encore :

$$n \ln n - n + 1 \leq \sum_{k=1}^n \ln k - \frac{\ln n}{2} + \frac{1}{8} \left(1 - \frac{1}{n} \right)$$

soit :

$$n \ln n - n + 1 \leq \ln(n!) - \frac{\ln(n)}{2} + \frac{1}{8} \left(1 - \frac{1}{n} \right)$$

puis :

$$\left(n + \frac{1}{2} \right) \ln n - n + 1 - \frac{1}{8} \left(1 - \frac{1}{n} \right) \leq \ln(n!)$$

soit :

$$\left(n + \frac{1}{2} \right) \ln n - n + \frac{7}{8} - \frac{1}{8n} \leq \ln(n!)$$

Par croissance de la fonction exponentielle, on en déduit alors :

$$\exp \left(\left(n + \frac{1}{2} \right) \ln n - n + \frac{7}{8} - \frac{1}{8n} \right) \leq n!$$

soit :

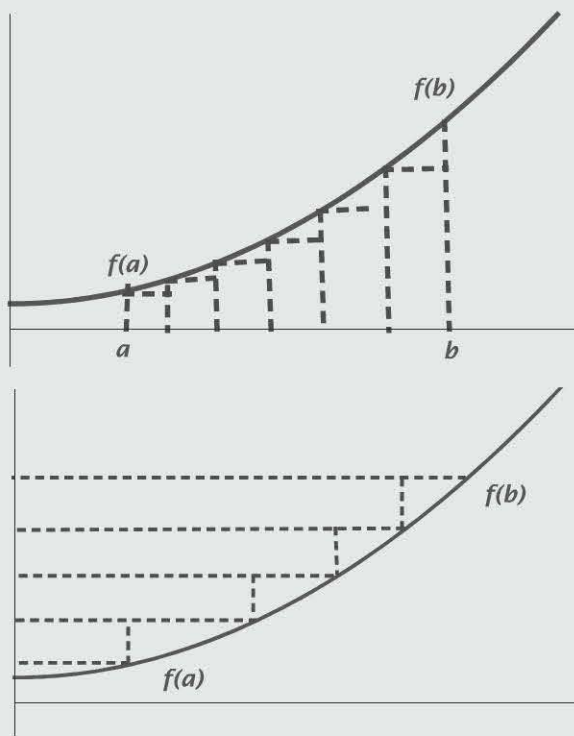
$$n^n \sqrt{n} e^{-n} e^{\frac{7}{8}} e^{-\frac{1}{8n}} \leq n!$$

On a donc obtenu l'encadrement suivant de $n!$:

$$n^n \sqrt{n} e^{-n} e^{\frac{7}{8}} e^{-\frac{1}{8n}} \leq n! \leq e n^n \sqrt{n} e^{-n}$$

Pour une fonction donnée f définie et continue sur un intervalle $[a, b]$ de \mathbb{R} , l'intégrale, dite de Riemann, de a à b de f , correspond à la valeur de l'aire comprise entre la courbe représentative de f sur $[a, b]$ et l'axe des abscisses.

Cette aire est calculée par l'intermédiaire d'un découpage, de plus en plus fin, de l'axe des abscisses, c'est-à-dire un découpage vertical de la forme suivante :



Intégrale de Riemann et intégrale de Lebesgue.

Du découpage vertical au découpage horizontal

Toutefois, un tel découpage n'est pas toujours adapté aux fonctions qui interviennent dans la vie de tous les jours du physicien, du mécanicien, du biologiste : dans ce cas, un découpage « horizontal », à partir donc de l'intervalle des valeurs prises par la fonction, est plus judicieux : c'est l'idée qu'eut le mathématicien Henri-Léon Lebesgue¹.

Bien sûr, toute fonction intégrable au sens de Riemann l'est au sens de Lebesgue. C'est la réciproque qui n'est pas vraie !

L'intégrale de Lebesgue apparaît alors comme un moyen de « mesurer » une fonction.

1. Henri-Léon Lebesgue (1875-1941), mathématicien français, dont la contribution à la théorie de l'intégration fut essentielle.

Étant donnée une partie bornée \mathcal{P} de \mathbb{R} , la fonction caractéristique $\mathbb{I}_{\mathcal{P}}$ de \mathcal{P} est définie par :

$$\mathbb{I}_{\mathcal{P}} : x \mapsto \begin{cases} 1 & \text{si } x \in \mathcal{P} \\ 0 & \text{sinon} \end{cases}$$

Lorsque \mathcal{P} est un intervalle $[a, b]$ de \mathbb{R} , sa mesure de Lebesgue $\mu([a, b])$ est égale à sa longueur $b - a$. Intégrales de Lebesgue et de Riemann coïncident alors ; si f est une fonction intégrable au sens de Riemann :

$$\int_a^b f(t) dt = \int_{[a,b]} f d\mu$$

où μ est la mesure de Lebesgue.

Lorsque \mathcal{P} est un ensemble dénombrable, fini ou infini, il est considéré comme étant « nul presque partout », et donc de mesure de Lebesgue nulle.

Grâce à l'intégrale de Lebesgue, il est donc possible d'intégrer des fonctions présentant des discontinuités comme, par exemple, la *fonction de Dirichlet*, qui est la fonction caractéristique de l'ensemble des rationnels $\mathbb{I}_{\mathbb{Q}}$:

$$\mathbb{I}_{\mathbb{Q}} : x \mapsto \begin{cases} 1 & \text{si } x \text{ rationnel} \\ 0 & \text{sinon} \end{cases}$$

L'ensemble des rationnels étant donc dénombrable, même s'il est infini, il est de mesure de Lebesgue nulle.

Le lecteur intéressé trouvera plus de précisions sur l'intégrale de Lebesgue dans [5], et sur la théorie de la mesure dans la bibliographie.

1. Intégrales impropres

► Intégrale impropre sur un intervalle de longueur infinie

Définition

Étant donné un réel a , et une fonction f , définie sur l'intervalle $[a, +\infty[$, continue par morceaux sur $[a, +\infty[$, la quantité :

$$\int_a^{+\infty} f(t) dt = \lim_{X \rightarrow +\infty} \int_a^X f(t) dt$$

est une intégrale dite « **impropre** ».

En effet, la quantité $\int_a^X f(t) dt$ n'a pas toujours une limite finie lorsque X tend vers $+\infty$.

- Si cette limite existe et est finie, on dit que l'intégrale $\int_a^{+\infty} f(t) dt$ est **convergente**.
- Dans le cas contraire, on dit que l'intégrale $\int_a^{+\infty} f(t) dt$ est **divergente**.

Exemples

1. L'intégrale $\int_0^{+\infty} e^{-t} dt$ est convergente. En effet :

$$\begin{aligned} \int_0^{+\infty} e^{-t} dt &= \lim_{X \rightarrow +\infty} \int_0^X e^{-t} dt \\ &= \lim_{X \rightarrow +\infty} [-e^{-t}]_0^X \\ &= \lim_{X \rightarrow +\infty} [1 - e^{-X}] \\ &= 1 \end{aligned}$$

2. L'intégrale $\int_1^{+\infty} \frac{dt}{t^2}$ est convergente. En effet :

$$\begin{aligned} \int_1^{+\infty} \frac{dt}{t^2} &= \lim_{X \rightarrow +\infty} \int_1^X \frac{dt}{t^2} \\ &= \lim_{X \rightarrow +\infty} \left[-\frac{1}{t} \right]_1^X \\ &= \lim_{X \rightarrow +\infty} \left[1 - \frac{1}{X} \right] \\ &= 1 \end{aligned}$$

3. L'intégrale $\int_1^{+\infty} \frac{dt}{t}$ est divergente. En effet :

$$\begin{aligned}\int_1^{+\infty} \frac{dt}{t} &= \lim_{X \rightarrow +\infty} \int_1^X \frac{dt}{t} \\ &= \lim_{X \rightarrow +\infty} [\ln t]_1^X \\ &= \lim_{X \rightarrow +\infty} \ln X \\ &= +\infty\end{aligned}$$



Ce n'est pas parce qu'une fonction est de limite nulle en l'infini que son intégrale va converger, comme on peut le voir avec l'exemple précédent.

► Intégrale impropre sur un intervalle de longueur finie

Définition

Étant donnés deux réels a et b , et une fonction f , définie sur l'intervalle $[a, b[$, continue par morceaux sur $[a, b[$, la quantité :

$$\int_a^b f(t) dt = \lim_{X \rightarrow b} \int_a^X f(t) dt$$

est une intégrale dite « **impropre** ».

En effet, la quantité $\int_a^X f(t) dt$ n'a pas toujours une limite finie lorsque X tend vers b .

- Si cette limite existe et est finie, on dit que l'intégrale $\int_a^b f(t) dt$ est **convergente**.
- Dans le cas contraire, on dit que l'intégrale $\int_a^b f(t) dt$ est **divergente**.

Exemples

1. L'intégrale $\int_0^1 \ln t dt$ est convergente. En effet :

$$\begin{aligned}\int_0^1 \ln t dt &= \lim_{X \rightarrow 0^+} \int_X^1 \ln t dt \\ &= \lim_{X \rightarrow 0^+} [t \ln t - t]_X^1 \\ &= \lim_{X \rightarrow 0^+} [X \ln X - X] \\ &= 0\end{aligned}$$

2. L'intégrale $\int_0^1 \frac{dt}{t}$ est divergente. En effet :

$$\begin{aligned}\int_0^1 \ln t dt &= \lim_{X \rightarrow 0^+} \int_X^1 \frac{dt}{t} \\ &= \lim_{X \rightarrow 0^+} [\ln t]_X^1 \\ &= \lim_{X \rightarrow 0^+} \ln X \\ &= -\infty\end{aligned}$$

2. Des intégrales généralisées remarquables : les intégrales de Riemann

► Intégrale de Riemann (sur un intervalle de longueur infinie)

Définition

Soit a un réel. On appelle **intégrale de Riemann** toute intégrale de la forme :

$$\int_a^{+\infty} \frac{dt}{t^\alpha} = \lim_{X \rightarrow +\infty} \int_a^X \frac{dt}{t^\alpha}$$

où α est un réel.

► Intégrale de Riemann (sur un intervalle de longueur finie)

Définition

Soit b un réel. On appelle **intégrale de Riemann** toute intégrale de la forme :

$$\int_0^b \frac{dt}{t^\alpha} = \lim_{X \rightarrow 0^+} \int_X^b \frac{dt}{t^\alpha}$$

où α est un réel.

► Condition nécessaire et suffisante de convergence d'une intégrale de Riemann (sur un intervalle de longueur infinie)

Soit a un réel. L'intégrale de Riemann :

$$\int_a^{+\infty} \frac{dt}{t^\alpha} = \lim_{X \rightarrow +\infty} \int_a^X \frac{dt}{t^\alpha}$$

converge si et seulement si $\alpha > 1$.

Démonstration : Il suffit de calculer, pour tout réel $X > a$:

$$\begin{aligned} \int_a^X \frac{dt}{t^\alpha} &= \int_a^X t^{-\alpha} dt \\ &= \left[\frac{t^{1-\alpha}}{1-\alpha} \right]_a^X \\ &= \left[\frac{a^{1-\alpha}}{1-\alpha} - \frac{X^{1-\alpha}}{1-\alpha} \right] \\ &= \frac{1}{1-\alpha} \left[\frac{1}{a^{\alpha-1}} - \frac{1}{X^{\alpha-1}} \right] \end{aligned}$$

Lorsque X tend vers $+\infty$, cette dernière expression n'admet de limite finie que si $\alpha > 1$. ■

Exemples

1. $\int_1^{+\infty} \frac{dt}{t^3}$ est une intégrale de Riemann convergente :

$$\begin{aligned} \lim_{X \rightarrow +\infty} \int_1^X \frac{dt}{t^3} &= \lim_{X \rightarrow +\infty} \int_1^X t^{-3} dt \\ &= \lim_{X \rightarrow +\infty} \left[\frac{t^{1-3}}{1-3} \right]_1^X \\ &= \lim_{X \rightarrow +\infty} \left[-\frac{t^{-2}}{2} \right]_1^X \\ &= \lim_{X \rightarrow +\infty} \left[\frac{1}{2} - \frac{X^{-2}}{2} \right] \\ &= \frac{1}{2} \end{aligned}$$

2. $\int_1^{+\infty} \frac{1}{\sqrt[4]{t}}$ est une intégrale de Riemann divergente :

$$\begin{aligned} \lim_{X \rightarrow +\infty} \int_1^X \frac{dt}{\sqrt[4]{t}} &= \lim_{X \rightarrow +\infty} \int_1^X t^{-\frac{1}{4}} dt \\ &= \lim_{X \rightarrow +\infty} \left[\frac{t^{1-\frac{1}{4}}}{1-\frac{1}{4}} \right]_1^X \\ &= \lim_{X \rightarrow +\infty} \left[\frac{t^{\frac{3}{4}}}{\frac{3}{4}} \right]_1^X \\ &= \lim_{X \rightarrow +\infty} \frac{4}{3} \left[t^{\frac{3}{4}} \right]_1^X \\ &= \lim_{X \rightarrow +\infty} \frac{4}{3} [X^{\frac{3}{4}} - 1] \\ &= +\infty \end{aligned}$$

► Condition nécessaire et suffisante de convergence d'une intégrale de Riemann (sur un intervalle de longueur finie)

Soit b un réel. L'intégrale de Riemann :

$$\int_0^b \frac{dt}{t^\alpha} = \lim_{X \rightarrow 0^+} \int_X^b \frac{dt}{t^\alpha}$$

converge si et seulement si $\alpha < 1$.

Démonstration : Il suffit de calculer, pour tout réel positif $X < b$:

$$\begin{aligned} \int_X^b \frac{dt}{t^\alpha} &= \int_X^b t^{-\alpha} dt \\ &= \left[\frac{t^{1-\alpha}}{1-\alpha} \right]_X^b \\ &= \left[\frac{b^{1-\alpha}}{1-\alpha} - \frac{X^{1-\alpha}}{1-\alpha} \right] \\ &= \frac{1}{1-\alpha} \left[\frac{1}{X^{\alpha-1}} - \frac{1}{b^{\alpha-1}} \right] \end{aligned}$$

Lorsque X tend vers 0^+ , cette dernière expression n'admet de limite finie que si $\alpha < 1$.

■

Exemples

1. $\int_0^1 \frac{1}{\sqrt{t}}$ est une intégrale de Riemann convergente :

$$\begin{aligned}\lim_{X \rightarrow 0^+} \int_X^1 \frac{dt}{\sqrt{t}} &= \lim_{X \rightarrow 0^+} \left[-2\sqrt{t} \right]_X^1 \\ &= \lim_{X \rightarrow 0^+} \left[2 - 2\sqrt{X} \right] \\ &= 2\end{aligned}$$

2. $\int_0^1 \frac{dt}{t^3}$ est une intégrale de Riemann divergente :

$$\begin{aligned}\lim_{X \rightarrow 0^+} \int_X^1 \frac{dt}{t^3} &= \lim_{X \rightarrow 0^+} \int_X^1 t^{-3} dt \\ &= \lim_{X \rightarrow 0^+} \left[\frac{t^{-2}}{-2} \right]_X^1 \\ &= \lim_{X \rightarrow 0^+} \left[-\frac{1}{2} + \frac{1}{2X^2} \right] \\ &= +\infty\end{aligned}$$

3. Propriétés

Les propriétés usuelles des intégrales (linéarité, positivité, croissance, relation de Chasles, changement de variable) sont encore vraies pour des intégrales impropres.

En ce qui concerne l'intégration par parties, le résultat peut être obtenu en se ramenant à un segment. Toutefois, pour alléger les écritures, et, bien évidemment, sous réserve de convergence des quantités en jeu, on pourra utiliser des notations de la forme :

$$\int_a^{+\infty} f(t) g'(t) dt = [f(t) g(t)]_a^{+\infty} - \int_a^{+\infty} f'(t) g(t) dt$$

4. Critères de convergence pour des intégrales positives

► **Comparaison avec une intégrale convergente (sur un intervalle de longueur infinie)**

Soient a un réel, f et g deux fonctions définies sur $[a, +\infty[$, positives, continues par morceaux, et telles que, pour tout réel t de $[a, +\infty[$:

$$f(t) \leq g(t)$$

Alors, si $\int_a^{+\infty} g(t) dt$ converge, il en est de même de $\int_a^{+\infty} f(t) dt$.

Ce résultat n'est vrai que si les fonctions sont à valeurs positives. En effet, si ce n'est pas le cas, rien ne dit que $\int_a^{+\infty} f(t) dt$ ne vaut pas $-\infty$, et $-\infty$ sera toujours plus petit que n'importe quelle quantité négative...



► **Comparaison avec une intégrale divergente (sur un intervalle de longueur infinie)**

Soient a un réel, f et g deux fonctions définies sur $[a, +\infty[$, positives, continues par morceaux, et telles que, pour tout réel t de $[a, +\infty[$:

$$f(t) \leq g(t)$$

Alors, si $\int_a^{+\infty} f(t) dt$ diverge, il en est de même de $\int_a^{+\infty} g(t) dt$.

Corollaire

Soient a un réel, f et g deux fonctions définies sur $[a, +\infty[$, positives, continues par morceaux, et telles que, lorsque t tend vers $+\infty$:

$$f(t) \sim g(t)$$

Alors, si $\int_a^{+\infty} g(t) dt$ converge, il en est de même de $\int_a^{+\infty} f(t) dt$.

De même, si $\int_a^{+\infty} g(t) dt$ diverge, il en est de même de $\int_a^{+\infty} f(t) dt$.



Ce n'est pas parce que, au voisinage de l'infini, vers $f(t) \sim g(t)$, que les intégrales $\int_a^{+\infty} f(t) dt$ et $\int_a^{+\infty} g(t) dt$ ont des valeurs équivalentes !

Exemple

On s'intéresse à la convergence de l'intégrale $\int_1^{+\infty} \frac{dt}{(t+1)^2}$.

Lorsque t tend vers $+\infty$: $\frac{1}{(t+1)^2} \sim \frac{1}{t^2}$

$\int_1^{+\infty} \frac{dt}{(t+1)^2}$ est une intégrale de Riemann convergente. On peut donc en déduire la convergence de $\int_1^{+\infty} \frac{dt}{(t+1)^2}$.

On peut alors calculer :

$$\begin{aligned} \int_1^{+\infty} \frac{dt}{(t+1)^2} &= \lim_{X \rightarrow +\infty} \int_1^X \frac{dt}{(t+1)^2} \\ &= \lim_{X \rightarrow +\infty} \left[-\frac{1}{t+1} \right]_1^X \\ &= \lim_{X \rightarrow +\infty} \left[\frac{1}{2} - \frac{1}{X+1} \right] \\ &= \frac{1}{2} \end{aligned}$$

sachant que :

$$\begin{aligned} \int_1^{+\infty} \frac{dt}{t^2} &= \lim_{X \rightarrow +\infty} \int_1^X \frac{dt}{t^2} \\ &= \lim_{X \rightarrow +\infty} \left[-\frac{1}{t} \right]_1^X \\ &= \lim_{X \rightarrow +\infty} \left[1 - \frac{1}{X+1} \right] \\ &= 1 \\ &\neq \frac{1}{2} \end{aligned}$$

► **Comparaison avec une intégrale convergente (sur un intervalle de longueur finie)**

Soient a et b deux réels, f et g deux fonctions définies sur $[a, b[$, positives, continues par morceaux, et telles que, pour tout réel t de $[a, b[$:

$$f(t) \leq g(t)$$

Alors, si $\int_a^b g(t) dt$ converge, il en est de même de $\int_a^b f(t) dt$.



Comme précédemment, ce résultat n'est vrai que si les fonctions sont à valeurs positives. En effet, si ce n'est pas le cas, rien ne dit que $\int_a^{+\infty} f(t) dt$ ne vaut pas $-\infty$, et $-\infty$ sera toujours plus petit que n'importe quelle quantité négative...

► **Comparaison avec une intégrale divergente (sur un intervalle de longueur finie)**

Soient a et b deux réels, f et g deux fonctions définies sur $[a, b[$, positives, continues par morceaux, et telles que, pour tout réel t de $[a, b[$:

$$f(t) \leq g(t)$$

Alors, si $\int_a^b f(t) dt$ diverge, il en est de même de $\int_a^b g(t) dt$.

Corollaire

Soient a et b deux réels, f et g deux fonctions définies sur $[a, b[$, positives, continues par morceaux, et telles que, lorsque t tend vers b :

$$f(t) \sim g(t)$$

Alors, si $\int_a^b g(t) dt$ converge, il en est de même de $\int_a^b f(t) dt$.

De même, si $\int_a^b g(t) dt$ diverge, il en est de même de $\int_a^b f(t) dt$.

Exemples

1. On s'intéresse à la convergence de l'intégrale $\int_0^1 \frac{e^{-t}}{\sqrt{t}} dt$.

Lorsque t tend vers 0^+ :

$$\frac{e^{-t}}{\sqrt{t}} \sim \frac{1}{\sqrt{t}}$$

$\int_0^1 \frac{dt}{\sqrt{t}}$ converge. Il en est donc de même de $\int_0^1 \frac{e^{-t}}{\sqrt{t}} dt$.

On aurait également pu utiliser le fait que, pour tout réel strictement positif t :

$$\frac{e^{-t}}{\sqrt{t}} \leq \frac{1}{\sqrt{t}}$$

et conclure, grâce à la convergence de $\int_0^1 \frac{dt}{\sqrt{t}}$.

2. On s'intéresse à la convergence de l'intégrale $\int_0^1 \frac{\sin t}{t^2}$.

Lorsque t tend vers 0^+ :

$$\frac{\sin t}{t^2} \sim \frac{t}{t^2}$$

soit :

$$\frac{\sin t}{t^2} \sim \frac{1}{t}$$

$\int_0^1 \frac{dt}{t}$ diverge. Il en est donc de même de $\int_0^1 \frac{\sin t}{t^2}$.

5. Comparaison série-intégrale

Théorème

Soit n_0 un entier naturel, et f une fonction définie sur $[n_0, +\infty[$, positive et décroissante. Alors, la série de terme général $f(n)$ converge si et seulement si f est intégrable sur $[n_0, +\infty[$. On a alors :

$$\int_{n_0+1}^{+\infty} f(t) dt \leq \sum_{k=n_0+1}^{+\infty} f(k) \leq \int_{n_0}^{+\infty} f(t) dt$$

1. Définitions

► Pavé (du plan \mathbb{R}^2)

On appelle **pavé (du plan \mathbb{R}^2)** tout domaine rectangulaire \mathcal{P} de la forme :

$$\mathcal{P} = [a, b] \times [c, d]$$

où a, b, c, d sont des réels tels que $a < b$ et $c < d$.

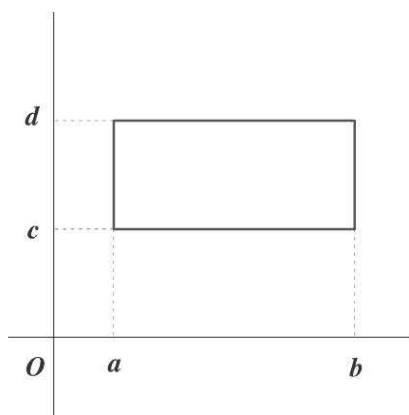


Figure 61.1 – Un pavé du plan \mathbb{R}^2 .

► Intégrale double sur un pavé (du plan \mathbb{R}^2)

Étant donnés quatre réels a, b, c, d tels que $a < b$ et $c < d$, et une fonction f continue sur le pavé $\mathcal{P} = [a, b] \times [c, d]$, à valeurs réelles, on appelle **intégrale double sur de f sur le pavé \mathcal{P}** le nombre :

$$\iint_{\mathcal{P}} f(x, y) dx dy = \int_a^b \left\{ \int_c^d f(x, y) dy \right\} dx = \int_c^d \left\{ \int_a^b f(x, y) dx \right\} dy$$

2. Propriétés

► Linéarité de l'intégrale double sur un pavé (du plan \mathbb{R}^2)

Théorème

Soient a, b, c, d quatre réels tels que $a < b$ et $c < d$, et f et g deux fonctions continues sur le pavé $\mathcal{P} = [a, b] \times [c, d]$, à valeurs réelles. Alors, pour tout réel λ :

$$\iint_{\mathcal{P}} \{f(x, y) + \lambda g(x, y)\} dx dy = \iint_{\mathcal{P}} f(x, y) dx dy + \lambda \iint_{\mathcal{P}} g(x, y) dx dy$$

► **Positivité de l'intégrale double sur un pavé (du plan \mathbb{R}^2)**

Théorème

Soient a, b, c, d quatre réels tels que $a < b$ et $c < d$, et f une fonction continue et positive sur le pavé $\mathcal{P} = [a, b] \times [c, d]$, à valeurs positives. Alors :

$$\iint_{\mathcal{P}} f(x, y) dx dy \geq 0 \quad (61.1)$$

► **Intégrale double sur un pavé (du plan \mathbb{R}^2) et valeur absolue**

Théorème

Soient a, b, c, d quatre réels tels que $a < b$ et $c < d$, et f une fonction continue sur le pavé $\mathcal{P} = [a, b] \times [c, d]$, à valeurs réelles. Alors :

$$\left| \iint_{\mathcal{P}} f(x, y) dx dy \right| \leq \iint_{\mathcal{P}} |f(x, y)| dx dy \quad (61.2)$$

► **Théorème de Fubini**

Théorème

Soient a, b, c, d quatre réels tels que $a < b$ et $c < d$, et f une fonction continue sur le pavé $\mathcal{P} = [a, b] \times [c, d]$, à valeurs réelles. Alors :

$$\iint_{\mathcal{P}} f(x, y) dx dy = \int_a^b \left\{ \int_c^d f(x, y) dy \right\} dx = \int_c^d \left\{ \int_a^b f(x, y) dx \right\} dy \quad (61.3)$$

► **Additivité de l'intégrale double (sur un pavé du plan \mathbb{R}^2) par rapport au domaine d'intégration**

Théorème

Soient $a, \alpha, b, c, \gamma, d$ six réels tels que $a < \alpha < b$ et $c < \gamma < d$, et f une fonction continue sur le pavé $\mathcal{P} = [a, b] \times [c, d]$, à valeurs réelles. Alors :

$$\iint_{\mathcal{P}} f(x, y) dx dy = \iint_{[a, \alpha] \times [c, d]} f(x, y) dx dy + \iint_{[\alpha, b] \times [c, d]} f(x, y) dx dy \quad (61.4)$$

et, de même :

$$\iint_{\mathcal{P}} f(x, y) dx dy = \iint_{[a, b] \times [c, \gamma]} f(x, y) dx dy + \iint_{[a, b] \times [\gamma, d]} f(x, y) dx dy \quad (61.5)$$

Exercices d'entraînement



Les corrigés sont disponibles en téléchargement sur le site dunod.com à partir de la page de présentation de l'ouvrage.

Continuité

1 On considère la fonction f définie par :

$$f(x) = \begin{cases} e^{-\frac{1}{x}} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

Étudier la continuité de f .

2 On considère la fonction g définie par :

$$g(x) = \begin{cases} \frac{1}{\ln|x|} & \text{si } x \notin \{0, -1, 1\} \\ 0 & \text{si } x = 1 \text{ ou } x = -1 \text{ ou } x = 0 \end{cases}$$

Étudier la continuité de g .

3 On considère la fonction h définie par :

$$h(x) = \begin{cases} \frac{1}{x^2 - 1} & \text{si } x \in \mathbb{R} \setminus \{-1, 1\} \\ 0 & \text{si } x = 1 \text{ ou } x = -1 \end{cases}$$

Étudier la continuité de h .

4 On considère la fonction φ définie par :

$$\varphi(x) = \begin{cases} x \sin\left(\frac{1}{x}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Étudier la continuité de φ .

5 On considère la fonction ψ définie par :

$$\psi(x) = E(x) + |x - E(x)|^4$$

Étudier la continuité de ψ .

6 Soient $\lambda \in]0, 1[$, et f une fonction continue sur l'intervalle $[\lambda, +\infty[$, telle que $f(\lambda) = 0$ et $\lim_{x \rightarrow +\infty} f(x) = 0$.

Soit F la fonction définie sur $]0, 1] \rightarrow \mathbb{R}$ par :

$$F(x) = f\left(\frac{1}{x} + \lambda - 1\right)$$

6.a) Montrer que F est bien définie sur $]0, 1]$.

6.b) Étudier la continuité de F sur $]0, 1]$.

6.c) Montrer que F est prolongeable par continuité sur $[0, 1]$.

7 Une équation fonctionnelle

Déterminer les fonctions f , définies sur \mathbb{R} , à valeurs dans \mathbb{R} , continues en 0 et en 1 telles que, pour tout réel x :

$$f(x^2) = f(x)$$

Dérivabilité

8 Donner la dérivée de la fonction $\phi : x \in \mathbb{R} \mapsto \ln(2 + \cos x)$.

9 On considère la fonction h_1 définie, pour tout réel x , par : $h_1(x) = e^{-x}$.

Montrer que h_1 est de classe C^∞ sur \mathbb{R} , et déterminer, pour tout entier naturel non nul n , l'expression de sa dérivée $n^{\text{ième}}$.

10 On considère la fonction h_2 définie, pour tout x de $\mathbb{R} \setminus \{1\}$, par :

$$h_2(x) = \frac{1}{1 - x}$$

Montrer que h_2 est dérivable sur $\mathbb{R} \setminus \{1\}$, et déterminer, pour tout entier naturel non nul n , l'expression de sa dérivée $n^{\text{ième}}$.

11 On considère la fonction f définie, pour tout réel x , par :

$$f(x) = \begin{cases} 1 - e^{-\frac{1}{x^2}} & \text{si } x \neq 0 \\ 1 & \text{si } x = 0 \end{cases}$$

Étudier la continuité et la dérivabilité de f .

12 On considère la fonction φ définie, pour tout réel x , par :

$$\varphi(x) = \begin{cases} x^3 \sin\left(\frac{1}{x}\right) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

Étudier la dérivabilité de φ .

13 On considère la fonction ψ définie, pour tout réel x , par :

$$\psi(x) = \begin{cases} 1 - e^{-\frac{1}{(x-1)^2}} & \text{si } x \neq 1 \\ 1 & \text{si } x = 1 \end{cases}$$

Étudier la continuité et la dérivabilité de ψ .

14 Des dérivées remarquables

On considère les fonctions ψ_1 et ψ_2 définies par :

$$\psi_1(x) = \ln\left(\tan\left(\frac{x}{2}\right)\right),$$

$$\psi_2(x) = \ln\left(\tan\left(\frac{x}{2} + \frac{\pi}{4}\right)\right)$$

Quel est le domaine de définition de ces deux fonctions ? Étudier la dérivabilité de ψ_1 et ψ_2 , puis donner l'expression de leurs dérivées respectives.

15 Des inégalités classiques

15.a) Montrer que, pour tout réel x :
 $e^x \geq 1 + x$.

15.b) Montrer que, pour tout réel positif x :
 $\sin x \leq x$.

15.c) Montrer que, pour tout réel $x > -1$:
 $\ln(1+x) \leq x$.

16 Des limites remarquables

16.a) Soit f_1 une fonction dérivable en zéro. Déterminer :

$$\lim_{x \rightarrow 0} \frac{f_1(2x) - f_1(0)}{2x}$$

$$\text{et } \lim_{x \rightarrow 0} \frac{f_1(2x) - f_1(x)}{x}$$

16.b) Soit f_2 une fonction définie sur \mathbb{R} , dérivable sur \mathbb{R} . Déterminer, pour tout réel a :

$$\lim_{x \rightarrow a} \frac{x f_2(a) - a f_2(x)}{x - a}$$

17 Règle de l'Hôpital¹

Soient f et g deux fonctions définies sur un intervalle $[a, b]$ de \mathbb{R} , à valeurs dans \mathbb{R} , continues sur $[a, b]$, dérivables sur $]a, b[$. On suppose que g' ne s'annule pas sur $]a, b[$. Montrer qu'il existe un réel c dans $]a, b[$ tel que :

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}$$

18 Une étude de fonction

Soit n un entier naturel supérieur ou égal à 2. On considère la fonction

$$f_n : \mathbb{R}_+ \rightarrow \mathbb{R} \\ x \mapsto \begin{cases} \left(1 - \frac{x}{n}\right)^n - e^{-x} & \text{si } x \in [0, n] \\ 0 & \text{sinon} \end{cases}$$

18.a) Montrer que f_n est dérivable sur $[0, n[$.

18.b) Calculer, pour tout x de $[0, n[$, $f'_n(x)$, et montrer qu'on peut l'écrire sous la forme

$$f'_n(x) = e^{-x} [1 - e^{h_n(x)}]$$

où h_n est une fonction définie sur $[0, n[$, dont on donnera l'expression.

18.c) Étudier les variations de h_n sur $[0, n[$, ainsi que son signe. On montrera qu'il existe un réel $\alpha_n \in]1, n[$ tel que :

$$h_n(\alpha_n) = 0$$

18.d) Étudier les variations de f_n sur $[0, n[$, et montrer que, pour tout réel x de $[0, n[$:

$$|f_n(x)| \leq \frac{\alpha_n e^{-\alpha_n}}{n}$$

18.e) Étudier les variations de la fonction ϕ qui, à tout réel x de \mathbb{R}_+ , associe $x e^{-x}$, et en déduire que, pour tout réel x de $[0, n[$:

$$|f_n(x)| \leq \frac{1}{n e}$$

Que peut-on en conclure ?

1. Guillaume François Antoine, marquis de L'Hôpital (1661-1704), mathématicien français, élève de Jean Bernoulli, spécialiste de calcul différentiel. Il est connu pour son ouvrage *Analyse des Infiniment Petits pour l'Intelligence des Lignes Courbes*, publié en 1696. Il travailla également sur les coniques.

Fonctions réciproques

19 Donner le domaine de définition de la fonction $\phi : x \mapsto \frac{1}{\arcsin(4x)}$.

20 Donner la dérivée de la fonction $\psi : x \in \mathbb{R}^* \mapsto \arctan x + \arctan\left(\frac{1}{x}\right)$, et en déduire que, pour tout réel strictement positif x :

$$\arctan x + \arctan\left(\frac{1}{x}\right) = \frac{\pi}{2}$$

21 On considère la fonction f définie, pour tout x de l'intervalle $\left[0, \frac{\pi}{2}\right]$, par :

$$f(0) = 1 \quad ; \quad f(x) = \frac{x}{\tan x} \\ \text{si } 0 < x < \frac{\pi}{2} \quad ; \quad f\left(\frac{\pi}{2}\right) = 0.$$

21.a) Montrer que, pour tout réel $x > 0$: $\sin x < x$.

21.b) Montrer que f est strictement décroissante sur $\left]0, \frac{\pi}{2}\right[$.

21.c) Déterminer l'image $J = f(I)$ de l'intervalle I par f , et montrer que f admet une application réciproque g définie sur J .

21.d) Montrer que $g\left(\frac{\pi}{4}\right) = \frac{\pi}{4}$.

21.e) Montrer que la fonction g est dérivable en $\frac{\pi}{4}$, et calculer $g'\left(\frac{\pi}{4}\right)$.

22 On considère la fonction f définie, pour tout réel x , par : $f(x) = \frac{e^{\sin x} - e^{-\sin x}}{2}$.

22.a) Montrer que f est dérivable sur \mathbb{R} , et calculer, pour tout réel x , $f'(x)$.

22.b) Déterminer l'image $J = f\left(\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\right)$ de $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ par f , et montrer que f admet une application réciproque f^{-1} définie sur J .

22.c) Montrer que la fonction f^{-1} est dérivable en 0, et calculer la valeur de sa dérivée en ce point.

Développements limités

23 Déterminer le développement limité à l'ordre 5, au voisinage de 0, de la fonction :

$$x \mapsto \varphi_1(x) = e^{\sin x}$$

24 Déterminer le développement limité à l'ordre 6, au voisinage de 0, de la fonction :

$$x \mapsto \varphi_2(x) = \ln(\cos x)$$

25 Déterminer le développement limité à l'ordre 6, au voisinage de 0, de la fonction :

$$x \mapsto \varphi_3(x) = (\cos x)^{\sin x}$$

26 Déterminer le développement limité à l'ordre 3, au voisinage de $\frac{\pi}{2}$, de la fonction :

$$x \mapsto \varphi_4(x) = \cos x - x \sin x$$

27 Déterminer le développement limité à l'ordre 4 au voisinage de 0 de la fonction :

$$x \in \mathbb{R} \mapsto \varphi_5(x) = e^{-x}$$

28 Déterminer le développement limité à l'ordre 3 au voisinage de 0 de la fonction :

$$x \in \mathbb{R} \setminus \left\{\frac{1}{2}\right\} \mapsto \varphi_6(x) = \frac{1}{1-2x}$$

29 Pour $p \in \mathbb{N}^*$, donner le développement limité de la fonction $x \mapsto \frac{1}{1+x^2}$ à l'ordre $2p$ en 0, et en déduire celui de la fonction $x \mapsto \arctan x$ à l'ordre $2p+1$ en 0.

30 Déterminer le développement limité à l'ordre 4, au voisinage de 0, de chacune des fonctions suivantes :

30.a) $x \mapsto f_1(x) = \frac{\cos x}{1-x} - e^x.$

30.b) $x \mapsto f_2(x) = \tan(\ln(1+x^2)).$

31 Déterminer le développement limité à l'ordre 4, au voisinage de 0, de chacune des fonctions suivantes :

31.a) $x \mapsto g_1(x) = \frac{1}{\sqrt{1-x^2}} + \arcsin(x).$

31.b) $x \mapsto g_2(x) = \ln(\cos^2 x - \sin^2 x).$

32 Déterminer, lorsqu'elles existent, les limites suivantes :

32.a) $\lim_{x \rightarrow 0} \frac{\sqrt{1+\sin x} + e^{-\frac{x}{2}} - 2}{x^3}.$

32.b) $\lim_{x \rightarrow 0} \frac{e^{\sqrt{1+\sin x}} - e}{\tan x}.$

32.c) $\lim_{x \rightarrow 1} \frac{x^x - x}{1 - x + \ln x}.$

32.d) $\lim_{x \rightarrow +\infty} \left(1 + \frac{\alpha}{x}\right)^x$, où α est un réel.

Équations différentielles linéaires du premier ordre

33 Résoudre, sur \mathbb{R} , l'équation différentielle (\mathcal{E}_1) :

$$y' + y = 2 \operatorname{sh} x$$

34 Résoudre, sur \mathbb{R} , l'équation différentielle (\mathcal{E}_2) :

$$y' + 2xy = 2x$$

35 Résoudre, sur \mathbb{R}_+^* , l'équation différentielle (\mathcal{E}_3) :

$$xy' - 3x^3y = 3x^3e^{2x^3}$$

36 Résoudre, sur l'intervalle $I =]0, +\infty[$, l'équation différentielle (\mathcal{E}_4)

$$y' = \frac{2y}{x} + x$$

37 Soit λ un réel. Résoudre, sur \mathbb{R} , l'équation différentielle (\mathcal{E}_5) :

$$(x^2 + 1)y' + y = \lambda$$

38 Déterminer les solutions polynomiales de l'équation différentielle (\mathcal{E}_6) :

$$x^2 y' - xy = x^3 + x$$

Suites

Étude de suites

39 Étudier la suite $(u_n)_{n \in \mathbb{N}}$ définie par :

$$u_0 = 1, \quad u_1 = 2 \quad \text{et}$$

$$\forall n \in \mathbb{N}^* : u_{n+1} = 2u_n + 3u_{n-1}$$

40 Étudier la suite $(v_n)_{n \in \mathbb{N}}$ définie par :

$$v_0 = 1, \quad v_1 = -1 \quad \text{et}$$

$$\forall n \in \mathbb{N} : v_{n+2} + v_{n+1} + v_n = 0$$

Démonstrations par récurrence

41 Somme des n premiers entiers naturels, $n \in \mathbb{N}$

Pour tout entier naturel non nul n , on pose :

$$S_n = \sum_{k=1}^n k.$$

41.a) En remarquant que :

$$S_n = 1 + 2 + 3 + \dots + (n-2) + (n-1) + n$$

$$S_n = n + (n-1) + (n-2) + \dots + 3 + 2 + 1$$

en déduire que : $2S_n = n(n+1)$, puis : $S_n = \frac{n(n+1)}{2}$.

41.b) Redémontrer ce résultat par récurrence.

42 Somme des n premiers carrés, $n \in \mathbb{N}$

42.a) Montrer que, pour tout entier naturel

$$\text{non nul } n : \sum_{k=1}^{n+1} k^3 = \sum_{k=1}^n (k+1)^3 + 1.$$

42.b) En déduire que, pour tout entier naturel n :

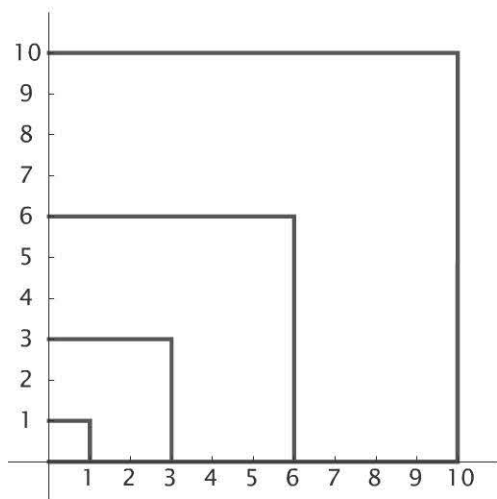
$$\sum_{k=0}^n k^2 = \frac{n(n+1)(2n+1)}{6}.$$

42.c) Redémontrer ce résultat par récurrence.

43 Somme des n premiers cubes, $n \in \mathbb{N}$

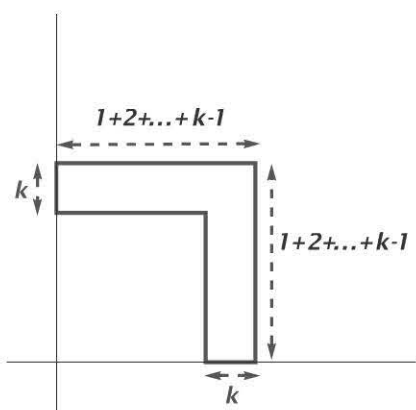
Considérons la suite suivante de carrés, emboîtés les uns dans les autres :

2. On rencontre, notamment, ce type de fonctions dans des équations aux dérivées partielles présentant une invariance d'échelle.



Le premier carré a un côté de longueur 1, le second, de longueur $1 + 2$, le troisième, $1 + 2 + 3$, etc. Le dernier (qui est le $n^{\text{ième}}$) – et plus grand des carrés – a donc pour côté $1 + 2 + \dots + n = \sum_{k=1}^n k = \frac{n(n+1)}{2}$. Son aire

\mathcal{A} est donc égale à $\mathcal{A} = \frac{n^2(n+1)^2}{4}$. Il est également possible de calculer cette aire en additionnant celle du carré de longueur 1 et des surfaces en forme d'« équerre » :



La $k^{\text{ième}}$ équerre, composée d'un carré de côté k et de deux rectangles de côtés respectifs k et $1 + 2 + \dots + k - 1 = \sum_{j=1}^{k-1} j$, a pour aire :

$$\begin{aligned} & k^2 + 2k(1 + 2 + \dots + k - 1) \\ &= k^2 + 2k \frac{(k-1)k}{2} \\ &= k^2 + k(k-1)k = k^3 \end{aligned}$$

43.a) Démontrer par récurrence que, pour tout entier naturel non nul n :

$$\begin{aligned} \sum_{k=1}^n k^3 &= 1^3 + 2^3 + \dots + n^3 \\ &= (1 + 2 + \dots + n)^2 = \left(\sum_{k=1}^n k \right)^2 \end{aligned}$$

43.b) Pour tout entier naturel non nul n , on pose :

$$\Sigma_n = 1 + 3 + 5 + \dots + (2n-1) = \sum_{k=0}^{n-1} (2k+1)$$

- En remarquant que, pour tout entier naturel n :

$$\begin{aligned} \Sigma_n &= 1 + 2 + 3 + 4 + 5 + \dots \\ &\quad + (2n-1) + 2n - (2 + 4 \end{aligned}$$

$$+ \dots + 2n) = \sum_{k=1}^{2n} k - 2 \sum_{k=1}^n k$$

en déduire : $\Sigma_n = n^2$.

- Redémontrer ce résultat par récurrence.

Suites : limites, convergence, applications

44 Des limites remarquables

44.a) Soit r un réel positif, strictement plus petit que 1. Déterminer

$$\lim_{n \rightarrow +\infty} \sum_{k=0}^n r^k$$

44.b) Soit ρ un réel positif, strictement plus petit que 1. Déterminer

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n \rho^k$$

44.c) Que vaut

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{1}{2^k} \quad ?$$

45 Le nombre d'or

Étudier la suite $(\varphi_n)_{n \in \mathbb{N}}$ définie par :

$$u_0 \in \mathbb{R}$$

et, pour tout entier naturel n :

$$\varphi_{n+1} = \sqrt{1 + \varphi_n}$$

46 La suite de Fibonacci

Le mathématicien italien Leonardo Pisano, dit Fibonacci, s'intéressa au problème suivant :

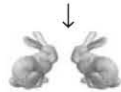
« Un homme met un couple de lapins dans un lieu isolé de tous les côtés par un mur. Combien de couples obtient-on en un an si chaque couple engendre tous les mois un nouveau couple à compter du troisième mois de son existence ? »

sous les hypothèses suivantes :

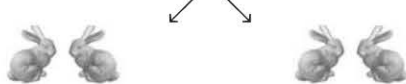
- au début du premier mois, il y a juste un couple de bébés lapins ;
- au bout d'un mois, ceux-ci sont devenus adultes, mais ils ne commenceront à avoir eux-mêmes des bébés lapins qu'à partir du début du troisième mois ;
- à chaque début de mois, tout couple de lapins adultes susceptibles de procréer engendre effectivement un nouveau couple de bébés lapins ;
- les lapins ne meurent jamais.



Génération 1 ; au total, un couple de lapins.



Génération 2 ; au total, 1+1=2 couples de lapins.



Génération 3 ; au total, 2+2=4 couples de lapins.

Pour résoudre celui-ci, on considère la suite $(\mathcal{F}_n)_{n \in \mathbb{N}}$, où, pour tout entier naturel n , \mathcal{F}_n représente le nombre de couples de lapins au début du $n^{\text{ième}}$ mois.

Il est naturel de poser $\mathcal{F}_0 = 0$, puisque, avant que l'on ne dépose les lapins au pied du mur, il n'y en avait aucun.

Comme il n'y a pas nouvelle naissance avant le début du troisième mois, on a ainsi $\mathcal{F}_1 = \mathcal{F}_2 = 1$. La naissance d'un nouveau couple de lapins au début de ce même troisième mois se traduit par : $\mathcal{F}_3 = 2$.

Considérons désormais un mois n quelconque, avec $n \geq 2$. Au mois $n+2$, le nombre total de couples de lapins est obtenu en ajoutant le nombre de ceux présents au mois n à celui qu'il y avait au mois $n+1$:

$$\mathcal{F}_{n+2} = \mathcal{F}_{n+1} + \mathcal{F}_n$$

La suite $(\mathcal{F}_n)_{n \in \mathbb{N}}$ vérifie donc une relation de récurrence linéaire d'ordre 2.

46.a) Exprimer, pour tout entier naturel n , \mathcal{F}_n en fonction de n .

46.b) Quelle est la limite de la suite $(\mathcal{F}_n)_{n \in \mathbb{N}}$?

47 Une somme télescopique

On considère la suite $(S_n)_{n \in \mathbb{N}^*}$ définie pour tout n de \mathbb{N}^* par : $S_n = \sum_{k=1}^n \frac{1}{k(k+1)}$.

47.a) Vérifier que, pour tout entier naturel non nul k :

$$\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$$

47.b) Quelle est la limite de la suite $(S_n)_{n \in \mathbb{N}^*}$?

48 La constante d'Euler

On considère la suite $(H_n)_{n \in \mathbb{N}^*}$ définie par :

$$\forall n \in \mathbb{N}^* : H_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} - \ln n$$

$$= \sum_{k=1}^n \frac{1}{k} - \ln n$$

48.a) Montrer que les suites $(H_n)_{n \in \mathbb{N}^*}$ et $\left(H_n - \frac{1}{n}\right)_{n \in \mathbb{N}^*}$ sont adjacentes, et en déduire qu'elles convergent vers une même limite, que l'on notera γ (On pourra utiliser l'étude des variations des fonctions $f : \mathbb{R}_+^* \rightarrow \mathbb{R}$,

$$t \mapsto \frac{1}{t+1} - \ln\left(1 + \frac{1}{t}\right) \text{ et } g : \mathbb{R}_+^* \rightarrow \mathbb{R},$$

$$t \mapsto \frac{1}{t} - \ln\left(1 + \frac{1}{t}\right).)$$

48.b) Montrer que : $0 < \gamma < 1$.

49 Fractions continues et suites homographiques

Soit a un réel différent de -1 . Quel sens peut-on donner à :

$$\frac{a}{1 + \frac{a}{1 + \frac{a}{1 + \dots}}}$$

(C'est ce que l'on appelle une « fraction continue »)

On pourra considérer la suite $(a_n)_{n \in \mathbb{N}}$ définie, pour tout entier naturel n , par :

$$a_0 = a, \quad a_{n+1} = \frac{a}{1 + a_n}$$

50 Le théorème de Cesàro

Soit $(u_n)_{n \in \mathbb{N}^*}$ une suite réelle, de limite finie $\ell \in \mathbb{R}$.

On considère la suite $(v_n)_{n \in \mathbb{N}^*}$ telle que :

$$\forall n \in \mathbb{N}^* : v_n = \frac{u_1 + u_2 + \dots + u_n}{n}$$

$$= \frac{1}{n} \sum_{k=1}^n u_k$$

50.a) Montrer que, pour tout $\varepsilon > 0$, il existe un rang n_0 tel que, pour tout entier naturel $n \geq n_0$:

$$|u_n - \ell| \leq \frac{\varepsilon}{2}$$

50.b) Vérifier que, pour tout entier naturel $n \geq n_0$:

$$|v_n - \ell| = \frac{1}{n} \left| \sum_{k=1}^n (u_k - \ell) \right|$$

$$= \frac{1}{n} \left| \sum_{k=1}^{n_0-1} (u_k - \ell) + \sum_{k=n_0}^n (u_k - \ell) \right|$$

50.c) Montrer qu'il existe un rang n_1 tel que, pour tout entier naturel $n \geq n_1$:

$$\frac{1}{n} \left| \sum_{k=1}^{n_0-1} (u_k - \ell) \right| \leq \frac{\varepsilon}{2}$$

50.d) En déduire qu'il existe un rang n_2 tel que, pour tout entier naturel $n \geq n_2$:

$$|v_n - \ell| \leq \varepsilon$$

Que peut-on en conclure ?

51 Étude de suites doubles

On considère les suites $(a_n)_{n \in \mathbb{N}}$ et $(b_n)_{n \in \mathbb{N}}$ définies par $0 < a_0 < b_0$, et, pour tout n de \mathbb{N} ,

$$\text{par : } a_{n+1} = \frac{a_n^2}{a_n + b_n}, \quad b_{n+1} = \frac{b_n^2}{a_n + b_n}.$$

51.a) Calculer, pour tout entier naturel n , la différence $a_{n+1} - b_{n+1}$, puis l'exprimer en fonction de a_0 et b_0 . On notera, dans ce qui suit : $d = a_0 - b_0$.

51.b) Calculer, pour tout entier naturel n , le rapport $\frac{a_{n+1}}{b_{n+1}}$.

En posant $r = \frac{a_0}{b_0}$, montrer que l'on a alors :

$$\frac{a_{n+1}}{b_{n+1}} = r^{2^{n+1}}$$

51.c) À l'aide des questions précédentes, exprimer, pour tout entier naturel n , a_n et b_n en fonction de n .

51.d) On considère la suite de terme général

$$P_n = \prod_{k=0}^n (1 + r^{2^k}).$$

Déterminer la limite de la suite $(P_n)_{n \in \mathbb{N}}$ lorsque n tend vers $+\infty$.

Développement asymptotique

52 Pour tout entier naturel non nul n , on

$$\text{pose : } u_n = \left(1 + \frac{1}{n}\right)^{2n}.$$

Déterminer, à l'aide d'un développement asymptotique, la limite de la suite $(u_n)_{n \in \mathbb{N}^*}$.

53 Pour tout entier naturel n , on pose :

$$v_n = \sin(\pi \sqrt{n^2 + 1}).$$

La suite $(v_n)_{n \in \mathbb{N}}$ est-elle convergente ?

54 Développement asymptotique et solution approchée d'une équation

Soit n un entier naturel non nul. Déterminer les trois premiers termes du développement généralisé de la $n^{\text{ième}}$ racine strictement positive de l'équation

$$\tan x = x$$

Intégration**55 Quelques calculs simples**

55.a) Calculer : $\int_{-4}^4 \sin t \cos t \, dt.$

55.b) Calculer : $\int_{-2\pi}^{2\pi} \cos^2 t \, dt.$

55.c) Calculer : $\int_0^1 t e^t \, dt.$

55.d) Calculer : $\int_1^2 \frac{\ln t}{t} \, dt.$

56 Calculs de primitives

Déterminer une primitive des fonctions suivantes :

56.a) $f_1 : \mathbb{R}_+^* \rightarrow \mathbb{R}, t \mapsto \ln t.$

56.b) $f_2 : \mathbb{R}_+^* \rightarrow \mathbb{R}, t \mapsto \frac{\ln t}{t}.$

56.c) $f_3 : \mathbb{R} \rightarrow \mathbb{R}, t \mapsto t \cos t.$

56.d) $f_4 : \mathbb{R} \rightarrow \mathbb{R}, t \mapsto t^2 e^t.$

56.e) $f_5 : \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi, k \in \mathbb{Z} \right\} \rightarrow \mathbb{R},$
 $t \mapsto \tan t.$

56.f) $f_6 : \mathbb{R} \setminus \{-1\} \rightarrow \mathbb{R}, t \mapsto \frac{t}{1+t^3}.$

56.g) $f_7 : \mathbb{R} \setminus \{-2, 0, 2\} \rightarrow \mathbb{R}, t \mapsto \frac{t+1}{t^3-4t}.$

57 Changement de variables

Calculer, à l'aide d'un changement de variable, les intégrales suivantes :

57.a) $I_1 = \int_0^{\frac{1}{\sqrt{2}}} \frac{1}{\sqrt{1-x^2}} dx.$

57.b) $I_2 = \int_0^1 \frac{e^t}{e^t+1} dt.$

57.c) $I_3 = \int_0^{\frac{\pi}{2}} (\sin^2 x - \cos^2 x) dx.$

58

58.a) Rappeler les formules d'addition pour le *cosinus*.

58.b) Soient m et n deux entiers naturels. Calculer :

$$\int_0^{2\pi} \cos(mt) \cos(nt) dt$$

(on pensera à distinguer le cas $m = n = 0$, et, lorsque m et n sont strictement positifs : $m = n$ et $m \neq n$)

59

59.a) Déterminer deux réels a et b tels que, pour tout réel t tel que $t \neq \frac{1}{2}$ et $t \neq -\frac{1}{2}$:

$$\frac{1}{4t^2-1} = \frac{a}{2t-1} + \frac{b}{2t+1}$$

59.b) Calculer :

$$\int_1^2 \frac{dt}{4t^2-1}$$

60 Intégrales de Wallis³

Pour tout entier naturel n , on pose :

$$I_n = \int_0^{\frac{\pi}{2}} \sin^n x dx, \quad J_n = \int_0^{\frac{\pi}{2}} \cos^n x dx$$

60.a) Calculer I_0, I_1, I_2 , et J_0, J_1, J_2 .

60.b) Montrer que, pour tout entier $n \geq 2$:

$$I_n = \frac{n-1}{n} I_{n-2}$$

60.c) Pour tout entier naturel p , donner l'expression de I_{2p} et I_{2p+1} en fonction de p .

60.d) Montrer que, pour tout entier strictement positif p , et tout x de $[0, \frac{\pi}{2}]$:

$$0 \leq \sin^{2p+1} x \leq \sin^{2p} x \leq \sin^{2p-1} x$$

et en déduire une inégalité portant sur I_{2p+1}, I_{2p} et I_{2p-1} .

60.e) En déduire que, pour tout entier strictement positif p :

$$1 \leq \frac{I_{2p}}{I_{2p+1}} \leq \frac{2p+1}{2p}$$

60.f) Que vaut $\lim_{p \rightarrow +\infty} \frac{I_{2p}}{I_{2p+1}}$?

60.g) En déduire que :

$$\pi = \lim_{p \rightarrow +\infty} \frac{1}{p} \frac{2^{4p} (p!)^4}{((2p)!)^2}$$

61

Pour tout entier naturel n , on pose :

$$I_n = \int_1^e (\ln x)^n dx$$

61.a) Calculer I_0 et I_1 .

61.b) Exprimer, pour tout entier naturel n , I_{n+1} en fonction de I_n .

61.c) Montrer que, pour tout entier naturel n :

$$0 \leq I_n \leq \frac{e}{n+1}$$

61.d) Déterminer la limite, lorsque l'entier n tend vers $+\infty$, de I_n .

62

Intégrales et convexité

Soit f une fonction de classe C^1 sur \mathbb{R}_+ , à valeurs positives, convexe.

On rappelle que le graphe d'une fonction convexe est situé au-dessous de ses cordes, mais au-dessus de ses tangentes.

62.a) Montrer que, pour tout entier $n \geq 2$:

$$\frac{f(1)}{2} + f(2) + \dots + f(n-1) + \frac{f(n)}{2} - \int_1^n f(t) dt \geq 0$$

3. John Wallis (1616-1703), mathématicien anglais, spécialiste de calcul différentiel et intégral. C'est lui qui introduisit la notation « ∞ ». À côté de son œuvre mathématique, il s'intéressa aussi à la phonétique, et est considéré comme un des précurseurs de l'orthophonie.

62.b) Montrer que, pour tout entier k de $\{1, \dots, n-1\}$:

$$\int_k^{k+\frac{1}{2}} f(t) dt \geq \frac{f(k)}{2} + \frac{1}{8} f'(k)$$

et :

$$\int_{k+\frac{1}{2}}^{k+1} f(t) dt \geq \frac{f(k+1)}{2} - \frac{1}{8} f'(k+1)$$

62.c) En déduire :

$$\frac{f(1)}{2} + f(2) + \dots + f(n-1) + \frac{f(n)}{2} - \int_1^n f(t) dt \leq \frac{f'(n) - f'(1)}{8}$$

63 Sommes de Riemann

63.a) Déterminer : $\lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{n}{k^2 + 3n^2}$.

63.b) Déterminer : $\lim_{n \rightarrow +\infty} \prod_{k=1}^n \left(1 + \frac{k}{n}\right)^{\frac{1}{n}}$.

63.c) Déterminer :

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n \frac{\sin\left(\frac{k\pi}{3n}\right) \cos\left(\frac{k\pi}{3n}\right)}{n}.$$

64 Limite d'une intégrale

64.a) Déterminer : $\int_0^1 t e^{-t^2} dt$.

64.b) Déterminer : $\lim_{X \rightarrow +\infty} \int_1^X t e^{-t^2} dt$.

64.c) Que peut-on en déduire ?

65 Une comparaison suite-intégrale

Soit f une fonction définie sur \mathbb{R}^+ , continue sur \mathbb{R}^+ , décroissante sur \mathbb{R}^+ .

On considère la suite $(u_n)_{n \in \mathbb{N}}$ définie par :

$$\forall n \in \mathbb{N} : u_n = f(n)$$

65.a) Montrer que, pour tout entier naturel non nul k :

$$\int_k^{k+1} f(t) dt \leq f(k) \leq \int_{k-1}^k f(t) dt$$

65.b) En déduire que, pour tout entier naturel non nul n :

$$\begin{aligned} \int_0^{n+1} f(t) dt &\leq \sum_{k=0}^n f(k) \\ &\leq \int_0^n f(t) dt + f(0) \end{aligned}$$

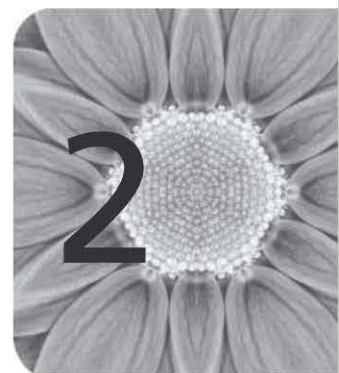
puis :

$$\int_0^{n+1} f(t) dt \leq \sum_{k=0}^n u_k \leq \int_0^n f(t) dt + f(0)$$

Algèbre

Partie

2




Introduction

La présentation est similaire à celle adoptée pour « Calcul Vectoriel » : après les nombres complexes, on présente les matrices ; les vecteurs arrivent ensuite très naturellement, le fait de disposer de la notion de déterminant simplifie les calculs pour la détermination d'équations de droites ou de plans dans l'espace à trois dimensions. Il est ensuite logique de passer à l'étude des transformations linéaires du plan et de l'espace, qui préfigurent celles en dimension quelconque $n \geq 2$. On passe ainsi facilement à la dimension $n \geq 2$, et aux espaces vectoriels.

Plan

Le plan complexe – Les nombres complexes	206
<i>Focus : Les nombres complexes</i>	209
<i>Focus : Transformations complexes, fractales,</i> <i>et représentations de la nature</i>	248
Matrices	252
<i>Focus : L'origine des matrices</i>	250
<i>Focus : Les matrices et leurs applications</i>	276
<i>Focus : Produit scalaire, espaces fonctionnels et calcul numérique</i>	297
<i>Focus : Géométrie euclidienne – ou non ? Encore des matrices !</i>	302
Transformations linéaires du plan	304
Transformations linéaires de l'espace	317
L'espace \mathbb{R}^n	330
<i>Focus : Groupe spécial orthogonal et cristallographie</i>	347
<i>Focus : Diagonalisation – La toupie de Lagrange (et de Michèle Audin)</i> ...	349
Espaces vectoriels	350
Exercices d'entraînement	359

Les bonus web sur Dunod.com

 Les corrigés des exercices sont consultables sur dunod.com sur la page de présentation de l'ouvrage.

Le corps des nombres complexes

Identifions \mathbb{R}^2 à un plan muni d'un repère orthonormé direct $(O; \vec{i}, \vec{j})$; c'est assez naturel, dans la mesure où un point du plan est repéré par deux grandeurs, ses deux coordonnées, abscisse et ordonnée : on est ainsi en dimension 2.

1. Le plan comme ensemble de nombres complexes

Une interprétation très intéressante et très naturelle pour introduire les nombres complexes est celle de Jean-Robert Argand ¹ [8], et dont Jos Leys, Étienne Ghys et Aurélien Alvarez [10] donnent une interprétation extrêmement claire.

Représentons (voir figure 62.1) l'axe des réels par une droite graduée $\mathcal{D}_{\mathbb{R}}$, d'origine O ; la multiplication de 1 par -1 envoie 1 sur -1 , qui est l'image de 1 par la symétrie de centre O, que l'on peut aussi considérer comme étant la rotation de centre O, d'angle π .

De même, la multiplication de -1 par -1 envoie -1 sur 1.

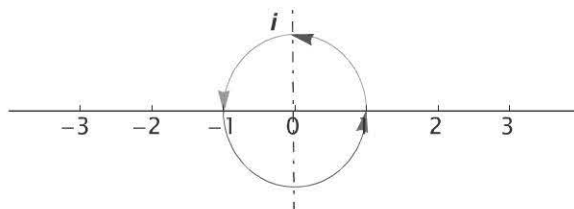


Figure 62.1 – La droite réelle

Notons i le point image de 1 par la rotation de centre O et d'angle $\frac{\pi}{2}$: ce point n'est plus sur la droite initiale, mais sur la perpendiculaire à celle-ci passant par l'origine.

Si on va un peu plus loin, et que l'on assimile la multiplication par i à l'opération résultant de la rotation de centre O et d'angle $\frac{\pi}{2}$, cela signifie que, si on applique cette même rotation au point i , c'est-à-dire qu'on le multiplie par lui-même, c'est-à-dire i , on obtient le point situé en -1 sur la droite $\mathcal{D}_{\mathbb{R}}$!

Considérer les points du plan comme des quantités sur lesquelles on peut définir une opération comme la multiplication permet donc de définir une racine carrée au nombre -1 , puisque l'on a alors :

$$i \times i = -1$$

Comme le point i a pour coordonnées $(0, 1)$, il est donc naturel de poser :

$$i = (0, 1)$$

1. Jean-Robert Argand (1768-1822), mathématicien suisse, célèbre pour son interprétation géométrique des nombres complexes comme points du plan. Il a également démontré le théorème de d'Alembert-Gauss, qui sera vu ultérieurement dans les pages qui suivent.

Ainsi, à chaque point du plan \mathbb{R}^2 , de coordonnées $(x, y) = x \times (1, 0) + y \times (0, 1)$, on peut associer le **nombre complexe**, ou **nombre imaginaire** :

$$z = x + iy$$

appelé affixe du point M .

Il est clair qu'il est plus facile de manipuler la grandeur $z = x + iy$ plutôt que $x \times (1, 0) + y \times (0, 1)$: on choisit donc, pour la suite, la notation la plus simple, c'est-à-dire la première !

2. Définitions et propriétés fondamentales

► Ensemble \mathbb{C}

On appelle **ensemble des nombres complexes**, que l'on note \mathbb{C} , l'ensemble :

$$\mathbb{C} = \{z = x + iy, (x, y) \in \mathbb{R}^2\}$$

(qui est aussi l'ensemble des couples de réels de la forme (x, y) , si on identifie \mathbb{R}^2 et \mathbb{C}).

► Écriture cartésienne

On appelle **écriture cartésienne** d'un nombre complexe z sa décomposition sous la forme :

$$z = x + iy$$

où x et y sont des réels.

► Racine



i est une **racine carrée complexe** de -1 , car il vérifie² :

$$i^2 = -1$$

(« Une » racine, car $-i$ est aussi une racine carrée complexe de -1 : $(-i)^2 = -1$.)

Théorème

Tout élément z de \mathbb{C} s'écrit, de manière unique $z = x + iy$, où x et y sont des réels.

► Partie réelle

On appelle **partie réelle** du nombre complexe $z = x + iy$, $(x, y) \in \mathbb{R}^2$, le nombre réel noté $\operatorname{Re}(z)$, et défini par :

$$\operatorname{Re}(z) = x$$

► Partie imaginaire

On appelle **Partie imaginaire** du nombre complexe $z = x + iy$, $(x, y) \in \mathbb{R}^2$, le nombre réel noté $\operatorname{Im}(z)$, et défini par :

$$\operatorname{Im}(z) = y$$

Tout nombre complexe dont la partie réelle est nulle, c'est-à-dire de la forme $z = iy$, $y \in \mathbb{R}$, est appelé **imaginaire pur**.

L'ensemble des nombres imaginaires purs est noté $i\mathbb{R}$.

2. Les règles de calcul dans \mathbb{C} seront développées au paragraphe suivant.

3. Règles de calcul dans \mathbb{C}

Théorème

L'ensemble \mathbb{C} peut être muni de deux lois, notées $+$ et \times , qui prolongent les lois $+$ et \times de \mathbb{R} .



On aura donc, pour tous nombres complexes $z = x + iy$ et $z' = x' + iy'$ de \mathbb{C} :

$$\begin{aligned} z z' &= (x + iy)(x' + iy') \\ &= x(x' + iy') + iy(x' + iy') \\ &= xx' + ix y' + iy x' - yy' \\ &= xx' - yy' + i(xy' + yx') \end{aligned}$$

4. Conjugué

On appelle **conjugué** du nombre complexe $z = x + iy$, $(x, y) \in \mathbb{R}^2$, le nombre complexe :

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

Propriété

Pour tout nombre complexe z :

$$\operatorname{Re}(z) = \operatorname{Re}(\bar{z})$$

Propriétés

1. Pour tout nombre complexe z :

$$\begin{cases} z + \bar{z} = 2\operatorname{Re}(z) \\ z - \bar{z} = 2i\operatorname{Im}(z) \end{cases}$$

2. Pour tout nombre complexe z :

$$z \in \mathbb{R} \Leftrightarrow z = \bar{z}$$

3. Pour tout nombre complexe z :

$$z \in i\mathbb{R} \Leftrightarrow z = -\bar{z}$$

4. Pour tout couple (z, z') de nombres complexes :

$$\begin{cases} \overline{z + z'} = \bar{z} + \bar{z}' \\ \overline{z z'} = \bar{z} \bar{z}' \end{cases}$$



Il résulte de la propriété précédente que, pour tout nombre complexe z :

$$\overline{-z} = -\bar{z}$$

et tout entier naturel n :

$$\overline{z^n} = \bar{z}^n$$



Une paternité controversée

Un peu d'histoire... C'est la fameuse Controverse de Cardan, au sujet de la résolution des équations du troisième degré, de la forme $x^3 + px = q$; Jérôme Cardan (1501-1576) publia, dans *Ars magna* en 1545, les formules donnant la solution de ces équations :

$$x = \sqrt[3]{\frac{q - \sqrt{q^2 + \frac{4p^3}{27}}}{2}} + \sqrt[3]{\frac{q + \sqrt{q^2 + \frac{4p^3}{27}}}{2}}$$

La controverse vint du fait que ces formules furent également trouvées par Nicolas Tartaglia (1499-1557), qui en revendiqua la paternité. Il semblerait, d'après ce qu'écrivit Cardan, que ces formules furent découvertes, en premier, par Scipione dal Ferro (1465-1526), qui, malheureusement, ne publia jamais ces résultats, et ne les confia qu'à un cercle restreint d'élèves.

À la fin du XVI^{ème} siècle, le mathématicien italien Rafael Bombelli (1526-1572) applique, dans son ouvrage *l'Algebra*, cette formule à l'équation $x^3 - 15x = 4$, et obtient :

$$x = \sqrt[3]{2 - 11\sqrt{-1}} + \sqrt[3]{2 + 11\sqrt{-1}}$$

où l'écriture « $\sqrt{-1}$ » désigne un nombre, a priori inconnu, dont le carré vaut -1 .

Une racine évidente entière de l'équation précédente est, bien sûr, 4. Mais si on recherche les autres racines, la formule obtenue par R. Bombelli prend un tout autre sens ; bien que la fonction $x \mapsto \sqrt[3]{x}$ ne soit définie que sur \mathbb{R}^+ , on constate, en utilisant les identités remarquables :

$$(a - b)^3 = a^3 - b^3 - 3a^2b + 3ab^2, \quad (a + b)^3 = a^3 + b^3 + 3a^2b + 3ab^2$$

que :

$$(2 - \sqrt{-1})^3 = 2^3 + \sqrt{-1} - 3 \times 4 \sqrt{-1} + 3 \times 2(-1) = 2 - 11\sqrt{-1}$$

et :

$$(2 + \sqrt{-1})^3 = 2^3 - \sqrt{-1} + 3 \times 4 \sqrt{-1} + 3 \times 2(-1) = 2 + 11\sqrt{-1}$$

Ainsi :

$$\sqrt[3]{2 - 11\sqrt{-1}} + \sqrt[3]{2 + 11\sqrt{-1}} = 4 \in \mathbb{R}$$

qui a un sens, et est bien solution de l'équation de départ :

$$4^3 - 15 \times 4 - 4 = 0$$

Le fait que -1 puisse être le carré d'un nombre, même « imaginaire », a ainsi commencé à faire son chemin.

Des nombres imaginaires bien pratiques

Leonhard Euler (1707-1783), s'intéressa également aux nombres complexes. On lui doit, notamment, la formule portant son nom. Jean le Rond D'Alembert (1717-1783) mit en évidence la propriété de clôture algébrique du corps des nombres complexes.

En 1799, Caspar Wessel (1745-1818), mathématicien danois et norvégien, publie un mémoire où il utilise les nombres complexes pour représenter des lignes géométriques, caractérisées par leur longueur et leur direction.

Les interprétations géométriques, et les applications qui en résultent, se développent, essentiellement, à partir du XIX^{ième} siècle, avec, tout d'abord, le chanoine Buée, puis Jean-Robert Argand (1768-1822), Carl Friedrich Gauss (1777-1855) et Augustin Cauchy (1789-1857). Depuis, la recherche sur les nombres complexes connaît un essor considérable : les nombres complexes sont au cœur de la géométrie algébrique et analytique moderne (avec, notamment, les travaux de Jean-Pierre Serre, Alexandre Grothendieck, et Hans Grauert), puis ceux d'Adrien Douady (1935-2006), professeur à l'Université d'Orsay, qui s'intéressa à l'application aux systèmes dynamiques des nombres complexes, les ensembles de Julia, les fractales et ensembles de Mandelbrot ...

Et c'est ainsi que sont nés les **nombres complexes**. Il faut les considérer comme des outils, extrêmement pratiques pour résoudre des problèmes qui, sinon, n'auraient pas de solution.

- Adrien-Quentin Buée, chanoine honoraire de Notre-Dame, est mort en 1825 à 80 ans ; versé dans les sciences, il publia des écrits mathématiques ; il est souvent qualifié « d'abbé » par confusion avec son frère l'Abbé Buée, chanoine titulaire de Notre-Dame [7].
- Jean-Pierre Serre (1926-) fut lauréat de la médaille Fields en 1954.
- Alexandre Grothendieck (1928-) fut lauréat de la médaille Fields en 1966, il apparaît comme un refondateur de la géométrie algébrique.
- Hans Grauert est un mathématicien allemand (1930-2011), il travailla beaucoup sur les variétés complexes. Ses travaux se placent dans la lignée de ceux d'Hermann Weyl, David Hilbert, Bernhard Riemann.
- Gaston Maurice Julia (1893-1978) est un mathématicien français.
- Benoît Mandelbrot (1924-2010) est un mathématicien franco-américain.

Représentation géométrique des nombres complexes

On se place, dans ce qui suit, dans le plan \mathbb{R}^2 rapporté à un repère orthonormé direct $(O; \vec{i}, \vec{j})$. À tout nombre complexe $z = x + iy$, on associe le point M de coordonnées (x, y) dans le plan.

► Image

On appelle **image** du nombre complexe $z = x + iy$, $(x, y) \in \mathbb{R}^2$, le point M de coordonnées (x, y) .

► Affixe

On appelle **affixe** du point M de coordonnées $(x, y) \in \mathbb{R}^2$, le nombre complexe $aff(M) = z_M = x + iy$.

On appelle **affixe** du vecteur $\vec{u} = \overrightarrow{AB}$ le nombre complexe $aff(\overrightarrow{AB}) = z_B - z_A$.



L'affixe du point M de coordonnées (x, y) est aussi celle du vecteur \overrightarrow{OM} .

Propriétés

1. Pour tout couple de vecteurs (\vec{u}, \vec{v}) du plan :

$$aff(\vec{u} + \vec{v}) = aff(\vec{u}) + aff(\vec{v})$$

2. Pour tout vecteur \vec{u} du plan, et tout réel λ :

$$aff(\lambda \vec{u}) = \lambda aff(\vec{u})$$

► Module

On appelle **module** du nombre complexe $z = x + iy$, $(x, y) \in \mathbb{R}^2$, et on note $|z|$, le réel positif $|z| = \sqrt{x^2 + y^2}$ (c'est aussi la distance du point d'affixe z à l'origine).

► Argument

On appelle **argument** du nombre complexe non nul $z = x + iy$, $(x, y) \in \mathbb{R}^2$, toute mesure, en *radians*, de l'angle orienté $\left(\vec{i}, \widehat{\overrightarrow{OM}}\right)$, où M est le point d'affixe z . On notera $arg(z)$ une telle mesure.

L'unique argument de z appartenant à l'intervalle $] -\pi, \pi]$ s'appelle **l'argument principal**.



1. Le nombre complexe nul 0 ne possède pas d'argument.
2. Un nombre complexe non nul possède une infinité d'arguments ! Si θ est un argument du nombre complexe $z \in \mathbb{C}^*$, les autres arguments de z sont exactement les réels de la forme $\theta + 2k\pi$, $k \in \mathbb{Z}$.

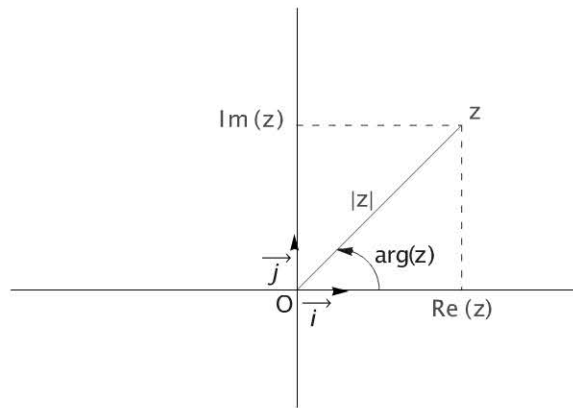


Figure 63.1 – Représentation géométrique d'un nombre complexe non nul.

Pour tout nombre réel θ , on pose :

$$e^{i\theta} = \cos \theta + i \sin \theta$$

Propriétés

1. Tout nombre complexe z de module 1 peut s'écrire sous la forme :

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

où θ est un réel, unique à 2π près, tel que :

$$\theta = \text{Arg}(z) \quad [2\pi]$$

où l'écriture $[2\pi]$ signifie **modulo** 2π , c'est-à-dire :

$$\theta = \arg(z) + 2k\pi, \quad k \in \mathbb{Z}$$

2. Pour tout nombre complexe z non nul, il existe un réel strictement positif, r , et un réel, θ , tels que z s'écrit sous la forme :

$$z = r e^{i\theta} = r \cos \theta + i r \sin \theta$$

où θ est un réel, unique à 2π près :

$$\theta = \arg(z) \quad [2\pi]$$

3. Pour tout nombre complexe z non nul, s'il existe un réel strictement positif, r , et un réel θ tels que :

$$z = r e^{i\theta} = r \cos \theta + i r \sin \theta$$

alors :

$$\begin{cases} r = |z| \\ \theta = \arg(z) \quad [2\pi] \end{cases}$$

► Exponentielle complexe

Pour tout nombre complexe $z = x + iy$, $(x, y) \in \mathbb{R}^2$, on pose :

$$e^z = e^{x+iy} = e^x e^{iy} = e^x (\cos y + i \sin y)$$

e désigne ainsi l'**exponentielle complexe**.



Il est clair que, pour $z = x \in \mathbb{R}$, on retrouve l'exponentielle réelle.

D'autre part, l'exponentielle complexe étant définie à partir de l'exponentielle réelle, on aura nécessairement, pour tout z de \mathbb{C} , comme $|e^z| = e^{\operatorname{Re}(z)}$:

$$e^z \neq 0$$

Propriétés

1. Pour tout couple de nombres complexes (z, z') :

$$e^{z+z'} = e^z e^{z'}$$

2. Pour tout nombre complexe z et tout entier naturel n :

$$(e^z)^n = e^{nz}$$

Inversion des nombres complexes

1. Inverse d'un nombre complexe

► Propriétés

1. Tout nombre complexe non nul z admet un unique inverse, noté z^{-1} , tel que :

$$z z^{-1} = z^{-1} z = 1$$

2. Pour tout nombre complexe z :

$$\frac{1}{e^z} = e^{-z}$$

3. Pour tout nombre complexe z et tout entier relatif k :

$$(e^z)^k = e^{kz}$$

► Calcul de l'inverse d'un nombre complexe

En pratique, pour calculer l'inverse d'un nombre complexe $z = x + iy$, $(x, y) \in \mathbb{R}^2$, on recherche l'inverse z^{-1} du nombre complexe z sous la forme $z^{-1} = x' + iy'$, puis on écrit :

$$x' + iy' = \frac{1}{x + iy} = \frac{x - iy}{(x + iy)(x - iy)} = \frac{x - iy}{x^2 + y^2}$$

(c'est la technique classique qui consiste à multiplier numérateur et dénominateur par l'expression conjuguée du dénominateur, ce qui permet donc de faire apparaître le module de celui-ci.)

Il ne reste plus qu'à identifier parties réelles et imaginaires, tout nombre complexe s'écrivant de manière unique :

$$x' = \frac{x}{x^2 + y^2} \quad , \quad y' = -\frac{y}{x^2 + y^2}$$

Il en résulte :

$$z^{-1} = \frac{\bar{z}}{|z|^2}$$

2. Anneau commutatif

On désigne, souvent, \mathbb{C} comme étant le **corps des nombres complexes**, parce qu'il est muni d'une structure algébrique, dite **structure de corps**, qui désigne un ensemble muni de deux opérations, notées « + » et « × », dans lequel tout élément non nul est inversible. Plus précisément, un corps est un **anneau commutatif**, dans lequel tout élément

non nul est inversible ; un **anneau** \mathcal{A} est un ensemble $(\mathcal{A}, +, \times)$, muni de deux *lois de composition internes*, $+$ et \times , telles que :

- $(\mathcal{A}, +)$ est un **groupe commutatif** (c'est-à-dire un ensemble non vide, muni d'une loi de composition interne, notée, en général, « $+$ », associative, commutative, admettant un élément neutre $0_{\mathcal{A}}$, et telle que tout élément soit symétrisable), d'élément neutre $0_{\mathcal{A}}$.

- la loi de composition interne \times est **associative** :

$$\forall (z_1, z_2, z_3) \in \mathcal{A}^3 :$$

$$z_1 \times (z_2 \times z_3) = (z_1 \times z_2) \times z_3$$

et **distributive à gauche et à droite par rapport à la loi $+$** :

$$\forall (z_1, z_2, z_3) \in \mathcal{A}^3 :$$

$$z_1 \times (z_2 + z_3) = z_1 \times z_2 + z_1 \times z_3 \quad , \quad (z_1 + z_2) \times z_3 = z_1 \times z_3 + z_2 \times z_3$$

- la loi de composition interne \times admet un **élément neutre**, distinct de $0_{\mathcal{A}}$, noté $1_{\mathcal{A}}$.

Si la loi \times est commutative, l'anneau est dit **commutatif**, ou **abélien**.

Propriétés fondamentales des nombres complexes

Pour tout nombre complexe non nul z :

$$\begin{cases} \arg(\bar{z}) = -\arg(z) \quad [2\pi] \\ \arg(-z) = \arg(z) + \pi \quad [2\pi] \\ \arg(-\bar{z}) = \pi - \arg(z) \quad [2\pi] \end{cases}$$

Pour tout nombre complexe non nul z , et tout réel strictement positif λ :

$$\arg(\lambda z) = \arg(z) \quad [2\pi]$$



Ces propriétés permettent de traduire des problèmes de géométrie par des relations entre nombres complexes.

Corollaire

- Pour tout couple de nombres complexes $(z, z') \in \mathbb{C}^* \times \mathbb{C}^*$:

$$\arg(zz') = \arg(z) + \arg(z') \quad [2\pi]$$

- Pour tout couple de nombres complexes $(z, z') \in \mathbb{C}^* \times \mathbb{C}^*$:

$$\arg\left(\frac{z}{z'}\right) = \arg(z) - \arg(z') \quad [2\pi]$$

Démonstration :

Il suffit d'utiliser la forme polaire. ■

Pour tout couple $(z, z') \in \mathbb{C} \times \mathbb{C}^* : \overline{\left(\frac{z}{z'}\right)} = \frac{\bar{z}}{\bar{z}'}$.

Pour tout nombre complexe $z = x + iy$, $(x, y) \in \mathbb{R}^2$:

$$|z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}$$

Pour tout couple de nombres complexes $(z, z') : |zz'| = |z||z'|$.

Corollaire

Pour tout nombre complexe $z \in \mathbb{C}$, et tout réel λ :

$$|\lambda z| = |\lambda| |z|$$

Pour tout couple $(z, z') \in \mathbb{C} \times \mathbb{C}^* : \left|\frac{z}{z'}\right| = \frac{|z|}{|z'|}$.

Pour tout nombre complexe non nul $z : \left|\frac{1}{z}\right| = \frac{1}{|z|}$.

Pour tout couple $(z, z') \in \mathbb{C}^2 : |z + z'| \leq |z| + |z'|$.



1. Dans \mathbb{C} , il n'y a plus la notion d'ordre usuelle « < », « > » : on ne peut donc comparer un nombre complexe à un autre, ou dire s'il est positif ou négatif, etc ...
2. Le symbole $\sqrt{}$ reste réservé aux nombres réels positifs.

► Formules d'Euler

Pour tout réel θ :

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2}, \quad \sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

► Formule de Moivre¹

Pour tout réel θ , et tout entier naturel n :

$$(\cos \theta + i \sin \theta)^n = \cos n\theta + i \sin n\theta$$

1. Abraham De Moivre (1667-1754). C'est un des premiers vrais « probabilistes ».

Complément : les polynômes de Tchebychev

Dans ce qui suit, on s'intéresse à la famille des polynômes de Tchebychev¹ $(T_n)_{n \geq 2}$, définie, pour tout entier naturel n , par $T_n(\theta) = \cos(n, \theta)$. On montre que chacun des T_n , $n \in \mathbb{N}$, est un polynôme de degré n en $\cos \theta$.

Soit θ un réel non nul.

1. Étape 1

Vérifions que $\cos(2\theta)$ est un polynôme de degré 2 en $\cos \theta$:

$$\cos(2\theta) = \cos^2 \theta - \sin^2 \theta = 2 \cos^2 \theta - 1 = P(\cos \theta),$$

où P est le polynôme défini sur \mathbb{C} par $P(z) = 2z^2 - 1$.

2. Étape 2

Montrons que $\cos(3\theta)$ est un polynôme de degré 3 en $\cos \theta$.

D'après la formule de Moivre :

$$\cos(3\theta) = \operatorname{Re}(e^{3i\theta})$$

Or :

$$e^{3i\theta} = (e^{i\theta})^3 = (\cos \theta + i \sin \theta)^3 = \cos^3 \theta + 3i \cos^2 \theta \sin \theta + 3 \cos \theta (i \sin \theta)^2 + (i \sin \theta)^3$$

soit :

$$e^{3i\theta} = \cos^3 \theta + 3i \cos^2 \theta \sin \theta - 3 \cos \theta \sin^2 \theta - i \sin^3 \theta$$

En identifiant parties réelles et parties imaginaires, on en déduit :

$$\begin{aligned} \cos(3\theta) &= \operatorname{Re}(e^{3i\theta}) = \cos^3 \theta - 3 \cos \theta \sin^2 \theta = \cos^3 \theta - 3 \cos \theta (1 - \cos^2 \theta) \\ &= 4 \cos^3 \theta - 3 \cos \theta \end{aligned}$$

3. Étape 3

Démontrons par récurrence que, pour tout entier naturel non nul n , $\cos(n\theta)$ est un polynôme, noté T_n , de degré n en $\cos \theta$, de coefficient dominant 2^{n-1} .

- **Au rang 0 :** $\cos(0 \times \theta) = 1 = \cos^0(\theta)$.
- **La propriété est vraie au rang 1 :** $\cos \theta$ est bien un polynôme de degré 1 en $\cos \theta$, de coefficient dominant $2^{1-1} = 1$.
- **La propriété est vraie au rang 2 :** $\cos(2\theta) = 2 \cos^2 \theta - 1$ est bien un polynôme de degré 2 en $\cos \theta$, de coefficient dominant $2^{2-1} = 2$.

1. Pafnouti Lvovitch Tchebychev (1821-1894), mathématicien russe, qui apporta de nombreuses contributions en probabilités et en statistiques.

• **Supposons la vraie jusqu'à un rang $n > 1$.**

Les formules d'addition permettent alors d'écrire :

$$\begin{aligned}\cos((n+1)\theta) &= \cos(n\theta) \cos \theta - \sin(n\theta) \sin \theta, \\ \text{et } \cos((n-1)\theta) &= \cos(n\theta) \cos \theta + \sin(n\theta) \sin \theta.\end{aligned}$$

Par suite, en additionnant membre à membre ces deux relations, on en déduit :
 $\cos((n+1)\theta) + \cos((n-1)\theta) = 2 \cos(n\theta) \cos \theta$, puis :

$$\cos((n+1)\theta) = 2 \cos(n\theta) \cos \theta - \cos((n-1)\theta). \quad (\star)$$

Par hypothèse de récurrence, $\cos(n\theta)$ est un polynôme de degré n en $\cos \theta$.
 $\cos(n\theta) \cos \theta$ est donc un polynôme de degré $n+1$ en $\cos \theta$. Comme $\cos((n-1)\theta)$ est un polynôme de degré $n-1$ en $\cos \theta$, $\cos((n+1)\theta)$ est bien un polynôme de degré $n+1$ en $\cos \theta$.

Le coefficient du terme de plus haut degré est : $2 \times 2^{n-1} = 2^n = 2^{n+1-1}$. La propriété est donc vraie au rang $n+1$.

- Comme elle est vraie au rang 1, elle est donc vraie pour tout entier naturel $n \geq 1$.

Ce résultat peut aussi se démontrer grâce à la formule de Moivre :

$$\cos(n\theta) = \operatorname{Re}(e^{in\theta})$$

$$\text{Or : } e^{in\theta} = (e^{i\theta})^n = (\cos \theta + i \sin \theta)^n = \sum_{k=0}^n C_n^k \cos^k \theta i^{n-k} \sin^{n-k} \theta = \sum_{k=0}^n C_n^k \cos^{n-k} \theta i^k \sin^k \theta$$

où C_n^k désigne le coefficient binomial $\binom{n}{k}$, soit :

$$\begin{aligned}e^{in\theta} &= \sum_{k=0, k \text{ pair}}^n C_n^k \cos^{n-k} \theta i^k \sin^k \theta + \sum_{k=0, k \text{ impair}}^n C_n^k \cos^{n-k} \theta i^k \sin^k \theta \\ &= \sum_{p=0}^{E(\frac{n}{2})} C_n^{2p} \cos^{n-2p} \theta i^{2p} \sin^{2p} \theta + \sum_{p=0}^{E(\frac{n-1}{2})} C_n^{2p+1} \cos^{n-2p-1} \theta i^{2p+1} \sin^{2p+1} \theta \\ &= \sum_{p=0}^{E(\frac{n}{2})} (-1)^p C_n^{2p} \cos^{n-2p} \theta (\sin^2 \theta)^p + i \sum_{p=0}^{E(\frac{n-1}{2})} (-1)^p C_n^{2p+1} \cos^{n-2p-1} \theta \sin^{2p+1} \theta \\ &= \sum_{p=0}^{E(\frac{n}{2})} (-1)^p C_n^{2p} \cos^{n-2p} \theta (1 - \cos^2 \theta)^p + i \sum_{p=0}^{E(\frac{n-1}{2})} (-1)^p C_n^{2p+1} \cos^{n-2p-1} \theta \sin^{2p+1} \theta\end{aligned}$$

Par suite :

$$\begin{aligned}\cos(n\theta) &= \sum_{p=0}^{E(\frac{n}{2})} (-1)^p C_n^{2p} \cos^{n-2p} \theta (1 - \cos^2 \theta)^p \\ &= \sum_{p=0}^{E(\frac{n}{2})} (-1)^p C_n^{2p} \cos^{n-2p} \theta \left\{ \sum_{k=0}^p (-1)^k C_p^k \cos^{2k} \theta \right\} \\ &= \sum_{p=0}^{E(\frac{n}{2})} \sum_{k=0}^p (-1)^{p+k} C_n^{2p} C_p^k \cos^{n+2k-2p} \theta\end{aligned}$$

Le terme de plus haut degré est obtenu lorsque $k = p$, et vaut donc :

$$\sum_{p=0}^{E(\frac{n}{2})} (-1)^{2p} C_n^{2p} C_p^p \cos^n \theta = \sum_{p=0}^{E(\frac{n}{2})} C_n^{2p} \cos^n \theta = \cos^n \theta \sum_{p=0}^{E(\frac{n}{2})} C_n^{2p}$$

Or :

$$2^n = (1 + 1)^n = \sum_{k=0}^n C_n^k 1^p 1^{n-k} = \sum_{k=0}^n C_n^k = \sum_{p=0}^{E(\frac{n}{2})} C_n^{2p} + \sum_{p=0}^{E(\frac{n-1}{2})} C_n^{2p+1}$$

$$0 = (1 - 1)^n = \sum_{p=0}^n C_n^k (-1)^k 1^{n-k} = \sum_{k=0}^n (-1)^k C_n^k = \sum_{p=0}^{E(\frac{n}{2})} C_n^{2p} - \sum_{p=0}^{E(\frac{n-1}{2})} C_n^{2p+1}$$

En additionnant membre à membre ces deux relations, on en déduit :

$$2^n = 2 \sum_{p=0}^{E(\frac{n}{2})} C_n^{2p}$$

et donc : $\sum_{p=0}^{E(\frac{n}{2})} C_n^{2p} = 2^{n-1}.$

4. Étape 4

Déterminons la relation de récurrence linéaire d'ordre 2 vérifiée par la suite de polynômes $(T_n)_{n \geq 2}$.

Si on pose, pour tout entier naturel non nul n , $T_n(\cos \theta) = \cos(n\theta)$, alors, pour tout entier $n \geq 2$, la relation (★) s'écrit aussi :

$$T_{n+1}(\cos \theta) = 2 \cos \theta T_n(\cos \theta) - T_{n-1}(\cos \theta)$$

T_n, T_{n-1}, T_{n-1} étant des polynômes, ils vérifient donc la relation de récurrence linéaire d'ordre 2 :

$$T_{n+1}(X) = 2X T_n(X) - T_{n-1}(X)$$

Cette relation permet, en particulier, d'obtenir l'expression du coefficient du terme de plus haut degré de T_n .

En effet, si on désigne par α_n ce coefficient, on a donc :

$$\alpha_{n+1} = 2 \alpha_n$$

(le degré du polynôme $X T_n(X)$ est différent de celui de $T_{n-1}(X)$).

En itérant, on en déduit, pour tout entier naturel n :

$$\alpha_{n+1} = 2 \alpha_n = 2 \times 2 \alpha_{n-1} = \dots = 2^n \alpha_1 = 2^n$$

La suite $(\alpha_n)_{n \in \mathbb{N}^*}$ est une suite géométrique de raison 2, de premier terme 1 ; ainsi, pour tout entier naturel n non nul : $\alpha_n = 2^{n-1}$.

Racines $n^{\text{ième}}$ de l'unité, racines $n^{\text{ièmes}}$ complexes

1. Racines $n^{\text{ième}}$ de l'unité

Définition

Étant donné un entier naturel non nul n , on appelle **racine $n^{\text{ième}}$ de l'unité** (ou **nombre de Moivre**) tout nombre complexe z solution de l'équation :

$$z^n = 1$$

Pour obtenir l'expression d'une *racine $n^{\text{ième}}$ de l'unité*, on cherche donc z dans \mathbb{C} tel que :

$$z^n = 1$$

On a donc, nécessairement : $z \neq 0$.

On peut donc chercher z sous la forme $z = r e^{i\theta}$, avec :

$$r = |z| \in \mathbb{R}_+^* \quad , \quad \theta = \text{Arg}(z) \quad [2\pi]$$

Il en résulte, grâce à la formule de Moivre :

$$(r e^{i\theta})^n = \underbrace{r^n}_{\in \mathbb{R}_+^*} e^{in\theta} = \underbrace{1}_{\in \mathbb{R}_+^*} \cdot e^{i0}$$

En identifiant les modules et arguments respectifs des deux membres, on en déduit :

$$\begin{cases} r^n = 1 \\ n\theta = 2k\pi, k \in \mathbb{Z} \end{cases}$$

Comme $r \in \mathbb{R}_+^*$, il en résulte :

$$r = 1 \quad , \quad \theta = \frac{2k\pi}{n}, k \in \mathbb{Z}$$

Si on effectue la division euclidienne de $k \in \mathbb{Z}$ par l'entier naturel non nul n , on obtient :

$$k = nq + r \quad \text{avec } r \in \llbracket 0, n-1 \rrbracket$$

Par 2π -périodicité du sinus et du cosinus, on a alors :

$$e^{\frac{2ik\pi}{n}} = e^{\frac{2i(nq+r)\pi}{n}} = e^{2iq\pi + \frac{2ir\pi}{n}} = e^{\frac{2ir\pi}{n}}$$

L'ensemble des racines $n^{\text{ièmes}}$ complexes de 1 est donc donné par :

$$\mathcal{S} = \left\{ e^{\frac{2ir\pi}{n}}, r \in \{0, 1, 2, \dots, n-1\} \right\} = \left\{ e^{\frac{2ik\pi}{n}}, k \in \{0, 1, 2, \dots, n-1\} \right\}$$

Dans \mathbb{C} , il existe exactement n racines $n^{\text{ièmes}}$ de l'unité.



Exemple

Réolvons, dans \mathbb{C} : $z^3 = 1$, soit :

$$z^3 = e^{2ik\pi}, \quad k \in \mathbb{Z}$$

En identifiant module et argument, on en déduit que les solutions sont données par :

$$z = e^{\frac{2ik\pi}{3}}, \quad k \in \{0, 1, 2\}$$

L'ensemble des solutions est donc :

$$\mathcal{S} = \left\{1, e^{\frac{2i\pi}{3}}, e^{\frac{4i\pi}{3}}\right\} = \{1, j, j^2\}$$

où :

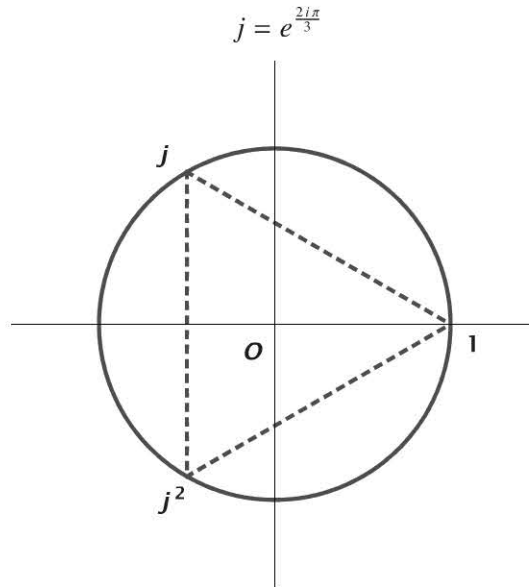


Figure 67.1 – Les racines cubiques complexes de l'unité.

Proposition

Pour tout entier naturel $n \geq 3$, les points dont les affixes sont les n racines $n^{\text{ièmes}}$ de l'unité, $e^{\frac{2ik\pi}{n}}$, $k \in \{0, 1, 2, \dots, n-1\}$, sont les sommets d'un polygone régulier, de centre O.

Démonstration : Pour tout k de $\{0, 1, 2, \dots, n-1\}$, désignons par M_k le point d'affixe $e^{\frac{2ik\pi}{n}}$, et par $M_n = M_0$ le point d'affixe 1.

Alors :

$$\begin{aligned} \left(\widehat{\overrightarrow{OM_k}, \overrightarrow{OM_{k+1}}} \right) &= \left(\widehat{\overrightarrow{OM_k}, \vec{i}} \right) + \left(\widehat{\vec{i}, \overrightarrow{OM_{k+1}}} \right) \quad [2\pi] \\ &= -\arg\left(e^{\frac{2ik\pi}{n}}\right) + \arg\left(e^{\frac{2i(k+1)\pi}{n}}\right) \quad [2\pi] \\ &= \arg\left(\frac{e^{\frac{2i(k+1)\pi}{n}}}{e^{\frac{2ik\pi}{n}}}\right) \quad [2\pi] \\ &= \arg\left(e^{\frac{2i\pi}{n}}\right) \quad [2\pi] \\ &= \frac{2\pi}{n} \quad [2\pi] \end{aligned}$$

Le polygone $M_0M_1 \dots M_n$, de centre O, est donc régulier. ■

2. Racine $n^{\text{ième}}$ complexe

Définition

Étant donné un entier naturel non nul n , et un nombre complexe z_0 , on appelle **racine $n^{\text{ième}}$ complexe** de z_0 tout nombre complexe z solution de l'équation :

$$z^n = z_0$$

En pratique : pour déterminer les racines $n^{\text{ièmes}}$ complexes du nombre complexe non nul z_0 , il faut déjà mettre ce dernier sous forme polaire :

$$z_0 = \rho_0 e^{i\theta_0} \quad , \quad \rho_0 > 0 \quad , \quad \theta_0 \in [0, 2\pi[$$

On recherche ensuite z sous la forme :

$$z = \rho e^{i\theta} \quad , \quad \rho > 0 \quad , \quad \theta \in [0, 2\pi[$$

Ainsi :

$$z^n = \rho^n e^{in\theta} = \rho_0 e^{i\theta_0} = \rho_0 e^{i(\theta_0 + 2k\pi)} \quad , \quad k \in \mathbb{Z}$$

Il en résulte :

$$\rho = \rho_0^{\frac{1}{n}}$$

et :

$$\theta = \frac{\theta_0 + 2k\pi}{n} \quad , \quad k \in \mathbb{Z}$$

Si on effectue la division euclidienne de $k \in \mathbb{Z}$ par l'entier naturel non nul n , on obtient :

$$k = nq + r$$

où $r \in \llbracket 0, n-1 \rrbracket$.

Par 2π -périodicité du sinus et du cosinus, on a alors :

$$e^{\frac{i(\theta_0 + 2k\pi)}{n}} = e^{\frac{i(\theta_0 + 2(nq+r)\pi)}{n}} = e^{\frac{i(\theta_0 + 2r\pi)}{n}} e^{2iq\pi} = e^{\frac{i(\theta_0 + 2r\pi)}{n}}$$

L'ensemble des racines $n^{\text{ièmes}}$ complexes de z_0 est donc donné par :

$$S = \left\{ \rho_0^{\frac{1}{n}} e^{\frac{i(\theta_0 + 2r\pi)}{n}} , r \in \{0, 1, 2, \dots, n-1\} \right\} = \left\{ \rho_0^{\frac{1}{n}} e^{\frac{i(\theta_0 + 2k\pi)}{n}} , k \in \{0, 1, 2, \dots, n-1\} \right\}$$

Exemple

Soit $n \in \mathbb{N}^*$. Résolvons, dans \mathbb{C} , l'équation suivante :

$$z^n = -1$$

En remarquant que $-1 = e^{i\pi}$, on en déduit :

$$z = e^{\frac{(2k+1)i\pi}{n}} \quad , \quad k \in \{0, \dots, n-1\}$$

Factorisation des polynômes dans le corps \mathbb{C}

1. Polynômes complexes

Définition

n étant un entier naturel, on appelle **polynôme complexe de degré n** une expression de la forme :

$$P(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = \sum_{k=0}^n a_k z^k$$

où a_n, a_{n-1}, \dots, a_0 sont des nombres complexes.

Par définition, le degré du polynôme P , noté $\deg P$, est donc celui de son terme (ou monôme) de plus haut degré, c'est-à-dire $a_n z^n$.

Par convention, le degré du polynôme nul sera noté $-\infty$.

Tout polynôme complexe de degré $n \in \mathbb{N}^*$ s'écrit de manière unique sous la forme :

$$P(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = \sum_{k=0}^n a_k z^k$$

$$(a_0, \dots, a_n) \in \mathbb{C}^n \times \mathbb{C}^*.$$

Notation

On notera $\mathbb{C}[X]$ l'ensemble des polynômes complexes.

Le sous-ensemble de $\mathbb{C}[X]$ constitué par les polynômes complexes de degré inférieur ou égal à n sera noté $\mathbb{C}_n[X]$.

2. Racine d'un polynôme complexe

Définition

On appelle **racine** d'un polynôme P , un nombre complexe z tel que :

$$P(z) = 0$$

Théorème

Si z_0 est une racine du polynôme P , supposé non nul, alors, il existe un polynôme Q , de degré $\deg P - 1$, tel que :

$$P(z) = (z - z_0) Q(z)$$

Démonstration : Ce résultat s'obtient par une simple division euclidienne par $z - z_0$. Le reste, qui est constant, est obtenu en évaluant la valeur de P en z_0 , et vaut donc zéro. ■

Exemple

Factorisons le polynôme P tel que :

$$\forall z \in \mathbb{C} : P(z) = z^3 + 4z = z(z^2 + 4)$$

$0, 2i$ et $-2i$ sont racines de P , et :

$$\forall z \in \mathbb{C} : P(z) = z(z + 2i)(z - 2i)$$

► Racine double**Définition**

On appelle **racine double** d'un polynôme P , un nombre complexe z_0 tel qu'il existe un polynôme Q vérifiant :

$$P(z) = (z - z_0)^2 Q(z)$$

avec $Q(z_0) \neq 0$.

On dit aussi que z_0 est une racine de P de *multiplicité* 2.

► Racine triple**Définition**

On appelle **racine triple** d'un polynôme P , un nombre complexe z_0 tel qu'il existe un polynôme Q vérifiant :

$$P(z) = (z - z_0)^3 Q(z)$$

avec $Q(z_0) \neq 0$.

On dit aussi que z_0 est une racine de P de *multiplicité* 3.

► Racine d'ordre n **Définition**

n étant un entier naturel non nul, on appelle **racine d'ordre n** d'un polynôme P , un nombre complexe z_0 tel qu'il existe un polynôme Q vérifiant :

$$P(z) = (z - z_0)^n Q(z)$$

avec $Q(z_0) \neq 0$.

On dit aussi que z_0 est une racine de P de *multiplicité* n .

► Théorème de d'Alembert-Gauss**Théorème**

Le corps \mathbb{C} est dit **algébriquement clos**, c'est-à-dire tout polynôme non constant, à coefficients dans le corps des nombres complexes, admet au moins une racine.

Par conséquent, tout polynôme à coefficients entiers, rationnels ou encore réels admet au moins une racine complexe, car ces nombres sont aussi des complexes.

a, b, c étant trois nombres complexes tels que $a \neq 0$:

$$a z^2 + b z + c = a(z - z_1)(z - z_2)$$

où :

$$z_1 = \frac{-b - \delta}{2a}, \quad z_2 = \frac{-b + \delta}{2a}$$

et où le nombre complexe δ vérifie $\delta^2 = b^2 - 4ac$, c'est-à-dire est une racine carrée complexe du discriminant Δ .



1. Dans \mathbb{C} , Δ admet deux racines carrées complexes, δ et $-\delta$. Choisir, comme racine, δ plutôt que $-\delta$ ne change en aucune façon le résultat.

En effet, si on considère la racine complexe $-\delta$ du discriminant Δ , on obtient alors, comme racines :

$$z'_1 = \frac{-b + \delta}{2a} = z_2, \quad z'_2 = \frac{-b - \delta}{2a} = z_1$$

Les racines sont donc les mêmes, mais obtenues dans un ordre différent.

2. Le discriminant Δ n'admet de racines réelles que si $b^2 - 4ac \geq 0$.

► Discriminant réduit

a, b, c étant trois nombres complexes tels que $a \neq 0$, on considère l'équation :

$$a z^2 + b z + c = 0$$

Lorsque b est de la forme $b = 2b'$, $b' \in \mathbb{C}$, les racines z_1, z_2 peuvent être obtenues en utilisant le **discriminant réduit** :

$$\Delta' = b'^2 - ac = \delta'^2$$

ce qui conduit à :

$$z_1 = \frac{-b' - \delta'}{a}, \quad z_2 = \frac{-b' + \delta'}{a}$$

où le nombre complexe δ' vérifie $\delta'^2 = b'^2 - ac$, c'est-à-dire est une racine carrée complexe du discriminant réduit Δ' .

Démonstration : Le discriminant est : $\Delta = b^2 - 4ac = 4(b'^2 - ac)$. Ainsi :

$$z_1 = \frac{-b - \delta}{2a} = \frac{-b' - \delta'}{a}, \quad z_2 = \frac{-b + \delta}{2a} = \frac{-b' + \delta'}{a} \quad \blacksquare$$



1. Comme précédemment, choisir, comme racine, δ' plutôt que $-\delta'$ ne change en aucune façon le résultat.
2. Le discriminant réduit Δ' n'admet de racines réelles que si

$$b'^2 - ac \geq 0$$

Exemple

Résoudre, dans \mathbb{C} $(1+i)z^2 + iz - 1 = 0$.

On obtient : $\Delta = 3 + 4i$.

Il s'agit alors de trouver un nombre complexe δ tel que $\delta^2 = \Delta$; on le cherche sous la forme $\delta = \alpha + i\beta$, ce qui conduit à :

$$\delta^2 = \alpha^2 - \beta^2 + 2i\alpha\beta = 3 + 4i$$

On en déduit : $\alpha^2 - \beta^2 = 3$, $\alpha\beta = 2$.

2 et 1, -2 et -1 sont racines évidentes :

$$\alpha = 2 \quad , \quad \beta = 1 \quad \text{ou} \quad \alpha = -2 \quad , \quad \beta = -1$$

Ces deux choix sont absolument équivalents, dans la mesure où :

$$\Delta = \delta^2 = (\alpha + i\beta)^2 = (-\delta)^2 = (-\alpha - i\beta)^2$$

Ainsi, les solutions de l'équation du second degré initiale sont :

$$\begin{cases} z_1 = \frac{-i - 2 - i}{2(1+i)} = -1 \\ z_2 = \frac{-i + 2 + i}{2(1+i)} = \frac{1}{1+i} = \frac{1-i}{2} \end{cases}$$

où l'on a utilisé, pour simplifier l'expression de z_2 , le passage à la forme conjuguée, afin d'obtenir un dénominateur réel.

3. Relations coefficients-racines**► Relations coefficients-racines pour un polynôme de degré 2**

Soit $(s, p) \in \mathbb{C}^2$. Les nombres complexes z_1 et z_2 vérifiant :

$$\begin{cases} z_1 z_2 = p \in \mathbb{C} \\ z_1 + z_2 = s \in \mathbb{C} \end{cases}$$

sont exactement les racines de $z^2 - sz + p = 0$.

Démonstration : Il suffit de calculer :

$$(z - z_1)(z - z_2) = z^2 - z(z_1 + z_2) + z_1 z_2 = z^2 - sz + p$$

■

► Relations coefficients-racines pour un polynôme complexe de degré n

n étant un entier naturel, on considère le *polynôme complexe de degré n* :

$$P(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = \sum_{k=0}^n a_k z^k$$

où $(a_0, a_1, \dots, a_n) \in \mathbb{C}^n \times \mathbb{C}^*$.

On suppose qu'il existe n nombres complexes z_1, z_2, \dots, z_n tels que :

$$P(z) = a_n (z - z_1)(z - z_2) \dots (z - z_n) = a_n \prod_{k=1}^n (z - z_k)$$

On appelle **fonctions symétriques des racines** :

$$\begin{cases} \sigma_1 = z_1 + z_2 + \dots + z_n = \sum_{i=1}^n z_i \\ \sigma_2 = z_1 z_2 + z_1 z_3 + \dots = \sum_{1 \leq i < j \leq n} z_i z_j \\ \sigma_3 = z_1 z_2 z_3 + z_1 z_2 z_4 + \dots = \sum_{1 \leq i < j < k \leq n} z_i z_j z_k \\ \vdots \\ \sigma_n = z_1 z_2 \dots z_n = \prod_{i=1}^n z_i \end{cases}$$

On dispose alors de relations entre les *coefficients du polynôme*, c'est-à-dire les a_i , $i = 0, \dots, n$, et les *fonctions symétriques* :

$$\forall k \in \{1, \dots, n\} : \sigma_k = (-1)^k \frac{a_{n-k}}{a_n}$$

Ces relations sont également appelées **relations de Viète**¹, et peuvent être extrêmement pratiques.

Démonstration : z_1, \dots, z_n étant les racines de $P : P(z) = a_n (z - z_1) \dots (z - z_n)$. En développant cette expression suivant les puissances décroissantes de z , on en déduit :

$$\begin{aligned} P(z) = & a_n z^n - a_n \left(\sum_{k=1}^n z_k \right) z^{n-1} + a_n \left(\sum_{1 \leq i < j \leq n} z_i z_j \right) z^{n-2} \\ & + \dots + (-1)^k a_n \left(\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \prod_{j=1}^k z_{i_j} \right) z^{n-k} + \dots + (-1)^n a_n z_1 \dots z_n \end{aligned}$$

Comme $P(z) = a_n z^n + \sum_{k=1}^n a_{n-k} z^{n-k}$, on en déduit, par identification, que, pour tout entier k de $\{1, \dots, n\}$:

$$a_{n-k} = (-1)^k a_n \left(\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} \prod_{j=1}^k z_{i_j} \right) \quad \blacksquare$$

ce qui conduit au résultat cherché.

1. François Viète (1540-1603), mathématicien, géomètre et astronome français, qui apporta de nombreuses contributions à l'algèbre ; il est à l'origine des prémices du calcul symbolique.

Fractions rationnelles et décomposition en éléments simples

La décomposition en « éléments simples », c'est-à-dire comme somme de fractions rationnelles avec, en dénominateur, des puissances de polynômes irréductibles, et, en numérateur, un polynôme de degré strictement inférieur à celui du polynôme irréductible qui intervient au dénominateur, est très utilisée en calcul intégral, pour déterminer des primitives, mais aussi en analyse spectrale, pour calculer des transformées de Laplace inverses (on trouvera, notamment, des exemples dans [44]).

1. Corps des fractions rationnelles

► À coefficients réels

On désigne par $\mathbb{R}(X)$ l'ensemble des fractions rationnelles à coefficients réels :

$$\forall R \in \mathbb{R}(X), \exists P \in \mathbb{R}[X] \text{ et } Q \neq 0 \in \mathbb{R}[X] : R(X) = \frac{P(X)}{Q(X)}$$

$\mathbb{R}(X)$ est le **corps des fractions rationnelles à coefficients réels**.

La notation « X » est liée au fait que l'on manipule, ici, des *polynômes formels*, qu'il faut bien distinguer des fonctions polynomiales.

► À coefficients complexes

On désigne par $\mathbb{C}(X)$ l'ensemble des fractions rationnelles à coefficients complexes :

$$\forall R \in \mathbb{C}(X), \exists P \in \mathbb{C}[X] \text{ et } Q \neq 0 \in \mathbb{C}[X] : R(X) = \frac{P(X)}{Q(X)}$$

$\mathbb{C}(X)$ est le **corps des fractions rationnelles à coefficients complexes**.



Attention, les polynômes P et Q ne sont en aucun cas uniques ! $2P$ et $2Q$ conviennent aussi, puisque :

$$\frac{P(X)}{Q(X)} = \frac{2P(X)}{2Q(X)}$$

Comme $\mathbb{R} \subset \mathbb{C}$, $\mathbb{R}(X) \subset \mathbb{C}(X)$: il est donc judicieux, dans un premier temps, de s'intéresser au cas des fractions rationnelles à coefficients complexes.

Proposition

Toute fraction rationnelle R , à coefficients dans le corps \mathbb{C} , et supposée non constante, peut s'écrire sous la forme

$$R(X) = \frac{P(X)}{Q(X)}$$

où P et Q sont deux polynômes complexes premiers entre eux :

$$P \wedge Q = 1$$

ce qui signifie qu'il n'existe pas de polynôme non constant divisant à la fois P et Q .

2. Pôle (d'une fraction rationnelle)

Définition

Soit R une fraction rationnelle, à coefficients dans le corps \mathbb{C} de la forme :

$$R(X) = \frac{P(X)}{Q(X)}$$

où P et Q sont deux polynômes complexes premiers entre eux.

On appelle **pôle** de R toute racine z_0 du polynôme Q , c'est-à-dire tout nombre complexe z_0 tel que :

$$Q(z_0) = 0$$

Exemple

-1 et 2 sont les pôles de la fraction rationnelle :

$$\frac{X^2 + 4}{(X + 1)(X - 2)}$$

► Pôle d'ordre p , $p \in \mathbb{N}^*$ (d'une fraction rationnelle)

Définition

Soit R une fraction rationnelle, à coefficients dans le corps \mathbb{C} de la forme

$$R(X) = \frac{P(X)}{Q(X)}$$

où P et Q sont deux polynômes complexes premiers entre eux.

On appelle **pôle d'ordre** p , $p \in \mathbb{N}^*$ de R , toute racine z_0 de multiplicité p du polynôme Q , c'est-à-dire tout nombre complexe z_0 tel qu'il existe un polynôme \tilde{Q} vérifiant :

$$Q(z) = (z - z_0)^p \tilde{Q}(z)$$

avec $\tilde{Q}(z_0) \neq 0$.

Exemple

3 est un pôle d'ordre 4 de la fraction rationnelle

$$\frac{X + 1}{(X - 3)^4 (X^2 + 1)}$$

3. Division euclidienne dans $\mathbb{C}[X]$

Étant donnés deux polynômes P et Q à coefficients dans \mathbb{C} , il existe un unique polynôme \tilde{P} , et un unique polynôme \tilde{R} , de degré strictement inférieur à celui de Q , tels que :

$$P(X) = \tilde{P}(X) Q(X) + \tilde{R}(X)$$

C'est la **division euclidienne de P par Q** . \tilde{P} est le quotient, et \tilde{R} le reste.



Ce résultat ne fait qu'étendre aux polynômes la notion de division euclidienne qui existe pour les réels. Le but est juste de simplifier le plus possible l'écriture d'une fraction rationnelle.

4. Décomposition d'une fraction rationnelle

► Décomposition en éléments simples dans $\mathbb{C}(X)$

Soit $R = \frac{P}{Q}$ une fraction rationnelle, à coefficients dans le corps \mathbb{C} , de pôles z_1, \dots, z_k , $k \geq 2$. On désigne par $n_1 \in \mathbb{N}, \dots, n_k \in \mathbb{N}$, les ordres respectifs de z_1, \dots, z_k .

La **division euclidienne** de P par le polynôme Q permet d'en déduire l'existence d'un unique polynôme \tilde{P} , appelé **partie entière** de R , et de complexes $\alpha_{z_i,j}$, $1 \leq i \leq k$, $1 \leq j \leq n_i$, tels que :

$$R(X) = \tilde{P}(X) + \frac{\tilde{R}(X)}{Q(X)} \quad \text{avec} \quad \deg \tilde{R} \leq \deg Q - 1$$

Comme $Q(X)$ est, à un facteur multiplicatif près, proportionnel à $\prod_{i=1}^k (X - z_i)^{n_i}$, cette relation peut aussi s'écrire sous la forme :

$$R(X) = \tilde{P}(X) + \frac{\tilde{\tilde{R}}(X)}{\prod_{i=1}^k (X - z_i)^{n_i}} \quad \text{avec} \quad \deg \tilde{\tilde{R}} \leq n_1 + \dots + n_k - 1$$

ou encore :

$$\begin{aligned} R(X) &= \tilde{P}(X) + \sum_{i=1}^k \left\{ \frac{\alpha_{z_i,1}}{X - z_i} + \frac{\alpha_{z_i,2}}{(X - z_i)^2} + \dots + \frac{\alpha_{z_i,k}}{(X - z_i)^{n_i}} \right\} \\ &= \tilde{P}(X) + \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{\alpha_{z_i,j}}{(X - z_i)^j} \end{aligned}$$

Pour tout entier i de $\{1, \dots, k\}$, $\sum_{j=1}^{n_i} \frac{\alpha_{z_i,j}}{(X - z_i)^j}$ est la **partie polaire de R associée au pôle z_i** .



Si la fraction rationnelle est à coefficients réels, et admet des pôles réels, alors sa décomposition en éléments simples est à coefficients réels.

Proposition

Soit R une fraction rationnelle, à coefficients dans le corps \mathbb{C} , de pôles z_1, \dots, z_k , $k \in \mathbb{N}^*$, qui se décompose sous la forme :

$$R(X) = \tilde{P}(X) + \frac{\tilde{R}(X)}{Q(X)}$$

où \tilde{P} , Q et \tilde{R} sont trois polynômes à coefficients complexes tels que $\deg \tilde{R} < \deg Q$.

S'il existe un pôle z_0 de multiplicité 1, la décomposition de R peut s'écrire sous la forme :

$$R(X) = \tilde{P}(X) + \frac{\tilde{R}(X)}{Q_0(X)(X - z_0)}$$

où Q_0 est un polynôme à coefficients complexes, de degré $\deg Q - 1$, et tel que :

$$Q_0(z_0) \neq 0$$

Alors, la partie polaire de R associée au pôle z_0 est $\frac{P(z_0)}{Q_0(z_0)(X - z_0)}$, ce qui signifie donc que le coefficient de $\frac{1}{X - z_0}$ dans la décomposition en éléments simples de R est $\frac{P(z_0)}{Q_0(z_0)}$.

Démonstration : La partie polaire de R associée au pôle z_0 est de la forme :

$$\frac{\alpha}{X - z_0}, \quad \alpha \in \mathbb{C}$$

Par suite :

$$R(X) = \tilde{P}(X) + \frac{\tilde{R}(X)}{Q_0(X)(X - z_0)} = \tilde{P}(X) + \frac{\alpha}{X - z_0} + \frac{P_1(X)}{Q_1(X)}$$

où P_1 et Q_1 sont deux polynômes complexes tels que $\deg P_1 < \deg Q_1$ et $Q_1(z_0) \neq 0$.

Il suffit alors de multiplier membre à membre par $X - z_0$:

$$(X - z_0)\tilde{P}(X) + \frac{\tilde{R}(X)}{Q_0(X)} = (X - z_0)\tilde{P}(X) + \alpha + (X - z_0)\frac{P_1(X)}{Q_1(X)}$$

puis d'évaluer l'expression obtenue en z_0 , ce qui conduit immédiatement à :

$$\frac{P(z_0)}{Q_0(z_0)} = \alpha$$

■

Exemple

Décomposons en éléments simples la fraction rationnelle :

$$R(X) = \frac{3X + 4}{X^2 - 1} = \frac{3X + 4}{(X - 1)(X + 1)}$$

1 et -1 sont des pôles simples. La partie entière de R est nulle, puisque :

$$\deg(3X + 4) < \deg(X^2 - 1)$$

R étant à coefficients dans \mathbb{R} , et ses pôles étant réels, on cherche donc des réels α_{-1} et α_1 tels que :

$$R(X) = \frac{\alpha_{-1}}{X + 1} + \frac{\alpha_1}{X - 1}$$

On aura donc :

$$\frac{3X + 4}{(X - 1)(X + 1)} = \frac{\alpha_{-1}}{X + 1} + \frac{\alpha_1}{X - 1}$$

En multipliant membre à membre (a) par $X + 1$, on en déduit :

$$\frac{3X + 4}{X - 1} = \alpha_{-1} + (X + 1)\frac{\alpha_1}{X - 1}$$

En évaluant cette dernière expression en -1 , on obtient :

$$\frac{-3+4}{-2} = \frac{-1}{2} = \alpha_{-1}$$

De même, en multipliant membre à membre (a) par $X-1$, on en déduit :

$$\frac{3X+4}{X+1} = (X-1) \frac{\alpha_{-1}}{X+1} + \alpha_1$$

En évaluant cette dernière expression en 1 , on obtient :

$$\frac{3+4}{2} = \frac{7}{2} = \alpha_1$$

La décomposition en éléments simples de R est donc :

$$R(X) = -\frac{1}{2} \frac{1}{X+1} + \frac{7}{2} \frac{1}{X-1}$$

► **Détermination de la partie polaire d'une fraction rationnelle pour un pôle d'ordre $p \in \mathbb{N}$, $p \geq 2$**

Proposition

Soit R une fraction rationnelle, à coefficients dans le corps \mathbb{C} , de la forme :

$$R(X) = \frac{P(X)}{Q(X)} = \frac{P(X)}{Q_0(X)(X-z_0)^p}$$

avec $Q_0(z_0) \neq 0$. On désigne par :

$$\frac{\alpha_j}{(X-z_0)^j}$$

la partie polaire de R associée au pôle z_0 .

Alors, le coefficient de $\frac{1}{(X-z_0)^p}$ dans la décomposition en éléments simples de R est :

$$\alpha_p = p! \frac{P(z_0)}{Q^{(p)}(z_0)}$$

Les autres termes peuvent être obtenus en effectuant un développement limité de

$$(X-z_0)^p R(X) = \alpha_1 (X-z_0)^{p-1} + \alpha_2 (X-z_0)^{p-2} + \dots + \alpha_{p-1} (X-z_0) + \alpha_p + (X-z_0)^p R_1(X)$$

où R_1 est une fraction rationnelle n'admettant pas z_0 pour pôle, au voisinage de z_0 .

Ce développement limité étant de la forme :

$$\alpha_1 (X-z_0)^{p-1} + \alpha_2 (X-z_0)^{p-2} + \dots + \alpha_{p-1} (X-z_0) + \alpha_p + o((X-z_0)^p)$$

on en déduit facilement, par unicité du développement limité, les valeurs des coefficients $\alpha_1, \dots, \alpha_p$.

Démonstration : Notons α_p le coefficient de $\frac{1}{(X - z_0)^p}$ dans la décomposition en éléments simples de R :

$$\alpha_p = [(X - z_0)^p R(X)]_{|X=z_0} = \frac{P(z_0)}{Q_0(z_0)}$$

La formule de Leibniz ¹ permet alors d'écrire :

$$\begin{aligned} Q^{(p)}(z_0) &= [(X - z_0)^p Q_0(X)]_{|X=z_0}^{(p)} = \left[\sum_{l=0}^p C_p^l ((X - z_0)^p)^{(l)} Q_0^{(p-l)}(X) \right]_{|X=z_0} \\ &= \left[C_p^p ((X - z_0)^p)^{(p)} Q_0(X) \right]_{|X=z_0} \\ &= p! Q_0(z_0) \end{aligned}$$

Il en résulte :

$$\alpha_p = \frac{P(z_0)}{Q_0(z_0)} = p! \frac{P(z_0)}{Q^{(p)}(z_0)} \quad \blacksquare$$

Les autres résultats sont admis.

Exemple

Décomposons dans $\mathbb{R}(X)$: $R(X) = \frac{X^2 + 2}{(X - 1)^3}$.

La décomposition de R en éléments simples réels est de la forme :

$$R(X) = \frac{\alpha_{11}}{X - 1} + \frac{\alpha_{12}}{(X - 1)^2} + \frac{\alpha_{13}}{(X - 1)^3}$$

Le développement limité à l'ordre 3 de $(X - 1)^3 R(X) = X^2 + 2$ au voisinage de 1 est :

$$\begin{aligned} & \left[(X^2 + 2)^{(0)} \right]_{X=1} + \left[(X^2 + 2)' \right]_{X=1} (X - 1) \\ & + \frac{\left[(X^2 + 2)^{(2)} \right]_{X=1}}{2!} (X - 1)^2 + \frac{\left[(X^2 + 2)^{(3)} \right]_{X=1}}{3!} (X - 1)^3 + o((X - 1)^3) \end{aligned}$$

c'est-à-dire :

$$3 + 2(X - 1) + (X - 1)^2 + o((X - 1)^3)$$

Comme :

$$(X - 1)^3 R(X) = \alpha_{11} (X - 1)^2 + \alpha_{12} (X - 1) + \alpha_{13}$$

on en déduit, par unicité du développement limité :

$$\alpha_{11} = 1, \quad \alpha_{12} = 2, \quad \alpha_{13} = 3$$

Ainsi :

$$R(X) = \frac{1}{X - 1} + \frac{2}{(X - 1)^2} + \frac{3}{(X - 1)^3}$$

1. Étant donné un entier naturel n , et deux fonctions f et g définies sur un même intervalle I de \mathbb{R} , n fois dérivables sur I , la fonction produit $f g$ est n fois dérivable sur I , et :

$$(f g)^{(n)} = \sum_{k=0}^n C_n^k f^{(k)} g^{(n-k)}$$

où, pour tout entier k de $\{0, \dots, n\}$, C_n^k désigne le coefficient binomial $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

► Décomposition en éléments simples dans $\mathbb{R}(X)$

Proposition

Soit $R = \frac{P}{Q}$ une fraction rationnelle, à coefficients dans le corps \mathbb{R} , où P et Q sont deux polynômes premiers entre eux, et où, en outre, Q est unitaire (c'est-à-dire le coefficient de son terme de plus haut degré vaut 1).

On désigne par $x_1, \dots, x_k, k \in \mathbb{N}$ les pôles de R , dont on note $n_1 \in \mathbb{N}, \dots, n_k \in \mathbb{N}$, les ordres respectifs.

Alors, Q peut être factorisé en produit de polynômes irréductibles sur \mathbb{R} , sous la forme :

$$Q(X) = \prod_{i=1}^k (X - x_i)^{n_i} \prod_{i=1}^r (X^2 + \lambda_i X + \mu_i)^{\ell_i}, \quad r \in \mathbb{N}, (\lambda_i, \mu_i) \in \mathbb{R}^2, \ell_i \in \mathbb{N}$$

et il existe un unique polynôme \tilde{R} , et des réels $\alpha_{x_i, j}, (i, j) \in \{1, \dots, k\} \times \{1, \dots, n_k\}, \beta_{ij}$ et $\gamma_{ij}, (i, j) \in \{1, \dots, r\} \times \{1, \dots, \ell_i\}$, tels que :

$$\begin{aligned} R(X) &= \tilde{R}(X) + \sum_{i=1}^k \left\{ \frac{\alpha_{x_i, 1}}{X - x_i} + \frac{\alpha_{x_i, 2}}{(X - x_i)^2} + \dots + \frac{\alpha_{x_i, k}}{(X - x_i)^{n_i}} \right\} \\ &+ \sum_{i=1}^r \left\{ \frac{\beta_{i, 1} X + \gamma_{i, 1}}{X^2 + \lambda_{i, 1} X + \mu_{i, 1}} + \dots + \frac{\beta_{i, \ell_i} X + \gamma_{i, \ell_i}}{(X^2 + \lambda_{i, \ell_i} X + \mu_{i, \ell_i})^{\ell_i}} \right\} \\ &= \tilde{R}(X) + \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{\alpha_{x_i, j}}{(X - x_i)^j} + \sum_{i=1}^r \sum_{j=1}^{\ell_i} \frac{\beta_{i, j} X + \gamma_{i, j}}{(X^2 + \lambda_{i, j} X + \mu_{i, j})^j} \end{aligned}$$



Les coefficients β_{ij} et $\gamma_{ij}, (i, j) \in \{1, \dots, r\} \times \{1, \dots, \ell_i\}$, pour les termes de la forme $\frac{\beta_{i, j} X + \gamma_{i, j}}{(X^2 + \lambda_{i, j} X + \mu_{i, j})^j}$, peuvent être obtenus en commençant par décomposer la fraction R dans $\mathbb{C}(X)$, puis en regroupant les termes deux à deux conjugués, les racines complexes du trinôme $X^2 + \lambda_{ij} X + \mu_{ij}$ étant deux à deux conjuguées.

Exemple

Décomposons dans $\mathbb{R}(X)$: $R(X) = \frac{X - 1}{X(X^2 + X + 1)}$.

Le seul pôle réel est 0 ; R admettant deux pôles complexes conjugués, $j = e^{\frac{2i\pi}{3}}$, et $j^2 = e^{\frac{4i\pi}{3}}$, on commence par décomposer R sur $\mathbb{C}(X)$.

La décomposition de R en éléments simples complexes est de la forme :

$$R(X) = \frac{\alpha_0}{X} + \frac{\alpha_j}{X - j} + \frac{\alpha_{j^2}}{X - j^2} \quad (\star)$$

Pour déterminer α_0 , on multiplie membre à membre (\star) par X , ce qui conduit à :

$$\frac{X - 1}{X^2 + X + 1} = \alpha_0 + X \frac{\alpha_j}{X - j} + X \frac{\alpha_{j^2}}{X - j^2}$$

En évaluant cette dernière expression en 0, on obtient :

$$-1 = \alpha_0$$

Pour déterminer α_j , on multiplie membre à membre (★) par $X - j$, ce qui conduit à :

$$\frac{X-1}{X(X-j^2)} = (X-j) \frac{\alpha_0}{X} + \alpha_j + (X-j) \frac{\alpha_{j^2}}{X-j^2}$$

En évaluant cette dernière expression en j , on obtient :

$$\frac{j-1}{j(j-j^2)} = \alpha_j$$

soit :

$$\frac{j-1}{j^2-1} = \frac{j-1}{(j-1)(j+1)} = \frac{1}{j+1} = -\frac{1}{j^2} = -\frac{j}{j^3} = -j = \alpha_j$$

puisque $j^3 = 1$, et $1 + j + j^2 = 0$.

On détermine de même α_{j^2} en multipliant membre à membre (★) par $X - j^2$, ce qui conduit à :

$$\frac{X-1}{X(X-j)} = (X-j^2) \frac{\alpha_0}{X} + (X-j^2) \frac{\alpha_j}{X-j} + \alpha_{j^2}$$

En évaluant cette dernière expression en j^2 , on obtient :

$$\frac{j^2-1}{j^2(j^2-j)} = \alpha_{j^2}$$

soit :

$$\frac{j^2-1}{j-1} = \frac{(j-1)(j+1)}{j-1} = j+1 = -j^2 = \overline{\alpha_j} = \alpha_{j^2}$$

puisque $j^3 = 1$, et $1 + j + j^2 = 0$. Ainsi :

$$\begin{aligned} R(X) &= -\frac{1}{X} \frac{j}{X-j} - \frac{j^2}{X-j^2} \\ &= -\frac{1}{X} - \frac{j(X-j^2) + j^2(X-j)}{(X-j)(X-j^2)} \\ &= -\frac{1}{X} - \frac{(j+j^2)X - 2j^3}{(X-j)(X-j^2)} \\ &= -\frac{1}{X} - \frac{X^2 + X + 1}{X^2 + X + 1} \\ &= -\frac{1}{X} + \frac{1}{X^2 + X + 1} \end{aligned}$$

► Décomposition en éléments simples sur $\mathbb{R}(X)$ d'une fraction rationnelle paire

Soit R une fraction rationnelle, à coefficients dans le corps \mathbb{R} , paire, c'est-à-dire telle que

$$R(-X) = R(X)$$

Soit

$$\begin{aligned} R(X) &= \tilde{R}(X) + \sum_{i=1}^k \left\{ \frac{\alpha_{x_i,1}}{X-x_i} + \frac{\alpha_{x_i,2}}{(X-x_i)^2} + \dots + \frac{\alpha_{x_i,k}}{(X-x_i)^{n_i}} \right\} \\ &\quad + \sum_{i=1}^r \left\{ \frac{\beta_{i,1}X + \gamma_{i1}}{X^2 + \lambda_{i1}X + \mu_{i1}} + \dots + \frac{\beta_{i,\ell_i}X + \gamma_{i,\ell_i}}{(X^2 + \lambda_{i,\ell_i}X + \mu_{i,\ell_i})^{\ell_i}} \right\} \\ &= \tilde{R}(X) + \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{\alpha_{x_i,j}}{(X-x_i)^j} + \sum_{i=1}^r \sum_{j=1}^{\ell_i} \frac{\beta_{i,j}X + \gamma_{ij}}{(X^2 + \lambda_{ij}X + \mu_{ij})^j} \end{aligned}$$

la décomposition en éléments simples de R dans $\mathbb{R}(X)$.

Alors :

$$\begin{aligned} & \tilde{R}(X) + \sum_{i=1}^k \left\{ \frac{\alpha_{x_i,1}}{X - x_i} + \frac{\alpha_{x_i,2}}{(X - x_i)^2} + \dots + \frac{\alpha_{x_i,k}}{(X - x_i)^{n_i}} \right\} \\ & + \sum_{i=1}^r \left\{ \frac{\beta_{i,1} X + \gamma_{i,1}}{X^2 + \lambda_{i,1} X + \mu_{i,1}} + \dots + \frac{\beta_{i,\ell_i} X + \gamma_{i,\ell_i}}{(X^2 + \lambda_{i,\ell_i} X + \mu_{i,\ell_i})^{\ell_i}} \right\} = \\ & \tilde{R}(-X) + \sum_{i=1}^k \left\{ \frac{\alpha_{x_i,1}}{-X - x_i} + \frac{\alpha_{x_i,2}}{(-X - x_i)^2} + \dots + \frac{\alpha_{x_i,k}}{(-X - x_i)^{n_i}} \right\} \\ & + \sum_{i=1}^r \left\{ \frac{\beta_{i,1} X + \gamma_{i,1}}{X^2 - \lambda_{i,1} X + \mu_{i,1}} + \dots + \frac{\beta_{i,\ell_i} X + \gamma_{i,\ell_i}}{(X^2 - \lambda_{i,\ell_i} X + \mu_{i,\ell_i})^{\ell_i}} \right\} \end{aligned}$$

ce qui permet d'obtenir des relations entre les coefficients $\alpha_{x_i,j}$, $(i, j) \in \{1, \dots, k\} \times \{1, \dots, n_k\}$, d'une part, $\beta_{i,j}$ et $\gamma_{i,j}$, $(i, j) \in \{1, \dots, r\} \times \{1, \dots, \ell_i\}$ d'autre part, et donc de simplifier le calcul de ceux-ci.

Exemple

Décomposons dans $\mathbb{R}(X)$: $R(X) = \frac{X^2}{X^4 - 1}$.

Comme $X^4 - 1 = (X^2 - 1)(X^2 + 1) = (X - 1)(X + 1)(X^2 + 1)$, la décomposition de R en éléments simples (réels) est de la forme :

$$R(X) = \frac{\alpha_1}{X - 1} + \frac{\alpha_{-1}}{X + 1} + \frac{\beta X + \gamma}{X^2 + 1}$$

Par parité de R , on en déduit :

$$\frac{\alpha_1}{X - 1} + \frac{\alpha_{-1}}{X + 1} + \frac{\beta X + \gamma}{X^2 + 1} = \frac{\alpha_1}{-X - 1} + \frac{\alpha_{-1}}{-X + 1} + \frac{-\beta X + \gamma}{X^2 + 1}$$

Il en résulte : $\alpha_1 = -\alpha_{-1}$, $\beta = -\beta$, ce qui conduit immédiatement à $\beta = 0$.

Pour déterminer α_1 , on multiplie membre à membre (\star) par $X - 1$, ce qui conduit à :

$$\frac{X^2}{(X + 1)(X^2 + 1)} = \alpha_1 + \frac{\alpha_{-1}(X - 1)}{X + 1} + \frac{\gamma(X - 1)}{X^2 + 1}$$

En évaluant cette dernière expression en 1, on obtient :

$$\frac{1}{4} = \alpha_1$$

On peut alors en déduire $\alpha_{-1} = -\alpha_1 = -\frac{1}{4}$.

La valeur de γ peut être obtenue, par exemple, en évaluant la valeur de $R(X)$ en 0 :

$$R(0) = 0 = -\alpha_1 + \alpha_{-1} + \gamma$$

ce qui conduit à :

$$\gamma = \alpha_1 - \alpha_{-1} = \frac{1}{2}$$

La décomposition en éléments simples cherchée est donc :

$$\frac{X^2}{X^4 - 1} = \frac{1}{4(X - 1)} - \frac{1}{4(X + 1)} + \frac{1}{2(X^2 + 1)}$$

► **Décomposition en éléments simples sur $\mathbb{R}(X)$ d'une fraction rationnelle impaire**

Soit R une fraction rationnelle, à coefficients dans le corps \mathbb{R} , impaire, c'est-à-dire telle que :

$$R(-X) = -R(X)$$

Soit alors :

$$\begin{aligned} R(X) &= \tilde{R}(X) + \sum_{i=1}^k \left\{ \frac{\alpha_{x_i,1}}{X - x_i} + \frac{\alpha_{x_i,2}}{(X - x_i)^2} + \dots + \frac{\alpha_{x_i,k}}{(X - x_i)^{n_i}} \right\} \\ &\quad + \sum_{i=1}^r \left\{ \frac{\beta_{i,1} X + \gamma_{i1}}{X^2 + \lambda_{i1} X + \mu_{i1}} + \dots + \frac{\beta_{i,\ell_i} X + \gamma_{i,\ell_i}}{(X^2 + \lambda_{i,\ell_i} X + \mu_{i,\ell_i})^{\ell_i}} \right\} \\ &= \tilde{R}(X) + \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{\alpha_{x_i,j}}{(X - x_i)^j} + \sum_{i=1}^r \sum_{j=1}^{\ell_i} \frac{\beta_{i,j} X + \gamma_{ij}}{(X^2 + \lambda_{ij} X + \mu_{ij})^j} \end{aligned}$$

la décomposition en éléments simples de R dans $\mathbb{R}(X)$.

Alors :

$$\begin{aligned} & -\tilde{R}(X) - \sum_{i=1}^k \left\{ \frac{\alpha_{x_i,1}}{X - x_i} + \frac{\alpha_{x_i,2}}{(X - x_i)^2} + \dots + \frac{\alpha_{x_i,k}}{(X - x_i)^{n_i}} \right\} \\ & - \sum_{i=1}^r \left\{ \frac{\beta_{i,1} X + \gamma_{i1}}{X^2 + \lambda_{i1} X + \mu_{i1}} + \dots + \frac{\beta_{i,\ell_i} X + \gamma_{i,\ell_i}}{(X^2 + \lambda_{i,\ell_i} X + \mu_{i,\ell_i})^{\ell_i}} \right\} = \\ & \tilde{R}(-X) + \sum_{i=1}^k \left\{ \frac{\alpha_{x_i,1}}{-X - x_i} + \frac{\alpha_{x_i,2}}{(-X - x_i)^2} + \dots + \frac{\alpha_{x_i,k}}{(-X - x_i)^{n_i}} \right\} \\ & + \sum_{i=1}^r \left\{ \frac{\beta_{i,1} X + \gamma_{i1}}{X^2 - \lambda_{i1} X + \mu_{i1}} + \dots + \frac{\beta_{i,\ell_i} X + \gamma_{i,\ell_i}}{(X^2 - \lambda_{i,\ell_i} X + \mu_{i,\ell_i})^{\ell_i}} \right\} \end{aligned}$$

ce qui permet d'obtenir des relations entre les coefficients $\alpha_{x_i,j}$, $(i, j) \in \{1, \dots, k\} \times \{1, \dots, n_k\}$, d'une part, β_{ij} et γ_{ij} , $(i, j) \in \{1, \dots, r\} \times \{1, \dots, \ell_i\}$ d'autre part, et donc de simplifier le calcul de ceux-ci.

Exemple

Décomposons dans $\mathbb{R}(X)$: $R(X) = \frac{X}{X^4 - 1}$.

Comme $X^4 - 1 = (X^2 - 1)(X^2 + 1) = (X - 1)(X + 1)(X^2 + 1)$, la décomposition de R en éléments simples (réels) est de la forme :

$$R(X) = \frac{\alpha_1}{X - 1} + \frac{\alpha_{-1}}{X + 1} + \frac{\beta X + \gamma}{X^2 + 1} \quad (\star)$$

Par imparité de R , on en déduit :

$$\frac{\alpha_1}{X - 1} + \frac{\alpha_{-1}}{X + 1} + \frac{\beta X + \gamma}{X^2 + 1} = -\frac{\alpha_1}{-X - 1} - \frac{\alpha_{-1}}{-X + 1} - \frac{-\beta X + \gamma}{X^2 + 1}$$

Il en résulte :

$$\alpha_1 = \alpha_{-1} \quad , \quad \gamma = -\gamma$$

ce qui conduit immédiatement à $\gamma = 0$.

Pour déterminer α_1 , on multiplie membre à membre (★) par $X - 1$, ce qui conduit à :

$$\frac{X}{(X+1)(X^2+1)} = \alpha_1 + \frac{\alpha_{-1}(X-1)}{X+1} + \frac{\beta X(X-1)}{X^2+1}$$

En évaluant cette dernière expression en 1, on obtient :

$$\frac{1}{4} = \alpha_1$$

On peut alors en déduire $\alpha_{-1} = \alpha_1 = \frac{1}{4}$.

La valeur de β peut être obtenue, par exemple, en multipliant membre à membre (★) par X , puis en passant à la limite lorsque $X \rightarrow +\infty$:

$$\lim_{X \rightarrow +\infty} X \frac{X}{X^4 - 1} = \lim_{X \rightarrow +\infty} X \left\{ \frac{\alpha_1}{X-1} + \frac{\alpha_{-1}}{X+1} + \frac{\beta X + \gamma}{X^2+1} \right\}$$

soit :

$$0 = \alpha_1 + \alpha_{-1} + \beta$$

ce qui conduit à :

$$\beta = -\alpha_1 - \alpha_{-1} = -\frac{1}{2}$$

La décomposition en éléments simples cherchée est donc :

$$\frac{X}{X^4 - 1} = \frac{1}{4(X-1)} + \frac{1}{4(X+1)} - \frac{X}{2(X^2+1)}$$

Transformations du plan : translations, homothéties

1. Translations

Définition

\vec{u} étant un vecteur du plan, on appelle **translation de vecteur \vec{u}** l'application $\tau_{\vec{u}}$, qui, à tout point M du plan associe le point M' tel que :

$$\overrightarrow{MM'} = \vec{u}.$$

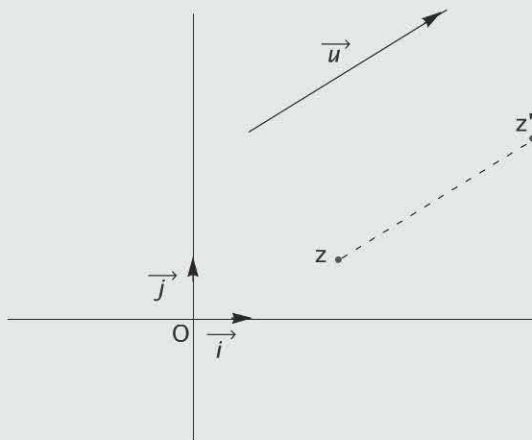


Figure 70.1 – La translation de vecteur \vec{u} .

► Expression analytique d'une translation de vecteur \vec{u}

Le point M' , image du point M d'affixe z , par la translation de vecteur \vec{u} , a pour affixe :

$$z' = z + z_{\vec{u}}$$

où $z_{\vec{u}}$ est l'affixe de \vec{u} .

Réciproquement, une application de la forme $z \in \mathbb{C} \mapsto z + z_{\vec{u}}$ est la translation de vecteur \vec{u} .

Démonstration : La relation : $\overrightarrow{MM'} = \vec{u}$ entraîne : $z_{\overrightarrow{MM'}} = z_{\vec{u}}$, soit : $z' - z = z_{\vec{u}}$ et donc : $z' = z + z_{\vec{u}}$

La démonstration de la réciprocity est laissée au lecteur. ■

2. Homothéties

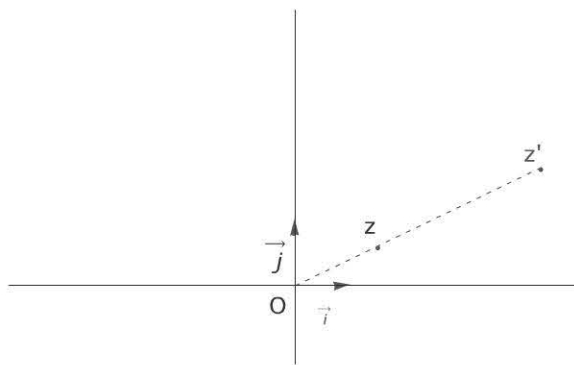
Définition

k étant un réel non nul, on appelle **homothétie de centre O , de rapport k** , l'application $h_{O,k}$, qui, à tout point M du plan, associe le point M' tel que : $\overrightarrow{OM'} = k \overrightarrow{OM}$.

► Expression analytique d'une homothétie de centre O , de rapport k

k étant un réel non nul, le point M' , image du point M d'affixe z , par l'homothétie de centre O , de rapport k , a pour affixe : $z' = kz$.

Réciproquement, une application de la forme $z \in \mathbb{C} \mapsto kz$ est l'homothétie de centre O , de rapport k .

Figure 70.2 – L'homothétie de centre O , de rapport k .

Démonstration : On traduit tout simplement, en termes d'affixes, le fait que :

$$\overrightarrow{OM'} = k \overrightarrow{OM}$$

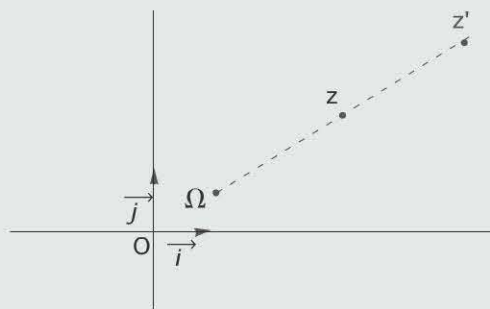
ce qui conduit à : $z' = kz$

La démonstration de la réciprocity est laissée au lecteur. ■

Définition

k étant un réel non nul, on appelle *homothétie de centre Ω , de rapport k* , l'application $h_{\Omega,k}$ qui, à tout point M du plan associe le point M' tel que :

$$\overrightarrow{\Omega M'} = k \overrightarrow{\Omega M}.$$

Figure 70.3 – L'homothétie de centre Ω , de rapport k .

► Expression analytique d'une homothétie de centre Ω , de rapport k

k étant un réel non nul, le point M' , image du point M d'affixe z , par l'homothétie de centre Ω , de rapport k , a pour affixe : $z' = z_{\Omega} + k(z - z_{\Omega})$ où z_{Ω} est l'affixe de Ω .

Réciproquement, une application de la forme $z \in \mathbb{C} \mapsto z_{\Omega} + k(z - z_{\Omega})$ est l'homothétie de centre Ω , de rapport k .

Démonstration : On traduit tout simplement, en termes d'affixes, le fait que :

$$\overrightarrow{\Omega M'} = k \overrightarrow{\Omega M}$$

soit :

$$z' - z_{\Omega} = k(z - z_{\Omega})$$

ce qui conduit au résultat cherché. La démonstration de la réciprocity est laissée au lecteur. ■

Transformations du plan : rotations

1. Rotation de centre O

Définition

α étant un réel, on appelle **rotation de centre O , d'angle α** , l'application $r_{O,\alpha}$ qui, à tout point M du plan distinct de O , associe le point M' tel que :

$$\begin{cases} OM' = OM \\ (\widehat{\overrightarrow{OM}, \overrightarrow{OM'}}) = \alpha \quad [2\pi] \end{cases}$$

en laissant le point O invariant.

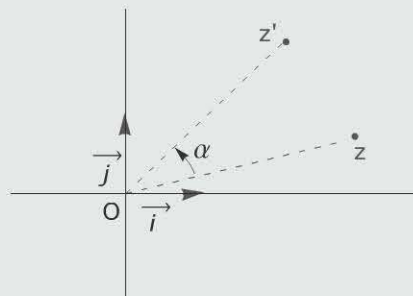


Figure 71.1 – La rotation de centre O , d'angle α .

► Expression analytique d'une rotation de centre O

Le point M' , image du point M d'affixe z , par la rotation de centre O , d'angle α , a pour affixe : $z' = z e^{i\alpha}$.

Réciproquement, une application de la forme $z \in \mathbb{C} \mapsto z e^{i\alpha}$ est la rotation de centre O , d'angle α .

Démonstration : Si M est distinct de O , la relation $(\widehat{\overrightarrow{OM}, \overrightarrow{OM'}}) = \alpha \quad [2\pi]$ entraîne :

$$\begin{aligned} \arg(z_{\overrightarrow{OM'}}) &= (\vec{i}, \overrightarrow{OM}) + (\widehat{\overrightarrow{OM}, \overrightarrow{OM'}}) \quad [2\pi] \\ &= \arg(z_{\overrightarrow{OM}}) + \alpha \quad [2\pi] \end{aligned}$$

$$\text{Comme : } z_{\overrightarrow{OM'}} = |z_{\overrightarrow{OM'}}| e^{i \arg(z_{\overrightarrow{OM'}})} = |z_{\overrightarrow{OM}}| e^{i(\arg(z_{\overrightarrow{OM}}) + \alpha)} = |z_{\overrightarrow{OM}}| e^{i \arg(z_{\overrightarrow{OM}})} e^{i\alpha}$$

$$\text{alors : } z_{\overrightarrow{OM'}} = z_{\overrightarrow{OM}} e^{i\alpha}, \text{ soit : } z' = z e^{i\alpha}$$

On remarque que la relation reste encore vraie lorsque M coïncide avec O .

La démonstration de la réciprocity est laissée au lecteur. ■

2. Rotation de centre Ω

Définition

θ étant un réel, on appelle **rotation de centre Ω , d'angle θ** , l'application $r_{\Omega,\theta}$, qui, à tout point M du plan, distinct de Ω , associe le point M' tel que :

$$\begin{cases} \Omega M' = \Omega M \\ (\widehat{\overrightarrow{\Omega M}, \overrightarrow{\Omega M'}}) = \theta \quad [2\pi] \end{cases}$$

en laissant le point Ω invariant.

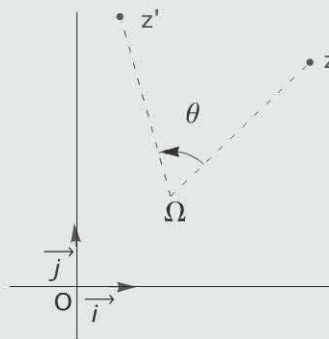


Figure 71.2 – La rotation de centre Ω , d'angle θ .

► Expression analytique d'une rotation de centre Ω

Le point M' , image du point M d'affixe z , par la rotation de centre Ω , d'angle θ , a pour affixe : $z' = z_\Omega + (z - z_\Omega) e^{i\theta}$ où z_Ω est l'affixe de Ω .

Réciproquement, une application de la forme $z \in \mathbb{C} \mapsto z_\Omega + (z - z_\Omega) e^{i\theta}$ est la rotation de centre Ω , d'angle θ .

Démonstration : Si M est distinct de Ω , la relation $(\overrightarrow{\Omega M}, \overrightarrow{\Omega M'}) = \theta \quad [2\pi]$ entraîne :

$$\begin{aligned} \arg(z_{\overrightarrow{\Omega M'}}) &= (\vec{i}, \overrightarrow{\Omega M}) + (\overrightarrow{\Omega M}, \overrightarrow{\Omega M'}) \quad [2\pi] \\ &= \arg(z_{\overrightarrow{\Omega M}}) + \theta \quad [2\pi] \end{aligned}$$

$$\text{Comme : } z_{\overrightarrow{\Omega M'}} = |z_{\overrightarrow{\Omega M'}}| e^{i \arg(z_{\overrightarrow{\Omega M'}})} = |z_{\overrightarrow{\Omega M}}| e^{i(\arg(z_{\overrightarrow{\Omega M}}) + \theta)} = |z_{\overrightarrow{\Omega M}}| e^{i \arg(z_{\overrightarrow{\Omega M}})} e^{i\theta}$$

$$\text{alors : } z_{\overrightarrow{\Omega M'}} = z_{\overrightarrow{\Omega M}} e^{i\theta}, \text{ soit : } z' - z_\Omega = (z - z_\Omega) e^{i\theta} \text{ et donc : } z' = z_\Omega + (z - z_\Omega) e^{i\theta}.$$

On remarque que la relation reste encore vraie lorsque M coïncide avec Ω .

La démonstration de la réciprocité est laissée au lecteur. ■

3. Composée de deux rotations

La composée de deux rotations de même centre Ω , d'angles respectifs α et β , est une rotation de centre Ω , d'angle $\alpha + \beta$.

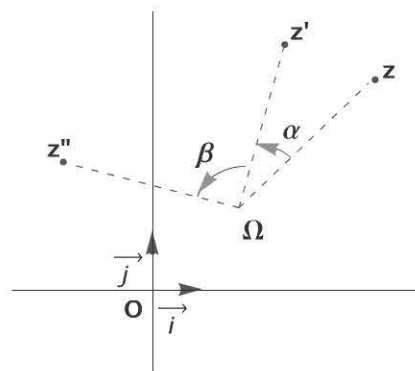


Figure 71.3 – La composée de deux rotations de même centre Ω , d'angles respectifs α et β .

Démonstration : Soit z' l'affixe du point M' , image du point M d'affixe $z \in \mathbb{C}$ par la rotation $r_{\Omega, \alpha}$, de centre Ω , d'angle α . On désigne par z_Ω l'affixe de Ω .

D'après ce qui précède :

$$z' = z_\Omega + (z - z_\Omega) e^{i\alpha}$$

Soit z'' l'affixe du point M'' , image du point M' d'affixe z' par la rotation $r_{\Omega, \beta}$, de centre Ω , d'angle β .

De même que précédemment, on montre que : $z'' = z_\Omega + (z' - z_\Omega) e^{i\beta}$.

Par suite :

$$z'' = z_\Omega + (z' - z_\Omega) e^{i\beta} = z_\Omega + (z_\Omega + (z - z_\Omega) e^{i\alpha} - z_\Omega) e^{i\beta}$$

$$\text{soit : } z'' = z_\Omega + (z - z_\Omega) e^{i(\alpha + \beta)}.$$

D'où le résultat, qui était naturellement prévisible, comme on peut le constater sur le dessin ! ■

Transformations du plan : similitudes

Et si on composait une rotation et une homothétie ? On fait tourner, et on agrandit ou on rétrécit ... ou l'inverse !

En géométrie euclidienne, une similitude est ainsi une transformation qui multiplie toutes les distances par une constante fixe (son rapport), et conserve ou inverse le sens des angles orientés. Ainsi, l'image de toute figure par une telle application est une figure « similaire », ou « de même forme ».

1. Similitudes directes

Considérons, dans un premier temps, l'homothétie de centre Ω , de rapport k ; si on désigne par z_Ω l'affixe du point Ω , le point M' , image du point M d'affixe z , a pour affixe : $z' = z_\Omega + k(z - z_\Omega)$.

Considérons, ensuite, la rotation de centre Ω , d'angle θ ; le point M'' , image du point M' , a pour affixe :

$$\begin{aligned} z'' &= z_\Omega + (z' - z_\Omega) e^{i\theta} \\ &= z_\Omega + (z_\Omega + k(z - z_\Omega) - z_\Omega) e^{i\theta} \\ &= z_\Omega + k(z - z_\Omega) e^{i\theta} \\ &= (1 - k e^{i\theta}) z_\Omega + k z e^{i\theta} \end{aligned}$$

Réciproquement, considérons une application de la forme $z \mapsto az + b$, $a \in \mathbb{C}^*$, $b \in \mathbb{C}$.

Lorsque $a \neq 1$, l'ensemble des points invariants est l'ensemble des points d'affixe z tels que $z = az + b$, soit : $z = \frac{b}{1-a}$.

Le point Ω , d'affixe $z_\Omega = \frac{b}{1-a}$, est donc le centre de l'application considérée.

D'autre part, si on considère un point M distinct de Ω , d'affixe z , et son image M' , d'affixe $z' = az + b$, alors :

$$\begin{aligned} \frac{M'\Omega}{M\Omega} &= \left| \frac{z_\Omega - z'}{z_\Omega - z} \right| \\ &= \left| \frac{\frac{b}{1-a} - az - b}{\frac{b}{1-a} - z} \right| \\ &= \left| \frac{b - a(1-a)z - b(1-a)}{b - (1-a)z} \right| \\ &= \left| \frac{a(b - (1-a)z)}{b - (1-a)z} \right| \\ &= |a| \end{aligned}$$

Ainsi, $M'\Omega = \Omega M' = |a| M\Omega = |a| \Omega M$. La relation reste valable lorsque M coïncide avec Ω .

L'application $z \mapsto az + b$ multiplie donc les distances par $|a|$.

Enfin, considérons quatre points M_1, M_2, M_3, M_4 , tels que $M_1 \neq M_2, M_3 \neq M_4$, d'affixes respectives z_1, z_2, z_3, z_4 , et leurs images M'_1, M'_2, M'_3, M'_4 , d'affixes respectives z'_1, z'_2, z'_3, z'_4 . On a alors $M'_1 \neq M'_2$ et $M'_3 \neq M'_4$.

De plus :

$$\begin{aligned}
 \widehat{(M'_1 M'_2, M'_3 M'_4)} &= \widehat{(M'_1 M'_2, i)} + \widehat{(i, M'_3 M'_4)} & [2\pi] \\
 &= -\arg\left(\frac{z_{M'_1 M'_2}}{i}\right) + \arg\left(\frac{z_{M'_3 M'_4}}{i}\right) & [2\pi] \\
 &= -\arg(z'_2 - z'_1) + \arg(z'_4 - z'_3) & [2\pi] \\
 &= -\arg(a(z_2 - z_1)) + \arg(a(z_4 - z_3)) & [2\pi] \\
 &= -\arg(a) - \arg(z_2 - z_1) + \arg(a) + \arg(z_4 - z_3) & [2\pi] \\
 &= -\arg(z_2 - z_1) + \arg(z_4 - z_3) & [2\pi] \\
 &= -\arg\left(\frac{z_{M_1 M_2}}{i}\right) + \arg\left(\frac{z_{M_3 M_4}}{i}\right) & [2\pi] \\
 &= \widehat{(M_1 M_2, M_3 M_4)} & [2\pi]
 \end{aligned}$$

Les angles sont donc conservés par l'application $z \mapsto az + b$.

Enfin :

$$\begin{aligned}
 \widehat{(\Omega M, \Omega M')} &= \widehat{(\Omega M, i)} + \widehat{(i, \Omega M')} & [2\pi] \\
 &= -\arg\left(\frac{z_{\Omega M}}{i}\right) + \arg\left(\frac{z_{\Omega M'}}{i}\right) & [2\pi] \\
 &= -\arg\left(z - \frac{b}{1-a}\right) + \arg\left(az + b - \frac{b}{1-a}\right) & [2\pi] \\
 &= -\arg((1-a)z - b) + \arg(a(1-a)z - ab) & [2\pi] \\
 &= -\arg((1-a)z - b) + \arg((1-a)z - b) + \arg(a) & [2\pi] \\
 &= \arg(a) & [2\pi]
 \end{aligned}$$

Lorsque $a = 1$, on retrouve l'expression analytique d'une translation.

Définition

On appelle **similitude (directe)**, une application de \mathbb{C} dans \mathbb{C} , de la forme :

$$z \mapsto az + b \quad \text{où } a \in \mathbb{C}^*, \text{ et } b \in \mathbb{C}$$

Le **rapport** de la similitude est $|a|$, l'angle, $\arg(a)$ $[2\pi]$.

► Conservation des angles orientés

Une **similitude directe** conserve les angles orientés.

La figure suivante montre l'action de similitudes successives sur cinq triangles, à qui on fait subir une rotation ayant pour centre le centre de gravité du pentagone central, puis une homothétie de même centre, de rapport strictement plus petit que un. Et on tourne ... Similarité à l'infini !

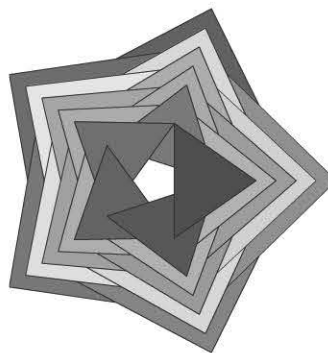


Figure 72.1 – Transformation de triangles par similitudes.

► Similitudes directes remarquables

- Lorsque $a = 1$: la similitude $z \mapsto z + b$, $b \in \mathbb{C}$, est une translation de vecteur \vec{u} , où le vecteur \vec{u} a pour affixe b .

- Lorsque $a = -1$: la similitude $z \mapsto -z + b$, $b \in \mathbb{C}$, est une symétrie centrale, dont le centre est le point d'affixe $\frac{b}{2}$.

2. Composée de deux similitudes

La composée de deux similitudes est une similitude.

Démonstration : Considérons les similitudes $\mathcal{S}_1 : z \mapsto a_1 z + b_1$, et $\mathcal{S}_2 : z \mapsto a_2 z + b_2$.

Soit M un point du plan, d'affixe z ; M_1 , d'affixe z_1 , son image par \mathcal{S}_1 .

L'affixe z_2 point M_2 , image de M_1 par \mathcal{S}_2 , est donc donnée par :

$$z_2 = a_2 z_1 + b_2 = a_2 (a_1 z + b_1) + b_2 = a_1 a_2 z + a_2 b_1 + b_2$$

$\mathcal{S}_2 \circ \mathcal{S}_1$ est donc bien une similitude, de rapport $|a_1 a_2| = |a_1| |a_2|$. ■

Considérons maintenant une application $z \mapsto a \bar{z} + b$, $a \in \mathbb{C}^*$, $b \in \mathbb{C}$.

L'ensemble des points invariants est l'ensemble des points d'affixe z tels que : $z = a \bar{z} + b$, ce qui entraîne : $\bar{z} = \bar{a} z + \bar{b}$. Par suite : $z = a \bar{a} z + a \bar{b} + b$.

Ainsi, si $|a| \neq 1$: $z = \frac{a \bar{b} + b}{1 - |a|^2}$.

On vérifie, que, réciproquement :

$$\begin{aligned} a \frac{a \bar{b} + b}{1 - |a|^2} + b &= \frac{a \bar{a} b + a \bar{b} + (1 - |a|^2) b}{1 - |a|^2} \\ &= \frac{a \bar{a} b + a \bar{b} + (1 - |a|^2) b}{1 - |a|^2} \\ &= \frac{a \bar{b} + b}{1 - |a|^2} \end{aligned}$$

L'application $z \mapsto a \bar{z} + b$ admet donc pour centre, dans le cas $|a| \neq 1$, le point Ω , d'affixe

$$z_\Omega = \frac{a \bar{b} + b}{1 - |a|^2}.$$

D'autre part, si on considère un point M distinct de Ω , d'affixe z , et son image M' , d'affixe $z' = a \bar{z} + b$, alors :

$$\begin{aligned} \frac{\Omega M'}{\Omega M} &= \frac{M' \Omega}{M \Omega} \\ &= \left| \frac{\frac{a \bar{b} + b}{1 - |a|^2} - a \bar{z} - b}{\frac{a \bar{b} + b}{1 - |a|^2} - z} \right| \\ &= \left| \frac{a \bar{b} + b - a(1 - |a|^2) \bar{z} - b(1 - |a|^2)}{a \bar{b} + b - (1 - |a|^2) z} \right| \\ &= \left| \frac{a \bar{b} - a(1 - |a|^2) \bar{z} + b |a|^2}{a \bar{b} + b - (1 - |a|^2) z} \right| \\ &= \left| \frac{a (\bar{b} - (1 - |a|^2) \bar{z} + b \bar{a})}{a \bar{b} + b - (1 - |a|^2) z} \right| \\ &= \left| \frac{\bar{a} (b - (1 - |a|^2) z + \bar{b} a)}{a \bar{b} + b - (1 - |a|^2) z} \right| \\ &= \frac{|\bar{a}|}{|a|} \\ &= \frac{|a|}{|a|} \end{aligned}$$

(un nombre complexe et son conjugué ayant même module).

Ainsi, $\Omega M' = |a|\Omega M$, et la relation reste vraie pour $\Omega = M$.

L'application $z \mapsto a\bar{z} + b$ multiplie donc les distances par $|a|$.

Enfin, considérons quatre points M_1, M_2, M_3, M_4 , tels que $M_1 \neq M_2, M_3 \neq M_4$, d'affixes respectives z_1, z_2, z_3, z_4 , et leurs images M'_1, M'_2, M'_3, M'_4 , d'affixes respectives z'_1, z'_2, z'_3, z'_4 .

On a alors $M'_1 \neq M'_2$ et $M'_3 \neq M'_4$.

De plus :

$$\begin{aligned}
 \left(\widehat{\overrightarrow{M'_1 M'_2}, \overrightarrow{M'_3 M'_4}} \right) &= \left(\widehat{\overrightarrow{M'_1 M'_2}, \vec{i}} \right) + \left(\widehat{\vec{i}, \overrightarrow{M'_3 M'_4}} \right) & [2\pi] \\
 &= -\arg\left(\overrightarrow{z_{M'_1 M'_2}}\right) + \arg\left(\overrightarrow{z_{M'_3 M'_4}}\right) & [2\pi] \\
 &= -\arg\left(z'_2 - z'_1\right) + \arg\left(z'_4 - z'_3\right) & [2\pi] \\
 &= -\arg\left(a(\overline{z_2 - z_1})\right) + \arg\left(a(\overline{z_4 - z_3})\right) & [2\pi] \\
 &= -\arg(a) - \arg(\overline{z_2 - z_1}) + \arg(a) + \arg(\overline{z_4 - z_3}) & [2\pi] \\
 &= \arg(z_2 - z_1) - \arg(z_4 - z_3) & [2\pi] \\
 &= -\arg\left(\overrightarrow{z_{M_1 M_2}}\right) + \arg\left(\overrightarrow{z_{M_3 M_4}}\right) & [2\pi] \\
 &= -\left(\widehat{\overrightarrow{M_1 M_2}, \overrightarrow{M_3 M_4}} \right) & [2\pi]
 \end{aligned}$$

Les angles orientés sont donc changés en leurs opposés par l'application $z \mapsto a\bar{z} + b$, un nombre complexe et son conjugué ayant des arguments opposés.

3. Similitude indirecte

Définition

On appelle **similitude indirecte**, une application de \mathbb{C} dans \mathbb{C} , de la forme :

$$z \mapsto a\bar{z} + b$$

où $a \in \mathbb{C}^*$, et $b \in \mathbb{C}$.

Le **rapport** de la similitude est $|a|$.

Une **similitude indirecte** inverse le sens des angles orientés.

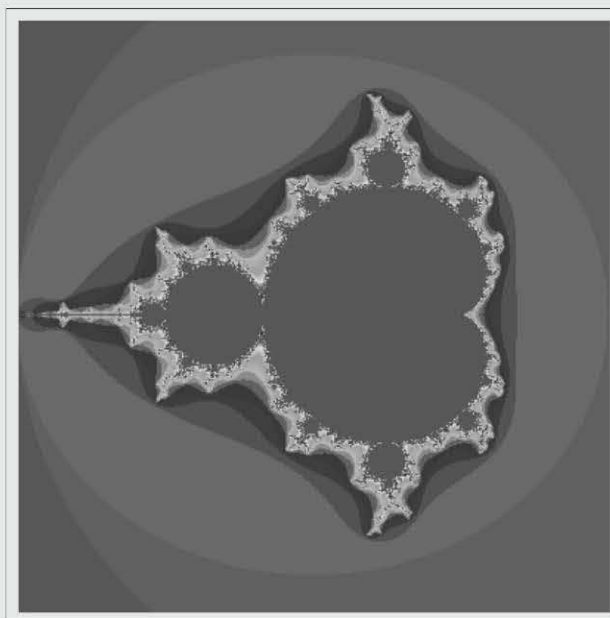
Nous avons présenté, dans ce qui précède, les transformations usuelles que sont les translations, les homothéties, rotations et similitudes. Il en existe une infinité d'autres.

L'ensemble de Mandelbrot

En particulier, considérons la transformation du plan complexe $z \in \mathbb{C} \mapsto z^2 + c$, $c \in \mathbb{C}$, et intéressons-nous à la suite $(z_n)_{n \in \mathbb{N}}$ définie par :

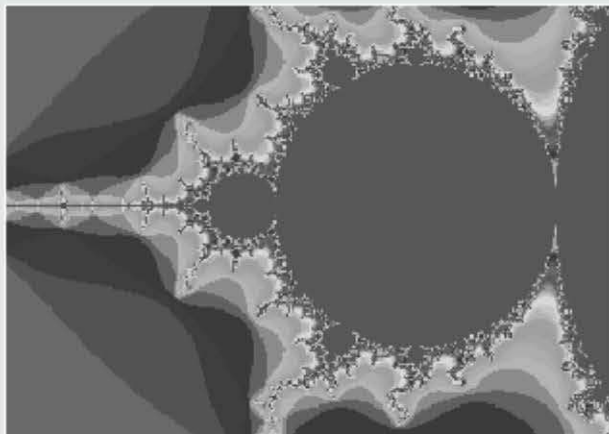
$$z_0 = 0 \quad , \quad z_{n+1} = z_n^2 + c$$

Si on représente, dans le plan complexe, les images successives des termes de cette suite, on obtient la figure suivante (les zones les plus claires sont celles avec les plus faibles concentrations de points, la couleur fonce en lien avec cette concentration).



L'ensemble de Mandelbrot.

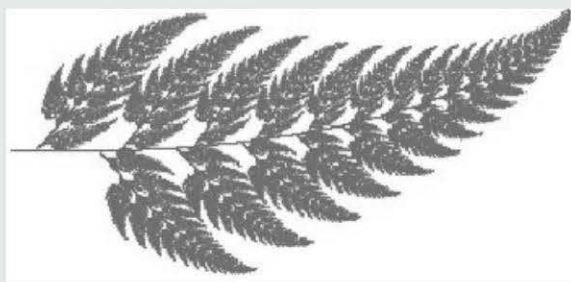
C'est l'ensemble de Mandelbrot, obtenu par le mathématicien Benoît Mandelbrot, qui, pour la désigner, créa le terme de **fractale** [12]. Depuis, on désigne par **fractale** une courbe de forme a priori irrégulière, mais dont la construction résulte en fait d'un processus itératif, parfois très simple, poussé à l'infini. Les propriétés de l'ensemble de Mandelbrot ont été étudiées de façon approfondie par Adrien Douady et John H. Hubbard [11], qui ont mis en évidence une propriété remarquable de ce type d'ensembles : leur auto-similarité, c'est-à-dire le fait que ces figures semblent se reproduire à l'intérieur d'elles-mêmes de façon infinie. En faisant un « zoom » sur une partie du dessin, on s'aperçoit ainsi qu'on retrouve exactement la même forme que la figure initiale, en plus petite certes, mais en tout point semblable.



Plan rapproché de l'ensemble de Mandelbrot.

Des structures fractales très naturelles

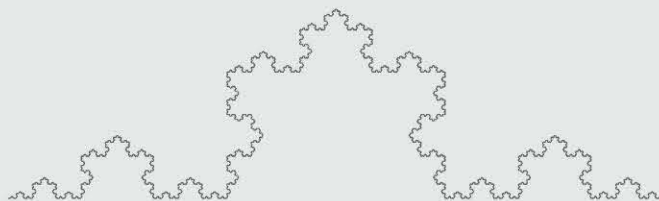
Ce phénomène d'auto-similarité existant aussi dans la nature, par exemple, une feuille est elle-même composée de plus petites feuilles de la même forme, elles-mêmes aussi, ... l'idée est alors venue d'utiliser les fractales pour représenter des formes naturelles comme, bien sûr, cette feuille de fougère, créée par le mathématicien australien Michael Barnsley [13], [14], [15] :



La fougère dite « de Barnsley ».

En biologie, notamment, les alvéoles pulmonaires présentent une structure fractale, de même que la ramification du réseau sanguin.

Il apparaît que les représentations ainsi obtenues sont très réalistes, et traduisent très bien de nombreux phénomènes naturels.



Une fractale obtenue à partir du flocon de Von Koch.

Des matrices pour la résolution de systèmes linéaires

Le concept de matrice est apparu au XVII^e siècle, afin de simplifier la résolution et l'étude des systèmes linéaires. En 1678, Gottfried Wilhelm Leibniz introduit une notation indicielle, dans le cas d'un système de trois équations à deux inconnues. Les matrices seront ensuite utilisées par Gabriel Cramer, toujours pour la résolution de systèmes linéaires (1774). Ses travaux sont poursuivis par les Français Alexandre Théophile Vandermonde, puis Pierre Simon Laplace, qui définissent par récurrence le déterminant d'une matrice carrée.

L'étude des transformations linéaires par Joseph Louis Lagrange et Carl Friedrich Gauss renvoient, à nouveau, aux matrices : ainsi, Gauss utilise pour la première fois une notation en tableau proche d'une écriture matricielle. Ses travaux seront poursuivis par Augustin-Louis Cauchy, qui définit le produit matriciel.

Le mot « matrice » est introduit par James Joseph Sylvester, pour désigner un tableau rectangulaire de nombres, mais c'est avec la publication par Arthur Cayley, en 1858, d'un article des *Philosophical Transactions* [17], que les matrices sont véritablement mises en place [18].

- Gottfried Leibniz (1646-1716) est un philosophe, scientifique, mathématicien, logicien, diplomate, juriste, bibliothécaire et philologue allemand.
- Gabriel Cramer (1704-1752) est un mathématicien suisse, connu essentiellement pour son traité sur les courbes algébriques publié en 1750, où il est le premier à démontrer qu'une courbe du $n^{\text{ième}}$ degré est déterminée par $\frac{n(n+3)}{2}$ de ses points. Il a également travaillé sur la résolution des systèmes linéaires, qui lui doivent les « formules de Cramer ».
- Théophile Vandermonde (1735-1796) est un mathématicien français, mais aussi économiste, musicien et chimiste.
D'après Lebesgue (Conférence d'Utrecht, 1937, [23]), le déterminant ne serait pas de lui...
- Pierre Laplace (1749-1827) est un mathématicien, astronome et physicien français.
- Joseph Lagrange (1736-1813) est un mathématicien, mécanicien et astronome italien, mais de famille française par son arrière-grand-père. Il s'est intéressé à l'algèbre, au calcul infinitésimal, aux probabilités, à la théorie des nombres, à la mécanique théorique, mais aussi la mécanique céleste, la mécanique des fluides, la cartographie...
- Carl Gauss (1777-1855) est un mathématicien, astronome et physicien allemand.
- Augustin-Louis Cauchy (1789-1857) est un mathématicien français, membre de l'Académie des sciences et professeur à l'École Polytechnique. Il est à l'origine de l'introduction des fonctions holomorphes en analyse, ainsi que des critères de convergence pour les suites et les séries entières (suites de Cauchy, critère de Cauchy).

- James Sylvester (1814-1897) est un mathématicien et géomètre anglais. Il a, notamment, travaillé avec Arthur Cayley sur les formes quadratiques et leurs invariants (théorème d'inertie de Sylvester), et les déterminants. Il a, également, introduit la fonction indicatrice d'Euler, qui, à un entier naturel n , associe le nombre d'entiers strictement positifs inférieurs ou égaux à n et premiers avec n , très utilisée en théorie des nombres.
- Arthur Cayley (1821-1895) est un mathématicien britannique, un des fondateurs de l'école britannique moderne de mathématiques pures.

Les applications du calcul matriciel dans l'industrie

Le calcul matriciel occupe, maintenant, une place fondamentale dans l'industrie, pour la modélisation et le calcul numérique. Ainsi, dans l'aéronautique, on utilise des modèles tridimensionnels de CFD (Computational Fluid Dynamics). L'ordinateur crée un modèle de la surface de l'avion à l'aide d'un réseau de « boîtes », situées soit à l'intérieur de la surface, soit à l'extérieur, soit à cheval entre les deux. Les « boîtes » qui intersectent (rencontrent) la surface de l'avion sont à nouveau divisées en un nouveau réseau de « boîtes » encore plus petites, où on ne retient que celles qui intersectent la surface. On itère le procédé jusqu'à ce que le réseau obtenu soit constitué de « boîtes » de très petite taille, afin d'obtenir le plus de précision possible. Pour déterminer la circulation de l'air autour de l'avion, et donc, les efforts s'exerçant sur celui-ci, l'ordinateur est, finalement, amené à résoudre un système d'équations linéaires de très grande taille, puisque celle-ci peut atteindre 2.10^6 , soit 2 millions d'équations – ou plus ! Le second membre du système est modifié à chaque calcul (ou itération), en fonction des résultats des calculs précédents, et des données du réseau.

Le stockage des données afférentes doit être fait de la façon la plus optimale possible, afin que l'ordinateur puisse effectuer, en un temps minimum, les calculs. Les données sont, ainsi, stockées sous forme de tableaux de nombres (des matrices, ou des vecteurs). D'autre part, afin d'optimiser les calculs, on essaye toujours de les exprimer sous une forme simplifiée au maximum. Pour une matrice, la présence de « zéros » judicieusement placés, regroupés en « blocs », facilitera ceux-ci. Mais, même simplifiés à l'extrême, ceux-ci, en raison de la taille des systèmes, restent toujours complexes. L'algèbre linéaire, la théorie de la réduction matricielle, la factorisation, sont des outils très puissants pour améliorer les performances et les résultats des simulations numériques. Le lecteur intéressé pourra trouver plus de précisions dans [19], sur le plan formel, et dans [20], pour une vision plus appliquée.

Matrices de taille 2×2

1. Définitions

On appelle **matrice de taille 2×2** une collection de $2 \times 2 = 4$ nombres, a_{11} , a_{12} , a_{21} , a_{22} , réels ou complexes, arrangés en tableau, sous la forme :

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ est le **premier vecteur colonne** de A , $\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ son **second vecteur colonne**.



1. Dans ce qui suit, on se limitera au cas de matrices à coefficients réels.
2. $\mathcal{M}_2(\mathbb{R})$ désigne l'ensemble des matrices de taille 2×2 à coefficients réels.

► Matrice triangulaire supérieure

Lorsqu'une matrice est de la forme :

$$A = \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$$

elle est dite **triangulaire supérieure**.

► Matrice triangulaire inférieure

Lorsqu'une matrice est de la forme :

$$A = \begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix}$$

elle est dite **triangulaire inférieure**.

► Matrice diagonale

Lorsqu'une matrice est de la forme :

$$A = \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix}$$

elle est dite **diagonale**.

► Transposée

Étant donnée une matrice de taille 2×2 , $A = (a_{ij})_{1 \leq i \leq 2, 1 \leq j \leq 2}$, on appelle **transposée** de la matrice A la matrice tA telle que :

$${}^tA = (a_{ji})_{1 \leq i \leq 2, 1 \leq j \leq 2} = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}$$

ce qui signifie que, par rapport à la matrice A , on a échangé les lignes et les colonnes.

► Trace

Étant donnée une matrice $A = (a_{ij})_{1 \leq i \leq 2, 1 \leq j \leq 2}$ de taille 2×2 , la **trace** de A , notée $tr A$, est la somme de ses termes diagonaux :

$$tr A = a_{11} + a_{22} = \sum_{i=1}^2 a_{ii}$$

2. Opérations élémentaires

Étant donnée une matrice de taille 2×2 , $A = (a_{ij})_{1 \leq i \leq 2, 1 \leq j \leq 2}$, on appelle **opération élémentaire** sur les colonnes de A l'une des opérations suivantes :

- **Transposition** : $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mapsto \begin{pmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{pmatrix}$.
- **Dilatation** : $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mapsto \begin{pmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{pmatrix} \quad \lambda \in \mathbb{R}^*$, ou encore

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & \lambda a_{12} \\ a_{21} & \lambda a_{22} \end{pmatrix} \quad \lambda \in \mathbb{R}^*$$

- **Transvection** : $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mapsto \begin{pmatrix} a_{11} + \lambda a_{12} & a_{12} \\ a_{21} + \lambda a_{22} & a_{22} \end{pmatrix} \quad \lambda \in \mathbb{R}$, ou encore :

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mapsto \begin{pmatrix} a_{11} & a_{12} + \lambda a_{11} \\ a_{21} & a_{22} + \lambda a_{21} \end{pmatrix} \quad \lambda \in \mathbb{R}$$

(À une colonne donnée, on ajoute l'autre colonne, multipliée par un facteur λ .)



1. L'opération élémentaire de transposition ne conduit pas du tout à la transposée d'une matrice.
2. On peut définir, de même, des opérations élémentaires sur les lignes de A .
3. Les opérations élémentaires sont **inversibles**, c'est-à-dire il est possible de revenir à la matrice de départ par d'autres opérations élémentaires.

Déterminant de matrices de taille 2×2

1. Définitions

► Déterminant d'une matrice

Étant donnée une matrice de taille 2×2 , $A = (a_{ij})_{1 \leq i \leq 2, 1 \leq j \leq 2}$, on appelle **déterminant** de la matrice A le nombre :

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} a_{22} - a_{21} a_{12}$$

Proposition

Une matrice de taille 2×2 et sa transposée ont le même déterminant :

$$\forall A \in \mathcal{M}_2(\mathbb{R}) : \det {}^t A = \det A$$

Démonstration : Il suffit de reprendre la formule donnée en définition. ■

► Déterminant de 2 vecteurs

On appelle **déterminant** de deux vecteurs $\vec{u} = (u_1, u_2)$ et $\vec{v} = (v_1, v_2)$ du plan \mathbb{R}^2 , le nombre :

$$\det(\vec{u}, \vec{v}) = \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} = u_1 v_2 - u_2 v_1$$

C'est donc, tout simplement, le déterminant de la matrice :

$$\begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$$

2. Propriétés

1. Étant donnée une matrice de taille 2×2 , $A = (a_{ij})_{1 \leq i \leq 2, 1 \leq j \leq 2}$, dont on désignera par C_1 et C_2 les vecteurs colonnes, le déterminant de la matrice A vérifie les propriétés suivantes :

- **Antisymétrie :**

$$\det A = \det(C_1, C_2) = -\det(C_2, C_1)$$

L'antisymétrie traduit le fait que le déterminant change de signe si on échange deux colonnes.

- **Homogénéité :**

$$\forall \lambda \in \mathbb{R} : \det(\lambda C_1, C_2) = \lambda \det(C_1, C_2)$$

- **Invariance par transvection :**

$$\forall \lambda \in \mathbb{R} : \det(C_1 + \lambda C_2, C_2) = \det(C_1, C_2)$$

La valeur du déterminant ne change pas si on ajoute, à une colonne donnée, l'autre colonne multipliée par un facteur λ .



Le déterminant d'une matrice et de sa transposée étant égaux, les propriétés d'invariance du déterminant par opérations élémentaires sur les vecteurs colonnes de la matrice sont également valables pour les lignes de la matrice.

2. Le déterminant d'une *matrice triangulaire* est égal au produit des termes diagonaux.

Exemples

1.

$$\begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} = 1 \times 3 - 2 \times 1 = 3 - 2 = 1$$

2.

$$\begin{vmatrix} -1 & 1 \\ 3 & 4 \end{vmatrix} = -1 \times 4 - 3 \times 1 = -4 - 3 = -7$$

3. Deux vecteurs $\vec{u} = (u_1, u_2)$ et $\vec{v} = (v_1, v_2)$ du plan sont non colinéaires si et seulement si leur déterminant est non nul :

$$\det(\vec{u}, \vec{v}) = \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} \neq 0$$

Démonstration : On démontre ce résultat par la contraposée : *deux vecteurs $\vec{u} = (u_1, u_2)$ et $\vec{v} = (v_1, v_2)$ du plan sont liés si et seulement si leur déterminant est nul.*

- Supposons \vec{u} et \vec{v} liés et tels que $\vec{u} = \vec{0}$ ou $\vec{v} = \vec{0}$; on a alors, clairement,

$$\det(\vec{u}, \vec{v}) = 0$$

- Supposons \vec{u} et \vec{v} liés et tels que $\vec{u} \neq \vec{0}$ et $\vec{v} \neq \vec{0}$, il existe alors un réel non nul λ tel que :

$$\vec{v} = \lambda \vec{u}$$

Par suite :

$$\det(\vec{u}, \lambda \vec{u}) = \lambda \det(\vec{u}, \vec{u}) = 0$$

- Réciproquement, si $\det(\vec{u}, \vec{v}) = 0$, alors :

$$u_1 v_2 - u_2 v_1 = 0$$

Si $v_2 \neq 0$:

$$u_1 = \frac{u_2}{v_2} v_1$$

Comme :

$$u_2 = \frac{u_2}{v_2} v_2$$

on a donc :

$$\vec{u} = \frac{u_2}{v_2} \vec{v}$$

les vecteurs \vec{u} et \vec{v} sont bien liés.

On démontre de même le résultat dans le cas où $u_2 \neq 0$.

Lorsque $u_2 = v_2 = 0$, u_1 et v_1 peuvent prendre chacun une valeur quelconque dans \mathbb{R} .

Les vecteurs \vec{u} et \vec{v} sont alors colinéaires, et donc liés.

■

Matrices de taille 3×3

1. Définitions

On appelle **matrice de taille 3×3** une collection de $3 \times 3 = 9$ nombres, réels ou complexes, $a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}, a_{32}, a_{33}$, arrangés en tableau, sous la forme :

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}$ est le **premier vecteur colonne** de A , $\begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}$ son **second vecteur colonne**, $\begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}$ son **troisième vecteur colonne**.



1. Dans ce qui suit, on se limitera au cas de matrices à coefficients réels.
2. $\mathcal{M}_3(\mathbb{R})$ désigne l'ensemble des matrices de taille 3×3 à coefficients réels.

► Matrice triangulaire supérieure

Lorsqu'une matrice est de la forme :

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}$$

elle est dite **triangulaire supérieure**.

► Matrice triangulaire inférieure

Lorsqu'une matrice est de la forme :

$$A = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

elle est dite **triangulaire inférieure**.

► Matrice diagonale

Lorsqu'une matrice est de la forme :

$$A = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$$

elle est dite **diagonale**.

► Transposée

Étant donnée une matrice de taille 3×3 , $A = (a_{ij})_{1 \leq i \leq 3, 1 \leq j \leq 3}$, on appelle **transposée** de la matrice A la matrice tA , de taille 3×3 , telle que :

$${}^tA = (a_{ji})_{1 \leq j \leq 3, 1 \leq i \leq 3} = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$$

ce qui signifie que, par rapport à la matrice A , on a échangé les lignes et les colonnes.

► Trace

Étant donnée une matrice $A = (a_{ij})_{1 \leq i \leq 3, 1 \leq j \leq 3}$ de $\mathcal{M}_3(\mathbb{R})$, la **trace** de A , notée $tr A$, est la somme de ses termes diagonaux :

$$tr A = a_{11} + a_{22} + a_{33} = \sum_{i=1}^3 a_{ii}$$

2. Opérations élémentaires

Étant donnée une matrice de taille 3×3 , $A = (a_{ij})_{1 \leq i \leq 3, 1 \leq j \leq 3}$, on appelle **opération élémentaire** sur les colonnes, respectivement désignées par C_1, C_2, C_3 , de A , l'une des opérations suivantes :

• Transposition :

$$A = (C_1, C_2, C_3) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \mapsto (C_2, C_1, C_3) = \begin{pmatrix} a_{12} & a_{11} & a_{13} \\ a_{22} & a_{21} & a_{23} \\ a_{32} & a_{31} & a_{33} \end{pmatrix}$$

ou encore :

$$A = (C_1, C_2, C_3) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \mapsto (C_1, C_3, C_2)$$

etc.

L'opération élémentaire de transposition consiste donc à échanger deux colonnes de la matrice.

• Dilatation :

$$A = (C_1, C_2, C_3) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \mapsto (\lambda C_1, C_2, C_3) = \begin{pmatrix} \lambda a_{11} & a_{12} & a_{13} \\ \lambda a_{21} & a_{22} & a_{23} \\ \lambda a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \lambda \in \mathbb{R}^*$$

ou encore :

$$A = (C_1, C_2, C_3) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \mapsto (C_1, \lambda C_2, C_3) = \begin{pmatrix} a_{11} & \lambda a_{12} & a_{13} \\ a_{21} & \lambda a_{22} & a_{23} \\ a_{31} & \lambda a_{32} & a_{33} \end{pmatrix} \quad \lambda \in \mathbb{R}^*$$

ou :

$$A = (C_1, C_2, C_3) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \mapsto (C_1, C_2, \lambda C_3) = \begin{pmatrix} a_{11} & a_{12} & \lambda a_{13} \\ a_{21} & a_{22} & \lambda a_{23} \\ a_{31} & a_{32} & \lambda a_{33} \end{pmatrix} \quad \lambda \in \mathbb{R}^*$$

L'opération élémentaire de dilatation, par un réel non nul λ , consiste donc à multiplier une colonne de la matrice par λ .

• **Transvection :**

$$A = (C_1, C_2, C_3) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \mapsto (C_1 + \lambda C_2, C_2, C_3) = \begin{pmatrix} a_{11} + \lambda a_{12} & a_{12} & a_{13} \\ a_{21} + \lambda a_{22} & a_{22} & a_{23} \\ a_{31} + \lambda a_{32} & a_{32} & a_{33} \end{pmatrix}, \quad \lambda \in \mathbb{R}$$

ou encore :

$$A = (C_1, C_2, C_3) = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \mapsto (C_1, C_2 + \lambda C_3, C_3), \quad \lambda \in \mathbb{R}$$

etc.

À une colonne donnée, on ajoute une combinaison linéaire des autres colonnes.



1. L'opération élémentaire de transposition ne conduit pas du tout à la transposée d'une matrice.
2. On peut définir, de même, des opérations élémentaires sur les lignes de A .
3. Les opérations élémentaires sont **inversibles**, c'est-à-dire qu'il est possible de revenir à la matrice de départ par d'autres opérations élémentaires.

1. Définitions

► Mineur

Soit $A \in \mathcal{M}_3(\mathbb{R})$.

Étant donné $(i, j) \in \{1, \dots, 3\}^2$, on appelle **mineur d'indice (i, j)** le déterminant, noté $\Delta_{ij}(A)$, de la matrice obtenue en enlevant à A la ligne i et la colonne j .

► Cofacteur

Soit $A \in \mathcal{M}_3(\mathbb{R})$.

Étant donné $(i, j) \in \{1, 2, 3\}^2$, on appelle **cofacteur d'indice (i, j)** :

$$(-1)^{i+j} \Delta_{ij}(A)$$

► Comatrice

Soit $A \in \mathcal{M}_3(\mathbb{R})$.

On appelle **comatrice** de la matrice A , et on note $ComA \in \mathcal{M}_3(\mathbb{R})$ la **matrice des cofacteurs** :

$$ComA = \left((-1)^{i+j} \Delta_{ij}(A) \right)_{1 \leq i \leq 3, 1 \leq j \leq 3}$$

► Déterminant

Étant donnée $A \in \mathcal{M}_3(\mathbb{R})$, on appelle **déterminant** de la matrice $A = (a_{ij})_{1 \leq i \leq 3, 1 \leq j \leq 3}$ le nombre, noté $\det A$, qui peut être calculé de la façon suivante :

- par un **développement suivant la ligne i** , $i \in \{1, \dots, 3\}$:

$$\det A = \sum_{j=1}^3 (-1)^{i+j} a_{ij} \Delta_{ij}(A)$$

- par un **développement suivant la colonne j** , $j \in \{1, \dots, 3\}$:

$$\det A = \sum_{i=1}^3 (-1)^{i+j} a_{ij} \Delta_{ij}(A)$$



On choisit la ligne ou la colonne que l'on veut ; quel que soit le choix, on obtient bien sûr le même résultat !

Proposition

Une matrice de taille 3×3 et sa transposée ont le même déterminant :

$$\forall A \in \mathcal{M}_3(\mathbb{R}) : \det {}^tA = \det A$$

Démonstration : Ce résultat s'obtient immédiatement à partir de la formule donnant le développement suivant une ligne ou une colonne. ■



1. si on développe par rapport à une ligne :

$$\begin{aligned}\det A &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= -a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{22} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{23} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{32} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\end{aligned}$$

2. et si on développe par rapport à une colonne :

$$\begin{aligned}\det A &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= -a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{22} \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} - a_{32} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \\ &= a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} - a_{23} \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} + a_{33} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\end{aligned}$$

2. Propriété

1. Étant donnée une matrice de taille 3×3 , $A = (a_{ij})_{1 \leq i \leq 3, 1 \leq j \leq 3}$, dont on désigne par C_1, C_2, C_3 les vecteurs colonnes, le déterminant de la matrice A vérifie les propriétés suivantes :

- **Antisymétrie :**

$$\det A = \det(C_1, C_2, C_3) = -\det(C_2, C_1, C_3) = -\det(C_1, C_3, C_2)$$

L'antisymétrie traduit le fait que le déterminant change de signe si on échange deux colonnes.

- **Homogénéité :**

$$\forall \lambda \in \mathbb{R} : \det(\lambda C_1, C_2, C_3) = \lambda \det(C_1, C_2, C_3)$$

- **Invariance par transvection :** $\forall (\lambda, \mu) \in \mathbb{R}^2 :$

$$\begin{aligned}\det(C_1, C_2 + \lambda C_1, C_3) &= \det(C_1, C_2, C_3 + \lambda C_1) \\ &= \det(C_1 + \lambda C_2, C_2, C_3) \\ &= \det(C_1 + \lambda C_2 + \mu C_3, C_2, C_3) \\ &= \dots \\ &= \det(C_1, C_2, C_3)\end{aligned}$$

La valeur du déterminant ne change pas si, à une colonne donnée, on ajoute une combinaison linéaire des autres colonnes.



Le déterminant d'une matrice et de sa transposée étant égaux, les propriétés d'invariance du déterminant par opérations élémentaires sur les vecteurs colonnes de la matrice sont également valables pour les lignes de la matrice.

2. Le déterminant d'une *matrice triangulaire* est égal au produit des termes diagonaux.

Démonstration : Il suffit de développer par rapport à la première colonne. ■

3. Le déterminant d'une *matrice diagonale* est égal au produit des termes diagonaux.

Démonstration : Une matrice diagonale est un cas particulier de matrice triangulaire. ■

Définition

On appelle **déterminant** de trois vecteurs $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$, $\vec{w} = (w_1, w_2, w_3)$ de l'espace \mathbb{R}^3 , le déterminant de la matrice :

$$\det(\vec{u}, \vec{v}, \vec{w}) = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix}$$

4. Trois vecteurs $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$, $\vec{w} = (w_1, w_2, w_3)$ de l'espace \mathbb{R}^3 , sont non coplanaires si et seulement si leur déterminant est non nul :

$$\det(\vec{u}, \vec{v}, \vec{w}) = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix} \neq 0$$

Démonstration : On démontre ce résultat par la contraposée : *trois vecteurs $\vec{u} = (u_1, u_2, u_3)$, $\vec{v} = (v_1, v_2, v_3)$, $\vec{w} = (w_1, w_2, w_3)$ sont liés si et seulement si leur déterminant est nul.*

- Supposons $\vec{u}, \vec{v}, \vec{w}$ liés ; alors, soit la famille (\vec{u}, \vec{v}) est liée, soit $\vec{w} \in \mathbb{R}\vec{u} + \mathbb{R}\vec{v}$.
Dans le premier cas, soit $\vec{v} = \vec{0}$, ce qui conduit immédiatement à $\det(\vec{u}, \vec{v}, \vec{w}) = 0$, soit il existe un réel α tel que $\vec{u} = \alpha\vec{v}$. On a alors :

$$\det(\vec{u}, \vec{v}, \vec{w}) = \det(\alpha\vec{v}, \vec{v}, \vec{w}) = \alpha \det(\vec{v}, \vec{v}, \vec{w}) = \alpha \det(\vec{v} - \vec{v}, \vec{v}, \vec{w}) = 0$$

en utilisant l'invariance du déterminant par transvection.

Dans le second cas, il existe deux réels λ et μ tels que $\vec{w} = \lambda\vec{u} + \mu\vec{v}$. Par suite :

$$\det(\vec{u}, \vec{v}, \vec{w}) = \det(\vec{u}, \vec{v}, \lambda\vec{u} + \mu\vec{v}) = \det(\vec{u}, \vec{v}, \vec{u}) = 0$$

(On utilise l'invariance du déterminant par transvection.)

- La réciproque se démontre par le calcul, de façon analogue à ce qui a été fait en dimension 2. ■

Matrices de taille $m \times n$

1. Coefficients d'une matrice

m et n étant des entiers naturels non nuls, on appelle **matrice de taille $m \times n$** une collection de $m \times n$ nombres, réels ou complexes, arrangés en tableau :

$$A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

Les a_{ij} sont appelés **coefficients de la matrice** : par convention, le premier indice désigne l'**indice de ligne**, le second, l'**indice de colonne**.



Dans ce qui suit, on se limitera au cas de matrices à coefficients réels.

m et n étant des entiers naturels non nuls, on désigne par $\mathcal{M}_{m,n}(\mathbb{R})$ l'ensemble des matrices de taille $m \times n$, à coefficients réels¹.

2. Matrice carrée d'ordre n

n étant un entier naturel non nul, on appelle **matrice carrée d'ordre n** une matrice de taille $n \times n$:

$$A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$$

n étant un entier naturel non nul, on désigne par $\mathcal{M}_n(\mathbb{R})$ l'ensemble des matrices carrées d'ordre n , à coefficients réels.

Étant donnée une matrice $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$ de $\mathcal{M}_n(\mathbb{R})$, les coefficients a_{ii} , $i = 1, \dots, n$, sont les **termes diagonaux** de A .

Étant donnée une matrice $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$ de $\mathcal{M}_n(\mathbb{R})$, la **trace** de A , notée $tr A$, est la somme de ses termes diagonaux :

$$tr A = a_{11} + \dots + a_{nn} = \sum_{i=1}^n a_{ii}$$

3. Matrice ligne et colonne

n étant un entier naturel non nul, on appelle **matrice ligne** une matrice de taille $1 \times n$, de la forme :

$$A = (a_1 \dots a_n)$$

1. On peut aussi trouver la notation $\mathcal{M}_{m \times n}(\mathbb{R})$.

m étant un entier naturel non nul, on appelle **matrice colonne**, ou encore vecteur colonne, une matrice de taille $m \times 1$, de la forme :

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$$

Étant donnée une matrice $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ de $\mathcal{M}_{m,n}(\mathbb{R})$, on appelle **transposée** de la matrice A la matrice tA , de taille $n \times m$, telle que :

$${}^tA = (a_{ji})_{1 \leq i \leq n, 1 \leq j \leq m} = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$$

ce qui signifie que, par rapport à la matrice A , on a échangé les lignes et les colonnes de A .

Opérations sur les matrices

Dans ce qui suit, m, n, p, q désignent des entiers naturels non nuls.

1. Définitions

► Addition de deux matrices

Étant données deux matrices A et B de $\mathcal{M}_{m,n}(\mathbb{R})$, on définit la matrice somme $A + B$ comme étant la matrice $C \in \mathcal{M}_{m,n}(\mathbb{R})$ telle que :

$$C = A + B = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

► Multiplication par un scalaire

Étant donnée une matrice A de $\mathcal{M}_{m,n}(\mathbb{R})$, on définit le produit de A par le réel λ comme étant la matrice :

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \dots & \lambda a_{mn} \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{R})$$

► Multiplication de deux matrices

La multiplication d'une matrice $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$, de taille $m \times n$, par une matrice $B = (b_{ij})_{1 \leq i \leq n, 1 \leq j \leq p}$, de taille $n \times p$, conduit à la matrice produit C , de taille $m \times p$, telle que, pour tout couple d'entiers (i, j) de $\{1, \dots, m\} \times \{1, \dots, p\}$:

$$C = AB = (c_{ij})_{1 \leq i \leq m, 1 \leq j \leq p}$$

avec :

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$$

Exemple

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \times 1 + 0 \times 0 + 1 \times 1 \\ 2 \times 1 + 1 \times 0 + 0 \times 1 \\ 1 \times 1 + 0 \times 0 + 0 \times 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$



Attention ! Pour pouvoir faire le produit de deux matrices, il est indispensable que **le nombre de colonnes de la première matrice soit égal au nombre de lignes de la seconde matrice**.
En pratique, le coefficient situé **ligne i , colonne j** de la matrice produit est obtenu en « multipliant la ligne i de la première matrice par la colonne j de la seconde matrice » :

Coefficient de AB ligne i , colonne j

||

$\sum_{\text{un par un}}$ coefficients de la ligne i de $A \times$ coefficients de la colonne j de B

ce qui permet de comprendre pourquoi le nombre de colonnes de la première matrice doit être égal au nombre de lignes de la seconde matrice : il faut tout simplement qu'il y ait le même nombre de coefficients sur cette ligne que sur la colonne.

La multiplication des matrices n'est pas commutative. En général,

$$A B \neq B A$$

2. Propriétés de la multiplication matricielle

La multiplication matricielle est :

- distributive à gauche :

$$\forall (A, B, C) \in \mathcal{M}_{m,n}(\mathbb{R}) \times (\mathcal{M}_{n,p}(\mathbb{R}))^2 : A(B + C) = AB + AC$$

- distributive à droite :

$$\forall (A, B, C) \in (\mathcal{M}_{m,n}(\mathbb{R}))^2 \times \mathcal{M}_{n,p}(\mathbb{R}) : (A + B)C = AC + BC$$

- associative :

$$\forall (A, B, C) \in \mathcal{M}_{m,n}(\mathbb{R}) \times \mathcal{M}_{n,p}(\mathbb{R}) \times \mathcal{M}_{p,q}(\mathbb{R}) : A(BC) = (AB)C$$



Si on désigne par C_1, \dots, C_n les vecteurs colonnes d'une matrice A de taille $m \times n$, alors, pour

tout vecteur $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, on peut aussi écrire :

$$AX = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 C_1 + \dots + x_n C_n$$

3. Formule du binôme de Newton

Étant données deux matrices A et B de $\mathcal{M}_n(\mathbb{R})$, qui commutent, c'est-à-dire $AB = BA$, alors, pour tout entier naturel p , la formule du binôme de Newton permet de calculer $(A + B)^p$:

$$(A + B)^p = \sum_{k=0}^p C_p^k A^k B^{p-k}$$

où, pour tout entier k de $\{0, \dots, p\}$, C_p^k désigne le coefficient binomial :

$$\binom{p}{k} = \frac{p!}{k!(p-k)!}$$

Matrices remarquables

1. Matrice identité

Définition

n étant un entier naturel non nul, on appelle **matrice identité d'ordre n** , la matrice carrée, de taille $n \times n$:

$$I_n = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix} = (\delta_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$$

Propriété

L'élément neutre de la multiplication dans $\mathcal{M}_n(\mathbb{R})$ est la matrice identité I_n :

$$\forall A \in \mathcal{M}_n(\mathbb{R}) : A I_n = I_n A = A$$

2. Base canonique

Définition

m et n étant des entiers naturels non nuls, on désigne par $\{E_{ij}\}_{1 \leq i \leq m, 1 \leq j \leq n}$ la **base canonique** de $\mathcal{M}_{m,n}(\mathbb{R})$, c'est-à-dire la famille des $m n$ matrices telles que :

$$\forall (i, j) \in \{1, \dots, m\} \times \{1, \dots, n\} : E_{ij} = (\delta_{ik} \delta_{jl})_{1 \leq k \leq m, 1 \leq l \leq n}$$

où les δ_{ij} désignent les **symboles de Kronecker**¹ :

$$\forall (i, j) \in \mathbb{N}^* \times \mathbb{N}^* : \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$$

Cela signifie que, pour i et j donnés, les coefficients de la matrice E_{ij} sont tous nuls, sauf celui situé sur la $i^{\text{ème}}$ ligne et la $j^{\text{ème}}$ colonne, qui vaut 1.

Ainsi, pour toute matrice $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ de $\mathcal{M}_{m,n}(\mathbb{R})$:

$$A = \sum_{i=1}^m \sum_{j=1}^n a_{ij} E_{ij}$$

1. Leopold Kronecker (1823-1891), mathématicien allemand. À l'origine, il était théoricien des nombres, mais apporta de nombreuses contributions à l'analyse algébrique.

Proposition

Lorsque l'on est dans $\mathcal{M}_n(\mathbb{R})$, c'est-à-dire lorsque les matrices sont carrées :

$$\forall (i, j, k, l) \in \{1, \dots, n\}^4 : E_{ij} E_{kl} = \delta_{jk} E_{il}$$

Démonstration : On a : $E_{ij} = (\delta_{im} \delta_{jp})_{1 \leq m \leq n, 1 \leq p \leq n}$, $E_{kl} = (\delta_{km} \delta_{lp})_{1 \leq m \leq n, 1 \leq p \leq n}$.

La formule du produit matriciel permet alors d'exprimer le terme d'indices m, p de la matrice produit $E_{ij} E_{kl}$:

$$\begin{aligned} (E_{ij} E_{kl})_{mp} &= \sum_{q=1}^n (E_{ij})_{mq} (E_{kl})_{qp} \\ &= \sum_{q=1}^n \delta_{im} \delta_{jq} \delta_{kq} \delta_{lp} \\ &= \delta_{im} \delta_{jj} \delta_{kj} \delta_{lp} \\ &= \delta_{im} \delta_{kj} \delta_{lp} \end{aligned}$$

(Le seul terme non nul de la somme est obtenu pour $q = j$.)

On reconnaît ainsi : $E_{ij} E_{kl} = \delta_{jk} E_{il}$. ■

3. Matrice de transposition**Définition**

n étant un entier naturel non nul, on appelle **matrice de transposition**, ou encore, **matrice de permutation**, toute matrice carrée d'ordre n , de la forme :

$$S_{ij} = \begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & 0 & 0 & \dots & 0 & 1 & 0 & \dots & \dots & 0 \\ & & & & 1 & & & & & & & \\ & & & & & \ddots & & & & & & \\ & & & & & & 1 & & & & & \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 & 0 & & & & \\ & & & & & & & 1 & & & & \\ & & & & & & & & \ddots & & & \\ & & & & & & & & & \ddots & & \\ & & & & & & & & & & 1 \end{pmatrix} = I_n - E_{ii} - E_{jj} + E_{ij} + E_{ji}$$

$$(i, j) \in \{1, \dots, n\}^2, i \neq j$$

(Les coefficients non écrits sont nuls.)

Proposition

Pour toute matrice A de $\mathcal{M}_n(\mathbb{R})$, le produit $A S_{ij}$ de la matrice A par la matrice de permutation S_{ij} , $(i, j) \in \{1, \dots, n\}^2, i \neq j$, est la matrice obtenue à partir de A en échangeant les colonnes i et j .

Démonstration : À l'aide de la formule du produit matriciel, on peut exprimer le terme situé ligne k et colonne l de la matrice $A S_{ij}$, qui vaut :

$$\begin{aligned} (A S_{ij})_{kl} &= \sum_{p=1}^n a_{kp} (S_{ij})_{pl} \\ &= \sum_{p=1}^n a_{kp} (\delta_{pl} - \delta_{ip} \delta_{il} - \delta_{jp} \delta_{jl} + \delta_{ip} \delta_{jl} + \delta_{jp} \delta_{il}) \\ &= a_{kl} - a_{ki} \delta_{il} - a_{kj} \delta_{jl} + a_{ki} \delta_{jl} + a_{kj} \delta_{il} \end{aligned}$$

Ainsi, pour toute colonne l différente de la colonne i et de la colonne j :

$$(A S_{ij})_{kl} = a_{kl}$$

et, pour la colonne i : $(A S_{ij})_{ki} = a_{ki} - a_{ki} - a_{kj} \delta_{ji} + a_{ki} \delta_{ji} + a_{kj} = a_{kj}$.

De même, pour la colonne j :

$$(A S_{ij})_{kj} = a_{kj} - a_{ki} \delta_{ij} - a_{kj} + a_{ki} \delta_{jj} + a_{kj} \delta_{ij} = a_{ki}$$

qui est le résultat cherché. ■

4. Matrice de dilatation

Définition

n étant un entier naturel non nul, et λ un réel non nul différent de 1, on appelle **matrice de dilatation**, toute matrice de $\mathcal{M}_n(\mathbb{R})$ de la forme :

$$D_i(\lambda) = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & \lambda & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} = I_n + (\lambda - 1) E_{ii} \quad , \quad i \in \{1, \dots, n\}$$

(Les coefficients non écrits sont nuls.)

Proposition

Pour toute matrice A de $\mathcal{M}_n(\mathbb{R})$, et tout réel non nul $\lambda \neq 1$, le produit $A D_i(\lambda)$ de la matrice A par la matrice de dilatation $D_i(\lambda)$, $i \in \{1, \dots, n\}$, est la matrice obtenue à partir de A en multipliant la $i^{\text{ème}}$ colonne par λ .

Démonstration : À l'aide de la formule du produit matriciel, on peut exprimer le terme situé ligne k et colonne l de la matrice $A D_i(\lambda)$, qui vaut :

$$\begin{aligned} (A D_i(\lambda))_{kl} &= \sum_{p=1}^n a_{kp} (D_i(\lambda))_{pl} \\ &= \sum_{p=1}^n a_{kp} (\delta_{pl} + (\lambda - 1) \delta_{pi} \delta_{li}) \\ &= a_{kl} + a_{ki} (\lambda - 1) \delta_{li} \end{aligned}$$

Ainsi, pour toute colonne $l \neq i$: $(A D_i(\lambda))_{kl} = a_{kl}$, pour la colonne i :

$$(A D_i(\lambda))_{ki} = a_{ki} + a_{ki} (\lambda - 1) = \lambda a_{ki}$$

qui est le résultat cherché. ■

5. Matrice de transvection

Définition

n étant un entier naturel non nul, et λ un réel, on appelle **matrice de transvection**, toute matrice de $\mathcal{M}_n(\mathbb{R})$ de la forme :

$$T_{ij}(\lambda) = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & 1 & 0 \dots 0 & \lambda & 0 \dots 0 & \\ & & & & 1 & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \end{pmatrix} = I_n + \lambda E_{ij}$$

$$(i, j) \in \{1, \dots, n\}^2, i \neq j$$

(Les coefficients non écrits sont nuls.)

6. Matrice élémentaire

Définition

On appelle **matrice élémentaire** une matrice de transposition, de dilatation, ou de transvection.

Introduction aux déterminants de matrices de taille $n \times n$

1. Définitions

► Déterminant

Étant donnée $A \in \mathcal{M}_n(\mathbb{R})$, on appelle **déterminant** de la matrice $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$ le nombre, noté $\det A$, qui peut être calculé, par récurrence, de la façon suivante :

- par un **développement suivant la ligne i** , $i \in \{1, \dots, n\}$:

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \Delta_{ij}(A)$$

- par un **développement suivant la colonne j** , $j \in \{1, \dots, n\}$:

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \Delta_{ij}(A)$$

où, pour tout couple d'indices $(i, j) \in \{1, \dots, n\}^2$, Δ_{ij} , appelé **mineur d'indice (i, j)** , est le déterminant de la matrice obtenue en enlevant à A la ligne i et la colonne j .

Quel que soit le mode de calcul retenu, on obtient, bien sûr, le même résultat.



Le déterminant d'une matrice de taille $n \times n$ se calcule donc par récurrence à partir de déterminants de matrices de taille $(n-1) \times (n-1)$, puis $(n-2) \times (n-2)$, ..., et, finalement, de déterminants de matrices de taille 2×2 .

► Cofacteur

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Étant donné $(i, j) \in \{1, \dots, n\}^2$, on appelle **cofacteur d'indice (i, j)** :

$$(-1)^{i+j} \Delta_{ij}$$

► Comatrice

Soit $A \in \mathcal{M}_n(\mathbb{R})$.

On appelle **comatrice** de la matrice A , et on note $\text{Com}A \in \mathcal{M}_n(\mathbb{R})$ la *matrice des cofacteurs* :

$$\text{Com}A = ((-1)^{i+j} \Delta_{ij}(A))_{1 \leq i \leq n, 1 \leq j \leq n}$$

Une matrice de taille $n \times n$ et sa transposée ont le même déterminant :

$$\forall A \in \mathcal{M}_n(\mathbb{R}) : \det {}^t A = \det A$$

Démonstration : Ce résultat est dû au fait que le déterminant peut être calculé, indifféremment, grâce au développement suivant une ligne ou une colonne. ■

2. Propriétés

Étant donnés n vecteurs u_1, \dots, u_n , le déterminant de la matrice $M = (u_1, \dots, u_n)$ vérifie les propriétés suivantes :

► Antisymétrie

$$\det(u_2, u_1, u_3, \dots, u_n) = -\det(u_1, u_2, u_3, \dots, u_n) = \dots = -\det(u_1, \dots, u_n)$$

L'antisymétrie traduit le fait que le déterminant change de signe si on échange deux colonnes.

► Homogénéité

$$\forall \lambda \in \mathbb{R}, \forall i \in \{1, \dots, n\} : \det(u_1, \dots, \lambda u_i, \dots, u_n) = \lambda \det(u_1, \dots, u_n)$$

► Invariance par transvection

Pour tout réel λ ,

$$\det(u_1 + \lambda u_2, u_2, \dots, u_n) = \det(u_1 + \lambda u_3, u_2, \dots, u_n) = \dots = \det(u_1, \dots, u_n)$$

La valeur du déterminant ne change pas si, à une colonne donnée, on ajoute une combinaison linéaire des autres colonnes.

Démonstration : Ce résultat s'obtient immédiatement à partir de la formule donnant le développement suivant une ligne ou une colonne. ■



Le déterminant d'une matrice et de sa transposée étant égaux, les propriétés d'invariance du déterminant par opérations élémentaires sur les vecteurs colonnes de la matrice sont également valables pour les lignes de la matrice.

► Autres propriétés

1. Le déterminant d'une *matrice triangulaire* est égal au produit des termes diagonaux.
2. Le déterminant d'une *matrice diagonale* est égal au produit des termes diagonaux.
3. Déterminant du produit de deux matrices d'ordre n :

$$\forall (A, B) \in (\mathcal{M}_n(\mathbb{R}))^2 : \det AB = \det A \det B$$

Inversion des matrices carrées

1. Conditions d'inversibilité

Une matrice carrée A d'ordre n est dite **inversible** ou **régulière** ou encore **non singulière**, s'il existe une matrice B d'ordre n telle que :

$$AB = BA = I_n$$

B est alors la matrice inverse de A . Elle est unique. On note, usuellement, A^{-1} la matrice inverse de A , qui est aussi de taille $n \times n$.

Une matrice carrée qui n'est pas inversible est dite **non inversible** ou **singulière**.

L'ensemble des matrices inversibles de $M_n(\mathbb{R})$ est appelé **groupe linéaire d'ordre n** , et noté $\mathcal{GL}_n(\mathbb{R})$ ¹.

Proposition

Une matrice est inversible si et seulement si son déterminant est non nul.

2. Inverse d'une matrice carrée

Théorème

$$\forall A \in M_n(\mathbb{R}) : A \cdot {}^t \text{Com} A = {}^t \text{Com} A \cdot A = \det A \cdot I_n$$

où ${}^t \text{Com} A$ désigne la transposée de la comatrice de A .

Démonstration : D'après la formule du produit matriciel :

$$(A \cdot {}^t \text{Com} A)_{ij} = \sum_{k=1}^n a_{ik} ({}^t \text{Com} A)_{kj} = \sum_{k=1}^n a_{ik} (\text{Com} A)_{jk}$$

Il faut donc calculer, pour tout couple d'indices $(i, j) : \sum_{k=1}^n a_{ik} (\text{Com} A)_{jk}$.

► Premier cas : $i \neq j$

Quitte à échanger i et j , on peut supposer $i < j$.

1. Il possède une structure de groupe, puisqu'il est stable par multiplication et passage à l'inverse.

Considérons la matrice obtenue en remplaçant la $j^{\text{ième}}$ ligne de A par sa $i^{\text{ième}}$ ligne :

$$\begin{pmatrix} a_{11} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{i,1} & \dots & a_{i,n} \\ \vdots & & \vdots \\ a_{j-1,1} & \dots & a_{j-1,n} \\ a_{i,1} & \dots & a_{i,n} \\ a_{j+1,1} & \dots & a_{j+1,n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{n,n} \end{pmatrix}$$

et développons son déterminant suivant la $j^{\text{ième}}$ ligne :

$$\begin{vmatrix} a_{11} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{i,1} & \dots & a_{i,n} \\ \vdots & & \vdots \\ a_{j-1,1} & \dots & a_{j-1,n} \\ a_{i,1} & \dots & a_{i,n} \\ a_{j+1,1} & \dots & a_{j+1,n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{n,n} \end{vmatrix} = \sum_{k=1}^n a_{ik} (ComA)_{jk} = \sum_{k=1}^n a_{ik} ({}^t ComA)_{kj}$$

Cette matrice ayant deux lignes identiques (la $i^{\text{ième}}$ et la $j^{\text{ième}}$), son déterminant est nul :

$$\sum_{k=1}^n a_{ik} ({}^t ComA)_{kj} = 0$$

► Deuxième cas : $i = j$

En développant $\det A$ suivant la $i^{\text{ième}}$ ligne, on obtient :

$$\det A = \begin{vmatrix} a_{11} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{i,n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{n,n} \end{vmatrix} = \sum_{k=1}^n a_{ik} (ComA)_{ik} = \sum_{k=1}^n a_{ik} ({}^t ComA)_{ki}$$

On a donc, finalement : $(A \cdot {}^t ComA)_{ij} = \det A \cdot \delta_{ij}$.

On montre de même que : $({}^t ComA \cdot A)_{ij} = \det A \cdot \delta_{ij}$.

D'où le résultat. ■

Corollaire

$$\forall A \in \mathcal{GL}_n(\mathbb{R}) : A^{-1} = \frac{1}{\det A} {}^t ComA$$

► Inverse de la transposée d'une matrice (de $\mathcal{GL}_n(\mathbb{R})$)

Corollaire

$$\forall A \in \mathcal{GL}_n(\mathbb{R}) : ({}^t A)^{-1} = {}^t (A^{-1})$$

► Autre méthode

Pour inverser une matrice de déterminant non nul, ou dont on connaît le caractère inversible, on peut aussi utiliser **l'algorithme Fang-Shen**, ou **algorithme de Gauss-Jordan**, ou **méthode du pivot de Gauss** ; si on pose $X = (x_1, \dots, x_n)$, inverser une matrice carrée A d'ordre n , de déterminant non nul, revient en effet à résoudre le système des n équations d'inconnues x_1, \dots, x_n :

$$A X = Y$$

où $Y = (y_1, \dots, y_n) \in \mathbb{R}^n$, ce système pouvant aussi s'écrire : $X = A^{-1} Y$.

Compte tenu de :

$$A X = Y = I_n Y, \quad I_n X = A^{-1} Y$$

on crée un tableau à n lignes et $2n$ colonnes en bordant la matrice A par la matrice identité I_n :

$$(A | I_n) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} & 0 & \cdots & 1 \end{pmatrix}$$

La transformation de Gauss-Jordan consiste à transformer ce système en un système équivalent dont le bloc gauche est l'identité, c'est-à-dire qu'il faut modifier la matrice $(A | I_n)$ pour qu'elle devienne de la forme $(I_n | A^{-1})$ en utilisant les propriétés de l'algorithme.

On désignera par :

- $A^{(i)}$ la matrice à l'itération i ;
- $L_j^{(i)}$ la $j^{\text{ème}}$ ligne de la matrice à l'itération i .

À l'itération i , il faut obtenir une matrice de la forme :

$$A^{(i)} = \begin{pmatrix} a_{11}^{(i)} & a_{12}^{(i)} & \cdots & a_{1i}^{(i)} & a_{1,i+1}^{(i)} & \cdots & a_{1n}^{(i)} \\ 0 & a_{22}^{(i)} & \cdots & a_{2i}^{(i)} & a_{2,i+1}^{(i)} & \cdots & a_{2n}^{(i)} \\ \vdots & \ddots & \ddots & \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{ii}^{(i)} & a_{i,i+1}^{(i)} & \cdots & a_{in}^{(i)} \\ 0 & \cdots & 0 & 0 & a_{i+1,i+1}^{(i)} & \cdots & a_{i+1,n}^{(i)} \\ 0 & \cdots & 0 & 0 & \vdots & \cdots & \vdots \\ 0 & \cdots & 0 & 0 & a_{n,i+1}^{(i)} & \cdots & a_{nn}^{(i)} \end{pmatrix}$$

avec $a_{11}^{(i)} \neq 0, \dots, a_{ii}^{(i)} \neq 0, a_{i+1,i+1}^{(i)} \neq 0$.

Pour passer de la matrice $A^{(i)}$ à la matrice $A^{(i+1)}$, il faut faire apparaître des 0 à partir de la ligne $i+2$ sur la colonne $i+1$, ce que l'on fait en remplaçant la ligne $L_j^{(i)}$, $j \geq i+2$, par la combinaison linéaire :

$$L_j^{(i)} - \frac{a_{j,i+1}^{(i)}}{a_{i+1,i+1}^{(i)}} L_{i+1}^{(i)}$$

Le nombre $a_{i+1,i+1}^{(i)}$ est appelé le *pivot de Gauss*.

Il existe plusieurs stratégies pour choisir le pivot, de la plus économique en temps de calcul à la plus robuste en terme d'erreurs numériques :

1. la méthode la plus rapide est, à l'itération i , de rechercher dans la colonne i le premier élément non nul.
2. Il est aussi possible de rechercher, dans la colonne i , le plus grand élément, et de le choisir comme pivot.

On est ainsi amené à effectuer des permutations de lignes uniquement. Il est à noter que le pivot doit nécessairement se trouver dans la partie inférieure de la matrice.

Exemple

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix}.$$

Pour inverser A par opérations élémentaires, on commence donc par écrire :

$$(A | I_2) = \left(\begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix} \middle| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

Les opérations à effectuer sont les suivantes :

$$\left(\begin{pmatrix} 1 & -2 \\ 0 & -1 \end{pmatrix} \middle| \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \right)$$

$$\text{Ligne 2} \leftarrow \text{Ligne 2} + 2 \times \text{Ligne 1}$$

$$\text{puis : } \left(\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \middle| \begin{pmatrix} 1 & 0 \\ -2 & -1 \end{pmatrix} \right)$$

$$\text{Ligne 2} \leftarrow -\text{Ligne 2}$$

$$\text{puis : } \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \middle| \begin{pmatrix} -3 & -2 \\ -2 & -1 \end{pmatrix} \right)$$

$$\text{Ligne 1} \leftarrow \text{Ligne 1} + 2 \times \text{Ligne 2}$$

Il ne faut pas oublier de vérifier que le calcul est correct :

$$\begin{pmatrix} 1 & -2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} -3 & -2 \\ -2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\text{On a donc bien : } A^{-1} = \begin{pmatrix} -3 & -2 \\ -2 & -1 \end{pmatrix}.$$

3. Inverse d'un produit de matrices (de $\mathcal{GL}_n(\mathbb{R})$)

Si A et B sont dans $\mathcal{GL}_n(\mathbb{R})$:

$$(A B)^{-1} = B^{-1} A^{-1}$$

Démonstration : A et B étant dans $\mathcal{GL}_n(\mathbb{R})$, leurs déterminants respectifs sont non nuls.

Il en résulte : $\det(A B) = \det A \det B \neq 0$.

La matrice produit AB est donc inversible. Sa matrice inverse $(A B)^{-1}$ est telle que :

$$AB(A B)^{-1} = I_n$$

En multipliant à gauche par A^{-1} , on en déduit : $A^{-1} A B (A B)^{-1} = B (A B)^{-1} = A^{-1}$.

De même, en multipliant à gauche par B^{-1} , on en déduit :

$$B^{-1} B (A B)^{-1} = (A B)^{-1} = B^{-1} A^{-1}.$$

On vérifie sans peine que :

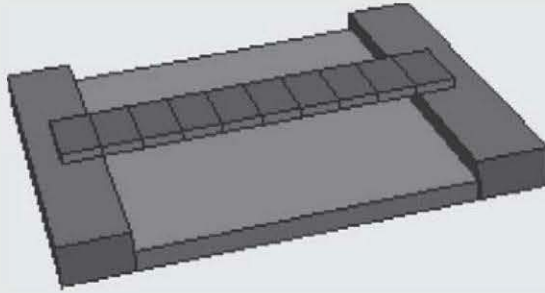
$$(A B) B^{-1} A^{-1} = B^{-1} A^{-1} A B = I_n$$

■



Calcul numérique

Étudions les vibrations verticales d'une passerelle soumise à un chargement vertical \vec{F} : à cet effet, on divise la passerelle en N éléments rectangulaires. Pour tout i de $\{1, \dots, N\}$, on désigne par u_i le déplacement de l'élément i , et par f_i l'effort subi par l'élément i .



La passerelle.

ω est la pulsation associée aux vibrations de la passerelle, et on considère le vecteur des N déplacements verticaux u_1, \dots, u_N : $\begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix}$, ainsi que le vecteur des efforts s'appliquant sur chacun des N éléments, $\begin{pmatrix} \text{colonnes } f_1 \\ \vdots \\ f_N \end{pmatrix}$.

Le problème des vibrations de la passerelle, modélisée par la réunion des N éléments rectangulaires, peut s'écrire sous la forme matricielle suivante :

$$M \omega^2 U - K U = F$$

ou encore :

$$(\omega^2 M - K) U = F$$

Ainsi, les N déplacements verticaux peuvent être calculés en résolvant le système linéaire précédent ! Plus l'entier N retenu sera grand, plus grande sera la précision.

Génétique

► La construction d'une matrice d'hérédité

Considérons une population pour laquelle la distribution, pour la génération n , des génotypes suivant p gènes donnés $\text{gène}_1, \text{gène}_2, \dots, \text{gène}_p$, $p \in \mathbb{N}^*$, est donnée par

le vecteur : $X_n = \begin{pmatrix} x_{n,1} \\ x_{n,2} \\ \vdots \\ x_{n,p} \end{pmatrix}$ avec : $\sum_{k=1}^p x_{n,k} = 1$ (c'est-à-dire une répartition de 100 %).

Les divers croisements de gènes pouvant avoir lieu peuvent être récapitulés dans une matrice de taille $p \times p$, \mathcal{H} , dite matrice d'hérédité, qui permet d'obtenir la distribution, pour la génération $n + 1$, des génotypes, en fonction de celle de l'année n :

$$X_{n+1} = \mathcal{H} X_n$$

Si on connaît le vecteur X_0 donnant la répartition initiale des génotypes, il est donc possible, par récurrence, d'exprimer explicitement le vecteur X_n donnant la répartition des génotypes pour la génération n : $X_n = \mathcal{H} X_{n-1} = \mathcal{H}^2 X_{n-2} = \dots = \mathcal{H}^n X_0$. Il est clair que si la matrice \mathcal{H} est diagonalisable, et donc semblable à une matrice diagonale D , le calcul sera facilité !

► **Comment disparaissent des gènes**

Prenons le cas où $p = 4$, et où la matrice \mathcal{H} est donnée par : $\mathcal{H} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 0 & 1 & 0 \\ \frac{3}{8} & -\frac{3}{8} & 0 & \frac{1}{8} \end{pmatrix}$

avec : $X_0 = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{6} \\ \frac{1}{12} \end{pmatrix}$. La matrice \mathcal{H} est diagonalisable. Ses valeurs propres sont $1, \frac{1}{2}, \frac{1}{4}$, et $\frac{1}{8}$; si on désigne par P la matrice de passage à la base des vecteurs propres (dans chaque colonne, on retrouve les coordonnées des vecteurs engendrant les sous-

espaces propres respectivement associés à chaque valeur propre) : $P = \begin{pmatrix} 0 & 1 & -3 & 0 \\ 0 & 0 & -3 & 0 \\ 1 & -\frac{1}{2} & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$

et D la matrice diagonale : $D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{8} \end{pmatrix}$ alors, compte-tenu de : $\mathcal{H} = P D P^{-1}$

on en déduit : $X_n = (P D P^{-1})^n X_0 = P D P^{-1} P D P^{-1} \dots P D P^{-1} X_0 = P D^n P^{-1} X_0$ soit :

$$X_n = P \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2^n} & 0 & 0 \\ 0 & 0 & \frac{1}{4^n} & 0 \\ 0 & 0 & 0 & \frac{1}{8^n} \end{pmatrix} P^{-1} X_0 = \begin{pmatrix} \frac{1}{2^{n+2}} + \frac{1}{2} \left(\frac{1}{4^n} - \frac{1}{2^n} \right) \\ \frac{1}{2^{2n+1}} \\ \frac{5}{24} + \frac{1}{4} \frac{1}{2^{n+1}} - \frac{1}{6} \frac{1}{4^n} \\ \frac{1}{3} \frac{1}{2^{3n}} - \frac{1}{4} \frac{1}{2^n} \end{pmatrix}$$

Il en résulte : $\lim_{n \rightarrow +\infty} X_n = \begin{pmatrix} 0 \\ 0 \\ \frac{5}{24} \\ 0 \end{pmatrix}$, ce qui signifie qu'au bout d'un très grand nombre de générations, il ne reste plus dans la population que des individus portant le gène 3.

1. Définitions

1. m et n étant deux entiers naturels non nuls, on appelle **système linéaire** de m équations, d'inconnues x_1, \dots, x_n :

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases}$$

Les a_{ij} , $1 \leq i \leq m$, $1 \leq j \leq n$ sont appelés **coefficients du système**.

Tout n -uplet (x_1, \dots, x_n) vérifiant les m équations précédentes est appelé **solution** du système.

2. Un système linéaire, d'inconnues x_1, \dots, x_n , de la forme :

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = 0 \end{cases}$$

est dit **homogène**¹.

3. Le système homogène associé au système linéaire : $\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases}$

est : $\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = 0 \end{cases}$

2. Propriétés

1. Un système linéaire, d'inconnues x_1, \dots, x_n , de la forme :

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases}$$

peut aussi s'écrire sous la forme : $AX = B$

où $A \in \mathcal{M}_{m,n}(\mathbb{R})$ est la matrice $(a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$, et B la matrice

colonne $\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$.

1. Le système est dit « homogène », car si le n -uplet (x_1, \dots, x_n) est solution, alors, pour tout réel λ , le n -uplet $(\lambda x_1, \dots, \lambda x_n)$ est également solution.

2. Un système linéaire, d'inconnues x_1, \dots, x_n , de la forme :

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases}$$

admet, suivant les cas : soit une solution unique (x_1, \dots, x_n) ; soit une infinité de solutions (qui seront exprimées en fonction d'un ou plusieurs paramètres) ; soit aucune solution.

3. Un système linéaire, d'inconnues x_1, \dots, x_n , de la forme :

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = b_n \end{cases}$$

admet une solution unique si et seulement si la matrice : $\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$

est inversible.

Corollaire

Un système linéaire, d'inconnues x_1, \dots, x_n , de la forme : $\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = b_n \end{cases}$

admet une solution unique si et seulement si : $\det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \neq 0$

Un tel système est appelé **système de Cramer**.

4. La solution d'un système de Cramer

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n = b_n \end{cases}$$

est obtenue grâce aux **formules de Cramer** :

$\forall i \in \{1, \dots, n\}$:

$$x_i = \frac{\det(C_1(A), \dots, C_{i-1}(A), B, C_{i+1}(A), \dots, C_n(A))}{\det A}$$

où, pour tout j de $\{1, \dots, n\}$, $C_j(A)$ désigne la $j^{\text{ième}}$ colonne de la matrice A .

Démonstration : Comme

$$\begin{aligned} & \det(C_1(A), \dots, C_{i-1}(A), B, C_{i+1}(A), \dots, C_n(A)) \\ &= \det(C_1(A), \dots, C_{i-1}(A), AX, C_{i+1}(A), \dots, C_n(A)) \\ &= \det(C_1(A), \dots, C_{i-1}(A), \sum_{i=1}^n x_i C_i(A), C_{i+1}(A), \dots, C_n(A)) \end{aligned}$$

on utilise tout simplement les propriétés d'*invariance par transvection du déterminant* ; le déterminant ne change pas si, à une colonne donnée, on ajoute une combinaison linéaire des autres colonnes :

$\forall i \in \{1, \dots, n\}$:

$$\begin{aligned} & \det(C_1(A), \dots, C_{i-1}(A), B, C_{i+1}(A), \dots, C_n(A)) \\ &= \det(C_1(A), \dots, C_{i-1}(A), x_i C_i(A), C_{i+1}(A), \dots, C_n(A)) \\ &= x_i \det(C_1(A), \dots, C_{i-1}(A), C_i(A), C_{i+1}(A), \dots, C_n(A)) \\ &= x_i \det A \end{aligned}$$

■

3. Réduction des systèmes linéaires

Théorème

Tout système linéaire de m équations à n inconnues, $AX = B$, $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $B \in \mathcal{M}_{m,1}(\mathbb{R})$, peut s'écrire sous la forme : $\tilde{A}X = \tilde{B}$

où $\tilde{B} \in \mathcal{M}_{m,1}(\mathbb{R})$, et où $\tilde{A} \in \mathcal{M}_{m,n}(\mathbb{R})$ est une matrice de la forme :

$$\tilde{A} = \begin{pmatrix} 1 & * & * & * & * & * & * & * \\ & & 1 & * & * & * & * & * \\ & & & 1 & * & * & * & * \\ & & & & \dots & \dots & \dots & \dots \\ & & & & & & \dots & \dots \end{pmatrix}$$

Les coefficients non écrits sont nuls, les autres sont désignés par des astérisques.



On peut qualifier la matrice \tilde{A} de « **matrice en escalier** », dans la mesure où les 1 forment un escalier ; cet escalier possède donc des marches de profondeurs non nécessairement égales. Au fond des « marches », se trouvent les « 1 ». Sous les « marches », les zéros (placard aux zéros). Au-dessus des marches, des coefficients absolument quelconques.

► Réduction des systèmes linéaires d'ordre n

Théorème

Tout système linéaire de n équations à n inconnues, $AX = B$, $A \in \mathcal{M}_n(\mathbb{R})$, $B \in \mathcal{M}_{n,1}(\mathbb{R})$, peut s'écrire sous la forme : $TX = B'$

où $T \in \mathcal{M}_n(\mathbb{R})$ est une matrice triangulaire supérieure dont les termes diagonaux sont soit nuls, soit égaux à 1, et où $B' \in \mathcal{M}_{n,1}(\mathbb{R})$.

Démonstration : La démonstration se fait par l'algorithme du pivot de Gauss. ■

Soient a , b et c trois réels strictement positifs. On considère la réaction chimique



La conservation des éléments conduit au système suivant :

$$\text{Cr} : 1 \times a + 0 \times b = 2 \times c$$

$$\text{O} : 0 \times a + 2 \times b = 3 \times c$$

On a alors :

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2c \\ 3c \end{pmatrix}$$

ce qui permet, en résolvant le système linéaire ainsi obtenu, de déterminer les expressions de a et b en fonction de c :

$$a = 2c, \quad b = \frac{3c}{2}$$

Pour des systèmes d'intérêt biologique indéniable, où les molécules en jeu présentent des formules autrement plus compliquées, il est très utile de vérifier un certain produit matriciel. Considérons la réaction entre le glucose et l'ATP (adénosine triphosphate) produisant du glucose-6-phosphate et de l'ADP (adénosine diphosphate) (il s'agit de la toute première étape dans le processus de glycolyse ; dégradation du glucose et production d'énergie) :



Les formules chimiques brutes de chacun de ces composés, respectivement, sont les suivantes : $\text{C}_6\text{H}_{12}\text{O}_6$, $\text{C}_{10}\text{H}_{16}\text{N}_5\text{O}_{13}\text{P}_3$, $\text{C}_6\text{H}_{13}\text{O}_9\text{P}$ et $\text{C}_{10}\text{H}_{15}\text{N}_5\text{O}_{10}\text{P}_2$.

On introduit alors la matrice C :

$$\begin{pmatrix} 6 & 10 & 6 & 10 \\ 12 & 16 & 13 & 15 \\ 6 & 13 & 9 & 10 \\ 0 & 3 & 1 & 2 \\ 0 & 5 & 0 & 5 \end{pmatrix}$$

composée du nombre d'éléments chimiques simples (en ligne, respectivement, C, H, O, P, N) contenus dans chaque molécule (en colonne, respectivement, Glucose, ATP, Glucose-6-phosphate et ADP), ainsi que le vecteur colonne \mathcal{V} :

$$\begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$$

constitué des coefficients algébriques stœchiométriques (négatifs pour les réactifs, positifs pour les produits) relatant simplement le fait qu'une molécule de glucose réagisse avec une molécule d'ATP pour donner une molécule de glucose-6-phosphate et une molécule d'ADP.

Alors, le produit matriciel $C\mathcal{V}$ doit être (par respect de la conservation de la matière) le vecteur colonne dont tous les éléments sont nuls ; si cela n'était pas le cas, cela voudrait dire qu'il y aurait eu erreur dans le décompte des nombres d'atomes.

1. Définitions et propriétés fondamentales

On appelle **vecteur à 3 composantes**, un triplet de réels $\vec{u} = (x, y, z)$, que l'on peut aussi écrire sous la forme $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$.

x, y et z sont appelées *composantes* du vecteur \vec{u} .

► Égalité de deux vecteurs

Le vecteur, dont les trois composantes sont nulles, est appelé **vecteur nul**, et noté $\vec{0}$.

Deux vecteurs sont égaux si et seulement s'ils ont les mêmes composantes.

► Somme de deux vecteurs

On considère les vecteurs $\vec{u} = (x_{\vec{u}}, y_{\vec{u}}, z_{\vec{u}})$ et $\vec{v} = (x_{\vec{v}}, y_{\vec{v}}, z_{\vec{v}})$ de l'espace \mathbb{R}^3 .

On définit le vecteur $\vec{u} + \vec{v} = \vec{w}$ comme l'unique vecteur de \mathbb{R}^3 , dont les composantes $(x_{\vec{w}}, y_{\vec{w}}, z_{\vec{w}})$ sont données par :

$$\begin{cases} x_{\vec{w}} = x_{\vec{u}} + x_{\vec{v}} \\ y_{\vec{w}} = y_{\vec{u}} + y_{\vec{v}} \\ z_{\vec{w}} = z_{\vec{u}} + z_{\vec{v}} \end{cases}$$

► Multiplication d'un vecteur par un réel

On considère le vecteur $\vec{u} = (x_{\vec{u}}, y_{\vec{u}}, z_{\vec{u}})$ de l'espace \mathbb{R}^3 .

Alors, pour tout réel λ , $\lambda \vec{u}$ est un vecteur de \mathbb{R}^3 , de composantes $(\lambda x_{\vec{u}}, \lambda y_{\vec{u}}, \lambda z_{\vec{u}})$.

2. Familles de vecteurs libres (ou famille libre de vecteurs)

► Combinaison de vecteurs

On appelle **combinaison linéaire de deux vecteurs \vec{u} et \vec{v}** , toute expression de la forme :

$$a \vec{u} + b \vec{v}$$

où a et b sont des réels.

n étant un entier naturel supérieur ou égal à 2, on appelle **combinaison linéaire de n vecteurs $\vec{u}_1, \dots, \vec{u}_n$** , toute expression de la forme :

$$\sum_{i=1}^n a_i \vec{u}_i$$

où a_1, \dots, a_n sont des réels.

► Vecteurs liés

Deux vecteurs \vec{u} et \vec{v} sont dits **liés** s'il existe une combinaison linéaire non triviale de ces vecteurs égale au vecteur nul, c'est-à-dire lorsque les coefficients de la combinaison sont non simultanément nuls, de la forme :

$$a\vec{u} + b\vec{v} = \vec{0} \quad (a, b) \neq (0, 0)$$



Dire que deux vecteurs \vec{u} et \vec{v} sont liés revient à dire qu'ils sont *proportionnels*, ou encore *colinéaires*.

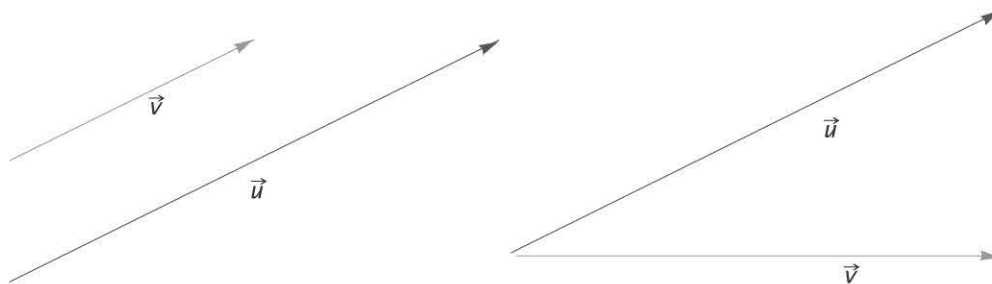


Figure 83.1 – Deux vecteurs liés, et deux vecteurs libres

Deux vecteurs \vec{u} et \vec{v} sont dits **libres** s'ils ne sont pas liés.

► Généralisation

Trois vecteurs \vec{u} , \vec{v} et \vec{w} sont dits **liés** s'il existe une combinaison linéaire non triviale de ces vecteurs égale au vecteur nul, c'est-à-dire lorsque les coefficients de la combinaison sont non simultanément nuls, de la forme :

$$a\vec{u} + b\vec{v} + c\vec{w} = \vec{0} \quad (a, b, c) \neq (0, 0, 0)$$

n étant un entier naturel supérieur ou égal à 2, n vecteurs $\vec{u}_1, \dots, \vec{u}_n$ sont dits **liés** s'il existe une combinaison linéaire non triviale de ces vecteurs égale au vecteur nul, c'est-à-dire lorsque les coefficients de la combinaison sont non simultanément nuls, de la forme :

$$\sum_{i=1}^n a_i \vec{u}_i = \vec{0}$$

où a_1, \dots, a_n sont des réels tels que $(a_1, \dots, a_n) \neq (0, \dots, 0)$.

Trois vecteurs \vec{u} , \vec{v} et \vec{w} sont dits **libres** s'ils ne sont pas liés.

n étant un entier naturel supérieur ou égal à 2, n vecteurs $\vec{u}_1, \dots, \vec{u}_n$ sont dits **libres** s'ils ne sont pas liés



Attention ! Trois vecteurs liés ne seront pas nécessairement deux à deux *colinéaires* : la seule chose que l'on peut dire, c'est qu'il existe une combinaison linéaire non triviale de ces vecteurs, qui est égale au vecteur nul.

Exemple

Soit un triangle non aplati¹ de sommets A, B, C :

$$\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CA} = \vec{0}$$

Les vecteurs \overrightarrow{AB} , \overrightarrow{BC} et \overrightarrow{CA} sont liés, mais non colinéaires ! Comme ils sont dans un même plan, ils sont **coplanaires**.

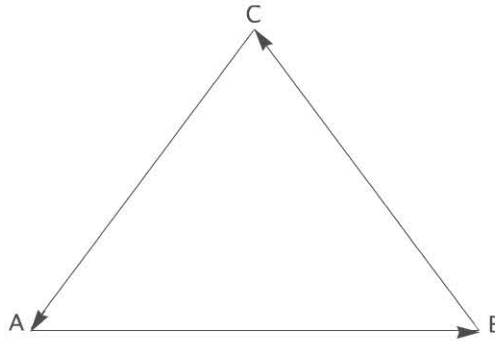


Figure 83.2 – Trois vecteurs liés

Propriété

Pour toute famille $(\vec{u}, \vec{v}, \vec{w})$ de vecteurs libres :

$$a\vec{u} + b\vec{v} + c\vec{w} = \vec{0} \Rightarrow (a, b, c) = (0, 0, 0)$$

ce qui signifie que, s'il existe une combinaison linéaire nulle de ces vecteurs, alors les coefficients de cette même combinaison linéaire sont **nécessairement** nuls.

Exemple

On considère les vecteurs $\vec{u} = (1, 2, 3)$, $\vec{v} = (3, 2, 1)$ et $\vec{w} = (-1, 2, 1)$.

Pour déterminer s'ils sont liés ou non, on cherche s'il existe des réels a, b et c tels que :

$$a\vec{u} + b\vec{v} + c\vec{w} = \vec{0}$$

soit :

$$(a + 3b - c, 2a + 2b + 2c, 3a + b + c) = (0, 0, 0)$$

ce qui implique donc, puisque deux vecteurs sont égaux si et seulement si leurs composantes sont égales :

$$\begin{cases} a + 3b - c = 0 \\ 2a + 2b + 2c = 0 \\ 3a + b + c = 0 \end{cases}$$

puis

$$a = b = c = 0$$

Les vecteurs $\vec{u} = (1, 2, 3)$, $\vec{v} = (3, 2, 1)$ et $\vec{w} = (-1, 2, 1)$ sont donc linéairement indépendants.

3. Repères

Définition

O étant un point de l'espace, et $\vec{i}, \vec{j}, \vec{k}$ trois vecteurs non coplanaires, on dit que $(O; \vec{i}, \vec{j}, \vec{k})$ est un **repère de l'espace**.

Proposition

Soit $(O; \vec{i}, \vec{j}, \vec{k})$ un repère de l'espace. Alors, pour tout point M de l'espace, il existe trois réels x, y, z , chacun de ces réels étant défini de manière unique, tels que :

$$\overrightarrow{OM} = x\vec{i} + y\vec{j} + z\vec{k}$$

(x, y, z) sont les coordonnées du vecteur \overrightarrow{OM} dans le repère $(O; \vec{i}, \vec{j}, \vec{k})$.

Proposition

Soit $(O; \vec{i}, \vec{j}, \vec{k})$ un repère de l'espace. Alors, tout vecteur \vec{u} de l'espace peut s'écrire de manière unique sous la forme :

$$\vec{u} = x\vec{i} + y\vec{j} + z\vec{k}$$

avec $(x, y, z) \in \mathbb{R}^3$.



x, y, z étant des réels, il est essentiel de faire la distinction entre le point M de l'espace, de coordonnées (x, y, z) , et le vecteur \vec{u} , de coordonnées (x, y, z) .

La position du point M est fixe, et entièrement déterminée par x, y et z , puisque l'origine du vecteur \overrightarrow{OM} , qui est le point O , est fixe.

Le vecteur \vec{u} , de coordonnées (x, y, z) , n'a pas d'origine fixe : pour le placer, on peut choisir n'importe quel point de l'espace comme origine.

1. Barycentre d'un système pondéré de deux points

Définition

Étant donnés deux réels α et β tels que $\alpha + \beta \neq 0$, et deux points $A(x_A, y_A, z_A)$ et $B(x_B, y_B, z_B)$ de l'espace \mathbb{R}^3 , on appelle **barycentre** du système pondéré $\{(A, \alpha), (B, \beta)\}$ le point G tel que :

$$\alpha \overrightarrow{GA} + \beta \overrightarrow{GB} = \vec{0}$$



Lorsque $\alpha = \beta = \frac{1}{2}$, le barycentre du système pondéré $\{(A, \alpha), (B, \beta)\}$ est bien sûr le milieu du segment $[A, B]$!

Théorème

Étant donnés deux réels α et β tels que $\alpha + \beta \neq 0$, et deux points $A(x_A, y_A, z_A)$ et $B(x_B, y_B, z_B)$ de l'espace \mathbb{R}^3 , un point G de l'espace est le barycentre du système pondéré $\{(A, \alpha), (B, \beta)\}$ si et seulement, pour tout point M de l'espace :

$$\alpha \overrightarrow{MA} + \beta \overrightarrow{MB} = (\alpha + \beta) \overrightarrow{MG}$$

Démonstration : Si G est le barycentre du système pondéré $\{(A, \alpha), (B, \beta)\}$, alors :

$$\alpha \overrightarrow{GA} + \beta \overrightarrow{GB} = \vec{0}$$

La relation de Chasles permet d'en déduire :

$$\alpha (\overrightarrow{GM} + \overrightarrow{MA}) + \beta (\overrightarrow{GM} + \overrightarrow{MB}) = \vec{0}$$

ce qui conduit bien à :

$$\alpha \overrightarrow{MA} + \beta \overrightarrow{MB} = (\alpha + \beta) \overrightarrow{MG}$$

Réciproquement, si le point G est tel que, pour tout point M de l'espace :

$$\alpha \overrightarrow{MA} + \beta \overrightarrow{MB} = (\alpha + \beta) \overrightarrow{MG}$$

on en déduit, pour $M = G$:

$$\alpha \overrightarrow{GA} + \beta \overrightarrow{GB} = \vec{0}$$

■

► Coordonnées cartésiennes du barycentre d'un système pondéré de deux points

Étant donnés deux réels α et β tels que $\alpha + \beta \neq 0$, et deux points $A(x_A, y_A, z_A)$ et $B(x_B, y_B, z_B)$ de l'espace \mathbb{R}^3 , le **barycentre** G du système pondéré $\{(A, \alpha), (B, \beta)\}$ a pour coordonnées :

$$\left(\frac{\alpha x_A + \beta x_B}{\alpha + \beta}, \frac{\alpha y_A + \beta y_B}{\alpha + \beta}, \frac{\alpha z_A + \beta z_B}{\alpha + \beta} \right)$$

2. Barycentre d'un système pondéré de n points, $n \in \mathbb{N}^*$

Définition

Étant donnés n réels $\alpha_1, \dots, \alpha_n$ tels que $\sum_{i=1}^n \alpha_i \neq 0$, et n points $A_1(x_1, y_1, z_1), \dots, A_n(x_n, y_n, z_n)$ de l'espace \mathbb{R}^3 , on appelle **barycentre** du système pondéré $\{(A_1, \alpha_1), \dots, (A_n, \alpha_n)\}$ le point G tel que :

$$\sum_{i=1}^n \alpha_i \overrightarrow{GA_i} = \vec{0}$$

Théorème

Étant donnés n réels $\alpha_1, \dots, \alpha_n$ tels que $\sum_{i=1}^n \alpha_i \neq 0$, et n points $A_1(x_1, y_1, z_1), \dots, A_n(x_n, y_n, z_n)$ de l'espace \mathbb{R}^3 , un point G de l'espace est le barycentre du système pondéré $\{(A_1, \alpha_1), \dots, (A_n, \alpha_n)\}$ si et seulement, pour tout point M de l'espace :

$$\sum_{i=1}^n \alpha_i \overrightarrow{MA_i} = \left(\sum_{i=1}^n \alpha_i \right) \overrightarrow{MG}$$

Démonstration : Si G est le barycentre du système pondéré $\{(A_1, \alpha_1), \dots, (A_n, \alpha_n)\}$, alors :

$$\sum_{i=1}^n \alpha_i \overrightarrow{GA_i} = \vec{0}$$

La relation de Chasles permet d'en déduire :

$$\sum_{i=1}^n \alpha_i (\overrightarrow{GM} + \overrightarrow{MA_i}) = \vec{0}$$

ce qui conduit bien à :

$$\sum_{i=1}^n \alpha_i \overrightarrow{MA_i} = \left(\sum_{i=1}^n \alpha_i \right) \overrightarrow{MG}$$

Réciproquement, si le point G est tel que, pour tout point M de l'espace :

$$\sum_{i=1}^n \alpha_i \overrightarrow{MA_i} = \left(\sum_{i=1}^n \alpha_i \right) \overrightarrow{MG}$$

on en déduit, pour $M = G$:

$$\sum_{i=1}^n \alpha_i \overrightarrow{GA_i} = \vec{0}$$

■

► **Coordonnées cartésiennes du barycentre d'un système pondéré de n points, $n \in \mathbb{N}^*$**

Étant donnés n réels $\alpha_1, \dots, \alpha_n$ tels que $\sum_{i=1}^n \alpha_i \neq 0$, et n points $A_1(x_1, y_1, z_1), \dots, A_n(x_n, y_n, z_n)$ de l'espace \mathbb{R}^3 , le **barycentre** G du système pondéré $\{(A_1, \alpha_1), \dots, (A_n, \alpha_n)\}$ a pour coordonnées :

$$\left(\frac{\sum_{i=1}^n \alpha_i x_i}{\sum_{i=1}^n \alpha_i}, \frac{\sum_{i=1}^n \alpha_i y_i}{\sum_{i=1}^n \alpha_i}, \frac{\sum_{i=1}^n \alpha_i z_i}{\sum_{i=1}^n \alpha_i} \right)$$

3. Isobarycentres

► **Isobarycentre d'un ensemble de n points, $n \in \mathbb{N}^*$**

Étant donnés n points $A_1(x_1, y_1, z_1), \dots, A_n(x_n, y_n, z_n)$ de l'espace \mathbb{R}^3 , on appelle **isobarycentre** des points A_1, \dots, A_n le point G tel que :

$$\sum_{i=1}^n \overrightarrow{GA_i} = \vec{0}$$

► **Isobarycentre d'un ensemble de 3 points, A, B, C**

Étant donnés 3 points A, B, C de l'espace \mathbb{R}^3 , leur isobarycentre est le centre de gravité du triangle ABC , point de concours des médianes.

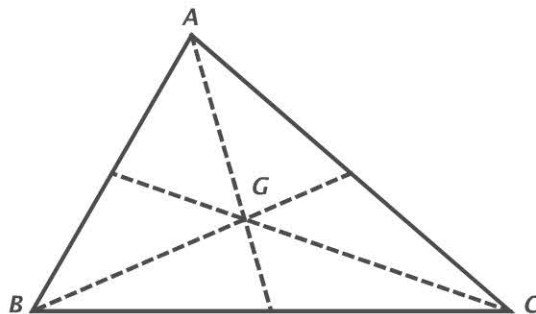


Figure 84.1 – Le centre de gravité G du triangle ABC .

4. Associativité du barycentre d'un système pondéré de trois points

On considère 3 réels $\alpha_1, \alpha_2, \alpha_3$ tels que $\sum_{i=1}^n \alpha_i \neq 0$, et 3 points A_1, A_2, A_3 de l'espace \mathbb{R}^3 . On désigne par $G_{1,2}$ le **barycentre** du système pondéré $\{(A_1, \alpha_1), (A_2, \alpha_2)\}$, par $G_{1,3}$ le

barycentre du système pondéré $\{(A_1, \alpha_1), (A_3, \alpha_3)\}$, et par $G_{2,3}$ le **barycentre** du système pondéré $\{(A_2, \alpha_2), (A_3, \alpha_3)\}$.

Alors, le **barycentre** G du système pondéré $\{(A_1, \alpha_1), (A_2, \alpha_2), (A_3, \alpha_3)\}$ est aussi celui du système pondéré $\{(G_{1,2}, \alpha_1 + \alpha_2), (A_3, \alpha_3)\}$, mais aussi aussi celui du système pondéré $\{(G_{1,3}, \alpha_1 + \alpha_3), (A_2, \alpha_2)\}$, ou encore aussi celui du système pondéré $\{(G_{2,3}, \alpha_2 + \alpha_3), (A_1, \alpha_1)\}$.



Ce résultat est, bien sûr, généralisable à un système pondéré de n points, $n \in \mathbb{N}^*$.

Exemple

Considérons un rectangle $ABCD$; on désigne par I le milieu du segment $[A, B]$, J le milieu du segment $[C, D]$; le centre de gravité du rectangle, qui est l'isobarycentre des points A, B, C et D , est aussi celui des points I et J , respectivement isobarycentres des points A, B d'une part, C et D d'autre part.

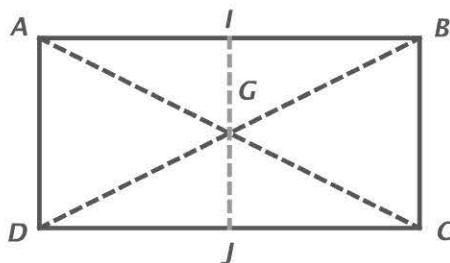


Figure 84.2 – Le rectangle $ABCD$, son centre de gravité G , et les milieux I et J des segments $[A, B]$ et $[C, D]$.

1. Droites

Définition

On appelle **droite vectorielle** engendrée par un vecteur non nul \vec{u} l'ensemble :

$$D = \mathbb{R} \vec{u} = \{\lambda \vec{u}, \lambda \in \mathbb{R}\}$$

Définition

On appelle **droite affine** passant par le point A et dirigée par le vecteur \vec{u} (ou encore par la droite vectorielle engendrée par \vec{u}) l'ensemble

$$\mathcal{D} = A + \mathbb{R} \vec{u} = \{A + \lambda \vec{u}, \lambda \in \mathbb{R}\}$$



Une *droite affine* passant par l'origine O peut aussi être considérée comme une *droite vectorielle*.

Exemple

La *droite affine* passant par le point A de coordonnées $(1, 1, 1)$ et dirigée par le vecteur $\vec{u} = (1, 2, 3)$ est donc l'ensemble

$$\begin{aligned} \mathcal{D} &= A + \mathbb{R} \vec{u} \\ &= \{A + \lambda \vec{u}, \lambda \in \mathbb{R}\} \\ &= \{M(x, y, z) / x = 1 + \lambda, y = 1 + 2\lambda, z = 1 + 3\lambda, \lambda \in \mathbb{R}\} \end{aligned}$$

c'est-à-dire l'ensemble des points M de coordonnées (x, y, z) tels que :

$$\begin{cases} x = 1 + \lambda \\ y = 1 + 2\lambda \\ z = 1 + 3\lambda \end{cases}, \lambda \in \mathbb{R}$$

On a ainsi obtenu une représentation paramétrique, ou paramétrage, de \mathcal{D} .

En éliminant le paramètre λ , par exemple, en utilisant la première relation, $\lambda = x - 1$, on en déduit :

$$\begin{cases} y = 1 + 2\lambda = 1 + 2x - 2 = 2x - 1 \\ z = 1 + 3\lambda = 1 + 3x - 3 = 3x - 2 \end{cases}$$

on obtient un système d'équations cartésiennes de \mathcal{D} .

Inversement, il serait aussi possible, à partir d'un *système d'équations cartésiennes* de \mathcal{D} , d'en déduire une représentation paramétrique.



Attention ! Dans l'espace, une droite, qui est aussi l'intersection de deux plans, est donc définie par deux équations, par contre, dans le plan rapporté à un repère $(O; \vec{i}, \vec{j})$, une droite est définie par une seule équation. On rappelle à cet effet qu'étant donnés trois réels a, b et c tels que $(a, b) \neq (0, 0)$, et une droite affine \mathcal{D} , d'équation cartésienne $ax + by + c = 0$, un vecteur directeur de \mathcal{D} est $\vec{u} = (-b, a)$.

2. Plans

Définition

On appelle **plan vectoriel** engendré par deux vecteurs libres \vec{u} et \vec{v} l'ensemble :

$$P = \mathbb{R} \vec{u} + \mathbb{R} \vec{v} = \{ \lambda \vec{u} + \mu \vec{v}, (\lambda, \mu) \in \mathbb{R}^2 \}$$

C'est donc l'ensemble de toutes les combinaisons linéaires possibles (c'est-à-dire donc une infinité !) des vecteurs \vec{u} et \vec{v} .



Un plan vectoriel est à comprendre comme une *direction commune de plans, parallèles deux à deux*.

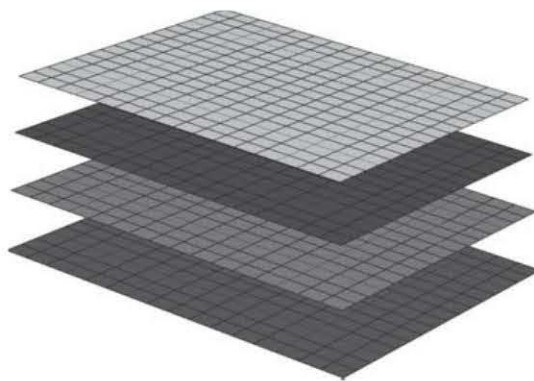


Figure 85.1 – Un plan vectoriel.

Définition

On appelle **plan affine** passant par le point A et dirigé par deux vecteurs libres \vec{u} et \vec{v} l'ensemble :

$$\mathcal{P} = A + \mathbb{R} \vec{u} + \mathbb{R} \vec{v} = \{ A + \lambda \vec{u} + \mu \vec{v}, (\lambda, \mu) \in \mathbb{R}^2 \}$$

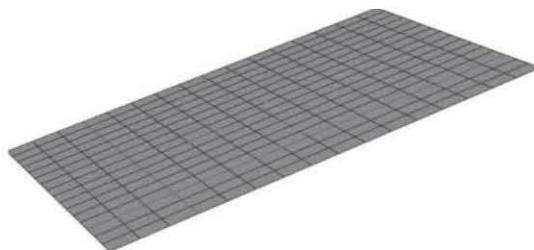


Figure 85.2 – Un plan affine.



Un *plan affine* passant par l'origine O peut aussi être considéré comme un *plan vectoriel*.

Exemple

Le *plan affine* passant par le point A de coordonnées $(1, 1, 1)$ et dirigé par les vecteurs $\vec{u} = (1, 2, 3)$ et $\vec{v} = (3, 2, 1)$ est donc l'ensemble :

$$\mathcal{P} = A + \mathbb{R}\vec{u} + \mathbb{R}\vec{v} = \{A + \lambda\vec{u} + \mu\vec{v}, (\lambda, \mu) \in \mathbb{R}^2\}$$

soit encore :

$$\mathcal{P} = \{M(x, y, z) / x = 1 + \lambda + 3\mu, y = 1 + 2\lambda + 2\mu, z = 1 + 3\lambda + \mu, (\lambda, \mu) \in \mathbb{R}^2\}$$

c'est-à-dire l'ensemble des points M de coordonnées (x, y, z) tels que :

$$\begin{cases} x = 1 + \lambda + 3\mu \\ y = 1 + 2\lambda + 2\mu \\ z = 1 + 3\lambda + \mu \end{cases}, (\lambda, \mu) \in \mathbb{R}^2$$

On a ainsi obtenu une représentation paramétrique, ou paramétrage, de \mathcal{P} .

En éliminant les paramètres λ et μ , par exemple, en utilisant les deux premières relations, on en déduit :

$$x - 2y + z = 0$$

On obtient alors une *équation cartésienne* de \mathcal{P} .

Inversement, il serait aussi possible, à partir d'une *équation cartésienne* de \mathcal{P} , d'en déduire une représentation paramétrique.

1. Définitions

► Produit scalaire

On appelle **produit scalaire** de deux vecteurs \vec{u} et \vec{v} , de composantes respectives (u_1, u_2, u_3) et (v_1, v_2, v_3) , le nombre, aussi appelé « scalaire » :

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

► Norme euclidienne

On appelle **norme euclidienne** d'un vecteur \vec{u} , de composantes (u_1, u_2, u_3) , le réel positif ou nul :

$$\|\vec{u}\| = \sqrt{u_1^2 + u_2^2 + u_3^2} = \sqrt{\vec{u} \cdot \vec{u}}$$

► Vecteur unitaire

Un vecteur est dit **unitaire** s'il est de norme 1.

► Vecteur orthogonaux

Deux vecteurs sont dits **orthogonaux** si leur produit scalaire est nul.

2. Propriétés

► Propriétés du produit scalaire

Le produit scalaire est une application

- **bilinéaire**, c'est-à-dire linéaire par rapport à chaque variable :

$$\forall (\vec{u}, \vec{v}, \vec{w}) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3, \forall \lambda \in \mathbb{R} : \begin{cases} \vec{u} \cdot (\vec{v} + \lambda \vec{w}) = \vec{u} \cdot \vec{v} + \lambda \vec{u} \cdot \vec{w} \\ (\vec{v} + \lambda \vec{w}) \cdot \vec{u} = \vec{v} \cdot \vec{u} + \lambda \vec{w} \cdot \vec{u} \end{cases}$$

- **symétrique** :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : \vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$$

- **définie positive** : d'une part, la positivité,

$$\forall \vec{u} \in \mathbb{R}^3 : \vec{u} \cdot \vec{u} \geq 0$$

et, d'autre part, le *caractère défini* :

$$\forall \vec{u} \in \mathbb{R}^3 : \vec{u} \cdot \vec{u} = 0 \Leftrightarrow \vec{u} = \vec{0}$$

► Inégalité de Cauchy-Schwarz

Théorème

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : |\vec{u} \cdot \vec{v}| \leq \|\vec{u}\| \|\vec{v}\|$$

et l'égalité a lieu si et seulement si les deux vecteurs sont liés.

Démonstration : On se place dans le cas où $\vec{u} \neq \vec{0}$ et $\vec{v} \neq \vec{0}$, l'inégalité étant toujours vérifiée lorsque \vec{u} ou \vec{v} est égal au vecteur nul (on a alors une égalité).

Pour tout réel t :

$$\|\vec{u} + t\vec{v}\|^2 = (\vec{u} + t\vec{v}) \cdot (\vec{u} + t\vec{v}) = \|\vec{u}\|^2 + 2t\vec{u} \cdot \vec{v} + t^2 \|\vec{v}\|^2$$

On obtient ainsi un trinôme du second degré en t , qui est toujours positif ; son discriminant réduit doit donc être négatif ou nul :

$$(\vec{u} \cdot \vec{v})^2 - \|\vec{u}\|^2 \|\vec{v}\|^2 \leq 0$$

ce qui conduit à :

$$|\vec{u} \cdot \vec{v}| \leq \|\vec{u}\| \|\vec{v}\|$$

Si le discriminant est nul, alors, le trinôme en t admet une racine double, et il existe alors un réel t tel que :

$$\|\vec{u} + t\vec{v}\|^2 = 0$$

soit :

$$\vec{u} + t\vec{v} = \vec{0}$$

Les vecteurs \vec{u} et \vec{v} sont alors liés.

Réciproquement, si $\vec{u} \neq \vec{0}$ et $\vec{v} \neq \vec{0}$ sont liés, il existe un réel λ tel que :

$$\vec{u} = \lambda \vec{v}$$

et on a alors :

$$|\vec{u} \cdot \vec{v}| = |\lambda| \|\vec{v}\|^2 = \|\vec{u}\| \|\vec{v}\|$$

■

► Propriétés d'une norme

Une *norme* sur \mathbb{R}^3 est une application à valeurs dans \mathbb{R}^+ qui doit vérifier les trois axiomes suivants :

- **séparation :**

$$\forall \vec{u} \in \mathbb{R}^3 : \|\vec{u}\| = 0 \Leftrightarrow \vec{u} = \vec{0}$$

- **homogénéité :**

$$\forall \vec{u} \in \mathbb{R}^3, \forall \lambda \in \mathbb{R} : \|\lambda \vec{u}\| = |\lambda| \|\vec{u}\|$$

- **inégalité triangulaire¹ :**

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : \|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|$$

1. « Le chemin le plus court est la ligne droite ! » Merci à Paul Pearce, étudiant PCME 2011 !

3. Plans

L'ensemble des vecteurs orthogonaux à un vecteur \vec{u} non nul est un *plan vectoriel*, appelé **plan orthogonal à \vec{u}** .

Démonstration : Soit \vec{u} un vecteur non nul, et P l'ensemble des vecteurs orthogonaux à \vec{u} .

On désigne par (u_1, u_2, u_3) les composantes de \vec{u} , et on suppose, dans un premier temps, $u_3 \neq 0$.

P est donc l'ensemble des vecteurs de composantes (x, y, z) telles que :

$$u_1 x + u_2 y + u_3 z = 0$$

ce qui conduit à :

$$z = -\frac{u_1 x + u_2 y}{u_3}$$

On peut alors écrire :

$$\begin{aligned} P &= \left\{ \left(x, y, -\frac{u_1 x + u_2 y}{u_3} \right) / (x, y) \in \mathbb{R}^2 \right\} \\ &= \left\{ x \left(1, 0, -\frac{u_1}{u_3} \right) + y \left(0, 1, -\frac{u_2}{u_3} \right) / (x, y) \in \mathbb{R}^2 \right\} \\ &= \mathbb{R} \left(1, 0, -\frac{u_1}{u_3} \right) + \mathbb{R} \left(0, 1, -\frac{u_2}{u_3} \right) \end{aligned}$$

Les vecteurs $\left(1, 0, -\frac{u_1}{u_3} \right)$ et $\left(0, 1, -\frac{u_2}{u_3} \right)$ étant linéairement indépendants, on obtient l'équation d'un plan vectoriel, ainsi que les vecteurs engendrant celui-ci, qui sont $\left(1, 0, -\frac{u_1}{u_3} \right)$ et $\left(0, 1, -\frac{u_2}{u_3} \right)$.

Le cas $u_3 = 0$ est laissé au lecteur. ■

► Vecteur normal à un plan vectoriel

Étant donné un plan vectoriel P orthogonal à un vecteur \vec{n} non nul, le vecteur \vec{n} est appelé **vecteur normal** au plan vectoriel P .

Proposition

Étant donnés trois réels non tous nuls a, b, c , et un plan vectoriel P , d'équation cartésienne $ax + by + cz = 0$, un vecteur normal à P est $\vec{n} = (a, b, c)$.

► Vecteur normal à un plan affine

On appelle **vecteur normal à un plan affine \mathcal{P}** tout vecteur directeur d'une droite perpendiculaire à \mathcal{P} .

Proposition

Étant donnés quatre réels a, b, c, d tels que $(a, b, c) \neq (0, 0, 0)$, et un plan affine \mathcal{P} , d'équation cartésienne $ax + by + cz = d$, un vecteur normal à \mathcal{P} est $\vec{n} = (a, b, c)$.

Démonstration : Soit $A(x_A, y_A, z_A)$ un point donné de \mathcal{P} , et $M(x_M, y_M, z_M)$ un point quelconque de \mathcal{P} .

Alors :

$$\overrightarrow{AM} \cdot \vec{n} = a(x_M - x_A) + b(y_M - y_A) + c(z_M - z_A)$$

A et M appartenant à \mathcal{P} , leurs coordonnées vérifient l'équation de \mathcal{P} :

$$\begin{cases} a x_A + b y_A + c z_A = d \\ a x_M + b y_M + c z_M = d \end{cases}$$

Par suite :

$$a(x_M - x_A) + b(y_M - y_A) + c(z_M - z_A) = 0$$

Les vecteurs \overrightarrow{AM} et \vec{n} sont donc bien orthogonaux.

Le point M ayant été choisi quelconque, on en déduit l'orthogonalité du plan \mathcal{P} et de la droite passant par M et dirigée par le vecteur \vec{n} , et donc, l'orthogonalité du vecteur \vec{n} au plan \mathcal{P} . ■

Pourquoi s'intéresse-t-on autant aux propriétés du produit scalaire ? Il s'agit, tout simplement, de pouvoir étendre ces notions géométriques, relatives à un espace concret, à des espaces plus abstraits, sur lesquels on va essayer de construire une géométrie.

Considérons, ainsi, un fluide en mouvement ; la plupart du temps, on ne peut pas déterminer la valeur exacte de la vitesse du fluide, on l'approche par des simulations numériques les plus précises possibles (soit à l'aide de différences finies, soit de volumes finis). Si on ne connaît certes pas la valeur exacte de cette vitesse, il faut cependant être sûr que l'approximation choisie soit la meilleure possible. Ainsi, il est indispensable de définir une « distance » entre la solution exacte et la solution approchée. On se place alors dans ce que l'on appelle un « espace fonctionnel », c'est-à-dire un espace de fonctions, dont les vecteurs sont, en fait, des fonctions. Si on considère que le mouvement du fluide est rectiligne, et a lieu entre des points d'abscisses respectives a et $b > a$, $(a, b) \in \mathbb{R}^2$, on peut ainsi considérer l'espace $C^1([a, b], \mathbb{R})$, c'est-à-dire l'espace des fonctions de classe C^1 sur $[a, b]$, à valeurs dans \mathbb{R} (dont la dérivée est continue sur $[a, b]$, pour des questions de régularité : l'accélération, qui est la dérivée de la vitesse, doit pouvoir être définie). Pour pouvoir définir une « distance » entre la solution exacte v_{exacte} et la solution approchée $v_{\text{approchée}}$, il faut donc un produit scalaire ; on choisit l'application :

$$\begin{aligned} \varphi : C^1([a, b], \mathbb{R}) \times C^1([a, b], \mathbb{R}) &\rightarrow \mathbb{R} \\ (f, g) &\mapsto \int_a^b f(x) g(x) dx \end{aligned}$$

L'application φ étant, clairement :

- **symétrique :**

$$\forall (f, g) \in C^1([a, b], \mathbb{R}) \times C^1([a, b], \mathbb{R}) : \varphi(g, f) = \varphi(f, g)$$

- **bilinéaire :**

- **définie positive :**

$$\forall f \in C^1([a, b], \mathbb{R}) : \varphi(f, f) \geq 0$$

et :

$$\forall f \in C^1([a, b], \mathbb{R}) : \varphi(f, f) = 0 \Leftrightarrow \int_a^b f^2(x) dx = 0 \Leftrightarrow f = 0$$

(Si l'intégrale sur un segment $[a, b]$ d'une fonction positive et continue est nulle, alors cette fonction est identiquement nulle sur $[a, b]$; réciproquement, et de façon évidente : $\int_a^b 0 dx = 0$.)

c'est un produit scalaire sur $C^1([a, b], \mathbb{R})$, à partir duquel on pourra mesurer la « distance » entre la solution exacte et la solution approchée :

$$d(v_{\text{exacte}}, v_{\text{approchée}}) = \sqrt{\int_a^b (v_{\text{exacte}} - v_{\text{approchée}})^2(x) dx}$$

Dans ce qui suit, on se place dans l'espace euclidien orienté, de dimension 3.

1. Définition géométrique

► Base directe de \mathbb{R}^3

Une famille libre de trois vecteurs $(\vec{u}, \vec{v}, \vec{w})$ de l'espace \mathbb{R}^3 est appelée **base directe** si :

$$\det(\vec{u}, \vec{v}, \vec{w}) > 0$$

► Angle entre deux vecteurs de \mathbb{R}^3

Soient \vec{u} et \vec{v} deux vecteurs de l'espace \mathbb{R}^3 . On choisit de placer leur origine en un point O de l'espace. Alors, par définition, l'angle (\vec{u}, \vec{v}) désigne l'angle orienté entre les droites passant par le point O , et de vecteurs directeurs respectifs \vec{u} et \vec{v} .

► Produit vectoriel

Le **produit vectoriel** de deux vecteurs $\vec{u} = (u_1, u_2, u_3)$ et $\vec{v} = (v_1, v_2, v_3)$ de l'espace \mathbb{R}^3 est le vecteur, noté $\vec{u} \wedge \vec{v}$, tel que :

- si \vec{u} et \vec{v} sont colinéaires : $\vec{u} \wedge \vec{v} = \vec{0}$
- si \vec{u} et \vec{v} ne sont pas colinéaires : $\vec{u} \wedge \vec{v} = \|\vec{u}\| \|\vec{v}\| \left| \sin(\widehat{(\vec{u}, \vec{v})}) \right| \vec{w}$ où \vec{w} est le vecteur unitaire orthogonal à \vec{u} et \vec{v} tel que la base $(\vec{u}, \vec{v}, \vec{w})$ soit directe.
 $\vec{u} \wedge \vec{v}$ est alors unique.



Pour déterminer si une base $(\vec{OA}, \vec{OB}, \vec{OC})$ de l'espace est directe, on peut utiliser la « règle du bonhomme d'Ampère » : si un personnage ayant les pieds en O , la tête en A , et regardant vers B , voit le point C à sa gauche, la base est directe, et indirecte dans le cas contraire.

Exemple

Dans l'espace euclidien de dimension 3 rapporté au repère orthonormé direct $(O; \vec{i}, \vec{j}, \vec{k})$: $\vec{i} \wedge \vec{i} = -\vec{k}$, $\vec{i} \wedge \vec{k} = -\vec{j}$, ... et $\vec{i} \wedge \vec{j} = \vec{k}$, $\vec{j} \wedge \vec{k} = \vec{i}$, $\vec{k} \wedge \vec{j} = -\vec{i}$



En pratique, dans l'espace euclidien de dimension 3 rapporté au repère orthonormé direct

$$(O; \vec{i}, \vec{j}, \vec{k}) : \vec{u} \wedge \vec{v} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \wedge \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{pmatrix}$$

2. Définition analytique

Définition

Le **produit vectoriel** de deux vecteurs $\vec{u} = (u_1, u_2, u_3)$ et $\vec{v} = (v_1, v_2, v_3)$ de l'espace \mathbb{R}^3 est l'unique vecteur, noté $\vec{u} \wedge \vec{v}$, tel que pour tout vecteur $\vec{w} = (w_1, w_2, w_3)$ de \mathbb{R}^3 , on ait : $\det(\vec{u}, \vec{v}, \vec{w}) = (\vec{u} \wedge \vec{v}) \cdot \vec{w}$.

Démonstration : Par le calcul ...

■

► Proposition

1. Le produit vectoriel est **antisymétrique** :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : \vec{u} \wedge \vec{v} = -\vec{v} \wedge \vec{u}$$

2. Le produit vectoriel est **bilinéaire** :

$$\forall (\vec{u}, \vec{v}, \vec{w}) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3, \forall \lambda \in \mathbb{R} : \begin{cases} \vec{u} \wedge (\vec{v} + \lambda \vec{w}) = \vec{u} \wedge \vec{v} + \lambda \vec{u} \wedge \vec{w} \\ (\vec{v} + \lambda \vec{w}) \wedge \vec{u} = \vec{v} \wedge \vec{u} + \lambda \vec{w} \wedge \vec{u} \end{cases}$$

3.

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : \vec{u} \wedge \vec{v} = \vec{0} \Leftrightarrow \vec{u} \text{ et } \vec{v} \text{ liés}$$

4.

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : (\vec{u} \wedge \vec{v}) \perp \vec{u}, (\vec{u} \wedge \vec{v}) \perp \vec{v}$$

5.

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : \|\vec{u} \wedge \vec{v}\| = \|\vec{u}\| \|\vec{v}\| \left| \sin(\widehat{(\vec{u}, \vec{v})}) \right|$$

6. **Double produit vectoriel**

$$\forall (\vec{u}, \vec{v}, \vec{w}) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 : (\vec{u} \wedge \vec{v}) \wedge \vec{w} = (\vec{u} \cdot \vec{w}) \vec{v} - (\vec{v} \cdot \vec{w}) \vec{u}$$

Démonstration : Par le calcul ... ■

► Application du produit vectoriel

Exemple

Déterminons un système d'équations cartésiennes de la droite \mathcal{D} passant par le point A de coordonnées $(1, 2, 3)$, et normale au plan P d'équation $x + y + z = 0$.

P est orthogonal au vecteur $\vec{u} = (1, 1, 1)$. Il suffit donc d'écrire que \mathcal{D} est l'ensemble des points M de coordonnées (x, y, z) tels que les vecteurs \overrightarrow{AM} et \vec{u} soient colinéaires, ce qui peut aussi s'écrire :

$$\overrightarrow{AM} \wedge \vec{u} = \vec{0}$$

soit :

$$\begin{pmatrix} x-1 \\ y-2 \\ z-3 \end{pmatrix} \wedge \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} y-2-z+3 \\ z-3-x+1 \\ x-1-y+2 \end{pmatrix} = \begin{pmatrix} y-z+1 \\ z-x-2 \\ x-y+1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ce qui conduit à :

$$\begin{cases} y - z + 1 = 0 \\ z - x - 2 = 0 \\ x - y + 1 = 0 \end{cases}$$

La somme de ces 3 équations donnant 0, celles-ci ne sont pas indépendantes (elles sont liées) ; le système précédent est donc équivalent, par exemple à :

$$\begin{cases} y - z + 1 = 0 \\ z - x - 2 = 0 \end{cases}$$

(les trois équations étant liées, il suffit d'en garder deux, à condition de bien les sélectionner, puisque la troisième n'est en fait qu'une combinaison linéaire des deux autres).

On a donc bien obtenu un système d'équations cartésiennes de la droite \mathcal{D} .

1. Aire d'un triangle

Soit \mathcal{P} un plan affine de \mathbb{R}^3 , et A, B, C trois points non alignés de ce plan.

Si on reprend la définition géométrique du produit vectoriel, il existe un réel λ , et un vecteur \vec{n} , unitaire, orthogonal à \mathcal{P} , tel que :

$$\vec{AB} \wedge \vec{AC} = \lambda \vec{n}$$

Soit H le projeté orthogonal de C sur la droite (AB) ; on a alors :

$$\vec{AB} \wedge \vec{AC} = \vec{AB} \wedge (\vec{AH} + \vec{HC}) = \vec{AB} \wedge \vec{HC}$$

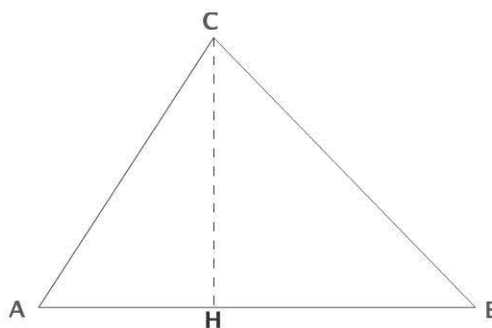


Figure 88.1 – Le triangle ABC .

Par suite : $\|\vec{AB} \wedge \vec{AC}\| = \|\vec{AB} \wedge \vec{HC}\|$, et donc :

$$|\lambda| = \|\vec{AB} \wedge \vec{AC}\| = \|\vec{AB} \wedge \vec{HC}\| = AB.HC$$

puisque \vec{AB} et \vec{HC} sont orthogonaux. On reconnaît alors le double de l'aire \mathcal{A}_{ABC} du triangle ABC :

$$\mathcal{A}_{ABC} = \frac{\text{base} \times \text{hauteur}}{2} = \frac{AB \times HC}{2}$$

2. Aire d'un parallélogramme

Soient A, B, C, D quatre points du plan \mathbb{R}^2 formant un parallélogramme non aplati. L'aire du parallélogramme de sommets A, B, C, D est :

$$\mathcal{A}_{ABCD} = \|\vec{AB} \wedge \vec{AD}\|$$

Définition

Dans le plan \mathbb{R}^2 , l'aire orientée du parallélogramme engendré par deux vecteurs $\vec{u} \neq \vec{0}$ et $\vec{v} \neq \vec{0}$ est donnée par :

$$\mathcal{A} = \det(\vec{u}, \vec{v})$$

Proposition

Dans le plan \mathbb{R}^2 , l'aire orientée du parallélogramme engendré par les vecteurs $\vec{u} = (u_1, u_2)$ et $\vec{v} = (v_1, v_2)$ est égale à l'aire du parallélogramme engendré par les vecteurs \vec{u} et $\vec{v} + \lambda \vec{u}$, où λ est un réel quelconque.

Démonstration : L'aire orientée du parallélogramme engendré par les vecteurs $\vec{u} = (u_1, u_2)$ et $\vec{v} = (v_1, v_2)$ est donnée par : $\mathcal{A} = \det(\vec{u}, \vec{v})$. L'aire orientée du parallélogramme engendré par les vecteurs \vec{u} et $\vec{v} + \lambda \vec{u}$ est donnée par : $\det(\vec{u}, \vec{v} + \lambda \vec{u}) = \det(\vec{u}, \vec{v}) = \mathcal{A}$, puisque le déterminant n'est pas modifié lorsque, à une colonne donnée, on ajoute une combinaison linéaire des autres colonnes. ■

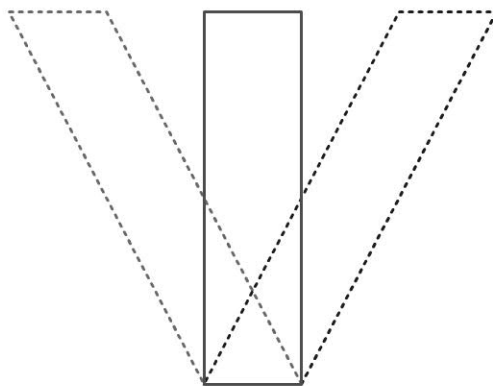


Figure 88.2 – Trois parallélogrammes de même aire.

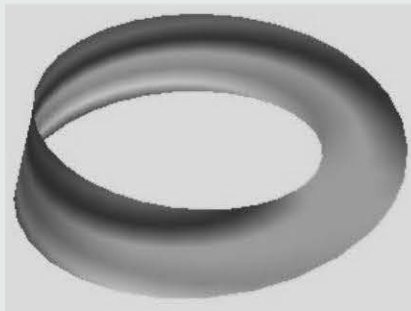
3. Volume orienté d'un parallélépipède**Définition**

Dans l'espace \mathbb{R}^3 , le volume orienté du parallélépipède engendré par trois vecteurs non coplanaires et non nuls \vec{u} , \vec{v} et \vec{w} est donné par :

$$\mathcal{V} = \det(\vec{u}, \vec{v}, \vec{w})$$

Il n'y a pas qu'Euclide

La géométrie usuelle est la géométrie euclidienne, ainsi nommée d'après le mathématicien grec Euclide. Il en existe toutefois de nombreuses autres ; ainsi, en géométrie différentielle, on s'intéresse à l'étude des courbes et des surfaces du point de vue de la longueur, la surface, la courbure ; un très joli exemple est donné par le ruban de Möbius.



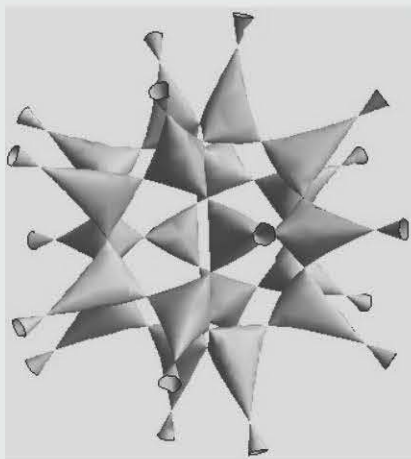
Le ruban de Möbius.

De façon encore plus spécifique, la géométrie riemannienne étudie des espaces courbes, les « variétés différentielles », sur lesquels existent des notions d'angles et de longueur, à l'aide de systèmes de coordonnées locales ; une variété n'est rien d'autre qu'un espace topologique, c'est-à-dire un espace sur lequel on a défini une « topologie », ou « cartographie ».

La géométrie algébrique s'intéresse quant à elle à la description des ensembles de zéros d'équations polynomiales, comme c'est le cas de la Sextique de Barth, dont une équation cartésienne, en fonction d'un paramètre réel m , est donnée par :

$$4(\varphi^2 x^2 - y^2)(\varphi^2 y^2 - z^2)(\varphi^2 z^2 - x^2) - (1 + 2\varphi)(x^2 + y^2 + z^2 - m^2)^2 m^2 = 0$$

où $\varphi = \frac{1 + \sqrt{5}}{2}$ est le nombre d'or.



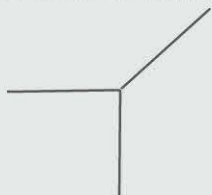
La sextique de Barth.

La géométrie tropicale pour des chemins optimaux

Il existe encore d'autres géométries, notamment, la géométrie tropicale, ainsi nommée en l'honneur de son inventeur, le mathématicien et informaticien brésilien Imre Simon (1943-2009). Cette géométrie est basée sur une re-définition de l'addition et de la multiplication :

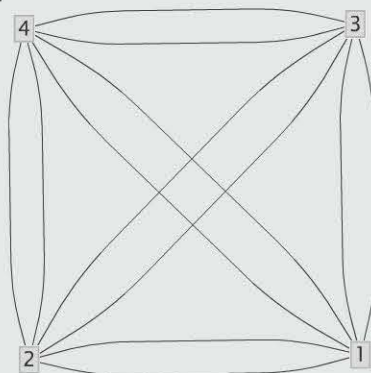
$$a \oplus b = \min \{a, b\} \quad , \quad a \otimes b = a + b$$

et où les objets mathématiques sont remplacés par des objets affines par morceaux ; ainsi, une droite « tropicale » est formée de trois demi-droites usuelles.



Une droite tropicale.

Elle connaît actuellement un essor croissant, en raison de son champ d'applications, notamment en théorie des graphes pour déterminer un chemin optimal, en cristallographie, ou encore en biologie quantitative [24],[25]. Prenons l'exemple d'une société de transport routier souhaitant optimiser non pas les distances parcourues, mais les coûts de transport [26] :



Les différents chemins possibles, dans le cas où $n = 4$.

Dans le cas de n arrêts possibles ($n \in \mathbb{N}^*, n \geq 2$), on commence par écrire les différents coûts possibles sous la forme d'une matrice de coût, $\mathcal{M}_{\text{coût}}$, de taille $n \times n$, le coefficient ligne i , $1 \leq i \leq n$, colonne j , $1 \leq j \leq n$, donnant le coût du transport entre l'arrêt numéro i et l'arrêt numéro j . Si l'on considère alors une puissance $k^{\text{ième}}$, $k \in \mathbb{N}$, de la matrice $\mathcal{M}_{\text{coût}}$, le coefficient ligne i , $1 \leq i \leq n$, colonne j , $1 \leq j \leq n$, donné par la formule du produit matriciel :

$$(\mathcal{M}_{\text{coût}})_{ij} = \bigoplus_{p=1}^n (\mathcal{M}_{\text{coût}})_{ip} \otimes (\mathcal{M}_{\text{coût}})_{pj}$$

correspond exactement à la somme des coûts des k trajets reliant l'arrêt numéro i et l'arrêt numéro j ! Ainsi, pour résoudre le problème donné, il suffit de calculer :

$$\bigoplus_{k=1}^{+\infty} \mathcal{M}_{\text{coût}}^k$$

Bases et transformations linéaires du plan

1. Bases

Définition

On appelle **base du plan** \mathbb{R}^2 toute famille libre de deux vecteurs de \mathbb{R}^2 .

Exemple

Déterminons si la famille $(\vec{u}, \vec{v}) = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right)$ est une base de \mathbb{R}^2 ; pour cela, il suffit de calculer son déterminant :

$$\det(\vec{u}, \vec{v}) = -2 \neq 0$$

(\vec{u}, \vec{v}) est une famille libre de deux vecteurs de \mathbb{R}^2 : c'est une base.

Définition

On appelle **base canonique du plan** \mathbb{R}^2 la famille (\vec{i}, \vec{j}) , avec :

$$\vec{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Proposition

Soit $\mathcal{B} = (\vec{e}_1, \vec{e}_2)$ une base de \mathbb{R}^2 .

Pour tout vecteur \vec{u} de \mathbb{R}^2 , il existe une unique combinaison linéaire des vecteurs de \mathcal{B} égale à \vec{u} : $\vec{u} = x_1 \vec{e}_1 + x_2 \vec{e}_2$ que l'on note aussi : $(\vec{u})_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. $x_1 \in \mathbb{R}$ et $x_2 \in \mathbb{R}$ sont les **coordonnées** (ou **composantes**) du vecteur \vec{u} dans la base \mathcal{B} .

Les composantes d'un vecteur (x, y) de \mathbb{R}^2 dans la base canonique (\vec{i}, \vec{j}) seront notées $\begin{pmatrix} x \\ y \end{pmatrix}$.

2. Transformations linéaires du plan

► Définitions

Définition

On appelle **transformation linéaire du plan** \mathbb{R}^2 , ou encore **application linéaire du plan** \mathbb{R}^2 , une application f , de \mathbb{R}^2 dans \mathbb{R}^2 , linéaire :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^2 \times \mathbb{R}^2, \forall \lambda \in \mathbb{R} : f(\vec{u} + \lambda \vec{v}) = f(\vec{u}) + \lambda f(\vec{v})$$



Pour toute **transformation linéaire** f du plan \mathbb{R}^2 :

$$f(\vec{0}) = \vec{0}$$

Exemples

1. L'application : $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto (x + 1, y)$ n'est pas linéaire.
2. L'application : $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto (3x, 3y)$ est linéaire.

Proposition

Étant donnée une matrice $A = (a_{ij})_{1 \leq i \leq 2, 1 \leq j \leq 2}$ de taille 2×2 , à coefficients réels, l'application :

$$f_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \begin{pmatrix} x \\ y \end{pmatrix} \mapsto A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y \\ a_{21}x + a_{22}y \end{pmatrix}$$

est **linéaire**, et définit une **transformation linéaire** de \mathbb{R}^2 .

► Matrice d'une transformation linéaire

Définition

On appelle **matrice**, dans une base $\mathcal{B} = (\vec{e}_1, \vec{e}_2)$ donnée, d'une **transformation linéaire** f de \mathbb{R}^2 , la matrice dont les vecteurs colonnes sont les coordonnées, dans \mathcal{B} , des images des vecteurs de \mathcal{B} : $\text{Matrice}_{\mathcal{B}}(f) = (f(\vec{e}_1), f(\vec{e}_2))$

Une application linéaire remarquable : l'application identité du plan \mathbb{R}^2

Définition

On appelle **application identité** du plan \mathbb{R}^2 l'application, notée $\text{Id}_{\mathbb{R}^2}$, de matrice associée $I_2 : \vec{u} \in \mathbb{R}^2 \mapsto \vec{u}$

Domaine de définition

Proposition

Une transformation linéaire de \mathbb{R}^2 est entièrement définie par ses valeurs sur une base donnée.

Démonstration : Soient $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ une transformation linéaire, et $\mathcal{B} = (\vec{e}_1, \vec{e}_2)$ une base de \mathbb{R}^2 .

On suppose que les images des vecteurs \vec{e}_1 et \vec{e}_2 sont données, de la forme :

$$\begin{cases} f(\vec{e}_1) = a_{11}\vec{e}_1 + a_{21}\vec{e}_2 \\ f(\vec{e}_2) = a_{12}\vec{e}_1 + a_{22}\vec{e}_2 \end{cases}$$

Alors, pour tout vecteur $\vec{u} = x_1\vec{e}_1 + x_2\vec{e}_2$ de \mathbb{R}^2 :

$$\begin{aligned} f(\vec{u}) &= f(x_1\vec{e}_1 + x_2\vec{e}_2) \\ &= x_1 f(\vec{e}_1) + x_2 f(\vec{e}_2) \\ &= x_1 (a_{11}\vec{e}_1 + a_{21}\vec{e}_2) + x_2 (a_{12}\vec{e}_1 + a_{22}\vec{e}_2) \\ &= (x_1 a_{11} + x_2 a_{12}) \vec{e}_1 + (x_1 a_{21} + x_2 a_{22}) \vec{e}_2 \end{aligned}$$

Ainsi,

$$(f(\vec{u}))_{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

D'où le résultat. ■

Proposition

Une transformation linéaire de \mathbb{R}^2 est entièrement définie par la matrice qui lui est associée dans une base donnée.

Démonstration : Le résultat découle immédiatement de la proposition précédente : une transformation linéaire de \mathbb{R}^2 est entièrement définie par ses valeurs sur une base donnée. ■

Définition

On appelle **matrice d'une transformation linéaire f de \mathbb{R}^2** , la matrice de f dans la base canonique (\vec{i}, \vec{j}) . Elle sera notée, dans ce qui suit, sous la forme A_f :

$$A_f = \text{Matrice}_{(\vec{i}, \vec{j})}(f) = (f(\vec{i}), f(\vec{j}))$$



Une transformation linéaire de \mathbb{R}^2 étant entièrement définie par la matrice qui lui est associée dans une base donnée, il est clair qu'il est préférable de se placer dans une base donnant une expression la plus agréable possible de la matrice : triangulaire, ou diagonale. Cela s'appelle *trigonaliser* ou *diagonaliser* une application linéaire.

► Propriétés

Composée de deux transformations linéaires

Proposition

La composée $f \circ g$ de deux transformations linéaires f et g de \mathbb{R}^2 , de matrices associées $A_{\mathcal{B}}(f)$ et $A_{\mathcal{B}}(g)$ dans une base $\mathcal{B} = (\vec{e}_1, \vec{e}_2)$ donnée, est une transformation linéaire de \mathbb{R}^2 , qui vérifie :

$$\forall \vec{u} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : (f \circ g)(\vec{u}) = f(g(\vec{u})) = A_{\mathcal{B}}(f) (A_{\mathcal{B}}(g) \vec{u}) = (A_{\mathcal{B}}(f) A_{\mathcal{B}}(g)) \vec{u} \quad (\star)$$

Ainsi, $f \circ g$ a pour matrice, dans la base \mathcal{B} : $A_{\mathcal{B}}(f \circ g) = A_{\mathcal{B}}(f) A_{\mathcal{B}}(g)$.

Démonstration : Par le calcul ...

D'après (\star) : $(f \circ g)(\vec{e}_1) = f(g(\vec{e}_1))$, puis on utilise $A_{\mathcal{B}}(g)$ pour obtenir $(f \circ g)(\vec{e}_1)$ comme combinaison linéaire de \vec{e}_1 et \vec{e}_2 , et on procède de même pour $(f \circ g)(\vec{e}_2)$. On peut alors construire la matrice de $f \circ g$ dans \mathcal{B} , et constater, par le calcul, qu'elle est égale au produit $A_{\mathcal{B}}(f) A_{\mathcal{B}}(g)$. ■

Noyau

Définition

On appelle **noyau** d'une transformation linéaire f de \mathbb{R}^2 , que l'on note $\text{Ker } f$ (de l'allemand *kern*) l'ensemble des vecteurs de \mathbb{R}^2 dont l'image par f est le vecteur nul :

$$\text{Ker } f = \{ \vec{u} \in \mathbb{R}^2 / f(\vec{u}) = \vec{0} \}$$

Par extension, le **noyau** d'une matrice $A \in \mathcal{M}_2(\mathbb{R})$ est le noyau de l'application linéaire associée à la matrice A .

Image

Définition

On appelle **image** d'une transformation linéaire f de \mathbb{R}^2 , que l'on note $Im f$ l'ensemble des images des vecteurs de \mathbb{R}^2 par f :

$$Im f = \{f(\vec{u}), \vec{u} \in \mathbb{R}^2\}$$

► Propriétés

Définition

Une transformation linéaire de \mathbb{R}^2 est dite **injective** si :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^2 \times \mathbb{R}^2 : f(\vec{u}) = f(\vec{v}) \Rightarrow \vec{u} = \vec{v}$$

Théorème

Une transformation linéaire de \mathbb{R}^2 est **injective** si son noyau est réduit au vecteur nul :

$$Ker f = \{\vec{0}\}$$

Démonstration : Ce résultat découle de la linéarité de f . ■

Définition

Une transformation linéaire de \mathbb{R}^2 est dite **surjective** si : $\forall \vec{u} \in \mathbb{R}^2, \exists \vec{v} \in \mathbb{R}^2 : \vec{u} = f(\vec{v})$.

► Automorphisme

Définition

Une transformation linéaire de \mathbb{R}^2 est un **automorphisme** de \mathbb{R}^2 si elle est *bijective*.

Théorème du rang dans \mathbb{R}^2 ou formule du rang

Théorème

Dans \mathbb{R}^2 : $\dim Im f + \dim Ker f = 2$.

$\dim Im f$ est le **rang** de f , noté aussi $rg f$.

Caractérisation des automorphismes de \mathbb{R}^2

Théorème

Soit f une transformation linéaire de \mathbb{R}^2 . Alors :

$$f \text{ injective} \Leftrightarrow f \text{ surjective} \Leftrightarrow f \text{ bijective}.$$

Pour qu'une transformation linéaire f de \mathbb{R}^2 soit un automorphisme de \mathbb{R}^2 , il faut et il suffit que le déterminant de sa matrice $A_{\mathcal{B}}(f)$ dans une base \mathcal{B} quelconque soit non nul. Sa réciproque f^{-1} est alors une transformation linéaire de \mathbb{R}^2 , de matrice dans \mathcal{B} :

$$A_{\mathcal{B}}(f^{-1}) = (A_{\mathcal{B}}(f))^{-1}$$

Changement de base en dimension 2, et déterminant d'une application linéaire

1. Matrice de passage

Soient $\mathcal{B} = (\vec{e}_1, \vec{e}_2)$ et $\mathcal{B}' = (\vec{e}'_1, \vec{e}'_2)$ deux bases de \mathbb{R}^2 .

On désigne par p_{11} et p_{21} les composantes de \vec{e}'_1 suivant \vec{e}_1 et \vec{e}_2 :

$$\vec{e}'_1 = p_{11} \vec{e}_1 + p_{21} \vec{e}_2$$

et par p_{12} et p_{22} les composantes de \vec{e}'_2 suivant \vec{e}_1 et \vec{e}_2 :

$$\vec{e}'_2 = p_{12} \vec{e}_1 + p_{22} \vec{e}_2$$

On désigne par $P_{\mathcal{B} \rightarrow \mathcal{B}'}$ la matrice $P_{\mathcal{B} \rightarrow \mathcal{B}'} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$, c'est-à-dire la matrice dont les vecteurs colonnes représentent les *coordonnées* (ou *composantes*) des vecteurs \vec{e}'_1 et \vec{e}'_2 dans la base \mathcal{B} .

$P_{\mathcal{B} \rightarrow \mathcal{B}'}$ est appelée **matrice de passage** de \mathcal{B} à \mathcal{B}' .

2. Formule de changement de base pour un vecteur

Soit \vec{u} un vecteur de \mathbb{R}^2 . On désigne par $(\vec{u})_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ ses coordonnées dans \mathcal{B} , et par

$(\vec{u})_{\mathcal{B}'} = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$ ses coordonnées dans \mathcal{B}' . Alors :

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$$

soit :

$$(\vec{u})_{\mathcal{B}} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} (\vec{u})_{\mathcal{B}'}$$

Démonstration :

$$\begin{aligned} \vec{u} &= x'_1 \vec{e}'_1 + x'_2 \vec{e}'_2 \\ &= x'_1 (p_{11} \vec{e}_1 + p_{21} \vec{e}_2) + x'_2 (p_{12} \vec{e}_1 + p_{22} \vec{e}_2) \\ &= (x'_1 p_{11} + x'_2 p_{12}) \vec{e}_1 + (x'_1 p_{21} + x'_2 p_{22}) \vec{e}_2 \end{aligned}$$

Comme $\vec{u} = x_1 \vec{e}_1 + x_2 \vec{e}_2$, alors, par unicité de la décomposition de \vec{u} dans \mathcal{B} :

$$\begin{cases} x_1 = x'_1 p_{11} + x'_2 p_{12} \\ x_2 = x'_1 p_{21} + x'_2 p_{22} \end{cases}$$

que l'on peut encore écrire :

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x'_1 p_{11} + x'_2 p_{12} \\ x'_1 p_{21} + x'_2 p_{22} \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \quad \blacksquare$$



Le fait que la formule de changement de base s'exprime sous la forme $(\vec{u})_{\mathcal{B}} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} (\vec{u})_{\mathcal{B}'}$ peut paraître déroutant : il n'en est rien. Une illustration très intéressante peut être obtenue pour la réduction de l'équation d'une conique.

Exemple

Considérons la courbe plane, d'équation $3x^2 + 3y^2 - 2xy = 8$ dans un repère orthonormé direct. La présence des « termes croisés » $-2xy$ ne facilite pas l'étude de la conique : un changement de base permet de faire disparaître celui-ci.

On effectue le changement de coordonnées :

$$x = \frac{X+Y}{\sqrt{2}}, \quad y = \frac{X-Y}{\sqrt{2}}$$

En injectant ces expressions dans l'équation de la conique, on obtient : $X^2 + 2Y^2 = 4$.

Il est alors plus facile de déterminer la nature de la courbe (une ellipse).

Il est clair que, connaissant l'équation dans le repère initial, la démarche naturelle est d'exprimer les coordonnées de départ (x, y) , en fonction des nouvelles coordonnées (X, Y) , et non le contraire :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}$$

Conjugaison – Matrices semblables de taille 2×2

Il a été vu précédemment que, pour une base donnée, il existe une unique matrice associée à une application linéaire f .

Il peut être intéressant de donner les relations permettant de passer de la matrice $A_{\mathcal{B}}(f)$ associée à f dans une base \mathcal{B} , à la matrice $A_{\mathcal{B}'}(f)$ associée à f dans une autre base \mathcal{B}' .

1. Matrices semblables

Considérons les bases $\mathcal{B} = (\vec{e}_1, \vec{e}_2)$ et $\mathcal{B}' = (\vec{e}'_1, \vec{e}'_2)$ de \mathbb{R}^2 .

On désigne par $A_{\mathcal{B}}(f) = (a_{ij})_{1 \leq i \leq 2, 1 \leq j \leq 2}$ la matrice associée à f dans la base \mathcal{B} , et par $A_{\mathcal{B}'}(f) = (\alpha_{ij})_{1 \leq i \leq 2, 1 \leq j \leq 2}$ la matrice associée à f dans la base \mathcal{B}' .

On appelle $P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}$ la matrice de passage de \mathcal{B} à \mathcal{B}' .

Soit \vec{u} un vecteur de \mathbb{R}^2 . On désigne par $(\vec{u})_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ ses coordonnées dans \mathcal{B} , et par $(\vec{u})_{\mathcal{B}'} = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$ ses coordonnées dans \mathcal{B}' ; de même, on désigne par $(f(\vec{u}))_{\mathcal{B}}$ les coordonnées de $f(\vec{u})$ dans \mathcal{B} , et par $(f(\vec{u}))_{\mathcal{B}'}$ les coordonnées de $f(\vec{u})$ dans \mathcal{B}' .

Alors :

$$\begin{aligned} (f(\vec{u}))_{\mathcal{B}} &= A_{\mathcal{B}}(f) (\vec{u})_{\mathcal{B}} \\ &= A_{\mathcal{B}}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} (\vec{u})_{\mathcal{B}'} \end{aligned}$$

Mais on a aussi :

$$(f(\vec{u}))_{\mathcal{B}'} = A_{\mathcal{B}'}(f) (\vec{u})_{\mathcal{B}'}$$

et :

$$(f(\vec{u}))_{\mathcal{B}} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} (f(\vec{u}))_{\mathcal{B}'}$$

Il en résulte :

$$(f(\vec{u}))_{\mathcal{B}} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} (f(\vec{u}))_{\mathcal{B}'} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) (\vec{u})_{\mathcal{B}'} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1} (\vec{u})_{\mathcal{B}}$$

Comme

$$(f(\vec{u}))_{\mathcal{B}} = A_{\mathcal{B}}(f) (\vec{u})_{\mathcal{B}}$$

on a donc :

$$A_{\mathcal{B}}(f) (\vec{u})_{\mathcal{B}} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1} (\vec{u})_{\mathcal{B}}$$

ou encore :

$$(A_{\mathcal{B}}(f) - P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1}) (\vec{u})_{\mathcal{B}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Le vecteur \vec{u} étant quelconque, on en déduit, en choisissant successivement pour \vec{u} chacun des vecteurs de \mathcal{B} , la nullité de chaque colonne de $A_{\mathcal{B}}(f) - P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1}$:

$$A_{\mathcal{B}}(f) - P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

ce qui conduit à :

$$A_{\mathcal{B}}(f) = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1}$$

ou encore :

$$A_{\mathcal{B}'}(f) = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1} A_{\mathcal{B}}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}$$

Définition

Deux transformations linéaires f et g de \mathbb{R}^2 , de matrices associées A_f et A_g dans une même base \mathcal{B} , sont dites **conjuguées** s'il existe une transformation linéaire bijective φ de \mathbb{R}^2 , de matrice associée A_φ , appelée **conjugaison**, telle que :

$$f \circ \varphi = \varphi \circ g$$

ce qui, matriciellement, se traduit par : $A_f A_\varphi = A_\varphi A_g$, ou encore, puisque A_φ est inversible :

$$A_g = A_\varphi^{-1} A_f A_\varphi$$

Les matrices A_f et A_g sont dites **semblables**.

Proposition

Deux matrices *semblables* ont *même déterminant*.

Démonstration : Soient A et B deux matrices semblables, de taille 2×2 .

D'après ce qui précède, il existe une matrice inversible P telle que :

$$P^{-1} A P = B$$

Il en résulte :

$$\det(B) = \det(P^{-1} A P) = \det(A P^{-1} P) = \det(A) \quad \blacksquare$$

2. Déterminant d'une application linéaire de \mathbb{R}^2

Définition

On appelle **déterminant** d'une transformation linéaire f de \mathbb{R}^2 le déterminant de sa matrice dans la base canonique, ou, de façon équivalente, dans une base quelconque \mathcal{B} .

La valeur du déterminant d'une transformation linéaire f de \mathbb{R}^2 ne dépend effectivement pas de la base choisie : si \mathcal{B} et \mathcal{B}' sont deux bases du plan \mathbb{R}^2 , les matrices $A_{\mathcal{B}}(f)$ et $A_{\mathcal{B}'}(f)$, qui représentent donc la même application linéaire dans deux bases différentes, sont semblables, et ont donc même déterminant.



Opérateurs orthogonaux en dimension 2

1. Définition

On appelle **opérateur orthogonal du plan** \mathbb{R}^2 , ou **transformation orthogonale du plan** \mathbb{R}^2 , une transformation linéaire f de \mathbb{R}^2 , de matrice A_f , qui vérifie l'une des trois propriétés équivalentes suivantes :

- f conserve le produit scalaire :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^2 \times \mathbb{R}^2 : f(\vec{u}) \cdot f(\vec{v}) = \vec{u} \cdot \vec{v}$$

- ${}^t A_f A_f = I_2$ ou $A_f {}^t A_f = I_2$
- les vecteurs colonnes C_1, C_2 de A_f constituent une base orthonormale, c'est-à-dire une famille libre de vecteurs deux à deux orthogonaux et de norme 1 :

$$\forall (i, j) \in \{1, 2\}^2 : C_i \cdot C_j = \delta_{ij}$$

On dit alors que la matrice A_f est **orthogonale**.



1. Dire que f conserve le produit scalaire s'explique donc de la façon suivante :

$$\begin{aligned} \forall \vec{u} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, \forall \vec{v} = \begin{pmatrix} x' \\ y' \end{pmatrix} \in \mathbb{R}^2 : A_f \vec{u} \cdot A_f \vec{v} &= A_f \begin{pmatrix} x \\ y \end{pmatrix} \cdot A_f \begin{pmatrix} x' \\ y' \end{pmatrix} \\ &= \vec{u} \cdot \vec{v} = x x' + y y' \end{aligned}$$

avec :

$$A_f \vec{u} \cdot A_f \vec{v} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y \\ a_{21}x + a_{22}y \end{pmatrix} \cdot \begin{pmatrix} a_{11}x' + a_{12}y' \\ a_{21}x' + a_{22}y' \end{pmatrix}$$

2. les vecteurs colonnes C_1, C_2 de $A_f = (a_{ij})_{1 \leq i \leq 2, 1 \leq j \leq 2} \in \mathcal{M}_2(\mathbb{R})$ sont :

$$C_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}, \quad C_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$$

Ainsi :

$$C_1 \cdot C_2 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} \cdot \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = a_{11} a_{12} + a_{21} a_{22}$$

La norme du premier vecteur colonne est : $\|C_1\| = \sqrt{a_{11}^2 + a_{21}^2}$.

La norme du deuxième vecteur colonne est : $\|C_2\| = \sqrt{a_{12}^2 + a_{22}^2}$.

L'ensemble des transformations orthogonales de \mathbb{R}^2 est appelé **groupe orthogonal** de \mathbb{R}^2 , et noté $O(\mathbb{R}^2)$.

2. Propriétés

Proposition

Si A est une matrice orthogonale, alors :

$$\det A = \pm 1$$

Démonstration :

$$A {}^t A = I_2 \Rightarrow \det(A {}^t A) = (\det A)^2 = 1$$

■

Proposition

Une transformation orthogonale f , de matrice associée A_f , conserve la norme (c'est une isométrie) :

$$\forall \vec{u} = (x, y) \in \mathbb{R}^2 : \|f(\vec{u})\| = \|\vec{u}\|$$

ce qui se traduit matriciellement par :

$$\forall \vec{u} = (x, y) \in \mathbb{R}^2 : \|A_f \vec{u}\| = \|\vec{u}\|$$

Démonstration : Il suffit d'utiliser le fait qu'une transformation orthogonale conserve le produit scalaire :

$$\forall \vec{u} = (x, y) \in \mathbb{R}^2 : A_f \vec{u} \cdot A_f \vec{u} = \|\vec{u}\| \cdot \|\vec{u}\| \quad \blacksquare$$

Proposition

Une matrice orthogonale A_f peut aussi être caractérisée par la propriété suivante :

$$\forall \vec{u} \in \mathbb{R}^2 : \|A_f \vec{u}\| = \|\vec{u}\|$$

qui signifie que l'application linéaire associée f conserve la norme (c'est une isométrie).

Démonstration : On montre que cette propriété est équivalente à i :

- Si f conserve le produit scalaire, on a vu que :

$$\forall \vec{u} \in \mathbb{R}^2 : \|A_f \vec{u}\| = \|\vec{u}\|$$

- Réciproquement, si f conserve la norme, on peut utiliser **l'identité de polarisation** :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^2 \times \mathbb{R}^2 : \vec{u} \cdot \vec{v} = \frac{1}{4} \{ \|\vec{u} + \vec{v}\|^2 - \|\vec{u} - \vec{v}\|^2 \}$$

et :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^2 \times \mathbb{R}^2 : A_f \vec{u} \cdot A_f \vec{v} = \frac{1}{4} \{ \|A_f \vec{u} + A_f \vec{v}\|^2 - \|A_f \vec{u} - A_f \vec{v}\|^2 \}$$

Par suite :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^2 \times \mathbb{R}^2 : A_f \vec{u} \cdot A_f \vec{v} = \frac{1}{4} \{ \|A_f (\vec{u} + \vec{v})\|^2 - \|A_f (\vec{u} - \vec{v})\|^2 \}$$

puis :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^2 \times \mathbb{R}^2 : A_f \vec{u} \cdot A_f \vec{v} = \frac{1}{4} \{ \|\vec{u} + \vec{v}\|^2 - \|\vec{u} - \vec{v}\|^2 \} = \vec{u} \cdot \vec{v}$$

D'où le résultat. ■

Proposition

Si A est une matrice orthogonale, alors A est de la forme :

$$\begin{pmatrix} \cos \theta & -\varepsilon \sin \theta \\ \sin \theta & \varepsilon \cos \theta \end{pmatrix}$$

où θ est un réel, et :

$$\varepsilon = \det A$$

Définition

L'ensemble des transformations orthogonales de \mathbb{R}^2 de déterminant égal à 1 est appelé **groupe spécial orthogonal** de \mathbb{R}^2 , et noté $SO(\mathbb{R}^2)$.

Rotations vectorielles du plan

1. Définition

Soient θ un réel, Ω un point donné, et \vec{u} un vecteur de \mathbb{R}^2 .

Soient M le point tel que $\overrightarrow{\Omega M} = \vec{u}$, et M' l'image de M par la rotation de centre Ω , d'angle θ .

On considère l'application :

$$R_\theta : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \\ \overrightarrow{\Omega M} \mapsto \overrightarrow{\Omega M'}$$

C'est une application linéaire.

L'image par R_θ d'un vecteur quelconque \vec{v} de \mathbb{R}^2 est ainsi le vecteur $\vec{v'}$ tel que :

$$\begin{cases} \|\vec{v'}\| = \|\vec{v}\| \\ (\vec{v}, \vec{v'}) = \theta \pmod{2\pi} \end{cases}$$

R_θ est une *rotation vectorielle*, à distinguer des rotations affines vues au chapitre 1. Le point Ω peut être choisi absolument quelconque. À la différence d'une rotation affine, une rotation vectorielle de \mathbb{R}^2 n'est ainsi définie que par son angle. On parlera, dans ce qui suit, de **rotation** pour désigner une telle application.

2. Propriétés

Proposition

La matrice d'une rotation de \mathbb{R}^2 , d'angle θ , dans une base orthonormale directe, est de la forme :

$$\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

Démonstration : Pour tout vecteur $\vec{u} = \begin{pmatrix} x \\ y \end{pmatrix}$ de \mathbb{R}^2 :

$$R_\theta(\vec{u}) = \cos \theta \vec{u} + \sin \theta \vec{u}^\perp$$

où \vec{u}^\perp désigne le vecteur directement orthogonal à \vec{u} , et de même norme que \vec{u} :

$$\vec{u}^\perp = \begin{pmatrix} -y \\ x \end{pmatrix}$$

Ainsi :

$$R_\theta(\vec{u}) = \cos \theta \begin{pmatrix} x \\ y \end{pmatrix} + \sin \theta \begin{pmatrix} -y \\ x \end{pmatrix}$$

D'où le résultat. ■



On vérifie aisément que le déterminant d'une matrice de rotation vaut +1.

Proposition

Tout élément de $SO(\mathbb{R}^2)$ est une **rotation** (de \mathbb{R}^2).

Démonstration : D'après ce qui précède, il est clair que toute rotation de \mathbb{R}^2 est dans $SO(\mathbb{R}^2)$.

Réciproquement, soit f un élément de $SO(\mathbb{R}^2)$, de matrice dans la base canonique A_f . A_f est de la forme :

$$A_f = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (a, b, c, d) \in \mathbb{R}^4$$

f étant dans $O(\mathbb{R}^2)$:

$${}^t A_f \cdot A_f = I_2$$

soit :

$$\begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

ce qui conduit à :

$$\begin{cases} a^2 + c^2 = 1 \\ ab + cd = 0 \\ b^2 + d^2 = 1 \end{cases}$$

Il existe donc deux réels φ et ψ tels que

$$\begin{cases} a = \cos \varphi \\ c = \sin \varphi \end{cases}, \quad \begin{cases} b = \cos \psi \\ d = \sin \psi \end{cases}$$

La condition :

$$ab + cd = 0$$

conduit à :

$$\cos \varphi \cos \psi + \sin \varphi \sin \psi = 0$$

soit :

$$\cos(\psi - \varphi) = 0$$

Il en résulte :

$$\psi - \varphi = \frac{\pi}{2} \quad [\pi]$$

ou encore :

$$\psi = \varphi + \frac{\pi}{2} + k\pi, \quad k \in \mathbb{Z}$$

Par suite :

$$\begin{cases} b = \cos \psi = \cos\left(\varphi + \frac{\pi}{2} + k\pi\right) = (-1)^{k+1} \sin \varphi \\ d = \sin \psi = \sin\left(\varphi + \frac{\pi}{2} + k\pi\right) = (-1)^k \cos \varphi \end{cases}$$

f étant dans $SO(\mathbb{R}^2)$, le déterminant de la matrice A_f doit être égal à 1, soit :

$$ad - bc = 1$$

ce qui conduit à :

$$(-1)^k (\cos^2 \varphi + \sin^2 \varphi) = 1$$

k est donc pair,

$$k \in 2\mathbb{Z}$$

On a donc, finalement :

$$\begin{cases} b = -\sin \varphi \\ d = \cos \varphi \end{cases}$$

et :

$$A_f = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

■



Le groupe spécial orthogonal de \mathbb{R}^2 , $SO(\mathbb{R}^2)$, est donc constitué :

- de l'application identité $id_{\mathbb{R}^2}$;
- des rotations d'angle $\theta \notin \pi\mathbb{Z}$;
- des symétries par rapport à une droite.

Définition

On appelle **base de l'espace \mathbb{R}^3** toute famille libre de trois vecteurs de \mathbb{R}^3 .

Exemple

Déterminons si la famille

$$(\vec{u}, \vec{v}, \vec{w}) = \left(\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right)$$

est une base de \mathbb{R}^3 ; pour cela, il suffit de calculer son déterminant :

$$\det(\vec{u}, \vec{v}, \vec{w}) = -4 \neq 0$$

$(\vec{u}, \vec{v}, \vec{w})$ est bien une base de \mathbb{R}^3 .

Définition

On appelle **base canonique de l'espace \mathbb{R}^3** la famille $(\vec{i}, \vec{j}, \vec{k})$, avec :

$$\vec{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Proposition

Soit $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ une base de \mathbb{R}^3 .

Pour tout vecteur \vec{u} de \mathbb{R}^3 , il existe une unique combinaison linéaire des vecteurs de \mathcal{B} égale à \vec{u} :

$$\vec{u} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + x_3 \vec{e}_3 = \sum_{i=1}^3 x_i \vec{e}_i$$

que l'on note aussi :

$$(\vec{u})_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$x_1 \in \mathbb{R}$, $x_2 \in \mathbb{R}$ et $x_3 \in \mathbb{R}$ sont les **coordonnées** (ou **composantes**) du vecteur \vec{u} dans la base \mathcal{B} .

La décomposition d'un vecteur suivant une base donnée est ainsi unique.

Les composantes d'un vecteur (x, y, z) de \mathbb{R}^3 dans la base canonique $(\vec{i}, \vec{j}, \vec{k})$ seront notées :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$



1. Définitions

Définition

On appelle **transformation linéaire de l'espace \mathbb{R}^3** , ou encore **application linéaire de l'espace \mathbb{R}^3** une application f , de \mathbb{R}^3 dans \mathbb{R}^3 , linéaire :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3, \forall \lambda \in \mathbb{R} : f(\vec{u} + \lambda \vec{v}) = f(\vec{u}) + \lambda f(\vec{v})$$



Pour toute **transformation linéaire f de l'espace \mathbb{R}^3** :

$$f(\vec{0}) = \vec{0}$$

Exemples

1. L'application :

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^3, (x, y, z) \mapsto (x + 1, y - 2, z)$$

n'est pas linéaire.

2. L'application :

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^3, (x, y, z) \mapsto (3x, 3y, 3z)$$

est linéaire.

Proposition

Étant donnée une matrice $A = (a_{ij})_{1 \leq i \leq 3, 1 \leq j \leq 3}$, de taille 3×3 , à coefficients réels, l'application :

$$f_A : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{pmatrix}$$

est *linéaire*, et définit une *transformation linéaire de \mathbb{R}^3* .

2. Matrice d'une transformation linéaire

Définition

On appelle **matrice, dans une base $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ donnée, d'une transformation linéaire f de \mathbb{R}^3** , la matrice dont les vecteurs colonnes sont les coordonnées, dans \mathcal{B} , des images des vecteurs de \mathcal{B} :

$$\text{Matrice}_{\mathcal{B}}(f) = (f(\vec{e}_1), f(\vec{e}_2), f(\vec{e}_3))$$

► Une application linéaire remarquable : l'application identité de l'espace \mathbb{R}^3

Définition

On appelle **application identité de l'espace \mathbb{R}^3** l'application, notée $Id_{\mathbb{R}^3}$, de matrice associée I_3 :

$$\vec{u} \in \mathbb{R}^3 \mapsto \vec{u}$$

► **Domaine de définition**

Proposition

Une transformation linéaire de \mathbb{R}^3 est entièrement définie par ses valeurs sur une base donnée.

Démonstration : Soit $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ une transformation linéaire, et $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ une base de \mathbb{R}^3 .

On suppose que les images des vecteurs \vec{e}_1, \vec{e}_2 et \vec{e}_3 sont données :

$$\forall i \in \{1, 2, 3\} : f(\vec{e}_i) = \sum_{j=1}^3 a_{ji} \vec{e}_j$$

Alors, pour tout vecteur $\vec{u} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + x_3 \vec{e}_3 = \sum_{i=1}^3 x_i \vec{e}_i$ de \mathbb{R}^3 :

$$\begin{aligned} f(\vec{u}) &= f\left(\sum_{i=1}^3 x_i \vec{e}_i\right) \\ &= \sum_{i=1}^3 x_i f(\vec{e}_i) \\ &= \sum_{i=1}^3 x_i \sum_{j=1}^3 a_{ji} \vec{e}_j \\ &= \sum_{i=1}^3 \sum_{j=1}^3 x_i a_{ji} \vec{e}_j \\ &= \sum_{i=1}^3 \sum_{j=1}^3 x_j a_{ij} \vec{e}_i \end{aligned}$$

Ainsi :

$$(f(\vec{u}))_{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

D'où le résultat. ■

Proposition

Une transformation linéaire de \mathbb{R}^3 est entièrement définie par la matrice qui lui est associée dans une base donnée.

Démonstration : Le résultat découle immédiatement de la proposition précédente : une transformation linéaire de \mathbb{R}^3 est entièrement définie par ses valeurs sur une base donnée. ■

Définition

On appelle **matrice d'une transformation linéaire f de \mathbb{R}^3** , la matrice de f dans la base canonique $(\vec{i}, \vec{j}, \vec{k})$. Elle sera notée, dans ce qui suit, sous la forme A_f :

$$A_f = \text{Matrice}_{(\vec{i}, \vec{j}, \vec{k})}(f) = (f(\vec{i}), f(\vec{j}), f(\vec{k}))$$



Une transformation linéaire de \mathbb{R}^3 étant entièrement définie par la matrice qui lui est associée dans une base donnée, il est clair qu'il est préférable de se placer dans une base donnant une expression la plus agréable possible de la matrice : triangulaire, ou diagonale. Cela s'appelle *triangulariser* ou *diagonaliser* une application linéaire (ou une matrice).

► Composée de deux transformations linéaires

Proposition

La composée $f \circ g$ de deux transformations linéaires f et g de \mathbb{R}^3 , de matrices associées $A_{\mathcal{B}}(f)$ et $A_{\mathcal{B}}(g)$ dans une base $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ donnée, est une transformation linéaire de \mathbb{R}^3 , qui vérifie :

$$\forall \vec{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : (f \circ g)(\vec{u}) = f(g(\vec{u})) = A_{\mathcal{B}}(f) (A_{\mathcal{B}}(g) \vec{u}) = (A_{\mathcal{B}}(f) A_{\mathcal{B}}(g)) \vec{u}$$

Ainsi, $f \circ g$ a pour matrice, dans la base \mathcal{B} :

$$A_{\mathcal{B}}(f \circ g) = A_{\mathcal{B}}(f) A_{\mathcal{B}}(g)$$

► Noyau

Définition

On appelle **noyau** d'une transformation linéaire f de \mathbb{R}^3 , que l'on note $\text{Ker } f$ (de l'allemand *kern*) l'ensemble des vecteurs de \mathbb{R}^3 dont l'image par f est le vecteur nul :

$$\text{Ker } f = \{ \vec{u} \in \mathbb{R}^3 / f(\vec{u}) = \vec{0} \}$$

Par extension, le **noyau** d'une matrice $A \in \mathcal{M}_3(\mathbb{R})$ est le noyau de l'application linéaire associée à la matrice A .

► Image

Définition

On appelle **image** d'une transformation linéaire f de \mathbb{R}^3 , que l'on note $\text{Im } f$ l'ensemble des images des vecteurs de \mathbb{R}^3 par f :

$$\text{Im } f = \{ f(\vec{u}), \vec{u} \in \mathbb{R}^3 \}$$

3. Propriétés

Définition

Une transformation linéaire de \mathbb{R}^3 est dite **injective** si :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : f(\vec{u}) = f(\vec{v}) \Rightarrow \vec{u} = \vec{v}$$

Théorème

Une transformation linéaire de \mathbb{R}^3 est **injective** si son noyau est réduit au vecteur nul :

$$\text{Ker } f = \{\vec{0}\}$$

Démonstration : Ce résultat découle de la linéarité de f . ■

Définition

Une transformation linéaire de \mathbb{R}^3 est dite **surjective** si :

$$\forall \vec{u} \in \mathbb{R}^3, \exists \vec{v} \in \mathbb{R}^3 : \vec{u} = f(\vec{v})$$

4. Automorphisme

Définition

Une transformation linéaire de \mathbb{R}^3 est un **automorphisme** de \mathbb{R}^3 si elle est *bijective*.

► Théorème du rang dans \mathbb{R}^3 (ou formule du rang)

Théorème

Dans \mathbb{R}^3 :

$$\dim \text{Im } f + \dim \text{Ker } f = 3$$

$\dim \text{Im } f$ est le **rang** de f , noté aussi $\text{rg } f$.

► Caractérisation des automorphismes de \mathbb{R}^3

Théorème

Soit f une transformation linéaire de \mathbb{R}^3 . Alors :

$$f \text{ injective} \Leftrightarrow f \text{ surjective} \Leftrightarrow f \text{ bijective}$$

Pour qu'une transformation linéaire f de \mathbb{R}^3 soit un automorphisme de \mathbb{R}^3 , il faut et il suffit que le déterminant de sa matrice $A_{\mathcal{B}}(f)$ dans une base \mathcal{B} quelconque soit non nul.

Sa réciproque f^{-1} est alors une transformation linéaire de \mathbb{R}^3 , de matrice dans \mathcal{B} :

$$A_{\mathcal{B}}(f^{-1}) = (A_{\mathcal{B}}(f))^{-1}$$

Changement de base en dimension 3

1. Matrice de passage

Définition

Soient $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ et $\mathcal{B}' = (\vec{e}'_1, \vec{e}'_2, \vec{e}'_3)$ deux bases de \mathbb{R}^3 .

On désigne par $P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} = (p_{ij})_{1 \leq i \leq 3, 1 \leq j \leq 3}$ la matrice dont les vecteurs colonnes représentent les *coordonnées* (ou *composantes*) des vecteurs \vec{e}'_1, \vec{e}'_2 et \vec{e}'_3 dans la base \mathcal{B} :

$$\forall i \in \{1, 2, 3\} : \vec{e}'_i = \sum_{j=1}^3 p_{ji} \vec{e}_j$$

$P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}$ est appelée **matrice de passage** de \mathcal{B} à \mathcal{B}' .

2. Formule de changement de base pour un vecteur

Soit \vec{u} un vecteur de \mathbb{R}^3 . On désigne par $(\vec{u})_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ ses coordonnées dans \mathcal{B} , et par

$(\vec{u})_{\mathcal{B}'} = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$ ses coordonnées dans \mathcal{B}' . Alors :

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$$

soit :

$$(\vec{u})_{\mathcal{B}} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} (\vec{u})_{\mathcal{B}'}$$

Démonstration :

$$\begin{aligned} \vec{u} &= \sum_{i=1}^3 x'_i \vec{e}'_i \\ &= \sum_{i=1}^3 x'_i \sum_{j=1}^3 p_{ji} \vec{e}_j \\ &= \sum_{i=1}^3 \sum_{j=1}^3 p_{ji} x'_i \vec{e}_j \\ &= \sum_{i=1}^3 \sum_{j=1}^3 p_{ij} x'_j \vec{e}_i \end{aligned}$$

(On a effectué un changement d'indices pour passer de l'avant-dernière à la dernière ligne.)

Comme :

$$\vec{u} = \sum_{i=1}^3 x_i \vec{e}_i$$

alors, par unicité de la décomposition de \vec{u} suivant \mathcal{B} , pour tout i de $\{1, 2, 3\}$:

$$x_i = \sum_{j=1}^3 p_{ij} x'_j$$

que l'on peut encore écrire :

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$$

■

Conjugaison – Matrices semblables de taille 3×3

Il a été vu précédemment que, pour une base donnée, il existe une unique matrice associée à une application linéaire f .

Il peut être intéressant de donner les relations permettant de passer de la matrice $A_{\mathcal{B}}(f)$ associée à f dans une base \mathcal{B} , à la matrice $A_{\mathcal{B}'}(f)$ associée à f dans une nouvelle base \mathcal{B}' .

1. Matrices semblables

Considérons les bases $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ et $\mathcal{B}' = (\vec{e}'_1, \vec{e}'_2, \vec{e}'_3)$ de \mathbb{R}^3 .

On désigne par $A_{\mathcal{B}}(f) = (a_{ij})_{1 \leq i \leq 3, 1 \leq j \leq 3}$ la matrice associée à f dans la base \mathcal{B} , et par $A_{\mathcal{B}'}(f) = (\alpha_{ij})_{1 \leq i \leq 3, 1 \leq j \leq 3}$ la matrice associée à f dans la base \mathcal{B}' .

On appelle $P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}$ la matrice de passage de \mathcal{B} à \mathcal{B}' .

Soit \vec{u} un vecteur de \mathbb{R}^3 . On désigne par $(\vec{u})_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ ses coordonnées dans \mathcal{B} , et par

$(\vec{u})_{\mathcal{B}'} = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$ ses coordonnées dans \mathcal{B}' ; de même, on désigne par $(f(\vec{u}))_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ les

coordonnées de $f(\vec{u})$ dans \mathcal{B} , et par $(f(\vec{u}))_{\mathcal{B}'} = \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$ les coordonnées de $f(\vec{u})$ dans \mathcal{B}' .

Alors :

$$\begin{aligned} (f(\vec{u}))_{\mathcal{B}} &= A_{\mathcal{B}}(f) (\vec{u})_{\mathcal{B}} \\ &= A_{\mathcal{B}}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} (\vec{u})_{\mathcal{B}'} \end{aligned}$$

Mais on a aussi :

$$(f(\vec{u}))_{\mathcal{B}'} = A_{\mathcal{B}'}(f) (\vec{u})_{\mathcal{B}'}$$

et :

$$(f(\vec{u}))_{\mathcal{B}} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} (f(\vec{u}))_{\mathcal{B}'}$$

Il en résulte :

$$(f(\vec{u}))_{\mathcal{B}} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} (f(\vec{u}))_{\mathcal{B}'} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) (\vec{u})_{\mathcal{B}'} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1} (\vec{u})_{\mathcal{B}}$$

Comme :

$$(f(\vec{u}))_{\mathcal{B}} = A_{\mathcal{B}}(f) (\vec{u})_{\mathcal{B}}$$

on a donc :

$$A_{\mathcal{B}}(f) (\vec{u})_{\mathcal{B}} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1} (\vec{u})_{\mathcal{B}}$$

ou encore :

$$(A_{\mathcal{B}}(f) - P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1}) (\vec{u})_{\mathcal{B}} = \vec{0}$$

Le vecteur \vec{u} étant quelconque, on en déduit, en choisissant successivement pour \vec{u} chacun des vecteurs de \mathcal{B} , la nullité de chaque colonne de $A_{\mathcal{B}}(f) - P A_{\mathcal{B}'}(f) P^{-1}$:

$$A_{\mathcal{B}}(f) = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1}$$

ou encore :

$$A_{\mathcal{B}'}(f) = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1} A_{\mathcal{B}}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}$$

Définition

Deux transformations linéaires f et g de \mathbb{R}^3 , de matrices associées A_f et A_g , sont dites **conjuguées** s'il existe une transformation linéaire bijective φ de \mathbb{R}^3 , de matrice associée A_φ , appelée **conjugaison**, telle que :

$$f \circ \varphi = \varphi \circ g$$

ce qui, matriciellement, se traduit par :

$$A_f A_\varphi = A_\varphi A_g$$

ou encore, puisque A_φ est inversible :

$$A_g = A_\varphi^{-1} A_f A_\varphi$$

Les matrices A_f et A_g sont dites **semblables**.

2. Déterminant d'une application linéaire de \mathbb{R}^3

Proposition

Deux matrices *semblables* ont même déterminant.

Démonstration : Soient A et B deux matrices semblables, de taille 3×3 .

D'après ce qui précède, il existe une matrice inversible P telle que :

$$P^{-1} A P = B$$

Il en résulte :

$$\det(B) = \det(P^{-1} A P) = \det(A P^{-1} P) = \det(A) \quad \blacksquare$$

Définition

On appelle **déterminant** d'une transformation linéaire f de \mathbb{R}^3 le déterminant de sa matrice dans la base canonique, ou, de façon équivalente, dans une base quelconque \mathcal{B} .



La valeur du déterminant d'une transformation linéaire f de \mathbb{R}^3 ne dépend effectivement pas de la base choisie : si \mathcal{B} et \mathcal{B}' sont deux bases de l'espace \mathbb{R}^3 , les matrices $A_{\mathcal{B}}(f)$ et $A_{\mathcal{B}'}(f)$, qui représentent donc la même application linéaire dans deux bases différentes, sont semblables, et ont donc même déterminant.

Opérateurs orthogonaux de l'espace \mathbb{R}^3

1. Définitions

On appelle **opérateur orthogonal de l'espace \mathbb{R}^3** , ou **transformation orthogonale de \mathbb{R}^3** , une transformation linéaire f de \mathbb{R}^3 , de matrice A_f , qui vérifie l'une des trois propriétés équivalentes suivantes :

- f conserve le produit scalaire :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : A_f \vec{u} \cdot A_f \vec{v} = \vec{u} \cdot \vec{v}$$

- ${}^t A_f A_f = I_3$ ou $A_f {}^t A_f = I_3$
- les vecteurs colonnes C_1, C_2, C_3 de A_f constituent une base orthonormale, c'est-à-dire une famille libre de vecteurs deux à deux orthogonaux et de norme 1 :

$$\forall (i, j) \in \{1, 2, 3\}^2 : C_i \cdot C_j = \delta_{ij}$$

On dit alors que la matrice A_f est **orthogonale**.

2. Isométrie

Proposition

Une transformation orthogonale f , de matrice associée A_f , *conserve la norme* (c'est une **isométrie**) :

$$\forall \vec{u} = (x, y, z) \in \mathbb{R}^3 : \|f(\vec{u})\| = \|\vec{u}\|$$

ce qui se traduit matriciellement par :

$$\forall \vec{u} = (x, y, z) \in \mathbb{R}^3 : \|A_f \vec{u}\| = \|\vec{u}\|$$

Démonstration : Il suffit d'utiliser le fait qu'une transformation orthogonale conserve le produit scalaire :

$$\forall \vec{u} = (x, y, z) \in \mathbb{R}^3 : A_f \vec{u} \cdot A_f \vec{u} = \|\vec{u}\| \cdot \|\vec{u}\| \quad \blacksquare$$

Proposition

Une matrice orthogonale A_f peut aussi être caractérisée par la propriété suivante :

$$\forall \vec{u} \in \mathbb{R}^3 : \|A_f \vec{u}\| = \|\vec{u}\|$$

qui signifie que f conserve la norme (l'application linéaire f est donc une isométrie).

Démonstration : On montre que cette propriété qu'elle est équivalente à i :

- Si f conserve le produit scalaire, alors :

$$\forall \vec{u} \in \mathbb{R}^3 : A_f \vec{u} \cdot A_f \vec{u} = \|\vec{u}\| \cdot \|\vec{u}\|$$

soit

$$\forall \vec{u} \in \mathbb{R}^3 : \|A_f \vec{u}\|^2 = \|\vec{u}\|^2$$

et donc :

$$\forall \vec{u} \in \mathbb{R}^3 : \|A_f \vec{u}\| = \|\vec{u}\|$$

- Réciproquement, si f conserve la norme, on peut utiliser l'**identité de polarisation** :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : \vec{u} \cdot \vec{v} = \frac{1}{4} \{ \|\vec{u} + \vec{v}\|^2 - \|\vec{u} - \vec{v}\|^2 \}$$

et :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : A_f \vec{u} \cdot A_f \vec{v} = \frac{1}{4} \{ \|A_f \vec{u} + A_f \vec{v}\|^2 - \|A_f \vec{u} - A_f \vec{v}\|^2 \}$$

Par suite :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : A_f \vec{u} \cdot A_f \vec{v} = \frac{1}{4} \{ \|A_f (\vec{u} + \vec{v})\|^2 - \|A_f (\vec{u} - \vec{v})\|^2 \}$$

puis :

$$\forall (\vec{u}, \vec{v}) \in \mathbb{R}^3 \times \mathbb{R}^3 : A_f \vec{u} \cdot A_f \vec{v} = \frac{1}{4} \{ \|\vec{u} + \vec{v}\|^2 - \|\vec{u} - \vec{v}\|^2 \} = \vec{u} \cdot \vec{v}$$

D'où le résultat. ■

3. Groupe orthogonal

Définition

L'ensemble des transformations orthogonales de \mathbb{R}^3 est appelé **groupe orthogonal** de \mathbb{R}^3 , et noté $O(\mathbb{R}^3)$.

Proposition

Si A est une matrice orthogonale, alors :

$$\det A = \pm 1$$

Démonstration :

$$A^t A = I_3 \Rightarrow \det(A^t A) = (\det A)^2 = 1 \quad \blacksquare$$

Proposition

Si A est une matrice orthogonale, alors A est semblable à une matrice de la forme :

$$A_\theta = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}$$

où θ est un réel, et :

$$\varepsilon = \det A \in \{-1, 1\}$$

Définition

L'ensemble des transformations orthogonales de \mathbb{R}^3 de déterminant égal à 1 est appelé **groupe spécial orthogonal** de \mathbb{R}^3 , et noté $SO(\mathbb{R}^3)$.

Rotations vectorielles de l'espace \mathbb{R}^3

1. Axe de rotation

Soient θ un réel, \vec{n} un vecteur de \mathbb{R}^3 , et P le plan vectoriel orthogonal à \vec{n} . On désigne par (\vec{i}_P, \vec{j}_P) une base de P telle que $(\vec{i}_P, \vec{j}_P, \vec{n})$ soit une base orthonormée directe de \mathbb{R}^3 . Dans cette base, un vecteur \vec{u} de \mathbb{R}^3 se décompose de manière unique sous la forme :

$$\vec{u} = x_{\vec{u}} \vec{i}_P + y_{\vec{u}} \vec{j}_P + z_{\vec{u}} \vec{n}$$

Soit alors :

$$\vec{u}_P = x_{\vec{u}} \vec{i}_P + y_{\vec{u}} \vec{j}_P$$

et R_θ^P la rotation (au sens vectoriel) de P , d'angle θ .

On considère l'application :

$$\begin{aligned} R_\theta : \mathbb{R}^3 &\rightarrow \mathbb{R}^3 \\ \vec{u} &\mapsto R_\theta^P(\vec{u}_P) + z_{\vec{u}} \vec{n} \end{aligned}$$

C'est une application linéaire, qui est une *rotation vectorielle de \mathbb{R}^3* , d'angle θ . On parlera, dans ce qui suit, de rotation pour désigner une telle application.

Si $\theta \notin 2\pi\mathbb{Z}$, l'ensemble de ses invariants est la droite vectorielle dirigée par le vecteur \vec{n} . Cette droite est appelée **axe de la rotation**.

2. Propriétés de l'axe de rotation

Proposition

Soit R_θ la rotation d'axe dirigé par le vecteur $\vec{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, d'angle θ .

Alors, pour tout vecteur \vec{u} de l'espace \mathbb{R}^3 :

$$R_\theta(\vec{u}) = \cos \theta \vec{u} + \sin \theta \vec{k} \wedge \vec{u} + (1 - \cos \theta) (\vec{u} \cdot \vec{k}) \vec{k}$$

Démonstration : Il suffit de vérifier la formule pour les trois vecteurs de la base canonique $(\vec{i}, \vec{j}, \vec{k})$. ■

Proposition

Soit R_θ la rotation axiale, d'angle θ , dont l'axe est dirigé par le vecteur unitaire \vec{n} .

Alors, pour tout vecteur \vec{u} de l'espace \mathbb{R}^3 :

$$R_\theta(\vec{u}) = \cos \theta \vec{u} + \sin \theta \vec{n} \wedge \vec{u} + (1 - \cos \theta) (\vec{u} \cdot \vec{n}) \vec{n}$$

Démonstration : On complète \vec{n} par deux vecteurs \vec{v}_1 et \vec{v}_2 , tels que $(\vec{v}_1, \vec{v}_2, \vec{n})$ soit une base orthonormale directe de \mathbb{R}^3 .

On a alors :

$$\begin{cases} R_\theta(\vec{v}_1) = \cos \theta \vec{v}_1 + \sin \theta \vec{v}_2 = \cos \theta \vec{v}_1 + \sin \theta \vec{n} \wedge \vec{v}_1 \\ R_\theta(\vec{v}_2) = -\sin \theta \vec{v}_1 + \cos \theta \vec{v}_2 = \cos \theta \vec{v}_2 + \sin \theta \vec{n} \wedge \vec{v}_2 \\ R_\theta(\vec{n}) = \vec{n} \end{cases}$$

d'où le résultat. ■

Proposition

La matrice d'une rotation axiale d'angle $\theta \in \mathbb{R}$, dans une base orthonormale directe dont le troisième vecteur dirige l'axe de la rotation, est de la forme :

$$A_\theta = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

A_θ est une matrice de rotation ; l'angle de la rotation est θ .

Démonstration : Soit R_θ la rotation d'axe dirigé par le vecteur $\vec{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, d'angle θ .

Alors, pour tout vecteur \vec{u} de l'espace \mathbb{R}^3 :

$$R_\theta(\vec{u}) = \cos \theta \vec{u} + \sin \theta \vec{k} \wedge \vec{u} + (1 - \cos \theta) (\vec{u} \cdot \vec{k}) \vec{k}$$

Il en résulte, pour les trois vecteurs de la base canonique $(\vec{i}, \vec{j}, \vec{k})$:

$$\begin{cases} R_\theta(\vec{i}) = \cos \theta \vec{i} + \sin \theta \vec{k} \wedge \vec{i} = \cos \theta \vec{i} + \sin \theta \vec{j} \\ R_\theta(\vec{j}) = \cos \theta \vec{j} + \sin \theta \vec{k} \wedge \vec{j} = -\sin \theta \vec{i} + \cos \theta \vec{j} \\ R_\theta(\vec{k}) = \vec{k} \end{cases}$$

d'où le résultat. ■



On vérifie aisément que le déterminant d'une matrice de rotation vaut +1.

Le cas d'une base quelconque est laissé au lecteur.

Le *groupe spécial orthogonal* de \mathbb{R}^3 , $SO(\mathbb{R}^3)$, est donc constitué :

1. de l'application identité $id_{\mathbb{R}^3}$;
2. des rotations axiales ;
3. des symétries par rapport à une droite (cas particulier de rotations, d'angle π).

Vecteurs en dimension $n, n \geq 2$

Dans ce qui précède, on s'est intéressé aux espaces usuels que sont le plan \mathbb{R}^2 , et l'espace \mathbb{R}^3 . \mathbb{R}^2 est de dimension 2, car pour se repérer dans le plan, il faut deux coordonnées. De même, \mathbb{R}^3 est de dimension 3, car pour se repérer dans l'espace, il faut trois coordonnées. Mais dans la vie réelle de l'ingénieur ou du scientifique, une grandeur n'est pas, en général, caractérisée par deux ou trois coordonnées, mais plus : ainsi, dès que l'on introduit une référence temporelle, il faut prendre en compte une donnée supplémentaire, le temps t . On se retrouve ainsi dans un espace de dimension 4, où chaque grandeur est caractérisée par ses trois coordonnées spatiales, et sa coordonnée temporelle.

De façon plus générale, il est donc utile de disposer de résultats et d'outils mathématiques permettant de gérer un nombre n de coordonnées, ce qui nous place ainsi dans un espace de dimension n ; \mathbb{R}^n , où on généralise les résultats déjà existants en dimension 2 ou 3, est l'exemple le plus naturel. Les résultats présentés dans ce qui suit s'appliquent encore pour un espace (vectoriel) E de dimension n .

Notation

Dans ce qui suit, n et N désignent des entiers supérieurs ou égaux à 2.

1. Définitions et propriétés fondamentales

Définition

On appelle **vecteur à n composantes**, un n -uplet de réels $x = (x_1, \dots, x_n)$, que l'on

peut aussi écrire sous la forme $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

x_1, \dots, x_n sont appelées *composantes* du vecteur x .

Propriété

Deux vecteurs sont égaux si et seulement si ils ont les mêmes composantes.

► Vecteur nul

Le vecteur, dont les n composantes sont nulles, est appelé **vecteur nul**, et noté 0 .

► Somme de deux vecteurs

Définition

On considère les vecteurs $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$ de l'espace \mathbb{R}^n .

On définit le vecteur $x + y = z$ comme l'unique vecteur de \mathbb{R}^n , dont les composantes (z_1, \dots, z_n) sont données par :

$$\forall i \in \{1, \dots, n\} : z_i = x_i + y_i$$

► Multiplication d'un vecteur par un réel

Définition

On considère le vecteur (x_1, \dots, x_n) de l'espace \mathbb{R}^n .

Alors, pour tout réel λ , λx est un vecteur de \mathbb{R}^n , de composantes $(\lambda x_1, \dots, \lambda x_n)$.

2. Familles de vecteurs libres (ou famille libre de vecteurs)

► Combinaison de vecteurs

On appelle **combinaison linéaire de deux vecteurs x et y** , toute expression de la forme :

$$\lambda x + \mu y$$

où λ et μ sont des réels.

On appelle **combinaison linéaire de N vecteurs x_1, \dots, x_N** , toute expression de la forme :

$$\sum_{i=1}^N \lambda_i x_i$$

où $\lambda_1, \dots, \lambda_N$ sont des réels.

► Vecteurs liés et vecteurs libres

Deux vecteurs x et y sont dits **liés** s'il existe une combinaison linéaire non triviale de ces vecteurs égale au vecteur nul, c'est-à-dire lorsque les coefficients de la combinaison sont non simultanément nuls, de la forme :

$$\lambda x + \mu y = 0 \quad (\lambda, \mu) \neq (0, 0)$$

Deux vecteurs x et y sont dits **libres** s'ils ne sont pas liés.

► Généralisation

N vecteurs x_1, \dots, x_N sont dits **liés** s'il existe une combinaison linéaire non triviale de ces vecteurs égale au vecteur nul, c'est-à-dire lorsque les coefficients de la combinaison sont non simultanément nuls, de la forme :

$$\sum_{i=1}^N \lambda_i x_i = 0$$

où $\lambda_1, \dots, \lambda_N$ sont des réels tels que $(\lambda_1, \dots, \lambda_N) \neq (0, \dots, 0)$.

N vecteurs x_1, \dots, x_N sont dits **libres** s'ils ne sont pas liés.

Propriété

Pour toute famille (x_1, \dots, x_N) de vecteurs libres :

$$\lambda_1 x_1 + \dots + \lambda_N x_N = 0 \Rightarrow (\lambda_1, \dots, \lambda_N) = (0, \dots, 0)$$

ce qui signifie que, s'il existe une combinaison linéaire nulle de ces vecteurs, alors les coefficients de cette même combinaison linéaire sont **nécessairement** nuls.

Espace engendré par une famille de vecteurs – Sous-espaces vectoriels de \mathbb{R}^n

1. Bases

Définition

On appelle **base de l'espace \mathbb{R}^n** toute famille libre de n vecteurs de \mathbb{R}^n .

► Base canonique

Définition

On appelle **base canonique de l'espace \mathbb{R}^n** la famille :

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

► Coordonnées d'un vecteur

Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de \mathbb{R}^n . Pour tout vecteur x de \mathbb{R}^n , il existe une unique combinaison linéaire des vecteurs de \mathcal{B} égale à x :

$$x = x_1 e_1 + \dots + x_n e_n = \sum_{i=1}^n x_i e_i$$

$x_1 \in \mathbb{R}, \dots, x_n \in \mathbb{R}$ sont les **coordonnées** (ou **composantes**) du vecteur x dans la base \mathcal{B} .

Les composantes d'un vecteur (x_1, \dots, x_n) de \mathbb{R}^n dans la base canonique seront notées

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

2. Famille génératrice

Étant donnée une famille (x_1, \dots, x_N) de N vecteurs de \mathbb{R}^n , on appelle espace engendré par la famille (x_1, \dots, x_N) l'ensemble de toutes les combinaisons linéaires possibles de cette famille, que l'on note $\text{Vect}\{x_1, \dots, x_N\}$.

Une famille (x_1, \dots, x_N) de n vecteurs de \mathbb{R}^n est dite **génératrice** si :

$$\text{Vect}\{x_1, \dots, x_N\} = \mathbb{R}^n$$

Une famille (x_1, \dots, x_N) de N vecteurs de \mathbb{R}^n est une **base de \mathbb{R}^n** si elle est libre et génératrice.

► Dimension d'une base de \mathbb{R}^n

Théorème

1. Toute base de l'espace \mathbb{R}^n admet exactement n éléments.
2. On appelle **dimension** de \mathbb{R}^n le nombre d'éléments de toute base de \mathbb{R}^n :

$$\dim \mathbb{R}^n = n$$

3. Toute famille libre de n éléments de l'espace \mathbb{R}^n est une base de \mathbb{R}^n .
4. Toute famille génératrice de n éléments de l'espace \mathbb{R}^n est une base de \mathbb{R}^n .

► Théorème de la base incomplète

Théorème

Toute famille libre $\{x_1, \dots, x_N\}$, $N \in \mathbb{N}$, $N < n$, peut être complétée en une base $\{x_1, \dots, x_N, x_{N+1}, \dots, x_n\}$ de \mathbb{R}^n .

3. Sous-espace vectoriel de \mathbb{R}^n

Définition

On appelle **sous-espace vectoriel** de \mathbb{R}^n un ensemble F de \mathbb{R}^n tel que :

- $0 \in F$;
- étant donnés deux éléments x et y de F , $x + y$ est aussi élément de F (ce qui signifie que F est stable par addition) ;
- étant donnés un élément x de F , et un réel λ , λx est aussi élément de F (ce qui signifie que F est stable par la multiplication par un scalaire).

► Caractérisation des sous-espaces vectoriels de \mathbb{R}^n

Théorème

Un ensemble F de \mathbb{R}^n est un sous-espace vectoriel de \mathbb{R}^n si et seulement si, étant donnés deux éléments x et y de F , et un réel λ , $x + \lambda y$ est aussi élément de F (ce qui signifie que F est stable par combinaisons linéaires) :

$$\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n, \forall \lambda \in \mathbb{R} : x + \lambda y \in F$$

► Dimension d'un sous-espace vectoriel

Théorème

On appelle **dimension** d'un sous-espace F de \mathbb{R}^n le nombre d'éléments de toute base de F .

4. Somme de sous-espaces vectoriels de \mathbb{R}^n

F_1 et F_2 étant deux sous-espaces vectoriels de \mathbb{R}^n , l'ensemble :

$$F_1 + F_2 = \{x_1 + x_2, x_1 \in F_1, x_2 \in F_2\}$$

est un sous-espace vectoriel de \mathbb{R}^n , appelé **somme de F_1 et F_2** .

► **Dimension d'une somme de sous-espaces de \mathbb{R}^n : la formule de Grassmann¹**

Théorème

F_1 et F_2 étant deux sous-espaces vectoriels de \mathbb{R}^n :

$$\dim (F_1 + F_2) = \dim F_1 + \dim F_2 - \dim (F_1 \cap F_2)$$

► **Somme directe de sous-espaces vectoriels de \mathbb{R}^n**

Deux sous-espaces vectoriels F_1 et F_2 de \mathbb{R}^n sont dits **en somme directe** si tout élément du sous-espace somme $F_1 + F_2$ s'écrit de manière unique comme somme d'un élément de F_1 et d'un élément de F_2 :

$$\forall x \in F_1 + F_2, \exists ! (x_1, x_2) \in F_1 \times F_2 : x = x_1 + x_2$$

La somme $F_1 + F_2$ se note alors $F_1 \oplus F_2$.

► **Caractérisation d'une somme directe**

Deux sous-espaces vectoriels F_1 et F_2 de \mathbb{R}^n sont **en somme directe** si et seulement si :

$$F_1 \cap F_2 = \{0\}$$

5. Supplémentarité

► **Sous-espaces vectoriels supplémentaires de \mathbb{R}^n**

Deux sous-espaces vectoriels F_1 et F_2 de \mathbb{R}^n sont dits **supplémentaires dans \mathbb{R}^n** si tout élément de \mathbb{R}^n s'écrit de manière unique comme somme d'un élément de F_1 et d'un élément de F_2 :

$$\forall x \in \mathbb{R}^n, \exists ! (x_1, x_2) \in F_1 \times F_2 : x = x_1 + x_2$$

On note alors : $F_1 \oplus F_2 = \mathbb{R}^n$.

► **Caractérisation de la supplémentarité dans \mathbb{R}^n**

Deux sous-espaces vectoriels F_1 et F_2 de \mathbb{R}^n sont **supplémentaires** si et seulement si

$$F_1 + F_2 = \mathbb{R}^n \quad \text{et} \quad F_1 \cap F_2 = \{0\}$$

1. Hermann Günther Grassmann (1809-1877), mathématicien, physicien et linguiste allemand. Il fut l'un des premiers à introduire la notion d'espace vectoriel.

1. Définitions

Définition

On appelle **transformation linéaire de l'espace \mathbb{R}^n** , ou encore **application linéaire de l'espace \mathbb{R}^n** une application f , de \mathbb{R}^n dans \mathbb{R}^n , linéaire :

$$\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n, \forall \lambda \in \mathbb{R} : f(x + \lambda y) = f(x) + \lambda f(y)$$

Pour toute **transformation linéaire f de l'espace \mathbb{R}^n** :

$$f(0) = 0$$

Exemples

1. L'application :

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^n, (x_1, \dots, x_n) \mapsto (x_1 + 1, \dots, x_n + 1)$$

n'est pas linéaire.

2. L'application :

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^n, (x_1, \dots, x_n) \mapsto 3(x_1, \dots, x_n) = (3x_1, \dots, 3x_n)$$

est linéaire.

Proposition

Étant donnée une matrice $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$, de taille $n \times n$, à coefficients réels, l'application :

$$f_A : \mathbb{R}^n \rightarrow \mathbb{R}^n, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{n1}x_1 + \dots + a_{nn}x_n \end{pmatrix}$$

est *linéaire*, et définit une *transformation linéaire de \mathbb{R}^n* .

2. Matrice d'une transformation linéaire

Définition

On appelle **matrice, dans une base $\mathcal{B} = (e_1, \dots, e_n)$ donnée, d'une transformation linéaire f de \mathbb{R}^n** , la matrice dont les vecteurs colonnes sont les coordonnées, dans \mathcal{B} , des images des vecteurs de \mathcal{B} :

$$\text{Matrice}_{\mathcal{B}}(f) = (f(e_1), \dots, f(e_n))$$

► Une application linéaire remarquable : l'application identité de \mathbb{R}^n

Définition

On appelle **application identité de \mathbb{R}^n** l'application, notée $Id_{\mathbb{R}^n}$, de matrice associée I_n :

$$x \in \mathbb{R}^n \mapsto x$$

► Domaine de définition

Proposition

Une transformation linéaire de \mathbb{R}^n est entièrement définie par ses valeurs sur une base donnée.

Démonstration : Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ une transformation linéaire, et $\mathcal{B} = (e_1, e_2, \dots, e_n)$ une base de \mathbb{R}^n .

On suppose que les images des vecteurs e_1, \dots, e_n sont données :

$$\forall i \in \{1, 2, \dots, n\} : f(e_i) = \sum_{j=1}^n a_{ji} e_j$$

Alors, pour tout vecteur $x = x_1 e_1 + \dots + x_n e_n = \sum_{i=1}^n x_i e_i$ de \mathbb{R}^n :

$$\begin{aligned} f(x) &= f\left(\sum_{i=1}^n x_i e_i\right) \\ &= \sum_{i=1}^n x_i f(e_i) \\ &= \sum_{i=1}^n x_i \sum_{j=1}^n a_{ji} e_j \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i a_{ji} e_j \\ &= \sum_{j=1}^n \sum_{i=1}^n x_i a_{ji} e_j \end{aligned}$$

Ainsi :

$$(f(x))_{\mathcal{B}} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

D'où le résultat. ■

Proposition

Une transformation linéaire de \mathbb{R}^n est entièrement définie par la matrice qui lui est associée dans une base donnée.

Démonstration : Le résultat découle immédiatement de la proposition précédente : une transformation linéaire de \mathbb{R}^n est entièrement définie par ses valeurs sur une base donnée. ■

Définition

On appelle **matrice d'une transformation linéaire f de \mathbb{R}^n** , la matrice de f dans la base canonique. Elle sera notée, dans ce qui suit, A_f .



Une transformation linéaire de \mathbb{R}^n étant entièrement définie par la matrice qui lui est associée dans une base donnée, il est clair qu'il est préférable de se placer dans une base donnant une expression la plus agréable possible de la matrice : triangulaire, ou diagonale. Cela s'appelle *trigonaliser* ou *diagonaliser* une application linéaire (ou une matrice).

► Composée de deux transformations linéaires

Proposition

La composée $f \circ g$ de deux transformations linéaires f et g de \mathbb{R}^n , de matrices associées $A_{\mathcal{B}}(f)$ et $A_{\mathcal{B}}(g)$ dans une base $\mathcal{B} = (e_1, \dots, e_n)$ donnée, est une transformation linéaire

de \mathbb{R}^n , qui vérifie : $\forall x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$:

$$(f \circ g)(x) = f(g(x)) = A_{\mathcal{B}}(f) (A_{\mathcal{B}}(g) x) = (A_{\mathcal{B}}(f) A_{\mathcal{B}}(g)) x$$

Ainsi, $f \circ g$ a pour matrice, dans la base \mathcal{B} :

$$A_{\mathcal{B}}(f \circ g) = A_{\mathcal{B}}(f) A_{\mathcal{B}}(g)$$

► Formule du binôme de Newton

Proposition

Étant données deux transformations linéaires f et g de \mathbb{R}^n qui commutent, c'est-à-dire telles que $f \circ g = g \circ f$, alors, pour tout entier naturel p , la formule du binôme de Newton permet de calculer $(f + g)^p$, où la puissance est au sens de la composition :

$$(f + g)^p = \sum_{k=0}^p C_p^k f^k \circ g^{p-k}$$

où, pour tout entier k de $\{0, \dots, p\}$, C_p^k désigne le coefficient binomial :

$$\binom{p}{k} = \frac{p!}{k! (p-k)!}$$

► Noyau

Définition

On appelle **noyau** d'une transformation linéaire f de \mathbb{R}^n , que l'on note $\text{Ker } f$ (de l'allemand *kern*) l'ensemble des vecteurs de \mathbb{R}^n dont l'image par f est le vecteur nul :

$$\text{Ker } f = \{x \in \mathbb{R}^n / f(x) = 0\}$$

Par extension, le **noyau** d'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est le noyau de l'application linéaire associée à la matrice A .

► Image

Définition

On appelle **image** d'une transformation linéaire f de \mathbb{R}^n , que l'on note $\text{Im } f$ l'ensemble des images des vecteurs de \mathbb{R}^n par f :

$$\text{Im } f = \{f(x), x \in \mathbb{R}^n\}$$

3. Propriétés

Définition

Une transformation linéaire de \mathbb{R}^n est dite **injective** si :

$$\forall (x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^n : f(x_1) = f(x_2) \Rightarrow x_1 = x_2$$

Théorème

Une transformation linéaire de \mathbb{R}^n est **injective** si son noyau est réduit au vecteur nul :

$$\text{Ker } f = \{0\}$$

Démonstration : Ce résultat découle de la linéarité de f . ■

Définition

Une transformation linéaire de \mathbb{R}^n est dite **surjective** si :

$$\forall y \in \mathbb{R}^n, \exists x \in \mathbb{R}^n : y = f(x)$$

4. Automorphisme

Définition

Une transformation linéaire de \mathbb{R}^n est un **automorphisme**, de \mathbb{R}^n si elle est *bijective*, c'est-à-dire injective et surjective.

L'ensemble des automorphismes de \mathbb{R}^n est appelé **groupe linéaire de \mathbb{R}^n** ¹, et noté $\mathcal{GL}(\mathbb{R}^n)$.

Proposition

Pour qu'une transformation linéaire f de \mathbb{R}^n soit un automorphisme de \mathbb{R}^n , il faut et il suffit que le déterminant de sa matrice $A_{\mathcal{B}}(f)$ dans une base \mathcal{B} quelconque soit non nul.

Sa réciproque f^{-1} est alors une transformation linéaire de \mathbb{R}^n , de matrice dans \mathcal{B}

$$A_{\mathcal{B}}(f^{-1}) = (A_{\mathcal{B}}(f))^{-1}$$

► Théorème du rang dans \mathbb{R}^n (ou formule du rang)

Théorème

Dans \mathbb{R}^n :

$$\dim \text{Im } f + \dim \text{Ker } f = n$$

$\dim \text{Im } f$ est le **rang** de f , noté aussi $\text{rg } f$.

► Caractérisation des automorphismes de \mathbb{R}^n

Théorème

Soit f une transformation linéaire de \mathbb{R}^n . Les propriétés suivantes sont équivalentes :

- f est injective ;
- f est surjective ;
- f est bijective.

1. car il possède une structure de groupe, c'est-à-dire est stable par produit, et tout élément non nul est inversible.

1. Matrice de passage

Définition

Soient $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (e'_1, \dots, e'_n)$ deux bases de \mathbb{R}^n .

On désigne par $P_{\mathcal{B} \rightarrow \mathcal{B}'} = (p_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$ la matrice dont les vecteurs colonnes représentent les *coordonnées* (ou *composantes*) des vecteurs e'_1, \dots, e'_n dans la base \mathcal{B} :

$$\forall i \in \{1, \dots, n\} : e'_i = \sum_{j=1}^n p_{ji} e_j$$

$P_{\mathcal{B} \rightarrow \mathcal{B}'}$ est appelée **matrice de passage** de \mathcal{B} à \mathcal{B}' .

2. Formule de changement de base pour un vecteur

Soit x un vecteur de \mathbb{R}^n . On désigne par $(x)_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ses coordonnées dans \mathcal{B} , et par

$(x)_{\mathcal{B}'} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ ses coordonnées dans \mathcal{B}' . Alors, si $P_{\mathcal{B} \rightarrow \mathcal{B}'} = (p_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$ est la matrice de passage de \mathcal{B} à \mathcal{B}' :

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P_{\mathcal{B} \rightarrow \mathcal{B}'} \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

soit :

$$(x)_{\mathcal{B}} = P_{\mathcal{B} \rightarrow \mathcal{B}'} (x)_{\mathcal{B}'}$$

Démonstration :

$$\begin{aligned} x &= \sum_{i=1}^n x'_i e'_i \\ &= \sum_{i=1}^n x'_i \sum_{j=1}^n p_{ji} e_j \\ &= \sum_{i=1}^n \sum_{j=1}^n p_{ji} x'_i e_j \\ &= \sum_{i=1}^n \sum_{j=1}^n p_{ij} x'_j e_i \end{aligned}$$

(On a effectué un changement d'indices pour passer de l'avant-dernière à la dernière ligne.)

Comme :

$$x = \sum_{i=1}^n x_i e_i$$

alors, par unicité de la décomposition de x suivant \mathcal{B} , pour tout i de $\{1, \dots, n\}$:

$$x_i = \sum_{j=1}^n p_{ij} x'_j$$

que l'on peut encore écrire :

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$$

■

Il a été vu précédemment que, pour une base donnée, il existe une unique matrice associée à une application linéaire f .

Il peut être intéressant de donner les relations permettant de passer de la matrice $A_{\mathcal{B}}(f)$ associée à f dans une base \mathcal{B} , à la matrice $A_{\mathcal{B}'}(f)$ associée à f dans une nouvelle base \mathcal{B}' .

1. Matrices semblables

Considérons les bases $\mathcal{B} = (e_1, \dots, e_n)$ et $\mathcal{B}' = (e'_1, \dots, e'_n)$ de \mathbb{R}^n .

On désigne par $A_{\mathcal{B}}(f) = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$ la matrice associée à f dans la base \mathcal{B} , et par $A_{\mathcal{B}'}(f) = (a'_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$ la matrice associée à f dans la base \mathcal{B}' .

On appelle $P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}$ la matrice de passage de \mathcal{B} à \mathcal{B}' .

Soit x un vecteur de \mathbb{R}^n . On désigne par $(x)_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ses coordonnées dans \mathcal{B} , et par

$(x)_{\mathcal{B}'} = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ ses coordonnées dans \mathcal{B}' ; de même, on désigne par $(f(x))_{\mathcal{B}}$ les coordonnées de $f(x)$ dans \mathcal{B} , et par $(f(x))_{\mathcal{B}'}$ les coordonnées de $f(x)$ dans \mathcal{B}' . Alors :

$$\begin{aligned} (f(x))_{\mathcal{B}} &= A_{\mathcal{B}}(f) (x)_{\mathcal{B}} \\ &= A_{\mathcal{B}}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} (x)_{\mathcal{B}'} \end{aligned}$$

Mais on a aussi :

$$(f(x))_{\mathcal{B}'} = A_{\mathcal{B}'}(f) (x)_{\mathcal{B}'}$$

et :

$$(f(x))_{\mathcal{B}} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} (f(x))_{\mathcal{B}'}$$

Il en résulte :

$$(f(x))_{\mathcal{B}} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} (f(x))_{\mathcal{B}'} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) (x)_{\mathcal{B}'} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1} (x)_{\mathcal{B}}$$

Comme

$$(f(x))_{\mathcal{B}} = A_{\mathcal{B}}(f) (x)_{\mathcal{B}}$$

on a donc :

$$A_{\mathcal{B}}(f) (x)_{\mathcal{B}} = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1} (x)_{\mathcal{B}}$$

ou encore :

$$(A_{\mathcal{B}}(f) - P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1}) (x)_{\mathcal{B}} = 0$$

Le vecteur x étant quelconque, on en déduit, en choisissant successivement pour \vec{u} chacun des vecteurs de \mathcal{B} , la nullité de chaque colonne de $A_{\mathcal{B}}(f) - P A_{\mathcal{B}'}(f) P^{-1}$:

$$A_{\mathcal{B}}(f) = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'} A_{\mathcal{B}'}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1}$$

ou encore :

$$A_{\mathcal{B}'}(f) = P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}^{-1} A_{\mathcal{B}}(f) P_{\mathcal{B} \rightsquigarrow \mathcal{B}'}$$

Définition

Deux transformations linéaires f et g de \mathbb{R}^n , de matrices associées A_f et A_g , sont dites **conjuguées** s'il existe une transformation linéaire bijective φ de \mathbb{R}^n , de matrice associée A_φ , appelée **conjugaison**, telle que :

$$f \circ \varphi = \varphi \circ g$$

ce qui, matriciellement, se traduit par :

$$A_f A_\varphi = A_\varphi A_g$$

ou encore, puisque A_φ est inversible :

$$A_g = A_\varphi^{-1} A_f A_\varphi$$

Les matrices A_f et A_g sont dites **semblables**.

Définition

Deux matrices A et B de $\mathcal{M}_n(\mathbb{R})$ sont dites **semblables** s'il existe une matrice inversible $P \in \mathcal{GL}_n(\mathbb{R})$ telle que :

$$B = P^{-1} A P$$



Les matrices A et B jouent des rôles symétriques.

2. Déterminant d'une application linéaire de \mathbb{R}^n

Proposition

Deux matrices *semblables* ont *même déterminant*.

Démonstration : Soient A et B deux matrices semblables, de taille $n \times n$.

D'après ce qui précède, il existe une matrice inversible P telle que :

$$P^{-1} A P = B$$

Il en résulte :

$$\det(B) = \det(P^{-1} A P) = \det(A P^{-1} P) = \det(A) \quad \blacksquare$$

Définition

On appelle **déterminant** d'une transformation linéaire f de \mathbb{R}^n le déterminant de sa matrice dans la base canonique, ou, de façon équivalente, dans une base quelconque \mathcal{B} .



La valeur du déterminant d'une transformation linéaire f de \mathbb{R}^n ne dépend effectivement pas de la base choisie : si \mathcal{B} et \mathcal{B}' sont deux bases de l'espace \mathbb{R}^n , les matrices $A_{\mathcal{B}}(f)$ et $A_{\mathcal{B}'}(f)$, qui représentent donc la même application linéaire dans deux bases différentes, sont semblables, et ont donc même déterminant.

1. Valeur propre, vecteur propre

Soit $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n}$ une matrice carrée d'ordre n , à coefficients réels. On appelle

valeur propre de A un réel λ tel qu'il existe un vecteur non nul $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ vérifiant :

$$AX = \lambda X$$

X est alors un **vecteur propre** de la matrice A .

L'ensemble des valeurs propres de A est appelé **Spectre** de A , et noté $Sp(A)$.

Théorème

Étant donnée une matrice $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} \in \mathcal{M}_n(\mathbb{R})$, les valeurs propres de A sont les racines du **polynôme caractéristique** de A :

$$\chi_A(\lambda) = \det(A - \lambda I_n)$$

Démonstration : S'il existe un vecteur X non nul tel que $AX = \lambda X$, $\lambda \in \mathbb{R}$, on a aussi :

$$(A - \lambda I_n)X = 0$$

Ce système linéaire admet alors une infinité de solutions ; en effet, si $X \neq 0$ est solution, alors, pour tout réel α :

$$(A - \lambda I_n)\alpha X = 0$$

Le déterminant du système est donc nul :

$$\det(A - \lambda I_n) = 0$$

Réciproquement, si λ est racine de $\det(A - \lambda I_n) = 0$, le noyau de l'application linéaire associée à la matrice $A - \lambda I_n$ n'est pas réduit au vecteur nul ; il existe donc un vecteur non nul X tel que :

$$(A - \lambda I_n)X = 0$$

λ est donc bien valeur propre de la matrice A . ■

► Multiplicité d'une valeur propre

Soient $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} \in \mathcal{M}_n(\mathbb{R})$, et λ une valeur propre de A . λ est dite **valeur propre, de multiplicité $n_\lambda \in \mathbb{N}$** , de A , si λ est racine d'ordre n_λ du polynôme caractéristique de A .

► Espace propre

Soient $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} \in \mathcal{M}_n(\mathbb{R})$, et λ une valeur propre de A .

On appelle **espace propre** associé à la valeur propre λ l'ensemble des vecteurs X tels que $AX = \lambda X$. C'est un sous-espace vectoriel de \mathbb{R}^n , que l'on note E_λ :

$$E_\lambda = \text{Ker}(A - \lambda I_n)$$

► Sous-espace caractéristique

Soient $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} \in \mathcal{M}_n(\mathbb{R})$, et λ une valeur propre de A , de multiplicité $n_\lambda \leq n$.

On appelle **sous-espace caractéristique** associé à la valeur propre λ le noyau $\text{Ker}(A - \lambda I_n)^{n_\lambda}$ de $(A - \lambda I_n)^{n_\lambda}$, c'est-à-dire l'ensemble des vecteurs X tels que :

$$(A - \lambda I_n)^{n_\lambda} X = 0$$

C'est un sous-espace vectoriel de \mathbb{R}^n .

2. Diagonalisabilité

Soit $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} \in \mathcal{M}_n(\mathbb{R})$. La matrice A est dite **diagonalisable** s'il existe une base (X_1, \dots, X_n) de \mathbb{R}^n de vecteurs propres.

Si P est la matrice de passage de la base canonique à la base (X_1, \dots, X_n) , alors la matrice $P^{-1}AP = D$ est diagonale. C'est, tout simplement, la matrice de l'application linéaire associée à A , mais exprimée dans la base (X_1, \dots, X_n) . Les coefficients diagonaux de la matrice D ainsi obtenue sont les valeurs propres de A .

Théorème

Pour qu'une matrice $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} \in \mathcal{M}_n(\mathbb{R})$, dont le polynôme caractéristique est scindé sur \mathbb{R} , soit diagonalisable, il faut et il suffit que la dimension de chaque espace propre soit égale à la multiplicité de la valeur propre associée :

$$\forall \lambda \in Sp(A) : \dim E_\lambda = \text{multiplicité}(\lambda)$$

Notation

Étant donné un polynôme P à coefficients réels, de la forme $\sum_{k=0}^N a_k X^k$, $N \in \mathbb{N}$, et une matrice $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} \in \mathcal{M}_n(\mathbb{R})$, $P(A)$ est la matrice carrée d'ordre n , à coefficients réels, donnée par :

$$P(A) = \sum_{k=0}^N a_k A^k = I_n + a_1 A + \dots + a_N A^N$$

(Par convention, $A^0 = I_n$.)

Théorème

Pour qu'une matrice $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} \in \mathcal{M}_n(\mathbb{R})$, soit diagonalisable, il faut et il suffit qu'il existe un polynôme scindé à racines simples $P \in \mathbb{R}[X]$ annulant A : $P(A) = 0$.

Théorème

Soit $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} \in \mathcal{M}_n(\mathbb{R})$. Si le polynôme caractéristique de A est scindé et à racines simples, alors A est diagonalisable.

3. Trigonalisabilité

Soit $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} \in \mathcal{M}_n(\mathbb{R})$. La matrice A est dite **trigonalisable** s'il existe une base (X_1, \dots, X_n) de \mathbb{R}^n dans laquelle la matrice de l'application linéaire associée à A est triangulaire. Les coefficients diagonaux de la matrice ainsi obtenue sont les valeurs propres de A .

Si P est la matrice de passage de la base canonique à la base (X_1, \dots, X_n) , alors la matrice $P^{-1} A P = T$ est triangulaire. C'est, tout simplement, la matrice de l'application linéaire associée à A , mais exprimée dans la base (X_1, \dots, X_n) .

Théorème

Soit $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq n} \in \mathcal{M}_n(\mathbb{R})$. Si le polynôme caractéristique de A est scindé, alors A est trigonalisable.

Exemple

Prenons le cas où $n = 3$, où le polynôme caractéristique de A est scindé, avec une seule racine λ , et où l'espace propre E_λ est de dimension 1.

Soit $A = (a_{ij})_{1 \leq i \leq 3, 1 \leq j \leq 3} \in \mathcal{M}_3(\mathbb{R})$ une matrice trigonalisable. On désigne par λ l'unique valeur propre de A , de multiplicité 3.

Trigonaliser A revient donc à déterminer une famille libre (X_1, X_2, X_3) de vecteurs tels que :

$$\begin{cases} A X_1 = \lambda X_1 \\ A X_2 = \lambda X_2 + t_{12} X_1 \\ A X_3 = \lambda X_3 + t_{13} X_1 + t_{23} X_2 \end{cases}, \quad (t_{12}, t_{13}, t_{23}) \in \mathbb{R}^3$$

Pour le premier vecteur, X_1 , il est naturel de prendre un vecteur propre associé à la valeur propre λ .

Pour le second vecteur, X_2 , il est intéressant de remarquer que la seconde relation s'écrit aussi :

$$(A - \lambda I_3) X_2 = t_{12} X_1$$

Si on multiplie à gauche par $A - \lambda I_3$, on obtient :

$$(A - \lambda I_3)^2 X_2 = t_{12} (A - \lambda I_3) X_1 = 0$$

puisque, par définition, X_1 est dans le noyau de $A - \lambda I_3$.

Il suffit donc de choisir pour X_2 un vecteur du noyau de $(A - \lambda I_3)^2$ qui soit indépendant de X_1 .

Pour le troisième vecteur, X_3 , il est intéressant de remarquer que la troisième relation s'écrit aussi :

$$(A - \lambda I_3) X_3 = t_{13} X_1 + t_{23} X_2$$

Si on multiplie à gauche par $(A - \lambda I_3)^2$, on obtient :

$$\begin{aligned} (A - \lambda I_3)^3 X_3 &= t_{13} (A - \lambda I_3)^2 X_1 + t_{23} (A - \lambda I_3)^2 X_2 \\ &= 0 \end{aligned}$$

puisque X_1 est dans le noyau de $A - \lambda I_3$, et, par construction, X_2 est dans le noyau de $(A - \lambda I_3)^2$.

Il suffit donc de choisir pour X_3 un vecteur du noyau de $(A - \lambda I_3)^3$, c'est-à-dire dans \mathbb{R}^3 (car on est en dimension 3) qui soit indépendant de X_1 et X_2 .

Il est à noter que le dernier vecteur, X_3 , peut être choisi quelconque du moment qu'il est indépendant de X_1 et X_2 : en effet, trigonaliser revient juste à exprimer la matrice de l'application linéaire associée à A dans une nouvelle base. Dans la mesure où le dernier vecteur colonne de la matrice obtenue par trigonalisation a des composantes suivant chacun des vecteurs X_1, \dots, X_3 , le choix de X_3 importe donc peu.

La réciproque se vérifie aisément.



Pour trigonaliser une matrice de $\mathcal{M}_n(\mathbb{R})$ dont le polynôme caractéristique est scindé, on procède de façon analogue.

Si $SO(\mathbb{R}^3)$ s'appelle le *groupe spécial orthogonal* de \mathbb{R}^3 , c'est, aussi, parce qu'il possède une **structure de groupe**, et constitue, donc, un ensemble muni d'une loi de composition interne, associative, admettant un élément neutre et telle que chaque élément de l'ensemble (du groupe, donc), admette un élément symétrique, ce qui est bien le cas :

si on munit $SO(\mathbb{R}^3)$ de la loi « \circ », l'élément neutre est l'*application identité* $id_{\mathbb{R}^3}$;

en composant une symétrie avec elle-même, on obtient l'identité ;

en composant une rotation de centre Ω , d'angle θ , avec la rotation de centre Ω , d'angle $-\theta$, on obtient également l'identité.

Application à la chimie moléculaire

En **chimie moléculaire**, un certain nombre de fonctions caractéristiques des molécules sont des éléments d'espaces vectoriels (c'est-à-dire un ensemble muni d'une structure permettant d'effectuer des combinaisons linéaires), qui doivent être invariantes par les opérations du groupe auquel la molécule appartient ; ces opérations, qui sont, tout simplement, des symétries ou des rotations, font en effet coïncider la molécule avec elle-même.

Un **espace vectoriel sur un corps** \mathbb{K} ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C} le plus souvent) est un ensemble E , muni d'une loi de composition interne, associative et commutative, notée, usuellement, $+$, telle que $(E, +)$ soit un groupe, et d'une loi de composition externe, appelée multiplication par un scalaire, et notée « \cdot », vérifiant les quatre **axiomes** suivants :

$$\forall (u, v) \in E \times E, \forall \lambda \in \mathbb{K} : \lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v ;$$

$$\forall (u, v) \in E \times E, \forall (\lambda, \mu) \in \mathbb{K}^2 : (\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot u ;$$

$$\forall (u, v) \in E \times E, \forall (\lambda, \mu) \in \mathbb{K}^2 : (\lambda\mu) \cdot u = \lambda \cdot (\mu \cdot u) ;$$

$$\forall u \in E : 1 \cdot u = u.$$

Ainsi, les molécules dites « **ballons de football** », ou *fullerène* C_{60} , possèdent, comme groupe de symétries, celui de l'**icosaèdre**¹.

Le groupe des rotations de l'icosaèdre est formé par les rotations de l'espace qui laissent invariante la position globale de l'icosaèdre, tout en permutant certaines faces.

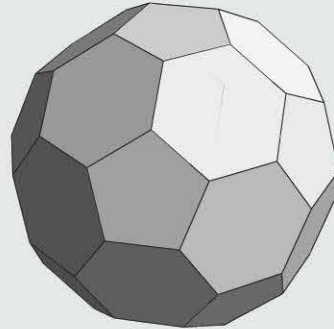
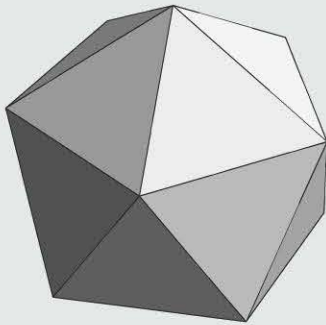
Les molécules de fullerène ont une structure d'icosaèdre tronqué, et possèdent donc les mêmes symétries que l'icosaèdre de départ.

L'importance de ces symétries est fondamentale, dans la mesure où ce sont les symétries qui peuvent aider à suggérer les formules de liaison chimiques, et donc, ensuite, de classer les molécules en fonction de leurs propriétés.

1. c'est-à-dire un *polyèdre*, solide de dimension 3, comportant, exactement, 20 faces.

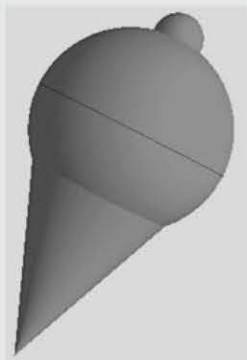


Une molécule de fullerène.



L'icosaèdre, et l'icosaèdre tronqué.

Considérons une toupie, de masse m , et lançons-la :



La toupie.

Désignons par \vec{M}_{toupie} le moment cinétique de la toupie, et $\vec{\Omega}_{toupie}$ son vecteur rotation instantanée, mesuré par rapport à son axe ; on a alors :

$$\vec{M}_{toupie} = J_{toupie} \vec{\Omega}_{toupie}$$

où J_{toupie} est la matrice d'inertie de la toupie par rapport à son axe, de la forme :

$$J_{toupie} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}$$

Si on considère un axe Δ différent de celui de la toupie, la nouvelle matrice d'inertie \tilde{J}_{toupie} comporte des termes non nuls en dehors de sa diagonale (désignés par des astérisques) :

$$\tilde{J}_{toupie} = \begin{pmatrix} \tilde{I}_1 & * & * \\ * & \tilde{I}_2 & * \\ * & * & \tilde{I}_3 \end{pmatrix}$$

Il est clair que prendre comme référence un axe différent de celui de la toupie complique donc les calculs ...

C'est à partir de cette constatation que Joseph-Louis Lagrange introduisit la notion de **diagonalisation** : diagonaliser une matrice revient, tout simplement, à en déterminer une expression équivalente (mais non égale), où les termes en dehors de la diagonale sont nuls.

On généralise ici les résultats dont on dispose sur l'espace \mathbb{R}^n , en considérant, de façon plus générale, des ensembles stables par combinaisons linéaires : les espaces vectoriels.

1. Définitions

► \mathbb{R} -espace vectoriel

On appelle **espace vectoriel** sur \mathbb{R} (ou \mathbb{R} -espace vectoriel) un triplet $(E, +, \times)$ tel que :

1. $(E, +)$ est un groupe commutatif, d'élément neutre 0_E , c'est-à-dire
 - étant donnés deux éléments x et y de E , $x + y = y + x$ est aussi élément de E , et $x + 0_E = 0_E + x = x$;
 - tout élément x de E admet l'élément $-x$ comme opposé :

$$x + (-x) = (-x) + x = 0_E$$

- étant donnés trois éléments x, y et z de E , $x + (y + z) = (x + y) + z$ (cela signifie que la loi « $+$ » est associative) ;
2. étant donnés deux éléments x et y de E , et un réel λ , $\lambda(x + y) = (\lambda x) + (\lambda y)$ est aussi élément de E ;
 3. étant donnés un élément x de E , et deux réels λ et μ , $(\lambda + \mu)x = (\lambda x) + (\mu x)$ est aussi élément de E ;
 4. étant donnés un élément x de E , et deux réels λ et μ :

$$(\lambda\mu)x = \lambda(\mu x)$$

5. étant donnés un élément x de E , le réel 1 est l'élément neutre de la multiplication par un scalaire :

$$1 \times x = x$$

Les éléments de $(E, +, \times)$ sont appelés **vecteurs**.

\mathbb{R} est le **corps de base** de $(E, +, \times)$.

Exemples

1. $(\mathbb{R}, +, \times)$ est un \mathbb{R} -espace vectoriel.
2. Étant donné un entier naturel non nul n , $(\mathbb{R}^n, +, \times)$ est un \mathbb{R} -espace vectoriel.
3. L'espace $C^1([0, 1], \mathbb{R})$ des fonctions de classe C^1 sur $[0, 1]$ et à valeurs dans \mathbb{R} est un \mathbb{R} -espace vectoriel.
4. L'espace $\mathbb{R}[X]$ des polynômes à coefficients réels \mathbb{R} est un \mathbb{R} -espace vectoriel.

► \mathbb{K} -espace vectoriel

De façon plus générale, si \mathbb{K} est un corps¹, on appelle **espace vectoriel** sur le corps \mathbb{K} (ou \mathbb{K} -espace vectoriel) un triplet $(E, +, \times)$ tel que :

1. $(E, +)$ est un groupe commutatif, d'élément neutre 0_E , c'est-à-dire :
 - étant donnés deux éléments x et y de E , $x + y = y + x$ est aussi élément de E , et $x + 0_E = 0_E + x = x$;
 - tout élément x de E admet l'élément $-x$ comme opposé :

$$x + (-x) = (-x) + x = 0_E$$

- étant donnés trois éléments x, y et z de E , $x + (y + z) = (x + y) + z$ (cela signifie que la loi « $+$ » est associative) ;
2. étant donnés deux éléments x et y de E , et λ dans \mathbb{K} , $\lambda(x + y) = (\lambda x) + (\lambda y)$ est aussi élément de E ;
 3. étant donnés deux éléments x et y de E , $x + y$ est aussi élément de E ;
 4. étant donné un élément x de E , et λ et μ dans \mathbb{K} , $(\lambda + \mu)x$ est aussi élément de E ;
 5. étant donné un élément x de E , et λ et μ dans \mathbb{K} :

$$(\lambda \mu)x = \lambda(\mu x)$$

6. étant donné un élément x de E , l'élément unité $1_{\mathbb{K}}$ de \mathbb{K} est l'élément neutre de la multiplication par un scalaire :

$$1_{\mathbb{K}} \times x = x$$

Les éléments de $(E, +, \times)$ sont appelés **vecteurs**.

\mathbb{K} est le **corps de base** de $(E, +, \times)$.

Notation

1. Pour alléger les écritures, un \mathbb{K} -espace vectoriel $(E, +, \times)$ sera noté E .
2. Les vecteurs d'un espace vectoriel seront notés « sans flèche » : x , pour les distinguer des vecteurs du plan ou de l'espace.

2. Vecteurs

► Combinaison linéaire

Soit n un entier naturel supérieur ou égal à 2. Étant donné un \mathbb{K} -espace vectoriel E , et n vecteurs x_1, \dots, x_n de E , on appelle combinaison linéaire de x_1, \dots, x_n tout vecteur de la forme :

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n = \sum_{i=1}^n \lambda_i x_i$$

Ainsi, un \mathbb{K} -espace vectoriel E est stable par combinaisons linéaires à coefficients dans \mathbb{K} : **toute combinaison linéaire d'éléments de E appartient à E .**

1. Par exemple : \mathbb{C} , \mathbb{Z} , etc.



► Vecteurs liés

Deux vecteurs x et y sont dits **liés** s'il existe une combinaison linéaire non triviale de ces vecteurs égale au vecteur nul, c'est-à-dire lorsque les coefficients de la combinaison sont non simultanément nuls, de la forme :

$$\lambda x + \mu y = 0 \quad (\lambda, \mu) \neq (0, 0)$$

► Vecteurs libres

Deux vecteurs x et y sont dits **libres** s'ils ne sont pas liés.

Généralisation

n vecteurs x_1, \dots, x_n sont dits **liés** s'il existe une combinaison linéaire non triviale de ces vecteurs égale au vecteur nul, c'est-à-dire lorsque les coefficients de la combinaison sont non simultanément nuls, de la forme :

$$\sum_{i=1}^n \lambda_i x_i = 0$$

où $\lambda_1, \dots, \lambda_n$ sont des réels tels que $(\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)$.

n vecteurs x_1, \dots, x_n sont dits **libres** s'ils ne sont pas liés

Propriété

Pour toute famille (x_1, \dots, x_n) de vecteurs libres :

$$\lambda_1 x_1 + \dots + \lambda_n x_n = 0 \Rightarrow (\lambda_1, \dots, \lambda_n) = (0, \dots, 0)$$

ce qui signifie que, s'il existe une combinaison linéaire nulle de ces vecteurs, alors les coefficients de cette même combinaison linéaire sont **nécessairement** nuls.

3. Dimension d'un espace vectoriel

► Espace vectoriel de dimension finie

Un \mathbb{K} -espace vectoriel E est dit **de dimension finie** s'il possède une partie génératrice finie.

Exemple

\mathbb{R}^3 est de dimension 3 (toute base de \mathbb{R}^3 possède trois éléments).

► Espace vectoriel de dimension infinie

Un \mathbb{K} -espace vectoriel E est dit **de dimension infinie** s'il ne possède pas de partie génératrice finie.

Exemple

L'espace $C^1([0, 1], \mathbb{R})$ des fonctions de classe C^1 sur $[0, 1]$ et à valeurs dans \mathbb{R} est de dimension infinie (on ne peut pas trouver de famille génératrice).

► Espace vectoriel produit

Théorème

n étant un entier naturel supérieur ou égal à 2, et E_1, \dots, E_n des \mathbb{K} -espaces vectoriels, on munit l'ensemble $E_1 \times \dots \times E_n$ des lois $+$ et \times en posant, pour tous n -uplets (x_1, \dots, x_n) et (y_1, \dots, y_n) d'éléments de $E_1 \times \dots \times E_n$, et tout λ de \mathbb{K} :

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n) \quad , \quad \lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$$

$(E_1 \times \dots \times E_n, +, \times)$ est alors un \mathbb{K} -espace vectoriel.

1. Définition

On appelle **sous-espace vectoriel** d'un espace vectoriel E un ensemble F de E tel que :

1. $0 \in F$;
2. étant donnés deux éléments x et y de F , $x+y$ est aussi élément de F (ce qui signifie que F est stable par addition) ;
3. étant donnés un élément x de F , et λ dans \mathbb{K} , λx est aussi élément de F (ce qui signifie que F est stable par la multiplication par un scalaire) ;

► Caractérisation des sous-espaces vectoriels

Théorème

Un ensemble F de E est un sous-espace vectoriel du \mathbb{K} -espace vectoriel E si et seulement si, étant donnés deux éléments x et y de F , et λ dans \mathbb{K} , $x + \lambda y$ est aussi élément de F (ce qui signifie que F est stable par combinaisons linéaires) :

$$\forall (x, y) \in F \times F, \forall \lambda \in \mathbb{K} : x + \lambda y \in F$$

2. Transformations linéaires d'un espace vectoriel E

► Définition

On appelle **transformation linéaire d'un espace vectoriel E** , ou encore **application linéaire d'un espace vectoriel E** une application f , de E dans E , linéaire :

$$\forall (x, y) \in E \times E, \forall \lambda \in \mathbb{K} : f(x + \lambda y) = f(x) + \lambda f(y)$$

► Une application linéaire remarquable : l'application identité

On appelle **application identité de l'espace vectoriel E** l'application, notée Id_E :

$$x \in E \mapsto x$$

► Propriétés

Une transformation linéaire f d'un espace vectoriel E est dite **injective** si :

$$\forall (x_1, x_2) \in E \times E : f(x_1) = f(x_2) \Rightarrow x_1 = x_2$$

Une transformation linéaire f d'un espace vectoriel E est dite **surjective** si :

$$\forall y \in E, \exists x \in E : y = f(x)$$

Une transformation linéaire d'un espace vectoriel E est un **automorphisme** de E si elle est *bijective*, c'est-à-dire injective et surjective.

L'ensemble des automorphismes d'un espace vectoriel E est appelé **groupe linéaire de E** ¹, et noté $\mathcal{GL}(E)$.

On appelle **noyau** d'une transformation linéaire f d'un espace vectoriel E , que l'on note $\text{Ker } f$ (de l'allemand *kern*) l'ensemble des vecteurs de E dont l'image par f est le vecteur nul :

$$\text{Ker } f = \{x \in E / f(x) = 0\}$$

Définition

On appelle **image** d'une transformation linéaire f d'un espace vectoriel E , que l'on note $\text{Im } f$ l'ensemble des images des vecteurs de E par f :

$$\text{Im } f = \{f(x), x \in E\}$$

Proposition

Étant donnée une transformation linéaire f d'un espace vectoriel E , $\text{Ker } f$ et $\text{Im } f$ sont deux sous-espaces vectoriels de E .

1. car il possède une structure de groupe, c'est-à-dire est stable par produit, et tout élément non nul est inversible.

Somme de sous-espaces vectoriels

F_1 et F_2 étant deux sous-espaces vectoriels d'un espace vectoriel E , l'ensemble

$$F_1 + F_2 = \{x_1 + x_2, x_1 \in F_1, x_2 \in F_2\}$$

est un sous-espace vectoriel de E , appelé **somme de F_1 et F_2** .

► Somme directe de sous-espaces vectoriels

Deux sous-espaces vectoriels F_1 et F_2 d'un espace vectoriel E sont dits **en somme directe** si tout élément du sous-espace somme $F_1 + F_2$ s'écrit de manière unique comme somme d'un élément de F_1 et d'un élément de F_2 :

$$\forall x \in F_1 + F_2, \exists! (x_1, x_2) \in F_1 \times F_2 : x = x_1 + x_2$$

On note alors :

$$F_1 \oplus F_2$$

► Caractérisation d'une somme directe

Deux sous-espaces vectoriels F_1 et F_2 d'un espace vectoriel E sont **en somme directe dans E** si et seulement si

$$F_1 \cap F_2 = \{0\}$$

► Sous-espaces vectoriels supplémentaires

Deux sous-espaces vectoriels F_1 et F_2 d'un espace vectoriel E sont dits **supplémentaires dans E** si tout élément de E s'écrit de manière unique comme somme d'un élément de F_1 et d'un élément de F_2 :

$$\forall x \in E, \exists! (x_1, x_2) \in F_1 \times F_2 : x = x_1 + x_2$$

On note alors :

$$F_1 \oplus F_2 = E$$

► Caractérisation de la supplémentarité

Deux sous-espaces vectoriels F_1 et F_2 d'un espace vectoriel E sont **supplémentaires dans E** si et seulement si :

$$F_1 + F_2 = E \quad \text{et} \quad F_1 \cap F_2 = \{0\}$$

1. Définitions

► Projection – Projecteurs

Soient E un espace vectoriel, et F_1 et F_2 deux sous-espaces vectoriels de E tels que :

$$E = F_1 \oplus F_2$$

Pour tout vecteur x de E , on considère l'unique couple de vecteurs (x_1, x_2) de $F_1 \times F_2$ tel que :

$$x = x_1 + x_2$$

On appelle **projection (ou projecteur) sur F_1 suivant la direction F_2** (ou parallèlement à F_2) l'application p_1 qui, au vecteur x , associe sa composante x_1 sur F_1 :

$$p_1(x) = p_1(x_1 + x_2) = x_1$$

On appelle **projection (ou projecteur) sur F_2 suivant la direction F_1** (ou parallèlement à F_1) l'application p_2 qui, au vecteur x , associe sa composante x_2 sur F_2 :

$$p_2(x) = p_2(x_1 + x_2) = x_2$$

► Symétrie

Soient E un espace vectoriel, et F_1 et F_2 deux sous-espaces vectoriels de E tels que :

$$E = F_1 \oplus F_2$$

Pour tout vecteur x de E , on considère l'unique couple de vecteurs (x_1, x_2) de $F_1 \times F_2$ tel que :

$$x = x_1 + x_2$$

On appelle **symétrie par rapport à F_1 parallèlement à F_2** l'application s_1 qui, au vecteur x , associe le vecteur $x_1 - x_2$:

$$s_1(x) = s_1(x_1 + x_2) = x_1 - x_2$$

On appelle **symétrie par rapport à F_2 parallèlement à F_1** l'application s_2 qui, au vecteur x , associe le vecteur $x_2 - x_1$:

$$s_2(x) = s_2(x_1 + x_2) = x_2 - x_1$$

2. Propriétés

► Caractérisation des projecteurs

Théorème

Soient E un espace vectoriel, et p une transformation linéaire de E .
Alors, p est un projecteur de E si et seulement si :

$$p \circ p = p$$

On dit alors que p est **idempotent**.

Théorème

Soient E un espace vectoriel, et p un projecteur de E . Alors :

$$E = \text{Ker } p \oplus \text{Im } p$$

Démonstration : Pour tout x de E :

$$x = x - p(x) + p(x)$$

Comme $p(x - p(x)) = p(x) - p(x) = 0$, on a donc : $x - p(x) \in \text{Ker } p$.

De façon évidente : $p(x) \in \text{Im } p$.

Par suite : $E = \text{Ker } p + \text{Im } p$.

Il reste à montrer que la somme est directe. À cet effet, on considère un élément y de $\text{Ker } p \cap \text{Im } p$.

Comme $y \in \text{Im } p$, il existe un élément x de E tel que : $y = p(x)$.

Or, comme $y \in \text{Ker } p$, $p(y) = 0$, soit $p(p(x)) = 0$, soit $p(x) = 0$.

Par suite : $x \in \text{Ker } p$, et $y = p(x) = 0$. La somme $\text{Ker } p + \text{Im } p$ est donc directe :

$$\text{Ker } p + \text{Im } p = \text{Ker } p \oplus \text{Im } p$$

■

► Caractérisation des symétries

Théorème

Soient E un espace vectoriel, et s une transformation linéaire de E .
Alors, s est une symétrie de E si et seulement si :

$$s \circ s = \text{Id}_E$$

On dit alors que s est **involutif**.

Théorème

Soient E un espace vectoriel, et s une symétrie de E . Alors :

$$E = \text{Ker } (s - \text{Id}_E) \oplus \text{Ker } (s + \text{Id}_E)$$

Exercices d'entraînement



Les corrigés sont disponibles en téléchargement sur le site dunod.com à partir de la page de présentation de l'ouvrage.

Le plan complexe

Racines $n^{\text{ièmes}}$

1 Résoudre, dans \mathbb{C} : $z^4 = 1$.

2 Résoudre, dans \mathbb{C} : $z^5 = 1$.

3 On considère l'équation :

$$(z+1)^{2n} = (z-1)^{2n}$$

où n est un entier naturel non nul.

3.a) Montrer que l'on ne peut pas avoir $z = 1$.

3.b) Montrer que l'équation donnée admet exactement $2n - 1$ racines, que l'on désignera par $\omega_1, \dots, \omega_{2n-1}$, et qui sont telles que :
pour $k \in \{1, \dots, 2n - 1\}$:

$$\omega_k = -i \cotan\left(\frac{k\pi}{2n}\right)$$

3.c) Pour aller plus loin :
que vaut le produit des racines non nulles ?

Trigonométrie

4 Linéarisation

θ étant un réel non nul, linéariser : $\cos^3 \theta$, $\sin^3 \theta$ et $\cos^4 \theta$.

5 Calculs de sommes

Soit $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$, et $n \in \mathbb{N}^*$. Donner, en fonction de θ , la valeur des sommes suivantes :

$$5.a) S_1(\theta) = \sum_{k=0}^n e^{ik\theta}.$$

$$5.b) S_2(\theta) = \sum_{k=-n}^n e^{2ik\theta}.$$

$$5.c) S_3(\theta) = \sum_{k=1}^n \cos(k\theta) \text{ et } S_4(\theta) = \sum_{k=1}^n \sin(k\theta).$$

Donner ensuite les relations entre $S_1(\theta)$, $S_2(\theta)$, $S_3(\theta)$, $S_4(\theta)$.

Transformations du plan

Translations, homothéties

Le plan euclidien orienté est rapporté à un repère orthonormé direct $(O; \vec{i}, \vec{j})$.

6 Expression analytique d'une translation

Donner l'expression de l'affixe z' du point M' image du point M d'affixe $z \in \mathbb{C}$ par :

6.a) la translation $\tau_{\vec{u}}$ de vecteur $\vec{u}(1, -2)$;

6.b) la translation $\tau_{\vec{v}}$ de vecteur $\vec{v}(3, 1)$.

7 Expression analytique d'une homothétie

Donner l'expression de l'affixe z' du point M' image du point M d'affixe $z \in \mathbb{C}$ par :

7.a) l'homothétie $h_{O,2}$, de centre O , de rapport 2 ;

7.b) l'homothétie $h_{A,-3}$, de centre $A(1, 1)$, de rapport -3 ;

7.c) l'homothétie $h_{B,4}$, de centre $B(-2, 0)$, de rapport 4.

Rotations

8 Expression analytique d'une rotation

8.a) Déterminer l'expression de l'affixe z' du point M' image du point M d'affixe $z \in \mathbb{C}$, distinct de A , par la rotation $r_{A, \frac{\pi}{4}}$, de centre $A(1, 1)$, d'angle $\frac{\pi}{4}$.

- 8.b) Déterminer l'expression de l'affixe z' du point M' image du point M d'affixe $z \in \mathbb{C}^*$ par la rotation $r_{O,\pi}$, de centre O , d'angle π .

9 Un peu de formalisme ...

- 9.a) Donner l'expression de l'affixe z' du point M' image du point M d'affixe $z \in \mathbb{C}$ par la rotation $r_{\Omega,\theta}$, de centre Ω , d'angle θ (on désignera par z_Ω l'affixe de Ω).
- 9.b) Donner l'expression de l'affixe z'' du point M'' image du point M d'affixe $z \in \mathbb{C}$ par la rotation $r_{\Omega',\theta'}$, de centre Ω' , d'angle θ' (on désignera par $z_{\Omega'}$ l'affixe de Ω').
- 9.c) Donner l'expression de l'affixe z''' du point M''' image du point M d'affixe $z \in \mathbb{C}$ par la composée $r = r_{\Omega',\theta'} \circ r_{\Omega,\theta}$ de $r_{\Omega,\theta}$ et $r_{\Omega',\theta'}$.

Et, de même, on détermine la composée de deux homothéties de centres distincts, la composée d'une rotation et d'une translation, ... : c'est à vous !

Introduction aux matrices

Calcul matriciel

10 Calculs pratiques

On considère : $X = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $Z = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$$

$$B = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{pmatrix}$$

Quels sont les produits matriciels possibles ? Les calculer.

11 Une matrice remarquable

Soit $(a, b, c) \in \mathbb{R}^3$, tel que $a^2 + b^2 + c^2 = 1$. On pose :

$$M = \begin{pmatrix} 1+a^2 & ab & ac \\ ab & 1+b^2 & bc \\ ac & bc & 1+c^2 \end{pmatrix} \text{ et } N = M - I_3.$$

- 11.a) Montrer que N peut s'écrire comme produit d'une matrice colonne et d'une matrice ligne.

- 11.b) Calculer N^n , où n désigne un entier naturel.

- 11.c) En déduire l'expression de M^n .

12 Puissances $n^{\text{ièmes}}$ d'une matrice

Soit a un réel quelconque, supposé non nul. On considère la matrice carrée d'ordre n ($n \geq 2$) :

$$A = \begin{pmatrix} 0 & a & \dots & \dots & a \\ a & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & a \\ a & \dots & \dots & a & 0 \end{pmatrix}$$

- 12.a) Montrer que l'on peut exprimer A^2 en fonction de A et I_n , où I_n désigne la matrice identité d'ordre n .

Pour aller plus loin :

- 12.b) On considère le polynôme P tel que :

$$P(X) = X^2 - (n-2)aX - (n-1)a^2$$

On note

$$P(A) = A^2 - (n-2)aA - (n-1)a^2I_n.$$

Vérifier que $P(A) = 0$.

- 12.c) Soit k un entier naturel non nul. Effectuer la division euclidienne de X^k par P , en montrant qu'elle conduit à un résultat de la forme :

$$X^k = P(X)Q(X) + \alpha X + \beta$$

où Q est un polynôme que l'on cherchera pas à expliciter, et α et β deux réels.

Comment peut-on calculer simplement α et β ?

- 12.d) En remarquant que

$A^k = P(A)Q(A) + \alpha A + \beta I_n$, en déduire une méthode de calcul de A^k .

Déterminants de matrices de taille $n \times n$

13 Un calcul pratique

Calculer le déterminant $n \times n$ suivant :

$$\begin{vmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 1 & \dots & 1 & 0 \end{vmatrix}.$$

14 n étant un entier naturel supérieur ou égal à 2, soient a et b deux réels distincts, non nuls, et (x_1, \dots, x_n) une famille de réels. Pour tout réel x , on pose :

$$\Delta_n(x) = \begin{vmatrix} x_1 + x & a + x & \dots & a + x \\ b + x & x_2 + x & \ddots & \vdots \\ \vdots & \ddots & \ddots & a + x \\ b + x & \dots & b + x & x_n + x \end{vmatrix}.$$

- 14.a) Montrer que la fonction $x \mapsto \Delta_n(x)$ est une fonction affine de x , c'est-à-dire de la forme $x \mapsto \alpha x + \beta$, où $(\alpha, \beta) \in \mathbb{R}^2$ (on ne cherchera pas à déterminer α et β dans cette question).
- 14.b) Que valent $\Delta_n(-a)$ et $\Delta_n(-b)$? En déduire les valeurs des coefficients α et β introduits à la question précédente.

15 Un déterminant remarquable : le Déterminant de Vandermonde¹

Soit (x_1, \dots, x_n) une famille de réels, et $x \in \mathbb{R}$.

On considère les déterminants :

$$V_n(x_1, \dots, x_n) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix},$$

$$\text{et } \widetilde{V}_n(x) = V_n(x_1, \dots, x_{n-1}, x) = \begin{vmatrix} 1 & 1 & \dots & 1 & 1 \\ x_1 & x_2 & \dots & x_{n-1} & x \\ x_1^2 & x_2^2 & \dots & x_{n-1}^2 & x^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_{n-1}^{n-1} & x^{n-1} \end{vmatrix}.$$

- 15.a) Montrer que $\widetilde{V}_n(x)$ est un polynôme de degré inférieur ou égal à $n-1$ en x .
- 15.b) On suppose, dans toute la suite, que les $x_i, i = 1, \dots, n$, sont deux à deux distincts. Montrer que $\widetilde{V}_n(x_1) = \widetilde{V}_n(x_2) = \dots = \widetilde{V}_n(x_{n-1})$, et en déduire qu'il existe une constante réelle C_n telle que, pour tout réel x :

$$\begin{aligned} \widetilde{V}_n(x) &= C_n (x - x_1) \dots (x - x_{n-1}) \\ &= C_n \prod_{i=1}^{n-1} (x - x_i) \end{aligned}$$

- 15.c) Déterminer la constante C_n , et en déduire l'expression de $V_n(x_1, \dots, x_n)$.
- 15.d) Que peut-on dire en ce qui concerne le cas où les $x_i, i = 1, \dots, n$, ne sont pas deux à deux distincts ?

Inversion des matrices carrées

16 Inversibilité et puissances $n^{\text{ièmes}}$

- 16.a) Après avoir vérifié leur inversibilité, inverser les matrices suivantes :

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

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

- 16.b) n étant un entier naturel non nul, soit $A \in \mathcal{M}_n(\mathbb{R})$ telle que $A^2 = A$, avec $A \neq I_n$. A est-elle inversible ?
- 16.c) Soit a un réel quelconque, supposé non nul. On considère la matrice carrée d'ordre n ($n \geq 2$) :

$$A = \begin{pmatrix} 0 & a & \dots & a \\ a & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & a \\ a & \dots & \dots & a & 0 \end{pmatrix}$$

1. Alexandre-Théophile Vandermonde (1735-1796), mathématicien français, mais aussi économiste, musicien et chimiste. D'après Lebesgue (Conférence d'Utrecht, 1937, [23]), le déterminant ne serait pas de lui ...

- Montrer que l'on peut exprimer A^2 en fonction de A et I_n , où I_n désigne la matrice identité d'ordre n .
- Pour quelles valeurs de a la matrice A est-elle inversible ? Déterminer, dans ce cas, sa matrice inverse A^{-1} .

17 On considère les matrices :

$$M = \begin{pmatrix} 5 & 3 \\ 1 & 3 \end{pmatrix}, \quad P = \begin{pmatrix} 3 & 1 \\ 1 & -1 \end{pmatrix}$$

- 17.a)** Montrer que P est inversible, et calculer son inverse.
17.b) Que vaut $P^{-1} M P$?
17.c) Montrer que, pour tout entier naturel non nul n :

$$(P^{-1} M P)^n = P^{-1} M^n P$$

et en déduire une méthode de calcul de M^n .

Systèmes linéaires

18 Résolution d'un système de taille 3×3

Résoudre :

$$\begin{cases} x - y - z = 4 \\ 2x + 2y + z = -5 \\ 3x - y - z = 6 \end{cases}$$

19 Un système paramétré, une solution à discuter !

Soit a un réel non nul. Résoudre :

$$\begin{cases} x + ay + a^2 z = a^4 \\ x + y + z = 1 \\ x - y + z = -1 \end{cases}$$

L'espace réel à 3 dimensions

Vecteurs

20 Déterminer si les vecteurs suivants sont linéairement indépendants ou non :

20.a) $\vec{u} = (1, 2, 3)$ et $\vec{v} = (-1, 3, 2)$:

20.b) $\vec{u} = (2, 1, 3)$, $\vec{v} = (4, 2, 6)$.

20.c) $\vec{u} = (3, 2, 1)$, $\vec{v} = (3, 0, 1)$.

20.d) $\vec{u} = (2, 4, 6)$, $\vec{v} = (4, 2, 6)$, $\vec{w} = (6, 4, 2)$.

20.e) $\vec{u} = (-1, 0, 1)$, $\vec{v} = (1, 1, 1)$, $\vec{w} = (0, 1, 2)$.

Droites et plans

21 Droites

Déterminer une représentation paramétrique, puis un système d'équations cartésiennes, de la droite \mathcal{D} passant par le point $A(-1, 2, 1)$, et dirigée par le vecteur $\vec{u} = (1, 0, 1)$.

De quel type de droite s'agit-il, affine ou vectorielle ?

22 Déterminer une représentation paramétrique, puis un système d'équations cartésiennes, de la droite \mathcal{D}' passant par le point $B(3, 0, 2)$, et orthogonale au plan \mathcal{P} d'équation $2x + y - z = 3$.

23 Déterminer une représentation paramétrique, puis un système d'équations cartésiennes, de la droite Δ passant par les points $C(-2, 1, 2)$, et $D(-4, 0, 1)$.

24 On considère la droite Δ' dont une représentation paramétrique est donnée par :

$$\begin{cases} x = 1 + 2\lambda \\ y = 3\lambda \\ z = 2 - 2\lambda \end{cases}, \quad \lambda \in \mathbb{R}$$

Donner un vecteur directeur de cette droite.

Plans

25 Déterminer une équation cartésienne du plan \mathcal{P}_1 passant par le point $A(-1, 2, 1)$, et engendré par les vecteurs $\vec{u} = (1, 1, 1)$ et $\vec{v} = (1, -1, 1)$.

26 Déterminer une équation cartésienne du plan \mathcal{P}_2 passant par le point $B(1, 1, 1)$, et orthogonal au vecteur $\vec{n} = (2, 3, 4)$.

27 On considère le plan vectoriel \mathcal{P}_4 dont une représentation paramétrique est donnée par :

$$\begin{cases} x = \lambda - \mu \\ y = \lambda + \mu \\ z = \lambda - \mu \end{cases}, \quad (\lambda, \mu) \in \mathbb{R}^2$$

Donner deux vecteurs engendrant \mathcal{P}_4 .

Aires et volumes

Déterminants et calculs d'aires

Dans ce qui suit, le plan euclidien orienté de dimension 2 est rapporté à un repère orthonormé direct $(O; \vec{i}, \vec{j})$.

28 Aire d'un parallélogramme

28.a) Soient $a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4$ des réels non nuls, vérifiant :

$$a_1 < a_2, \quad a_4 < a_3, \quad b_1 < b_4,$$

$$b_2 < b_3$$

On considère les points $A_1(a_1, b_1)$, $A_2(a_2, b_2)$, $A_3(a_3, b_3)$, $A_4(a_4, b_4)$: à quelle condition le quadrilatère $A_1A_2A_3A_4$ est-il un parallélogramme non aplati ?

28.b) Lorsque ces conditions sont vérifiées, calculer l'aire orientée, puis non orientée, du parallélogramme $A_1A_2A_3A_4$.

28.c) On suppose désormais que $\overrightarrow{A_1A_2} = \overrightarrow{A_1A_4} = \vec{a} > 0$, $b_1 = b_2$, $a_4 = a_1$. Que remarque-t-on ?

29 Aire d'un triangle

Soient a, b, c, d des réels non nuls. Calculer l'aire orientée, puis non orientée, du triangle de sommets $O, A(a, b), B(c, d)$.

Déterminants et calculs de volumes

Dans ce qui suit, l'espace euclidien orienté de dimension 3 est rapporté à un repère orthonormé direct $(O; \vec{i}, \vec{j}, \vec{k})$.

30 Soient a, b, c, d des réels non nuls.

30.a) On considère les points $A(1, a^2, a^4)$, $B(2, a^2 + c^2, a^4 + c^4)$, $C(1, c^2, c^4)$, $D(1, d^2, d^4)$: quelle est la nature du quadrilatère $OABC$?

30.b) On considère les points E, F, G , images respectives de A, B, C par la translation de vecteur \overrightarrow{OD} : quelle est la nature de $OABCDEFG$?

30.c) Calculer le volume orienté, puis le volume non orienté, de $OABCDEFG$.

Transformations linéaires du plan

31 Linéarité - ou non ?

Parmi les applications suivantes, lesquelles sont linéaires ? Donner, lorsque c'est possible, la matrice de l'application.

31.a) $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto (4x + y, 2y + x)$;

31.b) $g : \mathbb{R}^3 \rightarrow \mathbb{R}, (x, y) \mapsto \sin(xy)$;

32 Applications linéaires et matrices

On considère l'application linéaire :

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 \\ (x, y) \mapsto (3x - y, 2y + x)$$

32.a) Quelle est la matrice A_f associée à f dans la base canonique ?

32.b) Calculer $\det A_f$, et en conclure que f est un *automorphisme* de \mathbb{R}^2 , c'est-à-dire une application linéaire bijective de \mathbb{R}^2 dans \mathbb{R}^2 .

32.c) Déterminer le noyau de f , et retrouver ainsi le résultat du *i*.

32.d) Expliciter l'application réciproque f^{-1} .

Transformations linéaires de l'espace

33 Linéarité - ou non ?

Parmi les applications suivantes, lesquelles sont linéaires ? Donner, lorsque c'est possible, la matrice de l'application.

33.a) $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto (4x + y, 2y + x)$;

33.b) $g : \mathbb{R}^3 \rightarrow \mathbb{R}, (x, y, z) \mapsto \cos(xyz)$;

33.c) $h : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \\ (x, y, z) \mapsto (x + y + z, x - y + z, 0)$;

34 Applications linéaires et matrices

On considère les applications :

$$\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^3 \\ (x, y) \mapsto (3x - y, 2y + x, x - y)$$

$$\psi : \mathbb{R}^3 \rightarrow \mathbb{R}^2 \\ (x, y, z) \mapsto (x + z, y - z)$$

34.a) Quelles sont les matrices A_φ et A_ψ respectivement associées à φ et ψ dans la base canonique ?

- 34.b)** Expliciter l'application composée $\psi \circ \varphi$ de deux façons différentes (on donnera, notamment, la matrice $A_{\psi \circ \varphi}$ associée à $\psi \circ \varphi$ dans la base canonique). Que remarque-t-on ?

L'espace \mathbb{R}^n

Dans ce qui suit, n est un entier supérieur ou égal à 2.

Applications linéaires

- 35** On considère l'application linéaire f dont la matrice dans la base canonique de \mathbb{R}^n est :

$$A_f = \begin{pmatrix} 0 & 1 & \dots & \dots & 1 \\ 1 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & 1 \\ 1 & \dots & \dots & 1 & 0 \end{pmatrix}$$

- 35.a)** Montrer que f est un automorphisme de \mathbb{R}^n .
35.b) On désigne par $Id_{\mathbb{R}^n}$ l'application identité de \mathbb{R}^n . Montrer, de deux façons différentes, que $f + Id_{\mathbb{R}^n}$ n'est pas un automorphisme de \mathbb{R}^n .
35.c) Donner une base du noyau de $f - (n - 1)Id_{\mathbb{R}^n}$.

- 36** Soit φ une application linéaire de \mathbb{R}^n . Montrer l'équivalence suivante :

$$\text{Ker } \varphi = \text{Ker } \varphi^2 \Leftrightarrow \mathbb{R}^n = \text{Ker } \varphi \oplus \text{Im } \varphi$$

- 37** Soient f et g deux applications linéaires de \mathbb{R}^n telles que :

$$f \circ g = 0 \quad \text{et} \quad f + g \text{ injective}$$

Comparer $\text{Im } g$ et $\text{Ker } f$.

- 38** Soit f une application linéaire de \mathbb{R}^n nilpotente d'ordre n , c'est-à-dire telle que

$$f^n = 0 \quad \text{et} \quad f^{n-1} \neq 0$$

(la puissance est au sens de la composition)

Comparer $\text{Im } g$ et $\text{Ker } f$.

- 38.a)** Montrer qu'il existe une base \mathcal{B} de \mathbb{R}^n dans laquelle la matrice $A_{\mathcal{B}}(f)$ de f est triangulaire.

- 38.b)** Que vaut $\det(f + Id_{\mathbb{R}^n})$?

- 38.c)** Soit g une application linéaire de \mathbb{R}^n commutant avec f : que vaut $\det(f + g)$?

Réduction des matrices carrées

- 39** On considère la matrice

$$M_1 = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

- 39.a)** Déterminer les valeurs propres de M_1 .
39.b) Déterminer les sous-espaces propres de M_1 . M_1 est-elle diagonalisable ?
39.c) Déterminer la matrice de passage P_1 à une base de diagonalisation, ainsi que son inverse.
39.d) Donner une matrice diagonale D_1 semblable à M_1 . Que peut-on en déduire pour l'inversibilité de M_1 ?
39.e) Calculer, pour tout entier naturel non nul n , M_1^n .

- 40** On considère la matrice

$$M_2 = \begin{pmatrix} 0 & 4 \\ 1 & 3 \end{pmatrix}$$

- 40.a)** Déterminer les valeurs propres de M_2 .
40.b) Déterminer les sous-espaces propres de M_2 . M_2 est-elle diagonalisable ?
40.c) Déterminer la matrice de passage P_2 à une base de diagonalisation, ainsi que son inverse.
40.d) Donner une matrice diagonale D_2 semblable à M_2 . Que peut-on en déduire pour l'inversibilité de M_2 ?
40.e) Calculer, pour tout entier naturel non nul n , M_2^n .

- 41** On considère la matrice :

$$A = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ - & -1 & 3 \end{pmatrix}$$

- 41.a) Déterminer les valeurs propres de A .
 41.b) Déterminer les sous-espaces propres de A . A est-elle diagonalisable ?
 41.c) Déterminer la matrice de passage P_A à une base de diagonalisation, ainsi que son inverse.
 41.d) Donner une matrice diagonale D_A semblable à A .
 41.e) Calculer, pour tout entier naturel non nul n , A^n .

42 Soient a et b deux réels non nuls. On considère la matrice carrée d'ordre n ($n \geq 2$) :

$$A = \begin{pmatrix} b & a & \dots & \dots & a \\ a & b & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & b & a \\ a & \dots & \dots & a & b \end{pmatrix}$$

Déterminer les valeurs propres de A , puis ses espaces propres, et montrer que A est diagonalisable.

43 Soit α un réel non nul. On considère la matrice :

$$M = \begin{pmatrix} -1 & \alpha & -\alpha \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

- 43.a) Déterminer le polynôme caractéristique de M .
 43.b) Déterminer les espaces propres de M .
 43.c) M est-elle trigonalisable ? Si oui, la trigonaliser.

44 **Pour aller plus loin : Réduction de matrices blocs**

Soit n un entier supérieur ou égal à 2.

Soit A une matrice de taille $n \times n$, à coefficients réels. On demande d'étudier la diagonalisabilité de la matrice triangulaire supérieure par blocs de taille $2n \times 2n$:

$$B = \begin{pmatrix} A & A \\ 0 & A \end{pmatrix}$$

Espaces vectoriels

Sous-espaces vectoriels

45 On se place dans l'espace \mathbb{R}^3 . Montrer que l'ensemble \mathcal{D} des vecteurs de composantes (x, y, z) tels que :

$$\begin{cases} 2x - y + z = 0 \\ x + y - z = 0 \end{cases}$$

est un sous-espace vectoriel de \mathbb{R}^3 .

46 On se place dans l'espace vectoriel $\mathbb{R}[X]$ des polynômes à coefficients réels. Montrer que l'ensemble $\mathbb{R}_3[X]$ des polynômes de degré inférieur ou égal à 3, à coefficients réels, est un sous-espace vectoriel de $\mathbb{R}[X]$.

47 On se place dans l'espace vectoriel $C([0, 1], \mathbb{R})$ des fonctions continues sur l'intervalle $[0, 1]$, à valeurs dans \mathbb{R} . Montrer que l'ensemble des fonctions continues sur l'intervalle $[0, 1]$, à valeurs dans \mathbb{R} , s'annulant en 0, est un sous-espace vectoriel de $C([0, 1], \mathbb{R})$.

48 On se place dans l'espace vectoriel $C([0, 1], \mathbb{R})$ des fonctions continues sur l'intervalle $[0, 1]$, à valeurs dans \mathbb{R} . Montrer que l'ensemble $C^1([0, 1], \mathbb{R})$ des fonctions de classe C^1 sur l'intervalle $[0, 1]$, à valeurs dans \mathbb{R} , est un sous-espace vectoriel de $C([0, 1], \mathbb{R})$.

49 On se place dans l'espace vectoriel $C([-1, 1], \mathbb{R})$ des fonctions continues sur l'intervalle $[-1, 1]$, à valeurs dans \mathbb{R} . Montrer que les ensembles :

$$\begin{aligned} C^+ &= \{f \in C([-1, 1], \mathbb{R}), f \text{ paire}\}, \\ C^- &= \{f \in C([-1, 1], \mathbb{R}), f \text{ impaire}\} \end{aligned}$$

sont deux sous-espaces supplémentaires de $C([-1, 1], \mathbb{R})$.

50 Soit n un entier naturel supérieur ou égal à 2. On se place dans l'espace vectoriel $\mathcal{M}_n(\mathbb{R})$ des matrices de taille $n \times n$, à coefficients réels. Montrer que l'ensemble des matrices de taille $n \times n$, à coefficients réels, de trace nulle, est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$.

51 Soit E un \mathbb{R} -espace vectoriel, et f une application linéaire de E telle que :

$$f^2 - 5f + 6\text{id}_E = 0$$

Montrer que $\text{Ker}(f - 2\text{id}_E)$ et $\text{Ker}(f - 3\text{id}_E)$ sont des sous-espaces supplémentaires de E .

Projecteurs

52 On se place dans l'espace vectoriel $\mathbb{R}_4[X]$ des polynômes à coefficients réels, de degré inférieur ou égal à 4. On considère l'application π qui, à tout polynôme $P =$

$$\sum_{k=0}^4 \alpha_k X^k \text{ associe :}$$

$$\sum_{k=0}^2 \alpha_k X^k$$

Montrer que π est un projecteur de $\mathbb{R}_4[X]$, puis déterminer son image et son noyau.

53 Soit E un \mathbb{R} -espace vectoriel, et p et q deux projecteurs de E tels que :

$$\text{Im } p \subset \text{Ker } q$$

On pose : $r = p + q - p \circ q$.

Montrer que r est un projecteur de E , puis déterminer son image et son noyau.

54 Soient E et F deux \mathbb{R} -espaces vectoriels, f une application linéaire de E dans F , et g une application linéaire de F dans E .

On suppose que :

$$f \circ g \circ f = f, \quad g \circ f \circ g = g$$

54.a) Vérifier que $f \circ g$ et $g \circ f$ sont deux projecteurs, en précisant leurs espaces respectifs de départ.

54.b) Montrer que :

$$E = \text{Ker}(f) \oplus \text{Im}(g)$$

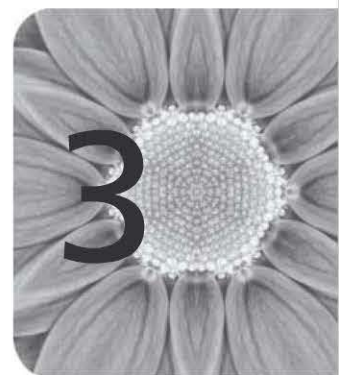
et :

$$F = \text{Ker}(g) \oplus \text{Im}(f)$$

Probabilités

Partie

3



Introduction


Compter ... Dénombrer, inventorier, répertorier : avant de pouvoir déterminer une probabilité, il est nécessaire de faire un « inventaire » des situations possibles. C'est le domaine de l'analyse combinatoire, où l'on s'intéresse aux configurations possibles, soit d'une collection d'objets, soit d'un ensemble de situations.

L'analyse combinatoire permet de mettre en place les axiomes des espaces probabilisés, et de construire les variables aléatoires discrètes, qui ne prennent qu'un nombre fini (ou dénombrable) de valeurs ; on arrive ainsi rapidement à la conclusion selon laquelle certains phénomènes aléatoires nécessitent l'utilisation de variables (aléatoires) continues, qui peuvent prendre toutes les valeurs réelles possibles dans un intervalle fini ou infini.

Plan

Analyse Combinatoire	368
Espaces probabilisés	376
Focus : Les jeux de hasard	380
Focus : Lancer de dés	382
Probabilités discrètes	389
Focus : Les tortues des Galápagos	406
Focus : Application de l'inégalité de Markov	427
Focus : Variations de température	429
Focus : Inégalité de Tchebychev et concentration de mesure	439
Probabilités continues	449
Focus : Le coût d'un déminage avec la loi uniforme	460
Focus : Dans l'attente d'un $n^{\text{ième}}$ client	487
Exercices d'entraînement	489

Les bonus web sur Dunod.com

 Les corrigés des exercices sont consultables sur dunod.com sur la page de présentation de l'ouvrage.

Définition

$$n! = 1 \times 2 \times 3 \times \dots \times n = \prod_{k=1}^n k$$
$$0! = 1$$
$$3! = 1 \times 2 \times 3 = 6 \quad , \quad 4! = 1 \times 2 \times 3 \times 4 = 24 \quad , \quad 5! = 1 \times 2 \times 3 \times 4 \times 5 = 120$$

$n!$ représente le nombre de façons de permuter n objets.

Quel est le nombre de façons de placer 125 étudiants dans un amphithéâtre de 125 places ? Il suffit de déterminer, de manière récursive, l'emplacement de chaque étudiant pour un ordre arbitraire fixe d'étudiants. Pour le premier étudiant, il y a 125 choix possibles. Il n'y en a plus que 124 pour le second. Et 123 pour le troisième. Ainsi de suite ... Ce qui donne, au final :

125 ! possibilités = $1 \times 2 \times 3 \times \dots \times 125$ possibilités

i.e.

1882677176888926099743767702491600857595403648714924258875982315083531
5633161359886688293288949592313364640544593005774063016191934138059781
88834575585470555243263755650071317708800000000000000000000000000000

On peut également représenter ce raisonnement de manière graphique, en représentant le nombre total de possibilités comme le nombre de feuilles d'un arbre avec 125 branches issues des racines, 124 branches issues de chacune des premières 125 branches, 123 branches issues des deuxièmes, etc. Étant donné qu'un tel dessin est impossible à représenter en réalité, vu le nombre astronomique de branches, on comprend aisément l'intérêt du raisonnement itératif non graphique, et de la définition formelle de la factorielle !

2. Formule de Stirling¹

$$\lim_{n \rightarrow +\infty} \frac{n! e^n}{n^n \sqrt{2\pi n}} = 1$$

que l'on peut aussi écrire, sous forme d'équivalent lorsque l'entier n tend vers l'infini :

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$

La formule de Stirling permet d'obtenir une estimation asymptotique de $n!$, lorsque l'entier n tend vers l'infini :

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3} + \frac{571}{2488320n^5} + \dots\right)$$

Toutefois nous n'aurons besoin, dans ce qui suit, que du premier terme de ce développement.

Démonstration : On ne donnera pas, ici, de démonstration complète. La formule de Stirling peut être obtenue à l'aide des intégrales de Wallis² :

$$\int_0^{\frac{\pi}{2}} \sin^n t \, dt = \int_0^{\frac{\pi}{2}} \cos^n t \, dt, \quad n \in \mathbb{N}$$

On obtient, pour tout entier naturel p :

$$\int_0^{\frac{\pi}{2}} \sin^{2p} t \, dt = \pi \frac{(2p)!}{2^{2p+1} (p!)^2}, \quad \int_0^{\frac{\pi}{2}} \sin^{2p+1} t \, dt = \frac{4^p (p!)^2}{(2p+1)!}$$

Le lecteur intéressé pourra trouver une très jolie démonstration dans [27]. ■

1. James Stirling (1692-1770), mathématicien écossais. Il apporta de nombreuses contributions à l'étude des séries numériques. En ce qui concerne la « formule de Stirling », Abraham de Moivre avait déjà montré que, lorsque l'entier n tend vers l'infini, $n! \sim C n^{n+\frac{1}{2}} e^{-n}$, où C est une constante positive. La contribution de James Stirling fut d'identifier la constante ($\sqrt{2\pi}$).

2. John Wallis (1616-1703), mathématicien anglais, spécialiste de calcul différentiel et intégral. C'est lui qui introduisit la notation « ∞ ». À côté de son œuvre mathématique, il s'intéressa aussi à la phonétique, et est considéré comme un des précurseurs de l'orthophonie.

1. Arrangements

Définition

Soit n un entier naturel non nul. Pour tout entier naturel $p \leq n$, le nombre de façons de choisir, de façon ordonnée, p éléments distincts parmi n est appelé **arrangement de p éléments parmi n** .

Il est aisé de calculer ce nombre en utilisant un raisonnement itératif, équivalant à l'usage d'un arbre. Pour le premier choix, on a n possibilités, pour le deuxième, il en reste $n - 1$, et ainsi de suite jusqu'au $p^{\text{ième}}$, pour lequel il y a $n - (p - 1)$ possibilités. Ce raisonnement permet de démontrer les résultats suivants, où on remarque également que $A_n^n = n!$.

Proposition

Pour tout couple d'entiers naturels (n, p) tels que $p \leq n \neq 0$,

$$A_n^p = n(n-1) \dots (n-p+1) = \frac{n!}{(n-p)!}.$$

2. Permutations

n étant un entier naturel non nul, le nombre de façons de permuter n objets est le nombre d'arrangements de tous ces entiers, i.e. le nombre de possibilités de choisir, de façon ordonnée, n éléments parmi n , et vaut donc $n!$. Par convention :

$$0! = 1$$

Exemple : Séquençage de l'ADN

Une séquence d'acide désoxyribonucléique (ADN) est constituée d'un enchaînement de quatre nucléotides : l'adénine, la cytosine, la guanine, et la thymine. L'information génétique correspond à l'ordre dans lequel s'enchaînent les quatre nucléotides, qui se regroupent par paires :

- l'adénine avec la thymine ;
- la thymine avec l'adénine ;
- la cytosine avec la guanine ;
- la guanine avec la cytosine.

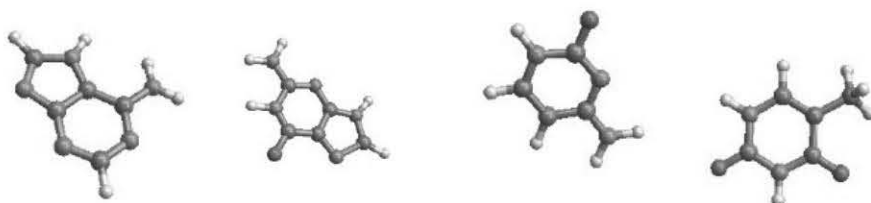


Figure 111.1 – La molécule d'adénine $C_5H_5N_5$, la molécule de guanine $C_5H_5N_5O$, la molécule de cytosine $C_4H_5N_3O$, et la molécule de thymine $C_5H_7N_2O_2$.

Il y a donc, au total, quatre combinaisons possibles par nucléotide.

Le nombre d'arrangements possibles de deux nucléotides distincts correspond au nombre de façons de choisir, de façon ordonnée, deux éléments parmi quatre :

$$A_4^2 = \frac{4!}{2!} = 3 \times 4 = 12$$

3. Arrangements avec répétitions

Soit n un entier naturel non nul. Pour tout entier naturel $p \leq n$, le nombre de façons de choisir p éléments, non nécessairement distincts, parmi n , vaut p^n . Notons que ce nombre est égal au nombre d'applications d'un ensemble à n éléments dans un ensemble à p éléments.

Démonstration : Pour le premier élément, il y a n choix possibles ; pour le second, il y a encore n choix possibles ; ... et, ce jusqu'au $p^{\text{ième}}$. ■

1. Combinaison ou coefficient binomial

Définition

Soit n un entier naturel non nul. Pour tout entier naturel $p \leq n$, le nombre de façons de choisir, dans un ordre quelconque, p éléments distincts parmi n , est appelé **combinaison de p éléments parmi n** , que l'on note $\binom{n}{p}$ ou C_n^p . Cette quantité s'appelle aussi **coefficient binomial** de paramètres n et p .

On peut utiliser la formule des arrangements pour calculer C_n^p . En effet, si l'ordre n'importe pas dans le nombre de choix de p éléments parmi n , on peut d'abord dénombrer le nombre de choix quand l'ordre importe, c'est-à-dire $A_n^p = \frac{n!}{(n-p)!}$, puis remarquer que chaque partie non ordonnée de p éléments sera dénombrée $p!$ fois dans les arrangements, car le nombre de façons d'ordonner les p éléments est le nombre de permutations de p . Ceci démontre le résultat suivant :

Proposition

Pour tout couple d'entiers naturels (n, p) tels que $p \leq n \neq 0$,

$$C_n^p = \binom{n}{p} = \binom{n}{n-p} = \frac{A_n^p}{p!} = \frac{n!}{p!(n-p)!}.$$



Le coefficient binomial est noté, indifféremment, C_n^p ou $\binom{n}{p}$. La notation $\binom{n}{p}$ est celle utilisée dans les pays anglo-saxons. En France, la norme ISO 31 préconise, sans la rendre obligatoire, cette dernière notation. L'un des auteurs de cet ouvrage préfère, par choix, garder la notation C_n^p , plus claire, et qui évite aussi beaucoup de confusions dès que l'on travaille avec des coefficients binomiaux et des matrices colonnes.

Exemple : Courses de chevaux...

Au tiercé, si l'ordre n'importe pas, le nombre de combinaisons possibles de trois chevaux parmi dix est :

$$C_{10}^3 = \binom{10}{3} = \frac{10!}{3!7!} = \frac{8 \times 9 \times 10}{3!} = \frac{8 \times 9 \times 10}{2 \times 3} = 4 \times 3 \times 10 = 120$$

2. Formule du binôme de Newton

Soit n un entier naturel non nul. Alors, pour tout couple de réels (a, b) :

$$(a + b)^n = \sum_{k=0}^n C_n^k a^k b^{n-k} = \sum_{k=0}^n \frac{n!}{k!(n-k)!} a^k b^{n-k}$$

Démonstration : Cette propriété est bien connue. Pour la démontrer, il suffit d'écrire :

$$(a + b)^n = \underbrace{(a + b)(a + b) \cdots (a + b)}_{n \text{ fois}},$$

puis de développer ce produit, en remarquant que le nombre de fois où l'on obtient le terme $a^k b^{n-k}$ est exactement égal au nombre de façons de choisir une partie non ordonnée de k éléments parmi les n termes du produit. ■

Corollaire

Soit n un entier naturel non nul. Alors :

$$2^n = \sum_{k=0}^n C_n^k$$

Démonstration : Il suffit d'appliquer la forme du binôme de Newton à $(1 + 1)^n$. ■

3. Le triangle de Pascal

Soit n un entier naturel non nul. Alors, pour tout entier naturel $k \leq n - 1$:

$$C_n^k + C_n^{k+1} = C_{n+1}^{k+1}$$

soit :

$$\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$$

Cette propriété peut être représentée à l'aide du triangle suivant :

$$\begin{array}{c} \{1\} \\ \{1, 1\} \\ \{1, 2, 1\} \\ \{1, 3, 3, 1\} \\ \{1, 4, 6, 4, 1\} \\ \{1, 5, 10, 10, 5, 1\} \\ \{1, 6, 15, 20, 15, 6, 1\} \\ \{1, 7, 21, 35, 35, 21, 7, 1\} \end{array}$$

Figure 112.1 – Le triangle de Pascal.

La première ligne du triangle donne le coefficient binomial $C_1^0 = \binom{1}{0} = 1$.

La seconde ligne du triangle donne les coefficients binomiaux $C_2^0 = \binom{2}{0}$ et $C_2^1 = \binom{2}{1}$.

La troisième ligne du triangle donne les coefficients binomiaux $C_2^0 = \binom{3}{0}$, $C_3^1 = \binom{3}{1}$ et $C_3^2 = \binom{3}{2}$, et ainsi de suite : la $n^{\text{ième}}$ ligne du triangle donne les coefficients binomiaux $C_n^p = \binom{n}{p}$ pour $p = 0, p = 1, \dots, p = n$.

Pour tout entier i de $\{2, \dots, n\}$, et tout j de $\{2, \dots, n\}$, le coefficient binomial situé ligne i , colonne j , s'obtient en ajoutant les coefficients binomiaux situés ligne $i - 1$, colonne $j - 1$, et ligne $i - 1$, colonne j .

Démonstration : Il suffit de calculer :

$$\begin{aligned}
 C_n^k + C_n^{k+1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!(n-k-1)!} \\
 &= \frac{n!(k+1)}{k!(k+1)(n-k)!} + \frac{n!(n-k)}{(k+1)!(n-k-1)!(n-k)} \\
 &= \frac{n!(k+1)}{(k+1)!(n-k)!} + \frac{n!(n-k)}{(k+1)!(n-k)!} \\
 &= \frac{n!(k+1+n-k)}{(k+1)!(n-k)!} \\
 &= \frac{n!(n+1)}{(k+1)!(n-k)!} \\
 &= \frac{(n+1)!}{(k+1)!(n+1-k-1)!} \\
 &= C_{n+1}^{k+1}
 \end{aligned}$$

■

4. Formule de Vandermonde¹

Soient m et n deux entiers naturels. Alors, pour tout entier naturel $k \leq m + n$:

$$\sum_{\substack{0 \leq i \leq m, 0 \leq j \leq n \\ i+j=k}} C_m^i C_n^j = C_{m+n}^k$$

Démonstration : La formule de Vandermonde se démontre à l'aide des identités polynomiales suivantes :

$\forall x \in \mathbb{R} :$

$$(1+x)^{m+n} = \sum_{\ell=0}^{m+n} C_{m+n}^{\ell} x^{\ell} 1^{m+n-\ell} = \sum_{\ell=0}^{m+n} C_{m+n}^{\ell} x^{\ell}$$

1. (1735-1796), mathématicien français, mais aussi économiste, musicien et chimiste.

Il est surtout connu pour le déterminant qui porte son nom. Mais d'après Lebesgue (Conférence d'Utrecht, 1937, [23]), le déterminant ne serait pas de lui...

et :

$$\begin{aligned}
 (1+x)^{m+n} &= (1+x)^m (1+x)^n = \left(\sum_{i=0}^m C_m^i x^i 1^{m-i} \right) \left(\sum_{j=0}^n C_n^j x^j 1^{n-j} \right) \\
 &= \left(\sum_{i=0}^m C_m^i x^i \right) \left(\sum_{j=0}^n C_n^j x^j \right) \\
 &= \sum_{i=0}^m \sum_{j=0}^n C_m^i C_n^j x^{i+j} \\
 &= \sum_{k=0}^{m+n} \sum_{\substack{0 \leq i \leq m, 0 \leq j \leq n \\ i+j=k}} C_m^i C_n^j x^k
 \end{aligned}$$

(On a commencé par effectuer un changement d'indices, en posant $i + j = k$; k varie donc de 0 à $m + n$.)

Il suffit ensuite d'identifier, pour tout entier i de $\{0, \dots, m+n\}$, les coefficients des puissances de x , ce qui conduit, pour tout entier k de $\{0, 1, \dots, m+n\}$ à :

$$C_{m+n}^k = \sum_{\substack{0 \leq i \leq m, 0 \leq j \leq n \\ i+j=k}} C_m^i C_n^j \quad \blacksquare$$

Dans ce qui suit, nous présentons les axiomes du calcul des probabilités, communs aux situations discrètes ou non.

La modélisation mathématique probabiliste utilise le langage des ensembles. Les correspondances avec les concepts probabilistes sont les suivantes :

Événement impossible	« Ensemble vide » \emptyset
Événement certain	« Ensemble plein » Ω
Épreuve, ou réalisation, ou issue, ou issue élémentaire	« élément ω », $\omega \in \Omega$
Événement	« Sous-ensemble de Ω », $A \subset \Omega$
L'épreuve ω est une réalisation possible de l'événement A	$\omega \in A$
L'événement A implique l'événement B	$A \subset B$
L'événement A ou l'événement B ont lieu	$A \cup B$
L'événement A et l'événement B ont lieu	$A \cap B$
L'événement A et l'événement B sont incompatibles	$A \cap B = \emptyset$
L'événement contraire de A	A^c , ou \bar{A}

De même qu'il est préférable de ne pas confondre l'élément ω avec le singleton $\{\omega\}$, il faut distinguer l'issue $\omega \in \Omega$ avec l'événement (élémentaire) $A = \{\omega\} \subset \Omega$.

1. Expérience aléatoire

Définition

On appelle **expérience aléatoire** une expérience renouvelable et qui, renouvelée dans des conditions identiques, ne donne pas nécessairement le même résultat à chaque fois.

Exemple

Lancer plusieurs fois une même pièce parfaitement équilibrée est une expérience aléatoire.

2. Événement, Univers

Définition

Étant donnée une expérience aléatoire, on appelle **épreuves élémentaires** les différents résultats possibles de cette expérience. Leur ensemble constitue l'**univers** Ω .

Exemple

Une enseignante de mathématiques doit concevoir un sujet d'examen. Elle veut que ses étudiants réussissent celui-ci le mieux possible. Elle est donc intéressée par les événements suivants :

- \mathcal{E}_1 : la note minimale à l'examen est supérieure ou égale à 90 sur 100 ;
- \mathcal{E}_2 : la note minimale à l'examen est supérieure ou égale à 95 sur 100 ;
- \mathcal{E}_3 : la note minimale à l'examen est supérieure ou égale à 99 sur 100.

L'enseignante considère alors la famille d'événements « la note minimale à l'examen est de N sur 100, où N varie continûment de 0 à 100. L'enseignante est ainsi amenée à considérer l'espace de toutes les épreuves possibles, Ω , et à l'identifier avec l'intervalle $[0, 100]$. On peut alors s'intéresser à l'épreuve ω : la note minimale à l'examen est 96,5 sur 100. Là, les événements \mathcal{E}_1 et \mathcal{E}_2 sont réalisés, alors que l'événement \mathcal{E}_3 ne l'est pas. Dans l'espace des possibilités Ω , l'épreuve ω est identifiée avec l'élément 96,5. L'événement \mathcal{E}_1 est représenté par l'intervalle $[90, 100]$, l'événement \mathcal{E}_2 est représenté par l'intervalle $[95, 100]$, et l'événement \mathcal{E}_3 par l'intervalle $[99, 100]$. L'appartenance « $\omega \in \mathcal{E}_1$ » est équivalente à « l'événement \mathcal{E}_1 est réalisé par l'épreuve ω ». Du fait que $\mathcal{E}_3 \subset \mathcal{E}_1$, l'appartenance « $\omega \in \mathcal{E}_3$ » implique « l'événement \mathcal{E}_1 est réalisé par l'épreuve ω », sans y être équivalent ; de même pour « $\omega \in \mathcal{E}_2$ ».

► Tribu (ou famille d'événements observables, ou σ -algèbre)

Étant donné un ensemble Ω , une **tribu** (aussi appelée famille d'événements observables, ou, encore, σ -algèbre) sur Ω est un sous-ensemble \mathcal{A} de l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω telle que :

- Ω appartient à \mathcal{A} ;
- Pour tout A de \mathcal{A} , son complémentaire, noté A^c , ou encore, \bar{A} , appartient lui aussi à \mathcal{A} .
- Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de \mathcal{A} , la réunion $\bigcup_{n \in \mathbb{N}} A_n$ appartient aussi à \mathcal{A} .

Les éléments de la tribu \mathcal{A} sont les **événements**.

3. Probabilité**Définition**

Soit Ω un ensemble (aussi appelé **univers**), et \mathcal{A} une tribu (ou famille d'événements observables) sur Ω . On appelle **probabilité sur** (Ω, \mathcal{A}) toute application P de \mathcal{A} dans $[0, 1]$ telle que :

- $P(\Omega) = 1$;
- pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements incompatibles (i.e. deux à deux disjoints) :

$$P\left(\bigcup_{n \in \mathbb{N}^*} A_n\right) = \sum_{n=1}^{+\infty} P(A_n)$$

(Ω, \mathcal{A}, P) est appelé **espace probabilisé**.



Définir une probabilité sur (Ω, \mathcal{A}) consiste, tout simplement, à attribuer un « poids » à chaque événement observable. L'événement certain a, bien évidemment, une probabilité de $1 = 100\%$. C'est le sens du premier point dans la définition ci-dessus. La probabilité de l'union d'un nombre fini ou dénombrable d'événements deux à deux incompatibles doit être égale à la somme des probabilités de ces événements (c'est le second point (sigma-additivité)).

Propriété

Soit Ω un ensemble (ou univers), et \mathcal{A} une tribu (ou famille d'événements observables) sur Ω .

Toute probabilité sur (Ω, \mathcal{A}) vérifie les propriétés suivantes :

1. **Probabilité de l'événement impossible :** $P(\emptyset) = 0$;
2. **Probabilité de l'événement complémentaire :**

$$P(A^c) = 1 - P(A).$$

3. **σ -additivité :**

Si $(A_n)_{n \in \mathbb{N}}$ est une suite d'événements disjoints, alors :

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n).$$

4. **Continuité décroissante :**

Si $(A_n)_{n \in \mathbb{N}}$ est une suite d'événements telle que, pour tout entier naturel n , $A_{n+1} \subset A_n$, alors :

$$\lim_{n \rightarrow +\infty} P(A_n) = P\left(\bigcap_{n=0}^{+\infty} A_n\right).$$

5. **Inégalité triangulaire :**

Si $(A_n)_{n \in \mathbb{N}}$ est une suite d'événements, alors :

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} P(A_n)$$

► Probabilité sur un univers fini

Une probabilité sur un univers fini Ω est une application P de $\mathcal{P}(\Omega)$ dans $[0; 1]$ telle que $P(\Omega) = 1$ et, pour toutes parties disjointes A et B de Ω :

$$P(A \cup B) = P(A) + P(B)$$

► Événement presque sûr

Soit (Ω, \mathcal{A}, P) un espace probabilisé. Un événement A est dit **presque sûr** si sa probabilité vaut 1 :

$$P(A) = P(\Omega) = 1$$

► Événement négligeable

Soit (Ω, \mathcal{A}, P) un espace probabilisé. Un événement A est dit **négligeable** si sa probabilité vaut 0 :

$$P(A) = 0$$



Nous rencontrerons, par la suite, des exemples d'événements négligeables qui ne sont pas pour autant impossibles (événement vide), même dans le contexte discret.

► Système complet d'événements

Soient A_1, \dots, A_n des événements d'un espace probabilisé (Ω, \mathcal{A}, P) . On dit que (A_1, \dots, A_n) est un **système complet d'événements** de (Ω, P) si les événements A_1, \dots, A_n sont deux à deux disjoints (i.e. $A_i \cap A_j = \emptyset$ si $i \neq j$) et si :

$$\bigcup_{i=1}^n A_i = \Omega$$

On dit aussi que A_1, \dots, A_n constituent une **partition** de Ω . On a, en particulier :

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) = 1$$

► Équiprobabilité de deux événements

Soit (Ω, \mathcal{A}, P) un espace probabilisé. Deux événements A_1 et A_2 sont dits **équiprobables** s'ils ont la même probabilité.

► Équiprobabilité de n événements, $n \in \mathbf{N}$

Soit (Ω, \mathcal{A}, P) un espace probabilisé, et n un entier naturel non nul. Des événements A_1, \dots, A_n sont dits **équiprobables** s'ils ont la même probabilité.

Propriété

Si les événements A_1, \dots, A_n d'un système complet (A_1, \dots, A_n) d'événements de (Ω, \mathcal{A}, P) sont équiprobables puisque fini. Il est très facile de calculer leur probabilité. Ils sont, par hypothèse, incompatibles entre eux. Il en résulte :

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

L'équiprobabilité signifie que :

$$P(A_1) = P(A_2) = \dots = P(A_n)$$

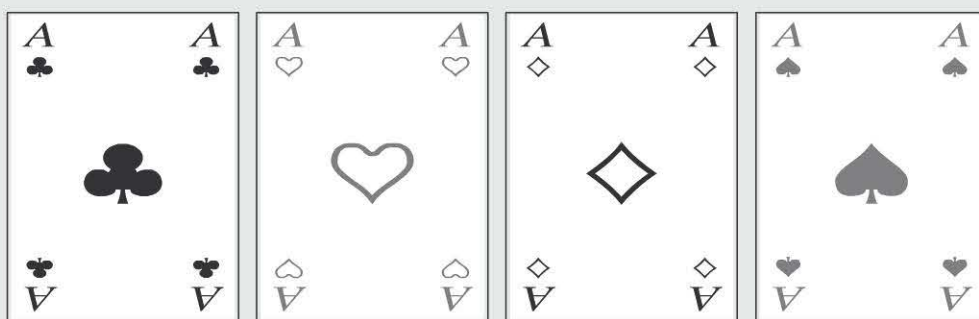
et :

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = P(\Omega) = 1$$

Par suite :

$$P(A_1) = P(A_2) = \dots = P(A_n) = \frac{1}{n}$$

Le hasard intervient, ou semble intervenir, dans de nombreux phénomènes, à commencer par les jeux dits de « de hasard »¹. De façon plus concrète, toute mesure physique est, elle-même, entachée de « bruit » (qui peut résulter, notamment, d'une préparation imparfaite du système à mesurer, d'une perturbation thermique, etc.) Aux échelles les plus infinitésimales, ce bruit est de nature quantique, puisque des particules comme les photons peuvent être concernées. Le « principe d'incertitude de Heisenberg »² [46] établit que l'on ne peut mesurer simultanément la position et l'impulsion d'une particule avec une précision infinie.



Un carré d'as...

Une modélisation dite « déterministe » où les événements sont régis par le principe de causalité³, doit alors laisser place à une modélisation « probabiliste », puisque, bien évidemment, lorsque l'on considère un phénomène aléatoire, on ne peut en prédire le résultat avec certitude. Il est toutefois intéressant d'établir des modèles mathématiques qui permettront une étude plus précise. À cet effet, on aura besoin d'un espace probabilisé, composé d'un ensemble (« l'univers » Ω), d'une famille de parties de cet univers (une « tribu » de sous-ensembles $A \subset \Omega$), et, bien évidemment, une probabilité, qui représente la fonction mathématique modélisant le phénomène probabiliste (nous renvoyons à [28] pour une étude plus générale des tribus).

1. Ces jeux de hasard furent, de fait à l'origine des premières théories probabilistes. C'est Blaise Pascal (1623-1662), qui fut un des premiers contributeurs. Il envoya à Pierre de Fermat, en 1654, une méthode de résolution du fameux « problème des partis », où deux joueurs jouent à un jeu de hasard en trois parties gagnantes ; chaque joueur mise la même somme d'argent. Mais le jeu est interrompu avant la fin. Comment faut-il partager les enjeux misés ?

2. Werner Karl Heisenberg (1901-1976), physicien allemand, à l'origine de la mécanique quantique, qui lui valut, en 1932, le prix Nobel de physique. Sa présentation de la mécanique quantique est extrêmement intéressante, car elle utilise, initialement, un formalisme matriciel.

3. Un fait – la cause – en engendre un autre.

Dans certaines situations, en particulier lorsqu'il s'agit de modéliser des nombres aléatoires diffus (ce que nous appellerons les « variables aléatoires continues », utiles pour comprendre les phénomènes naturels et les études sociales à grande échelle), il ne sera pas possible d'associer une probabilité à chaque issue élémentaire (i.e. à chaque $\omega \in \Omega$). En particulier, nous ne pourrons pas nécessairement considérer qu'un singleton $\{\omega\}$ est un événement A . Toutefois, il devra toujours être possible de définir la fonction de probabilité \mathbf{P} sur l'ensemble de tous les événements A qu'on a le droit de considérer (i.e. la tribu qu'on choisit d'utiliser). Nous n'insisterons pas sur ce point technique dans cet ouvrage, car il relève de la théorie mathématique dite « de la mesure » ; il sera néanmoins évident, en fonction du contexte dans lequel on se place, de déterminer les cas où les singletons sont, ou non, des événements. Dans le premier cas, on parle de situation discrète.

Considérons le lancer de deux dés. Le choix le plus simple pour désigner l'univers est $\Omega = \{(i, j); i, j = 1, 2, \dots, 6\}$. Cet univers est composé de 36 épreuves élémentaires. Supposons que les dés ne sont pas pipés, et que le lanceur ne contrôle pas les résultats, ce qui signifie que tous les résultats sont équiprobables. Chacun des 36 éléments $\omega = (i, j)$ de Ω peut être considéré comme un événement $\{\omega\}$, ces 36 événements forment un système complet et sont équiprobables, de sorte que $P(\{\omega\}) = \frac{1}{36}$. Calculons la probabilité de l'événement A : « la somme des deux chiffres qui sortent est supérieure ou égale à dix, mais strictement inférieure à douze ». Nous avons donc :

$$A = \{(5, 5), (5, 6), (6, 5), (6, 4), (4, 6)\}.$$

La probabilité cherchée vaut donc

$$P(A) = \sum_{\omega \in A} P(\{\omega\}) = \frac{5}{36}.$$

Mais, si l'on demande également que le résultat d'un des dés soit six, ceci signifie que le résultat $(5, 5)$ n'est pas autorisé ; il faut donc chercher la probabilité de l'événement :

$$C = \{(5, 6), (6, 5), (6, 4), (4, 6)\}.$$

Cette probabilité vaut :

$$P(C) = \frac{4}{36} = \frac{1}{9}.$$

Mais, d'un autre côté, si on considère que l'on a restreint l'ensemble des possibilités en demandant qu'un des dés ait le numéro 6, cela signifie que l'on est en train de changer l'univers, qui est donc :

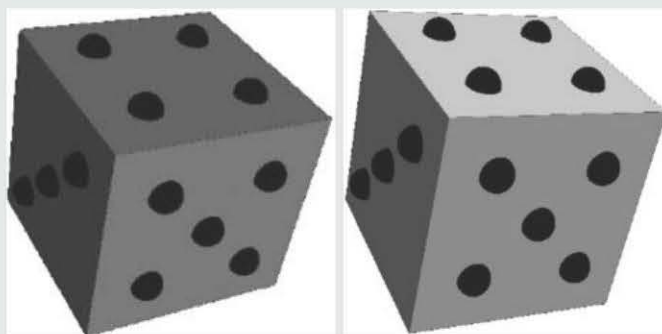
$$S = \{(6, 1); (1, 6); (6, 2); (2, 6); (6, 3); (3, 6); (6, 4); (4, 6); (5, 6); (6, 5); (6, 6)\}$$

Il n'y a que onze possibilités. La probabilité de l'événement A sachant que l'on doit obtenir au moins un 6 devrait donc devenir la probabilité de l'événement :

$$\{(6, 4); (4, 6); (5, 6); (6, 5)\}$$

lorsque l'univers est S , i.e. $\frac{4}{11}$. Nous allons voir comment raisonner ainsi de manière systématique et aisée.

La réponse à la question $P(C) = ?$ n'est pas identique à la réponse à celle qui demande « quelle est la probabilité que la somme des deux chiffres qui sortent soit supérieure ou égale à 10 et strictement inférieure à 12, si on a observé que l'un des deux des vaut 6 ? »



Un lancer de dés...

Imaginons qu'un ami vous propose un jeu où il faut s'acquitter de la somme d'un euro pour avoir le droit de lancer deux dés, un rouge et un bleu, et où l'on gagne deux euros si la somme des deux numéros obtenus est paire, i.e. si l'événement A s'est produit. Ce jeu semble équitable, nous verrons, ultérieurement, qu'il subsiste des choses intéressantes à ajouter sur le sujet. Mais supposons maintenant que, si l'on accepte de s'acquitter de cinq centimes de plus, on acquière le droit de savoir si le numéro sorti sur le dé bleu est, ou non, supérieur ou égal à deux, et, partant, d'abandonner et de récupérer sa mise. Comme l'événement B est indépendant de celui qui détermine le gain, il est alors possible de refuser de jouer ! En effet, l'information concernant l'événement B n'a pas d'influence sur l'événement A , et n'a donc aucune valeur. À une échelle beaucoup plus impressionnante, certaines compagnies prétendent vendre de l'information utile pour comprendre des opportunités d'investissement dans les marchés financiers. Les meilleures entreprises, comme Bloomberg, ont une réputation de donner des informations fortement liées aux mouvements des actions dans ces marchés, alors qu'il existe une multitude de petites compagnies qui parviennent à vendre des informations qui sont, en réalité, indépendantes de ces mouvements, à des investisseurs naïfs et avides de connaissances que n'auraient pas leurs compétiteurs.

La notion de conditionnement est un outil fondamental de la théorie des probabilités, dans la mesure où toute information supplémentaire concernant la réalisation d'un événement spécifique A modifie la vraisemblance d'un événement B .

1. Probabilité conditionnelle

Définition

Soient A et B deux événements d'un espace probabilisé (Ω, P) , avec $P(B) \neq 0$. On appelle probabilité conditionnelle de l'événement « A sachant B » la quantité, notée $P(A|B)$ telle que :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

2. Formule des probabilités composées

Soient A et B deux événements d'un espace probabilisé (Ω, P) . Alors :

$$P(A \cap B) = P(A|B) P(B) = P(B|A) P(A)$$

Démonstration : Si la probabilité de l'événement B n'est pas nulle :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

et donc :

$$P(A \cap B) = P(A|B) \cdot P(B)$$

Il est clair que si la probabilité de l'événement B est nulle : $P(A \cap B) = 0$.

Ainsi, dans tous les cas : $P(A \cap B) = P(A|B) P(B)$.

A et B jouant des rôles symétriques, on démontre de même l'autre égalité. ■

3. Formule des probabilités composées généralisée

Soit (A_1, \dots, A_n) un système d'événements d'un espace probabilisé (Ω, P) tel que :

$$P\left(\bigcap_{i=1}^n A_i\right) \neq 0$$

Alors :

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1) P(A_2|A_1) P(A_3|A_1 \cap A_2) \dots P(A_n|A_1 \cap \dots \cap A_{n-1})$$

Avec le même argument que ci-dessus, on peut s'affranchir de l'hypothèse $P\left(\bigcap_{i=1}^n A_i\right) \neq 0$ (voir page 106).

Démonstration : La démonstration se fait par récurrence sur l'entier n . ■

Exemple

Considérons un jeu de 52 cartes à jouer, qui contient quatre dames distinctes, et donc 48 cartes qui ne sont pas des dames. Calculons la probabilité qu'une main de 5 cartes ne contienne aucune dame, puis celle qu'elle contienne exactement une dame. À cet effet, on procède par itération, en désignant par A_i l'événement « la carte numéro i n'est pas une dame ». Dans le premier cas, on cherche :

$$\begin{aligned} P\left(\bigcap_{i=1}^5 A_i\right) &= P(A_1) P(A_2|A_1) P(A_3|A_1 \cap A_2) P(A_4|A_1 \cap A_2 \cap A_3) P(A_5|A_1 \cap \dots \cap A_4) \\ &= \frac{48}{52} \frac{47}{51} \frac{46}{50} \frac{45}{49} \frac{44}{48} \\ &= \frac{35\,673}{54\,145} \\ &\approx 0,66 \end{aligned}$$

Dans le second cas, il s'agit d'abord de déterminer laquelle des cinq cartes est une dame. Il y a donc cinq événements disjoints B_i correspondant, où B_i désigne l'événement « la carte numéro i est une dame, et les quatre autres cartes ne le sont pas ». Calculons, dans un premier temps :

$$\begin{aligned} P(B_1) &= P\left(A_1^c \cap \bigcap_{i=2}^5 A_i\right) \\ &= P(A_1^c) P(A_2|A_1^c) P(A_3|A_1^c \cap A_2) P(A_4|A_1^c \cap A_2 \cap A_3) P(A_5|A_1^c \cap A_2 \cap A_3 \cap A_4) \\ &= \frac{4}{52} \frac{48}{51} \frac{47}{50} \frac{46}{49} \frac{45}{48} = \frac{3243}{54\,145} \\ &\approx 0,060 \end{aligned}$$

En utilisant le même raisonnement, on trouve aussi

$$P(B_2) = \frac{48}{52} \frac{4}{51} \frac{47}{50} \frac{46}{49} \frac{45}{48} = P(B_1) = P(B_3) = P(B_4) = P(B_5)$$

La probabilité cherchée est donc

$$P\left(\bigcup_{i=1}^5 B_i\right) = 5 P(B_1) = \frac{5 \times 3243}{54\,145} = \frac{3243}{10\,829} \approx 0,30.$$

4. Formule des probabilités totales

Soit (A_1, \dots, A_n) un système complet d'événements d'un espace probabilisé (Ω, P) (i.e. les A_i , $i = 1, \dots, n$ sont deux à deux disjoints, et tels que $\bigcup_{i=1}^n A_i = \Omega$). Alors, pour tout événement B de (Ω, P) :

$$P(B) = P(B|A_1) P(A_1) + \dots + P(B|A_n) P(A_n) = \sum_{i=1}^n P(B|A_i) P(A_i)$$

Démonstration : Les ensembles $B \cap A_1, \dots, B \cap A_n$ constituent une partition de B , i.e. une réunion disjointe d'événements tels que :

$$B = \bigcup_{i=1}^n B \cap A_i$$

On en déduit alors, par additivité :

$$P(B) = \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P(B|A_i) P(A_i) \quad \blacksquare$$

Exemple

Une population de lapins comporte un tiers de lapins mâles et deux tiers de lapins femelles. 6 % des lapins mâles sont blancs, contre 4 % des lapins femelles. On cherche à déterminer la probabilité qu'un lapin choisi au hasard, et dont on ne connaît donc pas le sexe, soit blanc.

On dispose donc de deux systèmes complets d'événements :

- $(\mathcal{L}_{\text{mâle}}, \mathcal{L}_{\text{femelle}})$: « le lapin choisi est un mâle, le lapin choisi est une femelle » ;
- $(\mathcal{L}_{\text{blanc}}, \mathcal{L}_{\text{d'une autre couleur}})$: « le lapin choisi est blanc, le lapin choisi n'est pas blanc ».

On a alors :

$$\begin{aligned} P(\mathcal{L}_{\text{blanc}}) &= P(\mathcal{L}_{\text{blanc}}|\mathcal{L}_{\text{mâle}}) P(\mathcal{L}_{\text{mâle}}) + P(\mathcal{L}_{\text{blanc}}|\mathcal{L}_{\text{femelle}}) P(\mathcal{L}_{\text{femelle}}) \\ &= 0,06 \times \frac{1}{3} + 0,04 \times \frac{2}{3} \\ &= \frac{7}{150} \\ &\approx 0,047 \end{aligned}$$



L'utilisation de la formule des probabilités totales dans cet exemple est typique de questions que l'on rencontre en théorie de la décision.

De façon plus précise, cet exemple peut être considéré comme « précurseur » d'une méthodologie dite « de Bayes », pour évaluer la probabilité d'un événement non observé, sachant qu'un autre événement a été observé, connaissant les probabilités dites *a priori* des événements non conditionnés, à partir d'un *modèle* pour les probabilités conditionnelles (dans le sens opposé à ce que l'on cherche).

Corollaire (formule de Bayes)

Soient A et B deux événements d'un espace probabilisé (Ω, P) . Si $P(A) > 0$ et $P(B) > 0$, alors :

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)} = \frac{P(B|A) P(A)}{P(B|A) P(A) + P(B|A^c) P(A^c)}.$$

Démonstration : D'après la formule des probabilités conditionnelles :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Grâce à la formule des probabilités composées :

$$P(A \cap B) = P(A|B) P(B) = P(B|A) P(A) \quad (114.1)$$

Par suite :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A) P(A)}{P(B)}$$

La formule des probabilités totales permet de transformer le dénominateur de la première formule de Bayes, pour obtenir la seconde. ■

Corollaire (formule de Bayes généralisée)

Soit (A_1, \dots, A_n) un système complet d'événements d'un espace probabilisé (Ω, P) (i.e. les A_i , $i = 1, \dots, n$ sont deux à deux disjoints, et tels que $\bigcup_{i=1}^n A_i = \Omega$), de probabilités respectives non nulles. Alors, pour tout événement B de (Ω, P) tel que $P(B) \neq 0$:

$$P(A_i|B) = \frac{P(B|A_i) P(A_i)}{\sum_{k=1}^n P(B|A_k) P(A_k)}$$

Démonstration : Comme dans le corollaire précédent, il suffit d'appliquer la formule des probabilités totales :

$$\begin{aligned} P(A_i|B) &= \frac{P(A_i \cap B)}{P(B)} \\ &= \frac{P(B|A_i) P(A_i)}{P(B)} \\ &= \frac{P(B|A_i) P(A_i)}{\sum_{k=1}^n P(B|A_k) P(A_k)} \end{aligned}$$

Exemple

En Europe, environ 65 personnes sur mille possèdent une mutation du facteur de Leiden (Facteur V). Cette mutation affecte la coagulation du sang, avec des risques importants de thrombose et d'embolie. Un premier test de résistance plasmatique à la protéine C activée permet de détecter les personnes susceptibles de présenter cette mutation dans 90 % des cas. Par contre, il y a des faux positifs dans treize cas sur cent.

Pour déterminer la probabilité qu'une personne présentant un résultat positif à ce test soit porteuse de cette mutation, on commence par définir les événements

- M_V : la personne est porteuse de la mutation ;
- \mathcal{P}^+ : le test est positif.

On a alors :

$$P(M_V) = \frac{65}{1000}, \quad P(\mathcal{P}^+|M_V) = \frac{90}{100}, \quad P(\mathcal{P}^+|M_V^c) = \frac{13}{100}$$

On obtient alors, grâce à la formule de Bayes :

$$\begin{aligned} P(M_V|\mathcal{P}^+) &= \frac{P(\mathcal{P}^+|M_V) P(M_V)}{P(\mathcal{P}^+|M_V) P(M_V) + P(\mathcal{P}^+|M_V^c) P(M_V^c)} \\ &= \frac{\frac{90}{100} \frac{65}{1000}}{\frac{90}{100} \frac{65}{1000} + \frac{13}{100} \frac{935}{1000}} \\ &= \frac{90 \times 65}{90 \times 65 + 13 \times 935} \\ &= \frac{90}{277} \\ &\approx 0,325. \end{aligned}$$

Définition

Deux événements A et B sont dits **indépendants** si

$$P(A \cap B) = P(A) P(B).$$

Exemple

Reprenons l'exemple du jet de deux dés non pipés. On peut réinterpréter le fait que le joueur ne contrôle pas les numéros qui sortent, en remarquant en particulier que les événements liés à un dé sont indépendants des événements liés à l'autre. En particulier, comme chacun des deux dés peut être associé à son propre univers des possibles ayant six résultats élémentaires équiprobables, certains calculs en sont facilités.

Par exemple, on obtient alors une démonstration intuitive du fait que les résultats du jet des dés sont équiprobables, puisque, pour tout $\omega = (i, j)$ de Ω ,

$$\begin{aligned} P(\{(i, j)\}) &= P(\text{« le numéro du premier dé vaut } i \text{ et le numéro du second dé vaut } j \text{ »}) \\ &= P(\text{« le numéro du premier dé vaut } i \text{ »}) \times P(\text{« le numéro du second dé vaut } j \text{ »}) \\ &= \frac{1}{6} \times \frac{1}{6} \\ &= \frac{1}{36} \end{aligned}$$

Calculons aussi la probabilité de l'événement A : « la somme des deux numéros qui sortent est paire », et celle de B : « le numéro du second dé est inférieur ou égal à 2 ». A priori, comme l'événement A dépend simultanément des deux dés, le calcul de sa probabilité pourrait ne pas être évident. Mais, par symétrie, on comprend, intuitivement, que cette probabilité doit être égale à celle de son événement complémentaire, et donc que :

$$2 P(A) = 1$$

On peut également énumérer toutes les possibilités parmi les 36 pour en être sûr. Ainsi,

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{2}{6} = \frac{1}{3}$$

Calculons maintenant directement :

$$P(A \cap B) = P(\{(1, 1), (3, 1), (5, 1), (2, 2), (4, 2), (6, 2)\}) = \frac{1}{6}$$

Par suite,

$$P(A \cap B) = P(A) P(B)$$

Les événements A et B sont donc indépendants. Bien que cela soit très simple à déterminer, ce n'est pas évident a priori.

1. Variable aléatoire discrète

Définition

Soit (Ω, \mathcal{A}, P) un espace probabilisé. On appelle **variable aléatoire discrète** sur (Ω, \mathcal{A}, P) toute application $X : \Omega \rightarrow \mathbb{R}$, qui, à tout événement ω de Ω associe la quantité $X(\omega)$, et qui est telle que :

- l'ensemble des $X(\Omega) = \{X(\omega), \omega \in \Omega\}$ des images des épreuves de Ω est une partie au plus dénombrable de \mathbb{R} , i.e. ..., ce qui permet de « repérer » ses éléments par une indexation de la forme :

$$X(\Omega) = \{x_0, x_1, \dots, x_k, \dots\}$$

- pour tout x_k de $X(\Omega)$, l'ensemble $\{\omega \in \Omega, X(\omega) = x_k\}$ fait partie de la famille (ou tribu) \mathcal{A} d'événements auxquels on peut attribuer une probabilité par P .



Dès que l'univers Ω est fini ou dénombrable, toutes les fonctions X définies sur Ω sont des variables aléatoires discrètes. On se trouve, souvent, dans le cas d'univers dénombrables et non finis.

Notations

1. Dans ce qui suit, on se placera, implicitement, dans un espace probabilisé (Ω, \mathcal{A}, P) .
2. Dans ce qui suit, l'abréviation « v.a. » sera utilisée pour « variable aléatoire ».
3. L'antécédent (image réciproque) $X^{-1}(x_k)$ de la valeur x_k par la v.a. X , qui est, par définition, un événement, s'écrira $\{X = x_k\}$.

Propriété

La famille de tous les événements $\{X = x_k\}$ est telle que :

$$\sum_{x_k \in X(\Omega)} P(X = x_k) = 1$$



Pour représenter graphiquement une variable aléatoire discrète, il est aisé d'utiliser un « diagramme en bâtons ». Quand celui-ci correspond à une distribution de données réelles, on l'appellera aussi « histogramme ». Inversement, toute série de données réelles peut être représentée grâce à un histogramme, lequel peut, à son tour, être interprété comme la distribution d'une variable aléatoire. Pour une définition plus précise, on renvoie au concept de « loi » défini ci-dessous.

2. Loi d'une variable aléatoire discrète

Soit X une variable aléatoire discrète sur un espace probabilisé (Ω, \mathcal{A}, P) . L'ensemble $X(\Omega)$ des valeurs possibles de X est donc au plus dénombrable. On le note $\{x_k : k \in \mathbf{N}\}$. On considère la fonction p_X définie sur $X(\Omega)$ par :

$$p_X(x_k) = P(X = x_k)$$

Cette fonction s'appelle la **masse** ou la **fonction de probabilité** de X . On parle parfois, abusivement, de *loi* ou *distribution* de X .

La **loi**, ou encore **distribution**, de X est la fonction d'ensembles P_X définie sur toutes les parties B de \mathbb{R} par :

$$\forall B \subset \mathbb{R} : P_X(B) = \sum_{x_k \in B} P(X = x_k) = \sum_{x_k \in B} p_X(x_k).$$



1. Il est intéressant de remarquer que la loi P_X de X est déterminée par sa masse p_X , et inversement, d'où la possibilité de confondre les deux notions sans risque grave.
2. Ce n'est pas parce que deux variables aléatoires ont la même loi qu'elles sont égales !

Exemple

Ainsi, considérons le lancer de deux dés, un dé rouge, et un dé vert. Si X est le chiffre qui sort sur le dé rouge, et Y celui qui sort sur le dé vert, les variables aléatoires discrètes X et Y sont définies sur le même espace probabilisé $\Omega = \{1, 2, 3, 4, 5, 6\}$. Ainsi :

$$X(\Omega) = Y(\Omega) = \{1, 2, 3, 4, 5, 6\}$$

Si les deux dés sont non pipés, X et Y ont donc la même loi.

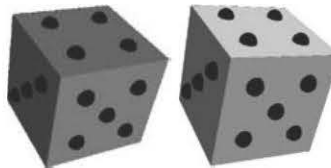


Figure 116.1 – Le lancer de dés...

Par contre, les variables aléatoires X et Y ne sont pas égales ; si tel était le cas, cela signifierait que, pour tout événement ω de l'univers des possibles Ω , on aurait $X(\omega) = Y(\omega)$, et donc que tout lancer des deux dés permettrait, à coup sûr, d'obtenir le même numéro, ce qui n'est pas le cas. De manière équivalente, cela signifie que le lancer du dé rouge permettrait de prédire à coup sûr le numéro qui sort sur le dé vert, ce qui est bien sûr choquant. La meilleure façon de comprendre pourquoi ce n'est pas possible est de faire l'hypothèse que les deux lancers sont indépendants ; on se retrouve alors dans le cas de l'exemple précédent, où les 36 résultats possibles sont équiprobables. Le lancer du dé rouge ne donne aucune information sur celui du dé bleu. Nous verrons ultérieurement que cet exemple permet de comprendre intuitivement la notion d'indépendance pour des variables aléatoires.

3. Fonction d'une variable aléatoire discrète

Soit (Ω, \mathcal{A}, P) un espace probabilisé, et f une fonction définie sur $X(\Omega)$, à valeurs réelles.

La variable aléatoire discrète $f(X)$ est appelée **fonction de la variable aléatoire X** .



La formule $Y = f(X)$ définit une autre variable aléatoire, puisque $\{Y = f(x_k)\} = \{X = x_k\}$ pour tout $x_k \in X(\Omega)$.

4. Variables aléatoires discrètes indépendantes

Deux variables aléatoires discrètes X et Y définies sur le même espace probabilisé (Ω, \mathcal{A}, P) sont dites **indépendantes** si, pour tous $x_k \in X(\Omega)$ et $y_\ell \in Y(\Omega)$, les événements $\{X = x_k\}$ et $\{Y = y_\ell\}$ sont indépendants.

Dans ce cas, la « masse » de la paire (X, Y) , qui peut être définie par :

$$p_{X,Y}(x_k, y_\ell) = P(\{X = x_k\} \cap \{Y = y_\ell\})$$

a la forme d'un produit :

$$p_{X,Y}(x_k, y_\ell) = p_X(x_k) p_Y(y_\ell)$$

On obtient ainsi une règle de produit pour les probabilités des intersections de toutes les paires d'événements basés sur X et Y .

Ces concepts se généralisent aisément à tout nombre fini de variables aléatoires discrètes. En particulier, si X_1, X_2, \dots, X_n sont n variables indépendantes, alors, pour tout x_k de $X_k(\Omega)$, $k = 1, 2, \dots, n$,

$$P(X_1 = x_1 \text{ et } X_2 = x_2 \text{ et } \dots \text{ et } X_n = x_n) = P(X_1 = x_1)P(X_2 = x_2) \dots P(X_n = x_n).$$

Exemple

Si X et Y sont respectivement les valeurs obtenues en lançant un dé rouge et un dé vert, et que les deux lancers sont indépendants, alors pour tout couple d'indices (i, j) de $\{1, \dots, 6\}^2$:

$$P(X = i \text{ et } Y = j) = P(X = i) P(Y = j) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

De surcroît,

$$\begin{aligned} P(\text{« } X \text{ est pair et } Y \text{ est un multiple de 3 »}) &= P(\text{« } X \text{ est pair »}) P(\text{« } Y \text{ est un multiple de 3 »}) \\ &= \frac{1}{2} \times \frac{1}{3} \\ &= \frac{1}{6} \end{aligned}$$

Définition

On appelle **fonction de répartition** d'une variable aléatoire X la fonction, notée F_X , définie pour tout réel t par :

$$F_X(t) = P_X([-\infty, t]) = P(X \leq t) = \sum_{x_k \in X(\Omega), x_k \leq t} P(X = x_k)$$

Exemples

1. Soit S la variable aléatoire donnant la somme des chiffres obtenus en lançant deux dés à six faces bien équilibrés. Sur l'espace de probabilités :

$$\Omega = \{(i, j), (i, j) \in \{1, \dots, 6\}^2\}$$

avec les 36 événements élémentaires équiprobables que nous connaissons,

$$S(i, j) = i + j$$

On peut réaliser un tableau pour lequel l'élément situé ligne i , colonne j , vaut $i + j$,

2	3	4	5	6	7
3	4	5	6	7	8
4	5	6	7	8	9
5	6	7	8	9	10
6	7	8	9	10	11
7	8	9	10	11	12

La loi de S est déterminée par sa masse $p_S(k)$, définie pour les 11 valeurs k dans :

$$S(\Omega) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

On calcule aisément :

$$p_S(2) = P(S = 2) = \frac{1}{36}, \quad p_S(3) = P(S = 3) = \frac{2}{36} = \frac{1}{18},$$

$$p_S(4) = P(S = 4) = \frac{3}{36} = \frac{1}{12}$$

$$p_S(5) = P(S = 5) = \frac{4}{36} = \frac{1}{9}, \quad p_S(6) = P(S = 6) = \frac{5}{36},$$

$$p_S(7) = P(S = 7) = \frac{6}{36} = \frac{1}{6}, \quad p_S(8) = P(S = 8) = \frac{5}{36},$$

$$p_S(9) = P(S = 9) = \frac{4}{36} = \frac{1}{9}$$

$$p_S(10) = P(S = 10) = \frac{3}{36} = \frac{1}{12}, \quad p_S(11) = P(S = 11) = \frac{2}{36} = \frac{1}{18},$$

$$p_S(12) = P(S = 12) = \frac{1}{36}$$

La fonction de répartition de la variable aléatoire S est définie par :

$$F_S(t) = P_S([-\infty, t]) = P(S \leq t) = \sum_{x_k \in S(\Omega), x_k \leq t} P(X = x_k)$$

Calculons $F_S(t)$ pour $t = 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$. Ainsi :

$$F_S(2) = p_S(2) = \frac{1}{36}, \quad F_S(3) = p_S(2) + p_S(3) = \frac{1}{36} + \frac{2}{36} = \frac{1}{12}$$

$$F_S(4) = p_S(2) + p_S(3) + p_S(4) = \frac{1}{36} + \frac{2}{36} + \frac{3}{36} = \frac{1}{6}, \quad \dots$$

et ainsi de suite. On note, en particulier, que F est croissante, et que

$$F_S(12) = 1$$

2. Prenons un autre exemple élémentaire et traitons le rapidement. On s'intéresse au lancer, de manière indépendante, de trois pièces de monnaie. On suppose que la probabilité de tomber sur « pile » vaut $\frac{1}{2}$ pour chacune des pièces. On attribue la valeur 1 au côté « pile », et 0 au côté « face ».

L'univers des possibles Ω possède huit éléments, qui sont les triplets de la forme (i, j, k) , où i, j et k prennent les valeurs 0 et 1. En raison de l'indépendance, les huit résultats sont équiprobables. On définit alors la variable aléatoire X , qui correspond au nombre de fois où les 3 pièces tombent sur « pile » non nécessairement de façon successive. X peut prendre les valeurs 0, 1, 2, 3. Les événements possibles sont :

$$\{X = 0\} = \{(0, 0, 0)\}$$

$$\{X = 1\} = \{(0, 0, 1), (0, 1, 0), (1, 0, 0)\}$$

$$\{X = 2\} = \{(0, 1, 1), (1, 0, 1), (1, 1, 0)\}$$

$$\{X = 3\} = \{(1, 1, 1)\}$$

La loi de X est donnée par :

$$p_X(0) = \frac{1}{8}, \quad p_X(1) = \frac{3}{8}, \quad p_X(2) = \frac{3}{8}, \quad p_X(3) = \frac{1}{8}$$

La fonction de répartition de X est donc telle que :

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{8} & \text{si } 0 \leq x < 1 \\ \frac{1}{2} & \text{si } 1 \leq x < 2 \\ \frac{7}{8} & \text{si } 2 \leq x < 3 \\ 1 & \text{si } x \geq 3 \end{cases}$$

Théorème

La fonction de répartition F_X d'une variable aléatoire discrète X est croissante sur \mathbb{R} , continue à droite, et bornée à gauche. Elle a pour limite 0 en $-\infty$ et 1 en $+\infty$.

La loi de Bernoulli, de paramètre $p \in [0, 1]$

Définition

On dit qu'une variable aléatoire discrète X suit la **loi de Bernoulli de paramètre** $p \in [0, 1]$ si elle ne prend que les valeurs 0 et 1, avec :

$$P(X = 1) = p, \quad P(X = 0) = 1 - p$$



On note alors : $X \sim \mathcal{B}(p)$

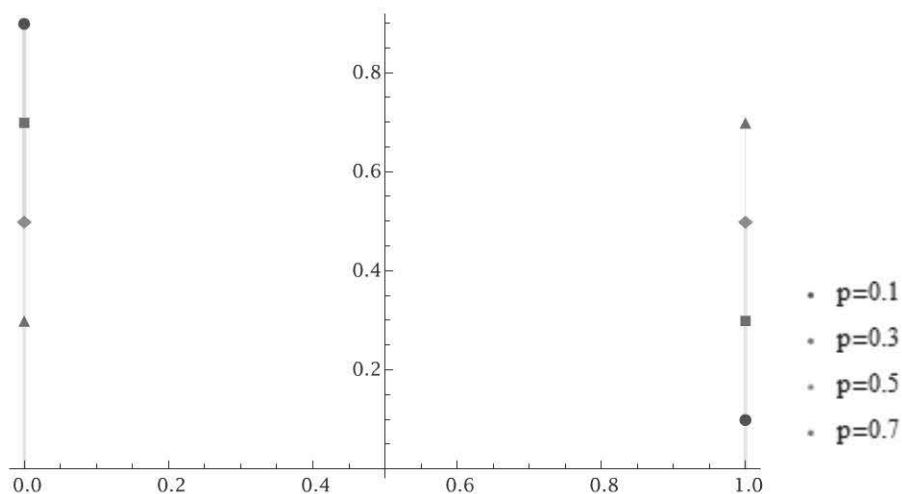


Figure 118.1 – Une représentation graphique de la fonction de probabilité pour la loi de Bernoulli, pour $p = 0,1$, $p = 0,3$, $p = 0,5$ et $p = 0,7$.

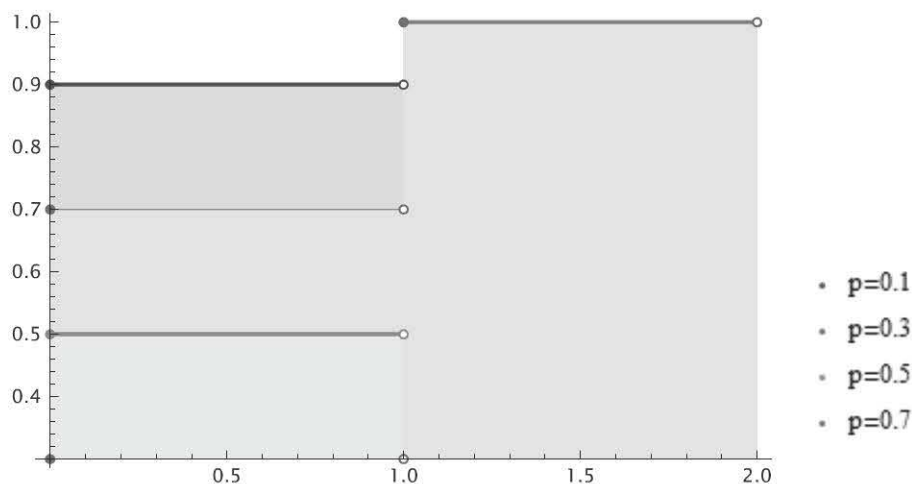


Figure 118.2 – Une représentation graphique de la fonction de répartition de la loi de Bernoulli, pour $p = 0,1$, $p = 0,3$, $p = 0,5$ et $p = 0,7$.

Pour tout événement A de probabilité p , on définit sa **fonction indicatrice** \mathbb{I}_A par :

$$\mathbb{I}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

i.e. :

$$\mathbb{I}_A(\omega) = \begin{cases} 1 & \text{si } A \text{ est réalisé} \\ 0 & \text{si } A \text{ n'est pas réalisé} \end{cases}$$



1. \mathbb{I}_A est une variable aléatoire de Bernoulli ; son paramètre p vaut $P(A)$.
2. La variable de Bernoulli est utilisée le plus souvent pour désigner une expérience avec un résultat binaire : un succès, à qui l'on attribue la valeur 1, et un échec, à qui l'on attribue la valeur 0.

La loi uniforme (sur un ensemble de réels)

Définition

On dit qu'une variable aléatoire discrète X suit la **loi uniforme** sur l'ensemble de réels $\{x_1, \dots, x_n\}$ si, pour tout entier k de $\{1, \dots, n\}$:

$$P(X = x_k) = \frac{1}{n}$$

ce qui signifie que la loi P_X est la loi d'équiprobabilité.



La loi de Bernoulli est la loi uniforme sur $\{0, 1\}$.

Exemples

1. Lors du lancer d'un dé à 6 faces non pipé, le chiffre qui sort suit la loi uniforme sur $\{1, 2, 3, 4, 5, 6\}$.

2. Au casino de Monte Carlo, à la roulette, la boule peut tomber sur les valeurs entières entre 0 et 36 de manière équiprobable, ce qui signifie que la probabilité de chaque valeur est $\frac{1}{37}$. La valeur obtenue est ainsi une variable uniforme sur $\{0, 1, 2, \dots, 36\}$.

En revanche, à Las Vegas, la roulette possède une case « 00 ». Chaque valeur a donc la probabilité $\frac{1}{38}$ de sortir ; la v.a. correspondante est uniforme sur l'ensemble $\{00, 0, 1, 2, \dots, 36\}$, bien que, à strictement parler, ce ne soit pas, si l'on se réfère à la définition précédente, une loi uniforme. Il suffit de remplacer la valeur 00 par -1 pour remédier à ce problème de notation.

L'apparition du double zéro 00 à Las Vegas peut être interprétée comme une indication de l'avidité des promoteurs du jeu dans cette ville du Far West, pour se donner un avantage par rapport aux traditions européennes de fair play. On pourrait croire que les joueurs ne seraient pas dupes, mais l'histoire pourrait donner raison à Las Vegas, si on mesure le succès d'un casino par le volume de jeu. Toutefois, bien que symbolique du succès des jeux d'argent, Las Vegas fait maintenant pâle figure dans ce domaine, en comparaison avec le succès de l'industrie du jeu à Macao, région au régime administratif spécial de la République Populaire de Chine.

Définition

On dit qu'une variable aléatoire discrète X suit la **loi binomiale de paramètres** $n \in \mathbb{N}^*$ et $p \in [0, 1]$, $\mathcal{B}(n, p)$, si l'ensemble des valeurs possibles est $X(\Omega) = \{0, 1, \dots, n\}$ et si, pour tout entier k de $\{0, 1, \dots, n\}$:

$$P(X = k) = C_n^k p^k (1 - p)^{n-k} = \binom{n}{k} p^k (1 - p)^{n-k}.$$



1. On note alors :

$$X \sim \mathcal{B}(n, p)$$

2. La loi binomiale correspond au renouvellement, de manière indépendante, de n épreuves de Bernoulli de paramètre p . Plus précisément, on dispose du résultat suivant :

1. De la loi de Bernoulli à la loi binomiale

Soit p un réel de l'intervalle $[0, 1]$, et X_1, \dots, X_n , des v.a. indépendantes suivant la loi de Bernoulli de paramètre p . Alors, la somme $X = X_1 + \dots + X_n$ suit la loi binomiale $\mathcal{B}(n, p)$.



Inversement, toute v.a. X suivant la loi binomiale $\mathcal{B}(n, p)$ peut être considérée comme la somme de n v.a. indépendantes suivant la loi de Bernoulli de paramètre p , ce qui signifie qu'une loi binomiale $\mathcal{B}(n, p)$ peut s'interpréter comme celle du nombre de succès dans une série de n expériences indépendantes pour lesquelles la probabilité de succès vaut p .

Démonstration : X vaut k si et seulement si on trouve, parmi les X_i , k variables qui valent 1, et $n - k$ qui sont nulles. Les variables étant indépendantes, la probabilité d'une telle configuration, où les k variables qui sont égales à 1 ont été choisies, est donnée par :

$$\underbrace{p \cdots p}_{k \text{ fois}} \cdot \underbrace{(1 - p) \cdots (1 - p)}_{n-k \text{ fois}} = p^k (1 - p)^{n-k}.$$

Le nombre de façons de choisir les k variables qui valent 1 est précisément le nombre de parties à k éléments pris parmi n , i.e. C_n^k . Chaque choix correspondant à un événement disjoint des autres, les probabilités correspondantes peuvent être additionnées, ce qui montre que :

$$P(X = k) = C_n^k p^k (1 - p)^{n-k}$$

et donc que la v.a. X est bien binomiale de paramètres n, p . ■

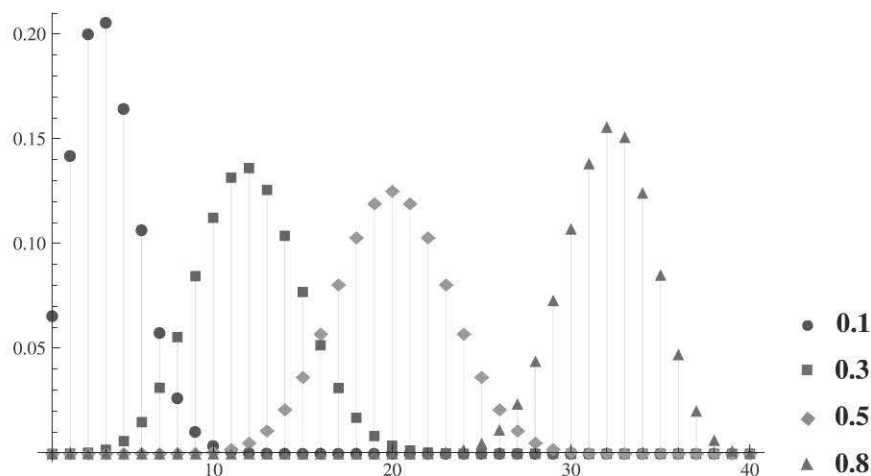


Figure 120.1 – Une représentation graphique de la distribution de probabilité pour la loi binomiale, pour $p = 0,1$, $p = 0,3$, $p = 0,5$, $p = 0,8$ et $n = 50$.

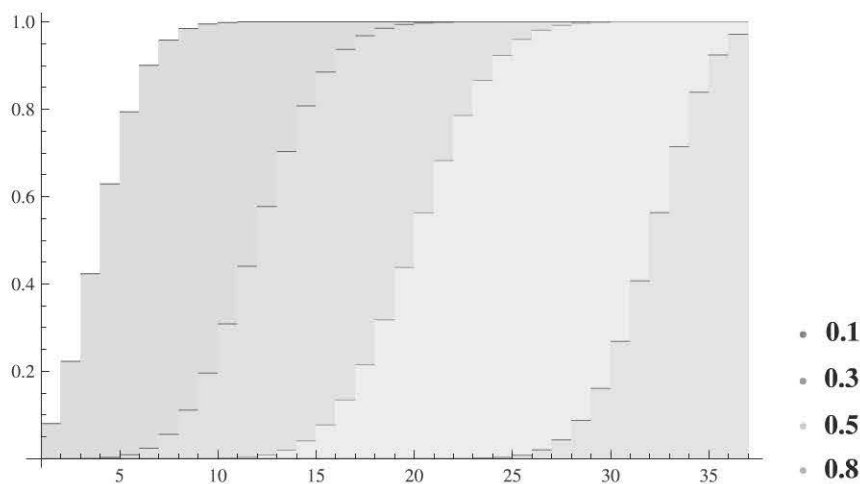


Figure 120.2 – Une représentation graphique de la fonction de répartition de la loi binomiale, pour $p = 0,1$, $p = 0,3$, $p = 0,5$, $p = 0,8$ et $n = 40$.



On vérifie bien que :

$$\sum_{k=0}^n P(X = k) = \sum_{k=0}^n C_n^k p^k (1-p)^{n-k} = (p + 1 - p)^n = 1$$

(c'est, tout simplement, la formule du binôme de Newton...)

2. Le théorème du jury de Condorcet

En l'an de grâce 1785, le marquis de Condorcet¹ produisit une justification formelle du suffrage universel dans le cas d'un choix binaire, grâce à un argument probabiliste. Son argument était le suivant : dans la mesure où le citoyen moyen a moins d'une chance sur deux de se tromper, la somme de tous les votes des citoyens a très peu de chance d'être erronée. Et, lorsque le nombre de participants augmente, la décision choisie tend à être

1. Marie Jean Antoine Nicolas de Caritat, 1743-1794.

la bonne. Mais cet argument ne tient que lorsque l'électeur n'a que deux choix, et non plus... [38, 39].

La formule permettant de calculer la probabilité que la décision choisie soit la bonne s'obtient à partir de la loi binomiale de paramètres n , p , où p est un nombre compris entre $\frac{1}{2}$ et 1, en sommant sur l'ensemble des votants (on ajoute les probabilités de « bon vote ») :

$$P\left(X \geq E\left[\frac{n}{2}\right], p\right) = \sum_{k=E\left[\frac{n}{2}\right]}^n C_n^k p^k (1-p)^{n-k} = \sum_{k=E\left[\frac{n}{2}\right]}^n \binom{n}{k} p^k (1-p)^{n-k}$$

avec $p = 0,5 + \varepsilon$, $\varepsilon \in \mathbb{R}^+$.

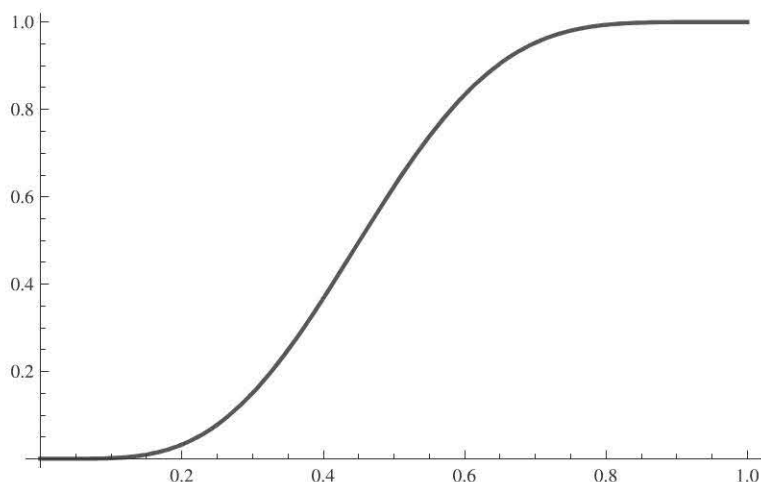


Figure 120.3 – Une première illustration graphique du théorème du jury de Condorcet : la probabilité que le vote tende à être le bon en fonction de p , pour un nombre de votants donné.

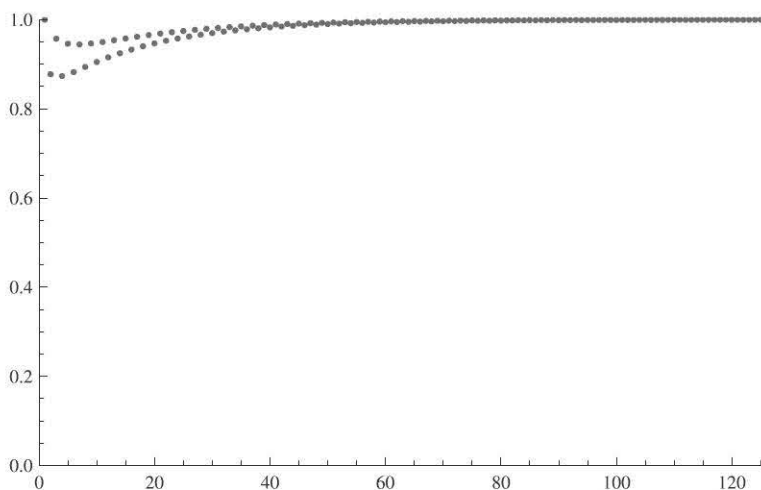


Figure 120.4 – Une seconde illustration graphique du théorème du jury de Condorcet : la probabilité que le vote tende à être le bon en fonction du nombre de votants n , pour une valeur de p donnée (ici : $p = 0,65$).

Définition

On dit qu'une variable aléatoire discrète X suit **la loi géométrique de paramètre** $p \in [0, 1[$ si l'ensemble des valeurs possibles est $X(\Omega) = \mathbb{N}^*$ et si, pour tout entier naturel n :

$$P(X = n) = p(1 - p)^{n-1}$$

On note alors :

$$X \sim \mathcal{G}(p)$$

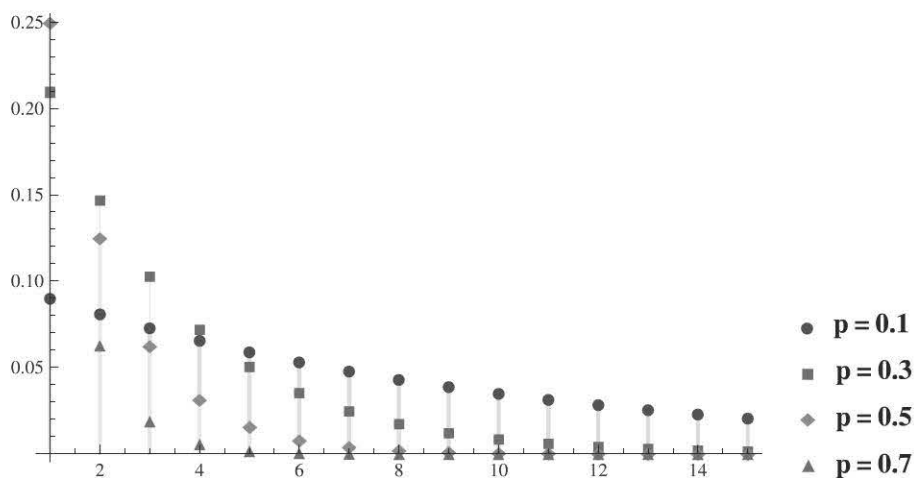


Figure 121.1 – Une représentation graphique de la fonction de probabilité pour la loi géométrique, pour $p = 0,1$, $p = 0,3$, $p = 0,5$ et $p = 0,7$.

1. Propriété

Pour la loi géométrique de paramètre p , les résultats sur les sommes de séries géométriques conduisent à :

$$\begin{aligned}
 P(X > n) &= \sum_{k=n+1}^{+\infty} p(1-p)^{k-1} \\
 &= \sum_{k=n}^{+\infty} p(1-p)^k \\
 &= \frac{p(1-p)^n}{1 - (1-p)} \\
 &= \frac{p(1-p)^n}{p} \\
 &= (1-p)^n.
 \end{aligned}$$

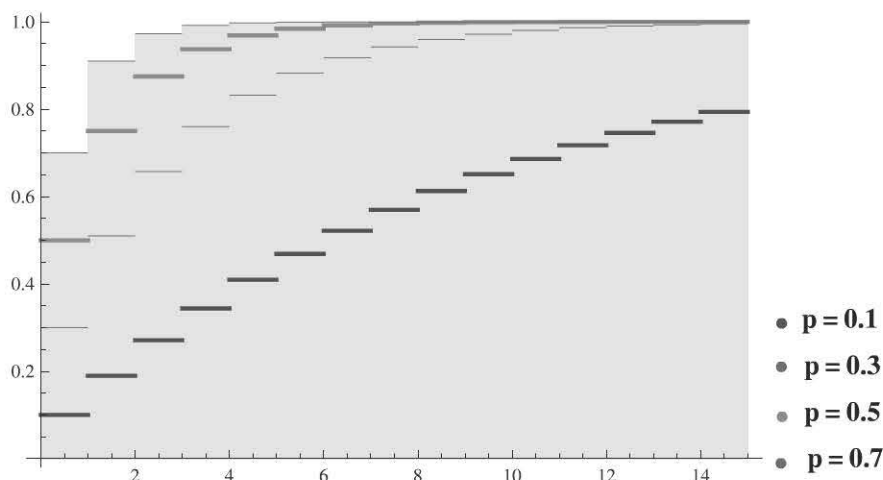


Figure 121.2 – Une représentation graphique de la fonction de répartition de la loi géométrique, pour $p = 0,1$, $p = 0,3$, $p = 0,5$ et $p = 0,7$.



Comme pour le cas de la loi binomiale, et bien d'autres, la loi géométrique peut également être définie grâce à une suite de variables de Bernoulli de paramètre p ; il faut alors considérer une suite infinie de telles variables $\{X_i : i = 1, 2, \dots\}$ indépendantes, et interpréter i comme une variable temporelle, ce qui signifie que X_i est le résultat obtenu (échec, 0, ou succès, 1) pour la tentative effectuée à l'instant i . On désigne par X le premier instant où l'on obtient un succès. L'événement $\{X = k\}$ est donné par :

$$\{X = k\} = \{X_1 = 0\} \cap \{X_2 = 0\} \cap \dots \cap \{X_{k-1} = 0\} \cap \{X_k = 1\}$$

En raison de l'indépendance des variables X_i , et donc des événements correspondant, on obtient immédiatement :

$$P(X = k) = (1 - p)^{k-1} p$$

La v.a. X est donc géométrique, de paramètre p .

La propriété précédente est facile à redémontrer par un raisonnement analogue :

$$\{X > k\} = \{X_1 = 0\} \cap \{X_2 = 0\} \cap \dots \cap \{X_k = 0\}$$

Par indépendance :

$$P(X > k) = (1 - p)^k$$

Le résultat suivant reprend cette discussion, qui rend, de fait, la définition et la proposition précédentes redondantes.

Proposition

La loi géométrique $\mathcal{G}(p)$ est la loi du premier instant de succès dans une suite infinie de tentatives indépendantes de probabilité de succès p (suite de variables de Bernoulli $\mathcal{B}(p)$).

Une variable aléatoire X , de loi $\mathcal{G}(p)$, vérifie :

$$P(X = k) = p (1 - p)^{k-1}$$

et :

$$P(X > k) = (1 - p)^k$$



1. La loi géométrique est donc utile lorsqu'il s'agit de modéliser des tentatives dont le résultat peut être traduit de façon binaire (un échec ou un succès), avec une dépendance temporelle, par exemple, des temps d'attente jusqu'à ce que se produise un événement attendu ou redouté, lorsque le temps est mesuré de manière discrète (nombre d'heures, de jours, etc.).
2. La loi géométrique correspond, précisément, à des temps de vie sans vieillissement. Considérons deux entiers naturels m et n , et une variable $X \sim \mathcal{G}(p)$. Grâce à la proposition précédente,

$$P(X > m) = (1 - p)^m$$

De plus, comme « $X > m + n$ » implique « $X > n$ », on a donc :

$$\{X > m + n\} \cap \{X > n\} = \{X > m + n\}$$

ce qui conduit alors, toujours grâce à la proposition précédente, à :

$$P(\{X > m + n\} \cap \{X > n\}) = (1 - p)^{m+n}$$

La définition de la probabilité conditionnelle permet d'en déduire :

$$P(X > m + n | X > n) = \frac{(1 - p)^{m+n}}{(1 - p)^n} = (1 - p)^m = P(X > m).$$

Nous venons de démontrer la propriété suivante.

2. Absence de mémoire pour une v.a. suivant la loi géométrique de paramètre $p \in [0, 1[$

Soit X une v.a. suivant la loi géométrique de paramètre $p \in [0, 1[$. La v.a. X vérifie la propriété d'**absence de mémoire** :

$$\forall (m, n) \in \mathbf{N} \times \mathbf{N} : P(X > m + n | X > n) = P(X > m)$$



Cette propriété est assez naturelle, dans la mesure où, dans une suite d'épreuves de Bernoulli, la loi de probabilité du nombre d'épreuves à répéter jusqu'à l'obtention du premier succès reste la même quel que soit le nombre d'échecs (les épreuves sont toutes identiques et indépendantes).

Exemples

1. Nous avons fait allusion à l'existence d'événements négligeables non vides. La loi géométrique donne une bonne illustration de ce phénomène. Si on lance, une infinité de fois, et de façons indépendantes les unes des autres, une pièce de monnaie, l'événement A : « on ne tombe jamais sur pile » n'est pas un ensemble vide, puisque :

$$A = \{0, 0, \dots\}$$

Si on désigne par X le nombre de lancers jusqu'à la première apparition de la face « pile », alors la v.a. X est géométrique, et $A = \{X = +\infty\}$. Par suite :

$$P[A] = P[X = +\infty] = 1 - P[X < +\infty] = 1 - \sum_{k=0}^{+\infty} P[X = k] = 1 - 1 = 0$$

L'événement A est bien non vide, mais négligeable.

2. Prenons l'exemple du légendaire Ty Cobb. On suppose que, pour chaque tentative de frapper la balle, sa probabilité de succès vaut 0,3664, et que les tentatives successives sont indépendantes. Le nombre X de tentatives jusqu'à l'obtention de son premier succès est donc une v.a. géométrique de paramètre p . La probabilité qu'il faille à Ty Cobb exactement 3 tentatives pour obtenir son premier succès est :

$$P(X = 3) = 0,3664 \times 0,6336^2 \approx 0,1471$$

La probabilité qu'il n'arrive pas à frapper la balle dans une partie où il a droit à 3 tentatives, est identique à celle qu'il lui faille plus de 3 tentatives pour frapper la balle, et vaut donc :

$$P(X > 3) = 0,6336^3 \approx 0,2544$$

Sachant que Ty Cobb est dans un mauvais jour, et qu'il vient de rater 3 tentatives de frapper la balle, la probabilité qu'il n'ait toujours pas frappé la balle au bout de la cinquième tentative est :

$$P(X > 5 | X > 3) = P(X > 2) = 0,6336^2 = 0,4014 \quad (121.1)$$

En fait, avec les hypothèses utilisées pour cet exemple, même si Ty Cobb avait eu une période extrêmement mauvaise dans laquelle il ne réussissait pas toujours à frapper la balle au bout d'un grand nombre de tentatives, on ne dispose d'aucune information sur ses chances de frapper celle-ci au cours des tentatives suivantes. Ce résultat peut être contre-intuitif, dans la mesure où les historiens du baseball estiment qu'il était doué d'une capacité de concentration hors du commun. Ceci dit, la modélisation géométrique permet néanmoins de très bonnes approximations dans de nombreux cas de tentatives répétées, et donc même pour notre champion Ty Cobb.

La loi binomiale négative

Définition

On dit qu'une variable aléatoire discrète X suit la **loi binomiale négative de paramètres** $n \in \mathbb{N}^*$ et $p \in [0, 1]$, $\mathcal{BN}(n, p)$ ¹, si l'ensemble $X(\Omega)$ des valeurs possibles est l'ensemble des entiers supérieurs ou égaux à n , et si, pour tout entier $k \geq n$:

$$P(X = k) = C_{k-1}^{n-1} p^n (1-p)^{k-n} = \binom{k-1}{n-1} p^n (1-p)^{k-n}$$

On note alors :

$$X \sim \mathcal{BN}(n, p)$$



La loi binomiale négative est une extension de la loi géométrique. Elle permet d'envisager les cas où le nombre n de succès (par rapport à un événement donné) est connu, et k est le nombre d'épreuves nécessaires pour obtenir ces n succès que l'on cherche. Comme d'habitude, p est la probabilité d'un succès.

Pour comprendre ce phénomène, supposons que la suite des tentatives soit une suite infinie de variables de Bernoulli $\mathcal{B}(p)$. La variable X que l'on recherche est en fait la somme de n variables géométriques indépendantes. En effet, pour obtenir exactement n succès, il suffit d'attendre le premier succès, puis de recommencer à compter jusqu'au second, etc., et ce, jusqu'au $n^{\text{ième}}$. Chaque fois que l'on remet ainsi le compteur à zéro, les tentatives qui ont permis d'atteindre le dernier succès sont indépendantes de celles qui permettront d'atteindre le prochain.

Réfléchissons maintenant à l'événement $\{X = k\}$, i.e. « il faut exactement k tentatives pour obtenir n succès ». Il faut bien sûr $k \geq n \geq 1$. De plus, pour définir cet événement, il suffit de décider où se situeront les n tentatives qui seront des succès, les $k - n$ autres étant des échecs, tout en réalisant que la dernière des k tentatives doit être un succès. On définit donc l'événement $\{X = k\}$ en choisissant $n - 1$ tentatives fructueuses restantes parmi les $k - 1$ tentatives qu'il reste à choisir. Le nombre de façons de choisir est, bien évidemment, C_{k-1}^{n-1} , alors que, pour un choix donné des instants des succès, la probabilité est $p^n (1-p)^{k-n}$ puisqu'il y a n succès et $k - n$ échecs. Ceci montre bien que :

$$P(X = k) = C_{k-1}^{n-1} p^n (1-p)^{k-n}$$

Nous résumons les résultats de cette discussion dans la proposition suivante, même si, comme dans le cas de la loi géométrique, il y a redondance avec la définition précédente.

1. Cette loi est aussi appelée Loi de Pascal, ou encore, Loi de Pólya.

Proposition

Une variable $X \sim \mathcal{BN}(n, p)$ peut s'écrire comme la somme de n variables géométriques indépendantes X_1, \dots, X_n de paramètre p :

$$X = \sum_{i=1}^n X_i$$

De plus, pour tout couple d'entiers naturels (n, k) tels que $k \geq n \geq 1$,

$$P(X = k) = C_{k-1}^{n-1} p^n (1-p)^{k-n}.$$



Pour étudier le comportement d'une population de tortues géantes des Galápagos, on insère, à travers leur carapace, des micro-capteurs thermiques dans leur cavité abdominale. Ces capteurs permettent de savoir si les tortues parviennent à échapper au stress thermique lié à la dégradation de leur environnement. Sur ces capteurs, seuls 80 % réussissent à fonctionner correctement et à recueillir les informations attendues. On considère qu'il faut un minimum de vingt tortues dont les capteurs fonctionnent pour que l'étude soit satisfaisante.

Soit X la v.a. correspondant au nombre de tortues dont les émetteurs fonctionnent. Si on considère que les perturbations qui peuvent affecter le fonctionnement des capteurs sont indépendantes les unes des autres, on en déduit, grâce à la proposition précédente, que X suit la loi binomiale négative $\mathcal{BN}\left(20, \frac{80}{100}\right)$.

La probabilité de devoir équiper exactement trente tortues de capteurs vaut donc :

$$\begin{aligned} P(X = 30) &= C_{30-1}^{20-1} \left(\frac{80}{100}\right)^{20} \left(1 - \frac{80}{100}\right)^{30-20} \\ &= C_{29}^{19} \left(\frac{4}{5}\right)^{20} \left(\frac{1}{5}\right)^{10} \\ &= \frac{29!}{19! 10!} \left(\frac{4}{5}\right)^{20} \left(\frac{1}{5}\right)^{10} \\ &= \frac{20 \times 21 \times \dots \times 29}{10!} \frac{4^{20}}{5^{30}} \\ &= \frac{4404\,645\,779\,893\,911\,552}{186\,264\,514\,923\,095\,703\,125} \\ &\approx 0,02365 \end{aligned}$$

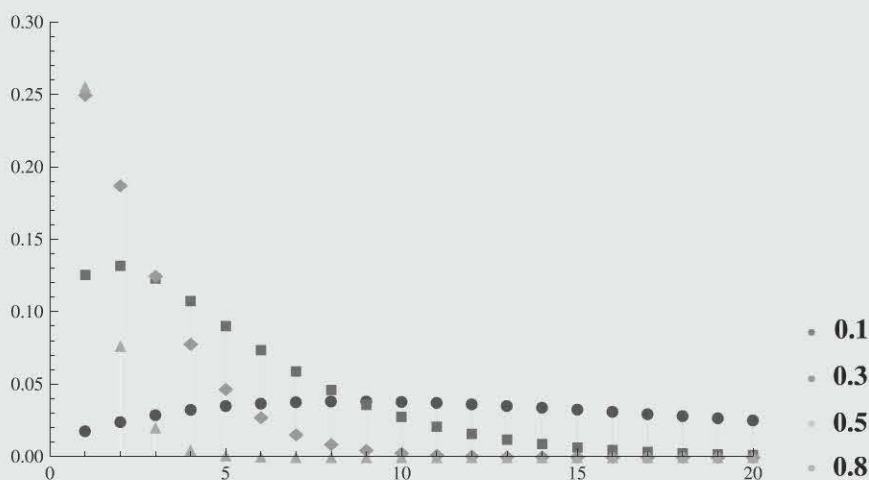
Il serait peut-être plus intéressant de connaître la probabilité qu'il faille équiper au moins 30 tortues. Étant donné que la variable X peut prendre des valeurs k arbitrairement grandes, on calculera plutôt la probabilité de son complémentaire. Ceci demande néanmoins de calculer 10 valeurs non triviales, ce qui, avec une calculatrice simple, est peu aisé, quoiqu'envisageable. Comme le nombre 10 n'est pas très grand au sens des théorèmes limites que l'on rencontre en théorie des probabilités (voir, par exemple, plus loin, le théorème central limite), il n'y a pas d'autre possibilité qu'un calcul direct :

$$\begin{aligned} P(X \geq 30) &= 1 - \sum_{k=20}^{29} P(X = k) \\ &= 1 - \sum_{k=20}^{29} \binom{k-1}{19} \left(\frac{4}{5}\right)^{20} \left(\frac{1}{5}\right)^{k-20} \\ &= \frac{367\,041\,806\,162\,986\,909}{7450\,580\,596\,923\,828\,125} \\ &\approx 0,04926. \end{aligned}$$

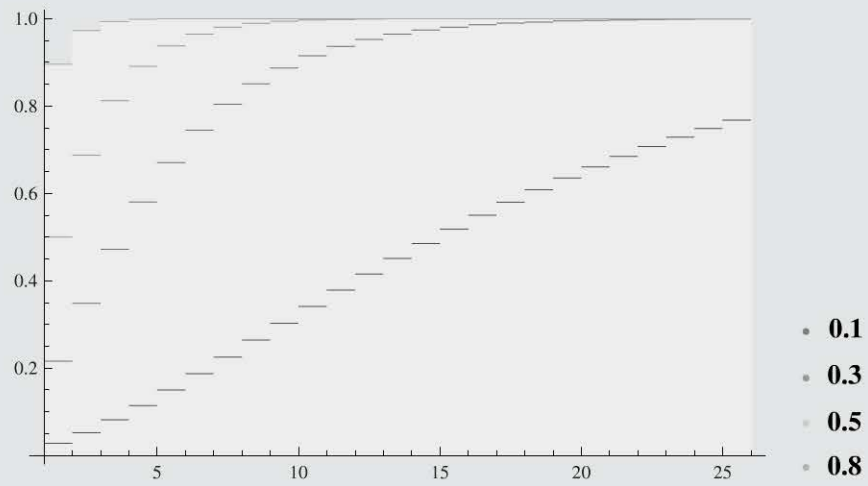
On peut remarquer que la probabilité de l'événement $\{X = 30\}$ représente près de la moitié de la masse de la queue entière de la distribution de X à partir de 30. Ceci indique que $P(X \geq k)$ décroît très vite, chose que l'on peut observer partiellement en faisant les calculs à la main. On peut également se demander quelle valeur de k maximise la probabilité $P(X = k)$.

Nous verrons, ultérieurement, l'importance de la notion d'espérance, qui correspond à celle, intuitive, de moyenne. Dans cet exemple, il se trouve que $P(X = k)$ est maximal pour $k = 25$. Cette valeur de k est également la moyenne intuitive de X . On s'attend en effet, puisque la chance de succès est de $\frac{4}{5}$, que l'on doit attendre en moyenne une durée de $\frac{5}{4}$ avant d'obtenir un succès (ou, de manière plus intuitive, on aurait en moyenne 4 succès sur 5 tentatives consécutives) ; de plus, comme X est la somme de 20 variables géométriques, pour obtenir 20 succès, on devrait attendre $20 \times \frac{5}{4} = 25$ tentatives successives en moyenne. Ce raisonnement intuitif sera rendu entièrement rigoureux dans ce qui suit lors des calculs d'espérance mathématique.

Les considérations de l'exemple précédent peuvent également s'appliquer à la loi binomiale, lorsque n est assez grand (disons $n \geq 30$), puisque cette loi est aussi celle de la somme de n variables de Bernoulli. Nous invitons le lecteur à fabriquer un tel exemple, peut-être avec notre ami Ty Cobb ? La seule différence notable entre les deux types de variables est que la variable binomiale prend ses valeurs k entre 0 et n , alors que la variable binomiale négative les prend au-delà de n . Mais lorsque n est grand, les effets de décroissance rapide de la probabilité $P(X = k)$, pour k grand, et de maximisation de $P(X = k)$ autour de la valeur moyenne intuitive de X , sont similaires dans les deux cas. Ce phénomène est dû au théorème central limite, qui sera vu ultérieurement.



Une représentation graphique de la fonction de probabilité pour la loi binomiale négative, pour $p = 0,1$, $p = 0,3$, $p = 0,5$ et $p = 0,8$.



Une représentation graphique de la fonction de répartition de la loi binomiale négative, pour $p = 0,1$, $p = 0,3$, $p = 0,5$ et $p = 0,8$.

Tirons simultanément n boules dans une urne contenant M boules rouges et $N-M$ boules noires. On peut utiliser comme univers Ω l'ensemble des parties à n éléments parmi N , sans que l'ordre importe. Si X désigne la variable aléatoire donnant le nombre de boules rouges extraites, sa loi de probabilité est :

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$$

Pour comprendre cette formule, il faut remarquer que, dans l'ensemble Ω défini ci-dessus, on fait l'hypothèse selon laquelle chaque boule a la même chance que les autres d'être tirée. Par conséquent, chaque partie à n éléments parmi N a la même chance d'être tirée. La probabilité de chaque événement élémentaire de Ω a une probabilité équiprobable égale à 1 sur le nombre de parties possibles, c'est-à-dire $\frac{1}{C_N^n}$.

Supposons maintenant qu'on s'intéresse seulement aux parties qui ont exactement k boules rouges. L'ensemble de ces parties correspond à exactement à $\{X = k\}$. Il nous suffit donc de compter le nombre de parties avec k boules rouges parmi les parties de n boules prises au sein des N boules. Pour déterminer une partie de n boules avec k boules rouges, il suffit d'abord de décider quelles boules, parmi les M rouges, seront utilisées ; il y a C_M^k tels choix. Il reste alors à déterminer lesquelles des $N-M$ boules noires seront les $n-k$ boules utilisées pour compléter la partie de n boules : il y a C_{N-M}^{n-k} tels choix. La probabilité $P(X = k)$ est donc bien celle annoncée ci-dessus. Nous sommes donc en mesure de donner la définition suivante, qui fait également office de théorème :

Définition

On dit qu'une variable aléatoire discrète X suit **la loi hypergéométrique de paramètres** $N \in \mathbb{N}^*$, $M \in \mathbb{N}^*$ et $n \in \mathbb{N}^*$, $\mathcal{H}(N, M, n)$, si l'ensemble des valeurs possibles est $X(\Omega) = \{0, 1, \dots, N\}$ et si, pour tout entier k de $\{0, 1, \dots, n\}$:

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$$

Cette loi est celle du nombre d'éléments de type \mathcal{T} dans un échantillon de taille n pris dans une population de N éléments comportant M éléments de type \mathcal{T} .

On note alors :

$$X \sim \mathcal{H}(N, M, n)$$



Comme l'indique la définition, cette loi est utile pour comprendre le nombre d'éléments d'un échantillon de taille n possédant une propriété donnée (le type \mathcal{T} de la définition), connaissant la proportion $\frac{M}{N}$ d'éléments ayant cette propriété, et la taille N de la population.

Ce qui rend cette distribution non triviale est le fait que le choix de l'échantillon soit réalisé sans remplacement ; ainsi, lorsqu'on tire successivement des éléments pour constituer l'échantillon, le type de chaque élément déjà tiré affecte les chances restantes de tirer des éléments de type \mathcal{T} . Les tirages ne sont pas indépendants.

Exemple

Un bol rempli de $N = 30$ cerises comporte six cerises moisisées, et donc $M = 24$ bonnes cerises. On prend une poignée de 5 cerises au hasard (en fermant les yeux et en évitant de sentir). Parmi ces $n = 5$ cerises, le nombre X de bonnes cerises a pour loi $\mathcal{H}(30, 24, 5)$. La probabilité que notre poignée contienne au moins 4 bonnes cerises est :

$$P(X \geq 4) = P(X = 4) + P(X = 5) = \frac{\binom{24}{4}\binom{6}{1}}{\binom{30}{5}} + \frac{\binom{24}{5}\binom{6}{0}}{\binom{30}{5}} = \frac{2530}{3393} \approx 0,7457$$

► Approximation binomiale de la loi hypergéométrique

On conçoit aisément que si N est beaucoup plus grand que n , et que $\frac{M}{N}$ n'est pas trop proche de 0 ou 1, alors les tirages seront approximativement indépendants, puisque le fait de retirer un petit nombre d'éléments d'une grande population a très peu d'influence sur la proportion d'éléments de type \mathcal{T} dans la population restante.

Si on considère alors qu'un succès correspond au tirage d'un élément de type \mathcal{T} , la probabilité d'un succès sera (approximativement) de $p = \frac{M}{N}$ à chaque tirage.

La variable aléatoire X sera donc, approximativement, binomiale de paramètres n et p .

Cette approximation n'est pas valable si p est proche de 0 ou de 1, ou si n n'est pas petit devant N .

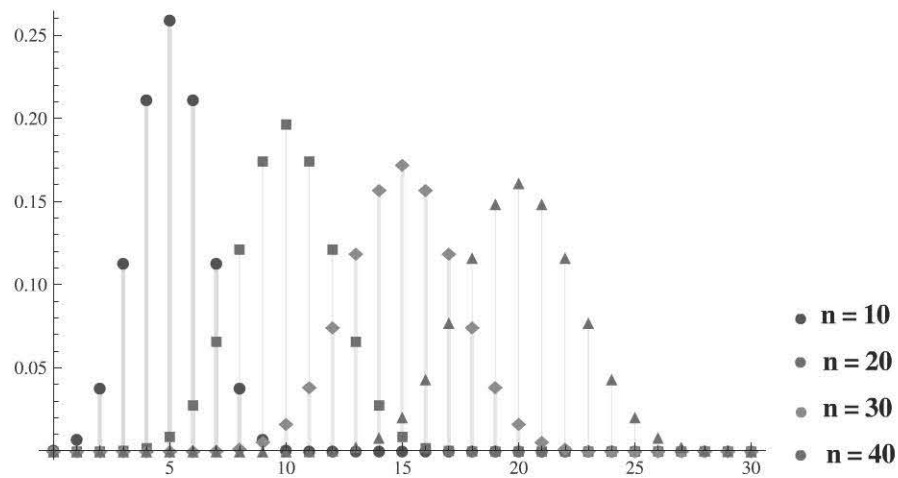


Figure 123.1 – Une représentation graphique de la fonction de probabilité pour la loi hypergéométrique, pour $n = 10$, $n = 20$, $n = 30$ et $n = 40$.

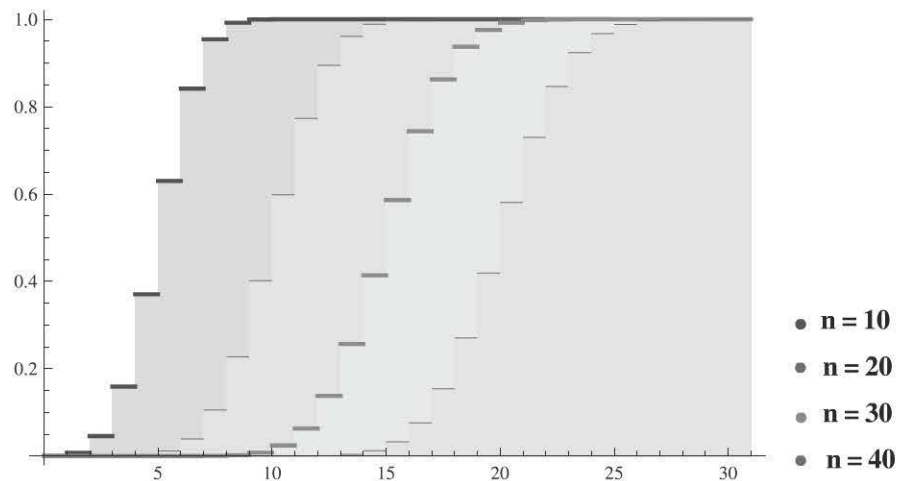


Figure 123.2 – Une représentation graphique de la fonction de répartition de la loi hypergéométrique, pour $n = 10$, $n = 20$, $n = 30$ et $n = 40$.

La loi de Poisson est utile pour compter le nombre d'événements qui ont lieu dans un intervalle de temps donné, si on connaît le nombre moyen d'événements par unité de temps. Cette loi est plus difficile à interpréter que les lois rencontrées précédemment faisant intervenir des variables de Bernoulli indépendantes, ou des tirages de parties.

Nous verrons, ultérieurement, que la somme de deux variables de Poisson indépendantes est encore une variable de Poisson. Nous avons déjà vu que ce n'est pas le cas pour les autres types de variables. Par exemple, une somme de variables de Bernoulli indépendantes est une variable binomiale, mais non de Bernoulli ; ou encore, une somme de variables géométriques indépendantes est une variable binomiale négative, mais non binomiale.

Cette propriété extraordinaire de stabilité par addition fait de la loi de Poisson un des fondements d'une branche importante du calcul moderne des probabilités, celle des processus stochastiques à sauts.

La seule autre loi que nous rencontrerons qui ait une propriété similaire de stabilité par addition est une loi non discrète, la loi normale, qui est le point de départ du théorème central limite, et de la théorie des processus stochastiques continus.

Définition

On dit qu'une variable aléatoire discrète X suit **la loi de Poisson de paramètre $\lambda > 0$** si l'ensemble des valeurs possibles est $X(\Omega) = \mathbb{N}$ et si, pour tout entier naturel k :

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

On note alors :

$$X \sim \mathcal{P}(\lambda)$$



Une façon de comprendre la loi de Poisson est d'imaginer un très grand nombre n de variables de Bernoulli, avec un paramètre $p = p_n$ qui dépend de n , tel que np_n reste constant. Ainsi, la probabilité de succès devient d'autant plus petite que n est grand ; plus précisément, elle est inversement proportionnelle au nombre n de tentatives. Le nombre de succès au sein des n variables de Bernoulli approche alors une loi de Poisson. En raison du lien entre les variables de Bernoulli et la loi binomiale, on obtient le théorème suivant, dont la démonstration donnée ci-dessous peut sembler très formelle ; une fois que l'on disposera de la notion d'espérance mathématique, il sera plus facile de comprendre, intuitivement, la relation entre n et p_n .

► Convergence de la loi binomiale vers la loi de Poisson

Théorème

Soient λ un réel strictement positif, et $(p_n)_{n \in \mathbb{N}^*}$ une suite de réels de $[0, 1]$ telle que :

$$\lim_{n \rightarrow +\infty} n p_n = \lambda$$

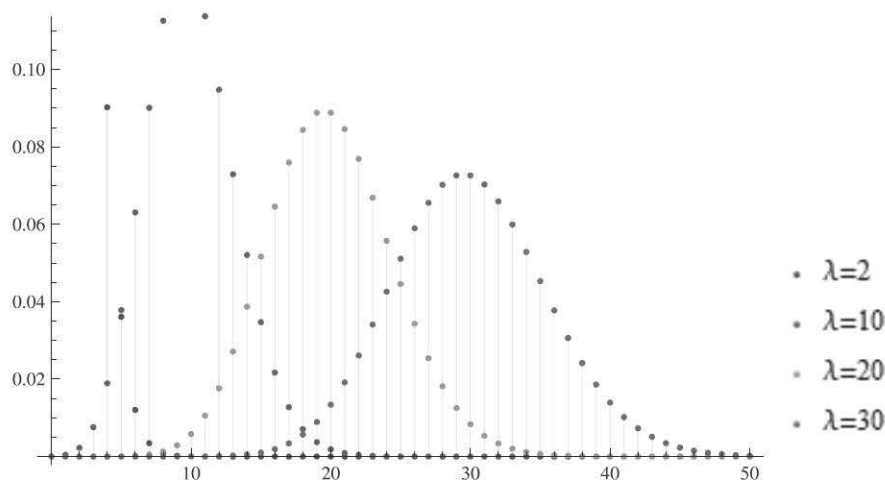


Figure 124.1 – Une représentation graphique de la fonction de probabilité pour la loi de Poisson, pour $\lambda = 2$, $\lambda = 10$, $\lambda = 20$ et $\lambda = 30$.

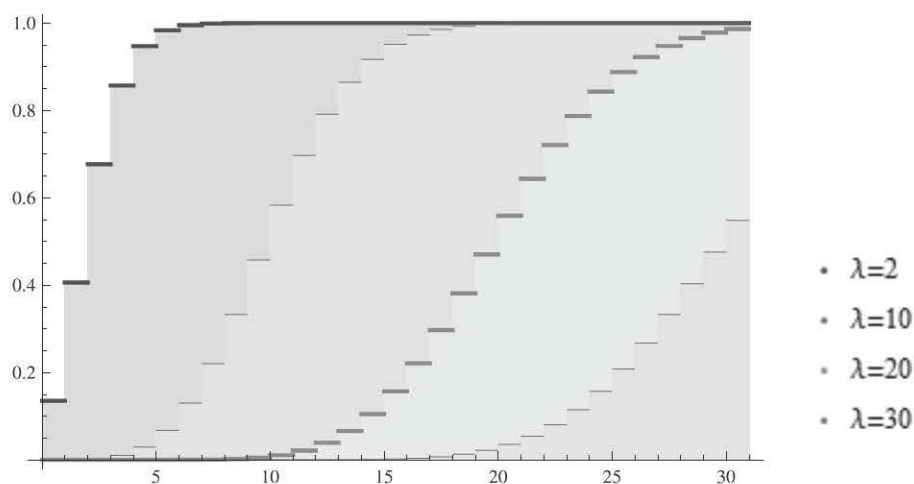


Figure 124.2 – Une représentation graphique de la fonction de répartition de la loi de Poisson, pour $\lambda = 2$, $\lambda = 10$, $\lambda = 20$ et $\lambda = 30$.

$(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires, de loi binomiale $\mathcal{B}(n, p_n)$, et X une variable aléatoire de loi de Poisson $\mathcal{P}(\lambda)$, alors, pour tout entier naturel k :

$$\lim_{n \rightarrow +\infty} P(X_n = k) = P(X = k)$$

Démonstration : Comme :

$$\lim_{n \rightarrow +\infty} n p_n = \lambda$$

on a donc, lorsque n tend vers $+\infty$: $n p_n = \lambda + o(1)$, et :

$$p_n = \frac{\lambda}{n} + o\left(\frac{1}{n}\right)$$

Par suite, pour tout entier naturel k :

$$\begin{aligned} \lim_{n \rightarrow +\infty} P(X_n = k) &= \lim_{n \rightarrow +\infty} C_n^k p_n^k (1 - p_n)^{n-k} \\ &= \lim_{n \rightarrow +\infty} \frac{n!}{k!(n-k)!} p_n^k (1 - p_n)^{n-k} \\ &= \lim_{n \rightarrow +\infty} \frac{n!}{k!(n-k)!} \lambda^k n^{-k} (1 + o(1))^k \left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^{n-k} \\ &= \lim_{n \rightarrow +\infty} \frac{n!}{k!(n-k)!} \lambda^k n^{-k} (1 + o(1)) e^{(n-k) \ln(1 - \frac{\lambda}{n} + o(\frac{1}{n}))} \\ &= \lim_{n \rightarrow +\infty} \frac{n!}{k!(n-k)!} \lambda^k n^{-k} (1 + o(1)) e^{-\frac{(n-k)\lambda}{n} + o(\frac{n-k}{n})} \\ &= \lim_{n \rightarrow +\infty} \frac{n!}{k!(n-k)!} \lambda^k n^{-k} (1 + o(1)) e^{-\lambda + \frac{k\lambda}{n} + o(\frac{n-k}{n})} \\ &= \lim_{n \rightarrow +\infty} \frac{n!}{k!(n-k)!} \lambda^k n^{-k} (1 + o(1)) e^{-\lambda} e^{\frac{k\lambda}{n} + o(\frac{n-k}{n})} \\ &= \lim_{n \rightarrow +\infty} \frac{n!}{k!(n-k)!} \lambda^k n^{-k} (1 + o(1)) e^{-\lambda} e^{o(1)} \\ &= \lim_{n \rightarrow +\infty} \frac{n!}{k!(n-k)!} \lambda^k n^{-k} e^{-\lambda} \\ &= \lim_{n \rightarrow +\infty} \frac{n!}{k!(n-k)!} \lambda^k e^{-\lambda} \\ &= \frac{\lambda^k e^{-\lambda}}{k!} \lim_{n \rightarrow +\infty} \frac{n!}{(n-k)! n^k} \\ &= \frac{\lambda^k e^{-\lambda}}{k!} \lim_{n \rightarrow +\infty} \frac{(n-k+1)(n-k+2) \dots n}{n^k} \\ &= \frac{\lambda^k e^{-\lambda}}{k!} \lim_{n \rightarrow +\infty} \left(1 + \frac{1-k}{n}\right) \left(1 + \frac{2-k}{n}\right) \dots (1 + 0) \\ &= \frac{\lambda^k e^{-\lambda}}{k!} \end{aligned}$$

■

Exemple

L'incidence d'une maladie génétique rare liée à l'appartenance au peuple Mapuche (le nom de ce peuple signifie « les Gens » (che) « de la Terre » (Mapu)) est estimée à 5 cas par million dans la population chilienne ; les racines Mapuche sont distribuées de façon très homogène dans les zones urbaines. La ville de Valparaíso, au Chili, compte 250 000 âmes. Si on considère que chaque habitant de Valparaíso a une chance $p = \frac{5}{10^6} = \frac{1}{200\,000}$ d'avoir la maladie, indépendamment des autres, le nombre X de personnes de Valparaíso qui ont cette maladie est une variable de loi $\mathcal{B}(n, p)$ avec $n = 250\,000$ et $p = \frac{1}{200\,000}$. On peut donc affirmer que cette loi est approximativement celle d'une variable de Poisson Y de paramètre $\lambda = np = \frac{5}{4}$. Ainsi, la probabilité qu'il y ait au moins une personne à Valparaíso avec cette maladie est approximativement égale à :

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - e^{-\frac{5}{4}} \approx 0,71.$$

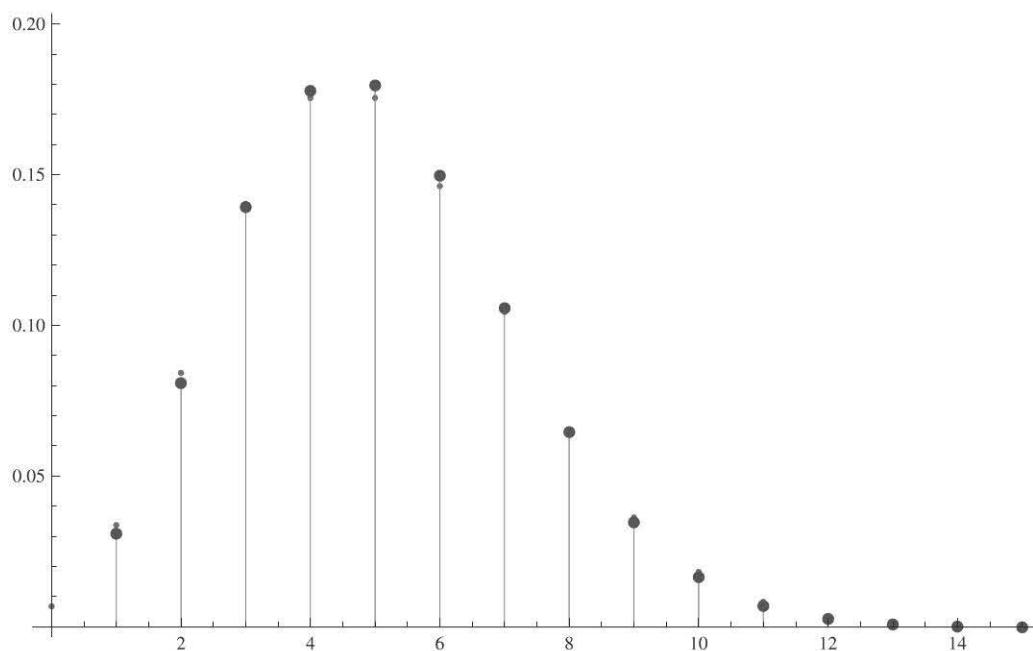


Figure 124.3 – Une illustration graphique de la convergence de la loi binomiale vers la loi de Poisson, dans le cas $\mathcal{B}(100, 0,05)$ et $\mathcal{P}(5)$.

Somme de variables aléatoires discrètes

1. Somme de deux variables aléatoires discrètes indépendantes

Étant données deux variables aléatoires discrètes indépendantes X et Y , à valeurs entières, la loi de la v.a. somme $X + Y$ est donnée, pour tout entier naturel k , par :

$$P(X + Y = k) = \sum_{i \in X(\Omega), j \in Y(\Omega), i+j=k} P(X = i) P(Y = j)$$

Démonstration : On calcule :

$$\begin{aligned} P(X + Y = k) &= P\left(\bigcup_{i \in X(\Omega), j \in Y(\Omega), i+j=k} \{X = i, Y = j\}\right) \\ &= \sum_{i \in X(\Omega), j \in Y(\Omega), i+j=k} P(X = i, Y = j) \\ &= \sum_{i+j=k} P(X = i) P(Y = j) \end{aligned}$$

(c'est grâce à l'indépendance des v.a. que $P(X = i, Y = j) = P(X = i) P(Y = j)$). ■

Exemple : Somme de deux v.a. indépendantes suivant respectivement les lois de Poisson $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$, $(\lambda, \mu) \in (\mathbb{R}_+^*)^2$

Soient λ et μ deux réels strictement positifs, et X et Y deux v.a. discrètes indépendantes suivant, respectivement, les lois de Poisson $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$.

La loi de la v.a. somme $X + Y = X_1 + \dots + X_m + Y_1 + \dots + Y_n$ est donc donnée, pour tout entier naturel k , par :

$$\begin{aligned} P(X + Y = k) &= \sum_{i \in \mathbb{N}, j \in \mathbb{N}, i+j=k} P(X = i, Y = j) \\ &= \sum_{i \in \mathbb{N}, j \in \mathbb{N}, i+j=k} P(X = i) P(Y = j) \\ &= \sum_{i=0}^k P(X = i) P(Y = k - i) \\ &= \sum_{i=0}^k \frac{e^{-\lambda} \lambda^i}{i!} \frac{e^{-\mu} \mu^{k-i}}{(k-i)!} \\ &= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{i=0}^k \frac{k! \lambda^i \mu^{k-i}}{i! (k-i)!} \\ &= \frac{e^{-(\lambda+\mu)}}{k!} \sum_{i=0}^k C_k^i \lambda^i \mu^{k-i} \\ &= \frac{e^{-(\lambda+\mu)}}{k!} (\lambda + \mu)^k \end{aligned}$$

On retrouve donc, pour la v.a. $X + Y$, la loi de Poisson $\mathcal{P}(\lambda + \mu)$.



Cette remarque permet de justifier certaines affirmations concernant la loi de Poisson. Supposons en effet, que, dans le cadre du nombre de clients qui pénètrent dans un magasin, l'on s'intéresse à la loi du nombre de clients qui arrivent pendant une période de n minutes, $n \in \mathbb{N}$. On considère que le nombre de clients qui arrivent au cours de toute période d'une minute suit une loi de Poisson de paramètre λ .

Supposons aussi que les variables de Poisson correspondant aux arrivées dans des périodes temporelles disjointes soient indépendantes. Cette dernière hypothèse se justifie grâce à un raisonnement utilisant une approximation binomiale. On trouve alors que le nombre d'arrivées

X au cours d'une période de n minutes vaut $X = \sum_{i=1}^n X_i$, où chaque X_i , $i = 1, \dots, n$, désigne le

nombre d'arrivées au cours de la minute numéro i , pendant la période des n minutes. Comme $X_i \sim \mathcal{P}(\lambda)$, on trouve bien, grâce au calcul de l'exemple précédent, appliqué itérativement, $X \sim \mathcal{P}(n\lambda)$. De surcroît, si l'on divise la période de longueur n minutes en deux sous-périodes disjointes de durées respectives n_1 et n_2 minutes ($n_1 \in \mathbb{N}$, $n_2 \in \mathbb{N}$, $n_1 + n_2 = n$), les nombres d'arrivées dans les deux sous-périodes seront indépendants, de lois de Poisson de paramètres respectifs $n_1 \lambda$ et $n_2 \lambda$.

2. Propriété

La discussion précédente peut être résumée à l'aide d'un théorème plus général, équivalent à la propriété d'*auto-indéfinie-divisibilité* de la loi de Poisson.

Théorème

Soit X une variable de Poisson de paramètre $\lambda > 0$. On suppose que la valeur de X correspond à des arrivées de deux types distincts, dits « type \mathcal{T}_1 » et « type \mathcal{T}_2 ». $p \in [0, 1]$ désigne la probabilité qu'une arrivée soit de type \mathcal{T}_1 , X_1 et X_2 sont, respectivement, les nombres d'arrivées de type \mathcal{T}_1 et \mathcal{T}_2 .

Alors, les v.a. X_1 et X_2 sont indépendantes. De plus,

$$X_1 \sim \mathcal{P}(\lambda_1) \quad \text{et} \quad X_2 \sim \mathcal{P}(\lambda_2)$$

avec $\lambda_1 = \lambda p$ et $\lambda_2 = \lambda (1 - p)$. En particulier, $X = X_1 + X_2$.

Démonstration : Il suffit de prendre la discussion ci-dessus comme modèle. ■



En pratique, il n'est pas toujours nécessaire de passer par la formule générale pour la loi d'une somme de variables indépendantes. C'est ainsi le cas lorsque les variables peuvent être considérées comme des sommes.

► Somme de deux v.a. indépendantes suivant respectivement les lois binomiales $\mathcal{B}(m, p)$ et $\mathcal{B}(n, p)$, $(m, n, p) \in (\mathbb{N}^*)^2 \times [0, 1]$

Soient m et n deux entiers naturels non nuls, p un réel de l'intervalle $[0, 1]$, et X et Y deux v.a. discrètes indépendantes suivant, respectivement, les lois binomiales $\mathcal{B}(m, p)$ et $\mathcal{B}(n, p)$.

Nous avons déjà vu que X peut être considéré comme somme de m v.a. indépendantes X_1, \dots, X_m , suivant chacune, la loi de Bernoulli $\mathcal{B}(p)$; de même, Y peut être considéré comme la somme de n v.a. indépendantes Y_1, \dots, Y_n , suivant chacune, la loi de Bernoulli

$\mathcal{B}(p)$. Les v.a. en jeu étant, toutes, indépendantes, la v.a. somme $X + Y = X_1 + \dots + X_m + Y_1 + \dots + X_n$ suit donc la loi binomiale $\mathcal{B}(m + n, p)$.

Ce résultat peut être retrouvé par un calcul direct (à partir de la formule générale pour les sommes de variables indépendantes, que nous ne présentons pas ici), grâce à la formule dite de Vandermonde :

$$\sum_{\substack{1 \leq i \leq m, 1 \leq j \leq n \\ i+j=k}} C_m^i C_n^j = C_{m+n}^k$$

► **Somme de deux v.a. indépendantes binomiales négatives de même paramètre $p \in]0, 1[$**

Un raisonnement analogue au précédent montre que la somme de deux v.a. indépendantes de lois binomiales négatives $\mathcal{BN}(n, p)$ et $\mathcal{BN}(m, p)$ est une v.a. qui suit également la loi binomiale négative $\mathcal{BN}(n + m, p)$. Il suffit en effet de considérer chacune des deux premières variables comme somme de variables géométriques indépendantes, de paramètre p .



Dans les deux exemples précédents, le fait que les variables que l'on somme aient le même paramètre (probabilité de succès) p est crucial. La somme de deux variables $\mathcal{N}(n, p)$ et $\mathcal{N}(m, q)$ lorsque $p \neq q$ ne peut être reliée à des lois classiques.

Nous avons déjà évoqué la notion d'espérance pour des v.a., lorsque l'on a considéré la valeur moyenne « intuitive » d'une variable. L'espérance, ou espérance mathématique, formalise ce concept intuitif.

Définition

Soit X une variable aléatoire discrète telle que la famille $(|x_k| P(X = x_k))_{x_k \in X(\Omega)}$ soit sommable, i.e.

$$\sum_{x_k \in X(\Omega)} |x_k| P(X = x_k) < +\infty$$

On appelle **espérance mathématique de X** le réel $\mathbb{E}(X)$ tel que :

$$\mathbb{E}(X) = \sum_{x_k \in X(\Omega)} x_k P(X = x_k)$$



1. L'espérance mathématique peut donc être considérée comme le barycentre des valeurs possibles de la variable aléatoire X , pondérées par leurs probabilités de réalisation. Ce concept fonctionne encore si X est une v.a. vectorielle. On se limitera ici au cas scalaire.
2. Du fait de la condition d'absolue sommabilité dans la définition de l'espérance, on n'a pas le droit de parler d'espérance si la somme $\sum_{x_k \in X(\Omega)} x_k P(X = x_k)$ est semi-convergente et non absolument convergente. Le lecteur vérifiera qu'un exemple d'une telle situation non autorisée est le suivant :

$$x_k = (-1)^k k \quad \text{et} \quad P(X = x_k) = 6 k^{-2} \pi^{-2}$$

1. Propriétés

► Espérance d'une v.a. constante

Soit X une variable aléatoire discrète constante, prenant la valeur $C > 0$. Alors :

$$\mathbb{E}(X) = C$$

Démonstration :

$$\mathbb{E}(X) = \sum_{x_k \in X(\Omega)} x_k P(X = x_k) = \sum_{x_k \in X(\Omega)} C P(X = x_k) = C P(X = C) = C$$

(puisque $X(\Omega) = \{C\}$). ■

► Espérance et valeur absolue

Soit X une variable aléatoire discrète admettant une espérance $\mathbb{E}(X)$. Alors :

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|) < +\infty$$

► Linéarité de l'espérance

Soient X et Y deux variables aléatoires discrètes sur un espace probabilisé, admettant chacune une espérance mathématique ($\mathbb{E}(X)$ et $\mathbb{E}(Y)$ respectivement). Alors :

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

et, pour tout réel a :

$$\mathbb{E}(aX) = a\mathbb{E}(X).$$

Plus généralement, si $(X_n)_{n \in \mathbb{N}}$ est une suite de variables aléatoires sur un espace probabilisé, admettant chacune une espérance, et si $(a_n)_{n \in \mathbb{N}}$ est une suite de réels, alors pour tout entier $N \geq 0$,

$$\mathbb{E} \left[\sum_{i=1}^N a_i X_i \right] = \sum_{i=1}^N a_i \mathbb{E}[X_i]$$

Si la famille $|a_i| \mathbb{E}[|X_i|]$ est sommable, la formule précédente reste encore valable en passant à la limite lorsque N tend vers $+\infty$.

Démonstration : Considérons la variable aléatoire $Z = X + Y$. On choisit, dans un premier temps, de s'intéresser au cas où $X(\Omega)$ et $Y(\Omega)$ sont finis. $Z(\Omega)$ est donc aussi fini, et l'existence de l'espérance mathématique de Z est assurée.

La loi de Z est donnée par :

$$\forall z \in Z(\Omega) : P(Z = z_k) = \sum_{x_i + y_j = z_k} P(X = x_i, Y = y_j)$$

Il en résulte :

$$\begin{aligned} \mathbb{E}(Z) &= \sum_{z_k \in Z(\Omega)} z_k P(Z = z_k) \\ &= \sum_{z_k \in Z(\Omega)} \sum_{x_i + y_j = z_k} (x_i + y_j) P(X = x_i, Y = y_j) \\ &= \sum_{x_i \in X(\Omega)} \sum_{y_j \in Y(\Omega)} (x_i + y_j) P(X = x_i, Y = y_j) \\ &= \sum_{x_i \in X(\Omega)} x_i \left\{ \sum_{y_j \in Y(\Omega)} P(X = x_i, Y = y_j) \right\} \\ &\quad + \sum_{y_j \in Y(\Omega)} y_j \left\{ \sum_{x_i \in X(\Omega)} P(X = x_i, Y = y_j) \right\} \\ &= \sum_i x_i P(X = x_i) + \sum_j y_j P(Y = y_j) \\ &= \mathbb{E}(X) + \mathbb{E}(Y) \end{aligned}$$

(On passe sans souci de

$$\sum_{z_k \in Z(\Omega)} \sum_{x_i + y_j = z_k} (x_i + y_j) P(X = x_i, Y = y_j)$$

à :

$$\sum_{x_i \in X(\Omega)} \sum_{y_j \in Y(\Omega)} (x_i + y_j) P(X = x_i, Y = y_j)$$

car dire que z_k parcourt $Z(\Omega)$ revient à sommer toutes les valeurs possibles de X et de Y ; la v.a. Z a, en effet, été définie directement comme somme des v.a. X et Y .)

La convergence absolue étant assurée, on peut sommer « par paquets » et réarranger :

$$\begin{aligned} \mathbb{E}(Z) &= \sum_{z_k \in Z(\Omega)} z_k P(Z = z_k) \\ &= \sum_{z_k \in Z(\Omega)} \sum_{x_i + y_j = z_k} (x_i + y_j) P(X = x_i, Y = y_j) \\ &= \sum_{x_i \in X(\Omega)} \sum_{y_j \in Y(\Omega)} (x_i + y_j) P(X = x_i, Y = y_j) \\ &= \sum_{x_i \in X(\Omega)} x_i \left\{ \sum_{y_j \in Y(\Omega)} P(X = x_i, Y = y_j) \right\} \\ &\quad + \sum_{y_j \in Y(\Omega)} y_j \left\{ \sum_{x_i \in X(\Omega)} P(X = x_i, Y = y_j) \right\} \\ &= \sum_i x_i P(X = x_i) + \sum_j y_j P(Y = y_j) \\ &= \mathbb{E}(X) + \mathbb{E}(Y) \end{aligned}$$

Les assertions relatives aux suites de variables X_i et de coefficients a_i se démontrent grâce aux propriétés précédentes par itération. ■

► Positivité de l'espérance

Soit X une variable aléatoire discrète, à valeurs positives, admettant une espérance mathématique $\mathbb{E}(X)$. Alors :

$$\mathbb{E}(X) \geq 0$$

► Croissance de l'espérance

Soient X et Y deux variables aléatoires discrètes telles que $X \leq Y$, admettant chacune une espérance mathématique ($\mathbb{E}(X)$ et $\mathbb{E}(Y)$ respectivement). Alors :

$$\mathbb{E}(X) \leq \mathbb{E}(Y)$$

Théorème de transfert

Soit X une variable aléatoire discrète, et f une fonction définie sur $X(\Omega)$, à valeurs réelles. Alors, la v.a. $f(X)$, définie par :

$$f(X)(\omega) = f(X(\omega)) \quad \forall \omega \in \Omega$$

est d'espérance finie si et seulement si la famille $(f(x_k) P(X = x_k))_{x_k \in X(\Omega)}$ est sommable. On a alors :

$$\mathbb{E}[f(X)] = \sum_{x_k \in X(\Omega)} f(x_k) P(X = x_k)$$

2. Calcul des espérances

Dans ce qui suit, on calcule les espérances mathématiques des v.a. classiques présentées précédemment. Pour celles qui peuvent être considérées comme des sommes d'autres v.a., le calcul sera particulièrement aisé.

► Calcul de l'espérance dans le cas de la loi de Bernoulli

Considérons une variable aléatoire X suivant la loi de Bernoulli de paramètre p . Alors :

$$\mathbb{E}(X) = 0 \times P(X = 0) + 1 \times P(X = 1) = 0 \times (1 - p) + p = p$$

► Calcul de l'espérance dans le cas de la loi uniforme

Considérons une variable aléatoire X suivant la loi uniforme sur l'ensemble fini $\{1, \dots, n\}$. Alors :

$$\mathbb{E}(X) = \sum_{k=1}^n x_k P(X = x_k) = \sum_{k=1}^n \frac{k}{n} = \frac{1}{n} \sum_{k=1}^n k = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}$$

► Calcul de l'espérance dans le cas de la loi binomiale $\mathcal{B}(n, p)$

Considérons une variable aléatoire X suivant la loi binomiale $\mathcal{B}(n, p)$. On a alors $X(\Omega) = \{0, 1, \dots, n\}$. Étant donné que X ne prend qu'un nombre fini de valeurs, on sait qu'elle admet une espérance. X peut s'exprimer comme $\sum_{i=1}^n X_i$, où X_1, \dots, X_n sont des v.a. de Bernoulli de paramètre p indépendantes. Elles vérifient donc, pour tout entier i de $\{1, \dots, n\}$, $\mathbb{E}(X_i) = p$. Sans même utiliser la propriété d'indépendance, et grâce, tout simplement, à la linéarité, on obtient immédiatement :

$$\mathbb{E}(X) = \mathbb{E}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbb{E}(X_i) = np.$$

Le lecteur vérifiera qu'on retrouve le même résultat, au prix d'un calcul bien plus complexe, si on part de la formule $P(X = k) = C_n^k p^k (1-p)^{n-k}$ et de la définition de l'espérance.

► **Calcul de l'espérance dans le cas de la loi géométrique de paramètre $p > 0$**

Considérons une variable aléatoire X suivant la loi géométrique de paramètre $p \in]0, 1[$. On a alors $X(\Omega) = \mathbb{N}$, et, pour tout entier naturel non nul k :

$$P(X = k) = p(1 - p)^{k-1}$$

Comme $0 \leq p < 1$, la série de terme général $k P(X = k) = k p (1 - p)^{k-1}$ est clairement convergente (comme série dérivée d'une série géométrique convergente ; on peut également invoquer le fait, vérifiable par le lecteur grâce à une simple étude de fonctions, que pour k suffisamment grand,

$$k(1 - p)^{k-1} \leq \left(1 - \frac{p}{2}\right)^k$$

où $\left(1 - \frac{p}{2}\right)^k$ est le terme général d'une série géométrique convergente ; on peut, aussi, utiliser le critère de d'Alembert).

Il en résulte :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=1}^{+\infty} k P(X = k) \\ &= \sum_{k=1}^{+\infty} k p (1 - p)^{k-1} \\ &= p \sum_{k=1}^{+\infty} \left[\frac{d x^k}{d x} \right]_{x=1-p} \\ &= p \left[\frac{d}{d x} \left(\sum_{k=1}^{+\infty} x^k \right) \right]_{x=1-p} \\ &= p \left[\frac{d}{d x} \left(\frac{x}{1-x} \right) \right]_{x=1-p} \\ &= p \left[\frac{d}{d x} \left(\frac{x-1+1}{1-x} \right) \right]_{x=1-p} \\ &= p \left[\frac{d}{d x} \left(\frac{1}{1-x} \right) \right]_{x=1-p} \\ &= p \left[\frac{1}{(1-x)^2} \right]_{x=1-p} \\ &= p \left[\frac{1}{p^2} \right]_{x=1-p} \\ &= \frac{1}{p} \end{aligned}$$

(On applique le théorème de dérivation terme à terme des séries entières, puisque l'on est bien à l'intérieur du disque de convergence dont le rayon est 1.)

► **Calcul de l'espérance dans le cas de la loi binomiale négative de paramètres $p \in]0, 1[$, et $n \in \mathbb{N}^*$**

Considérons une variable aléatoire X suivant la loi binomiale négative de paramètre $p \in]0, 1[$ et $n \in \mathbb{N}^*$. On a alors $X(\Omega) = \mathbb{N}$. X peut aussi être considérée comme la

somme $\sum_{i=1}^n X_i$, où X_1, \dots, X_n sont des v.a. géométriques indépendantes, de paramètre p . Elles vérifient donc, pour tout entier i de $\{1, \dots, n\}$, $\mathbb{E}(X_i) = p^{-1}$. Par simple linéarité, sans utiliser l'indépendance, on obtient immédiatement :

$$\mathbb{E}(X) = \mathbb{E}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbb{E}(X_i) = \frac{n}{p}.$$

Le lecteur vérifiera qu'on retrouve le même résultat, au prix d'un calcul beaucoup plus complexe, si on part de la formule

$$P(X = k) = \binom{k-1}{n-1} p^n (1-p)^{k-n}$$

et de la définition de l'espérance.

► Calcul de l'espérance dans le cas de la loi de Poisson de paramètre $\lambda \in \mathbb{R}$

Considérons une variable aléatoire X suivant la loi de Poisson de paramètre $\lambda \in \mathbb{R}$. On a alors $X(\Omega) = \mathbb{N}$, et, pour tout entier naturel non nul k : $P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$.

La série de terme général $k P(X = k) = k \frac{e^{-\lambda} \lambda^k}{k!}$ est clairement convergente. On peut, par exemple, utiliser le critère de d'Alembert, puisque :

$$\lim_{k \rightarrow +\infty} \frac{(k+1) \frac{e^{-\lambda} \lambda^{k+1}}{(k+1)!}}{k \frac{e^{-\lambda} \lambda^k}{k!}} = \lim_{k \rightarrow +\infty} \frac{\frac{\lambda}{k!}}{\frac{1}{(k-1)!}} = \lim_{k \rightarrow +\infty} \frac{\lambda}{k} = 0$$

On peut donc calculer :

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=1}^{+\infty} k P(X = k) \\ &= \sum_{k=1}^{+\infty} k \frac{e^{-\lambda} \lambda^k}{k!} \\ &= \sum_{k=1}^{+\infty} \frac{e^{-\lambda} \lambda^k}{(k-1)!} \\ &= \sum_{k=0}^{+\infty} \frac{e^{-\lambda} \lambda^{k+1}}{k!} \\ &= \lambda e^{-\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} \\ &= \lambda e^{-\lambda} e^{\lambda} \\ &= \lambda \end{aligned}$$

Moment d'ordre $r, r \in \mathbb{N}^*$, d'une variable aléatoire discrète

Définition

Soient r un entier naturel non nul, et X une variable aléatoire discrète. On appelle **moment d'ordre r de la variable aléatoire X** la quantité, notée $\mathbb{E}(X^r)$, donnée par :

$$\mathbb{E}(X^r) = \sum_{x_k \in X(\Omega)} (x_k)^r P(X = x_k)$$

sous réserve que $\mathbb{E}(|X|^r) = \sum_{x_k \in X(\Omega)} |x_k|^r P(X = x_k)$ soit fini.

1. Existence du moment d'ordre

Si une variable aléatoire X admet un moment d'ordre $r, r \in \mathbb{N}^*$, alors, pour tout entier positif $n \leq r$, le moment d'ordre n , $\mathbb{E}(X^n)$, existe aussi.

Démonstration : La série de terme général $|x_k|^n P(X = x_k)$ est une série positive.

Considérons, dans un premier temps, le cas d'un univers fini (comportant N éléments, $N \in \mathbb{N}^*$). Une sommation « par paquets » conduit à :

$$\begin{aligned} \mathbb{E}(|X|^n) &= \sum_{x_k \in X(\Omega)} |x_k|^n P(X = x_k) \\ &= \sum_{x_k \in X(\Omega), |x_k| \leq 1} |x_k|^n P(X = x_k) + \sum_{x_k \in X(\Omega), |x_k| > 1} |x_k|^n P(X = x_k) \\ &\leq \sum_{x_k \in X(\Omega), |x_k| \leq 1} P(X = x_k) + \sum_{x_k \in X(\Omega), |x_k| > 1} |x_k|^n P(X = x_k) \\ &\leq P(X \leq 1) + \sum_{x_k \in X(\Omega), |x_k| > 1} |x_k|^r P(X = x_k) \\ &\leq 1 + \mathbb{E}(X^r) \\ &< +\infty \end{aligned}$$

Ceci étant vrai pour toute somme finie comportant N termes, on en déduit, par passage à la limite lorsque N tend vers $+\infty$, le résultat dans le cas général. ■

Cette propriété est aussi une conséquence de la fameuse inégalité de Jensen (prononcer « Yennsenne » ; nous ne la démontrerons pas), qui dit que, pour toute variable aléatoire positive Y admettant une espérance, et toute fonction ϕ convexe positive sur $[0, +\infty[$, alors :

$$\phi(\mathbb{E}(Y)) \leq \mathbb{E}[\phi(Y)]$$

Il suffit en effet d'appliquer cette inégalité à $Y = |X|^n$ et à la fonction $y \mapsto \phi(y) = y^{\frac{r}{n}}$, qui est convexe puisque $\frac{r}{n} \geq 1$.



2. Inégalité de Markov

Soit X une variable aléatoire positive admettant une espérance mathématique. Alors, pour tout réel $t > 0$:

$$P(X \geq t) \leq \frac{\mathbb{E}(X)}{t}.$$

Démonstration : Il suffit de sommer « par paquets » :

$$\mathbb{E}(X) = \sum_{x_k \in X(\Omega)} x_k P(X = x_k) = \sum_{x_k \in X(\Omega), x_k < t} x_k P(X = x_k) + \sum_{x_k \in X(\Omega), x_k \geq t} x_k P(X = x_k)$$

Il en résulte, puisque la v.a. X est à valeurs positives :

$$\begin{aligned} \mathbb{E}(X) &\geq \sum_{x_k \in X(\Omega), x_k \geq t} x_k P(X = x_k) \\ &\geq \sum_{x_k \in X(\Omega), x_k \geq t} t P(X = x_k) \\ &= t \sum_{x_k \in X(\Omega), x_k \geq t} P(X = x_k) \\ &= t P(X \geq t) \end{aligned}$$

ce qui conduit directement au résultat. ■



Cette inégalité est utile pour estimer la vitesse de décroissance vers 0 de la queue $P(X \geq t)$ d'une v.a. X lorsque t tend vers l'infini. Elle signifie que la décroissance doit être au moins de l'ordre de t^{-1} . Si des moments d'ordre plus grand que 1 existent, on obtient une vitesse plus rapide, comme l'indique le corollaire suivant, qui s'obtient immédiatement en appliquant l'inégalité de Markov à l'événement $\{X \geq t\} = \{X^r \geq t^r\}$.

Corollaire

Soit X une variable aléatoire positive admettant un moment d'ordre $r \in]0, +\infty[$. Alors, pour tout réel $t > 0$:

$$P(X \geq t) \leq \frac{\mathbb{E}(X^r)}{t^r}.$$



L'inégalité de Markov, ainsi que son corollaire, sont particulièrement utiles lorsqu'on ne connaît que très peu de choses sur la loi d'une v.a. X , par exemple seulement l'existence d'un moment donné. Pour les lois classiques, il est préférable d'utiliser des estimations plus directes.

1. Dans l'attente d'un premier client rue de la Paix

Supposons que le temps d'attente X pour l'arrivée du premier client après l'ouverture, dans un magasin de luxe rue de la Paix, mesuré en minutes entières, suive la loi géométrique $\mathcal{G}(p)$, avec $p = \frac{1}{10}$. Le temps d'attente moyen est donc de $\mathbb{E}(X) = \frac{1}{p} = 10$ minutes. Mais qu'en est-il de la probabilité qu'il faille attendre plus de dix minutes, plus de vingt minutes, ou plus d'une heure, pour qu'arrive le premier client ? Pour ces deux probabilités, le calcul direct donne :

$$P(X > 10) = \left(\frac{9}{10}\right)^{10} \approx 0,3487,$$

$$P(X > 20) = \left(\frac{9}{10}\right)^{20} \approx 0,1216,$$

$$P(X > 60) = \left(\frac{9}{10}\right)^{60} \approx 0,001797.$$

Supposons maintenant qu'on n'ait aucune raison de croire que la loi de X soit $\mathcal{G}(p)$, par exemple, à cause du fait qu'on doute fort que la propriété d'absence de mémoire soit vraie. Une raison pour un tel doute peut être la suivante : on a observé que si on a attendu longtemps le premier client, c'est le signe d'une mauvaise journée, et alors, en raison de cette attente, le temps d'attente restant moyen devient nettement plus élevé que 10 minutes. Néanmoins, on est toujours d'accord pour dire que le temps d'attente moyen, en l'absence d'autres informations, est de 10 minutes. On peut alors utiliser l'inégalité de Markov, puisque $\mathbb{E}(X) = 10$, ce qui permet d'affirmer :

$$P(X > 10) \leq \frac{10}{10} = 1,$$

$$P(X > 20) \leq \frac{10}{20} = 0,5,$$

$$P(X > 60) \leq \frac{10}{60} = 0,1666 \dots$$

Il est intéressant de remarquer que si X suivait la loi $\mathcal{G}(p)$, ces estimations deviendraient de plus en plus mauvaises au fur et à mesure que l'on s'approche du comportement asymptotique de la queue de X .

Poursuivons encore un peu l'analyse. Nous verrons, par la suite, que le second moment de la loi $\mathcal{G}(p)$ se calcule aussi, et vaut :

$$\mathbb{E}(X^2) = \frac{2-p}{p^2}$$

Dans notre exemple, ceci donne $\mathbb{E}(X^2) = 100(2 - 10^{-1}) = 190$. Imaginons que l'on trouve, empiriquement, que nous sommes d'accord avec cette valeur. Si nous ne croyons toujours pas que X est géométrique, et en l'absence de toute autre information en dehors du second moment de X , on peut toujours utiliser le corollaire de l'inégalité de Markov, appliquée au second moment de X :

$$\begin{aligned} P(X > 10) &\leq \frac{19}{10} = 1,9, \\ P(X > 20) &\leq \frac{19}{40} = 0,475, \\ P(X > 60) &\leq \frac{19}{360} \approx 0,05278. \end{aligned}$$

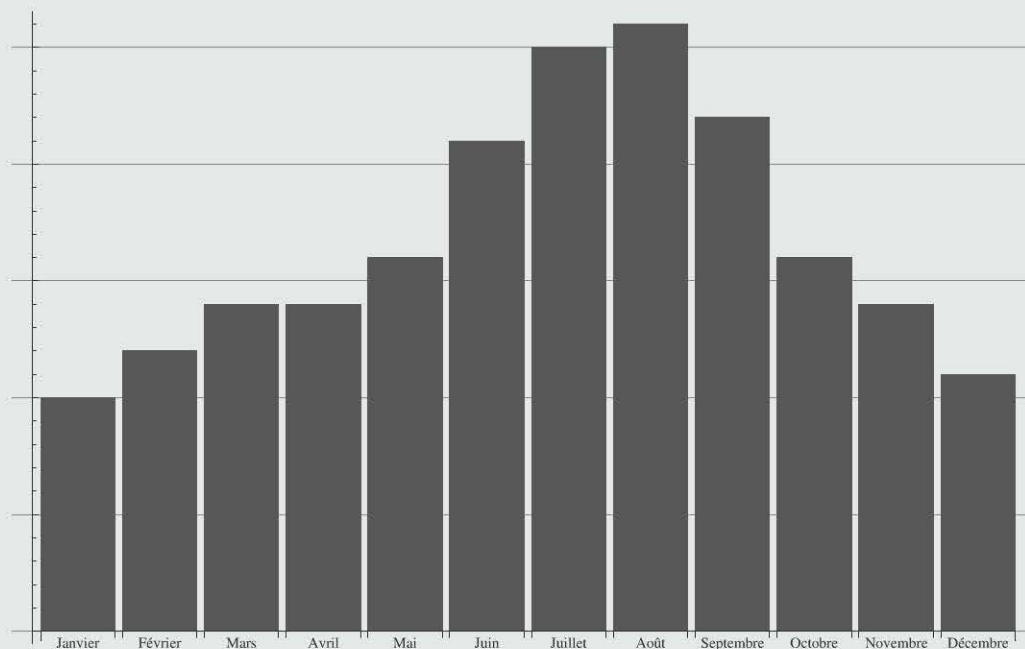
On constate que la première estimation est moins bonne qu'avec le premier moment (elle est même stupide...), mais les deux autres sont meilleures.

La situation de l'exemple ci-dessus est typique de l'usage délicat de l'inégalité de Markov : lorsque t n'est pas très grand, il est préférable de se servir du premier moment, mais lorsque t devient très grand, il devient progressivement avantageux d'utiliser des moments d'ordre plus élevés. Dans tous les cas, si on a plus d'informations que celles données par les valeurs de quelques moments uniquement, il faut éviter de recourir à cette inégalité.

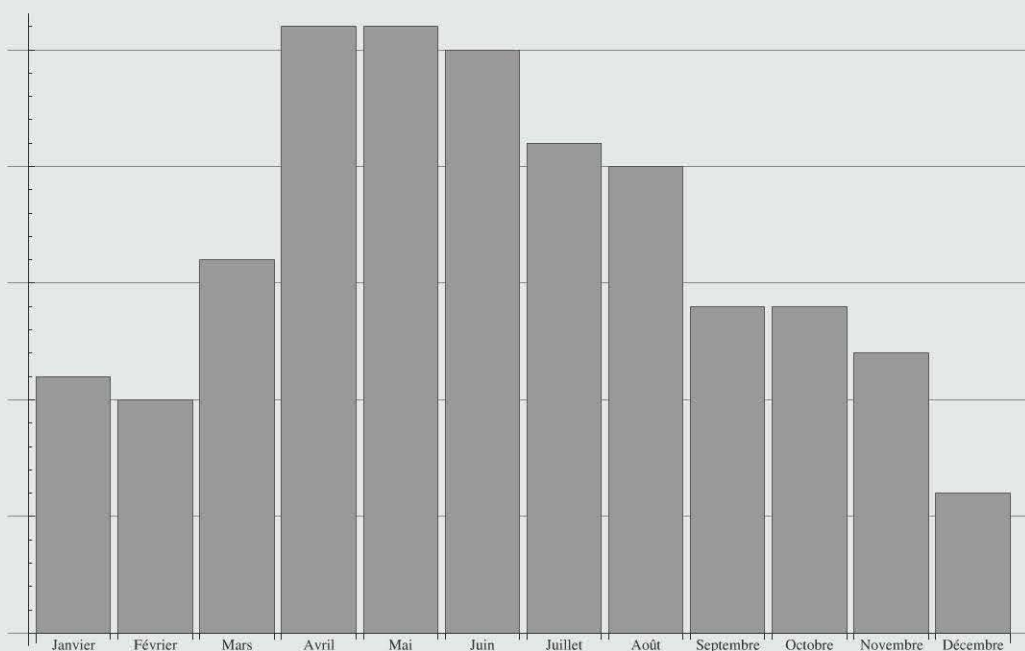


FOCUS Variations de température

Considérons la répartition des températures moyennes mensuelles observées, respectivement, dans une ville \mathcal{V}_1 , et une ville \mathcal{V}_2 . On dispose donc de deux séries statistiques dont les individus sont les mois de l'année, et la variable la température mensuelle.



La répartition des températures dans la ville \mathcal{V}_1 .



La répartition des températures dans la ville \mathcal{V}_2 .

En supposant, pour simplifier, que les douze mois ont tous la même durée, on calcule, grâce aux histogrammes, que les moyennes annuelles des températures dans les deux villes sont exactement les mêmes :

$$\frac{67}{4} \simeq 16,75$$

Par contre, la répartition des températures dans les deux villes n'est pas du tout la même ! Un outil mathématique permettant de mesurer la « dispersion » autour des moyennes annuelles est donc indispensable.

Calculons, pour la ville \mathcal{V}_1 , la quantité :

$$\frac{\sum_{\text{Premier mois}}^{\text{Dernier mois}} (\text{Température du mois} - \text{Moyenne annuelle})^2}{\text{Nombre de mois}} \simeq 27,020$$

Faisons le même calcul pour la ville \mathcal{V}_2 :

$$\frac{\sum_{\text{Premier mois}}^{\text{Dernier mois}} (\text{Température du mois} - \text{Moyenne annuelle})^2}{\text{Nombre de mois}} \simeq 87,64$$

On s'aperçoit que ces grandeurs sont très différentes ! Elles permettent de mesurer la dispersion des températures autour de la moyenne annuelle.

Toutefois, ces grandeurs se mesurent en degrés au carré. Il est justifiable de considérer qu'on peut préférer une quantité qui permette de mesurer la dispersion dans les mêmes unités que les séries de départ. Une telle quantité pourrait s'interpréter comme l'écart moyen entre la température mensuelle et la température moyenne. La convention adoptée en probabilités et en statistiques est de considérer tout simplement la racine carrée des grandeurs précédentes ; on obtient, respectivement, environ 5,19816 pour la première ville, et 9,36166 pour la seconde. Ceci permet toujours d'affirmer que les températures de la seconde ville sont plus dispersées que celles de la première, mais aussi de dire que les valeurs obtenues en prenant la racine carrée sont des mesures des écarts moyens. Si on se souvient que les formules données avant de prendre les racines carrées font intervenir des moyennes de carrés des écarts, on s'aperçoit que la terminologie « écarts moyens » n'est pas tout à fait représentative de la réalité des calculs ; il s'agit, plus précisément, de « racine carrée de la moyenne des carrés des écarts ». On pourrait définir une autre quantité qui correspondrait de façon plus précise à des « écarts moyens » dans la bonne unité : la moyenne de la valeur absolue des écarts, c'est-à-dire en l'occurrence, la quantité :

$$\frac{\sum_{\text{Premier mois}}^{\text{Dernier mois}} |\text{Température du mois} - \text{Moyenne annuelle}|}{\text{Nombre de mois}}.$$

Il se trouve que cette définition est bien moins pratique, mathématiquement, que celle faisant intervenir la racine de la moyenne des carrés des écarts. Nous verrons, en particulier, une propriété d'additivité des dispersions définies avec les carrés, très utile pour comprendre la dispersion des sommes de variables aléatoires indépendantes, qui devient fautive si on remplace le carré par la valeur absolue.

1. Définitions

► Variance

Soit X une variable aléatoire discrète admettant un moment d'ordre 2. On appelle **variance de la variable aléatoire** X la quantité, notée $Var(X)$, donnée par :

$$Var(X) = \mathbb{E}[(X - \mathbb{E}(X))^2]$$

► Écart-type

Soit X une variable aléatoire discrète admettant un moment d'ordre 2. On appelle **écart-type de la variable aléatoire** X la quantité, notée $\sigma(X)$, donnée par :

$$\sigma(X) = \sqrt{Var(X)}$$

2. Propriétés

► Changement d'échelle

Soit X une variable aléatoire réelle admettant un moment d'ordre 2. Alors, pour tout réel λ :

$$Var(\lambda X) = \lambda^2 Var(X)$$

Démonstration : On a :

$$\begin{aligned} Var(\lambda X) &= \mathbb{E}[(\lambda X - \mathbb{E}(\lambda X))^2] \\ &= \mathbb{E}[(\lambda X - \lambda \mathbb{E}(X))^2] \quad (\text{linéarité de l'espérance}) \\ &= \mathbb{E}[\lambda^2 (X - \mathbb{E}(X))^2] \\ &= \lambda^2 \mathbb{E}[(X - \mathbb{E}(X))^2] \quad (\text{linéarité de l'espérance}) \\ &= \lambda^2 Var(X) \end{aligned}$$

► Invariance par translation

Soit X une variable aléatoire réelle admettant un moment d'ordre 2. Alors, pour tout réel μ :

$$Var(X + \mu) = Var(X)$$

Démonstration : On a :

$$\begin{aligned} Var(X + \mu) &= \mathbb{E}[(X + \mu - \mathbb{E}(X + \mu))^2] \\ &= \mathbb{E}[(X + \mu - \mathbb{E}(X) - \mu)^2] \quad (\text{espérance d'une v.a. constante}) \\ &= \mathbb{E}[(X - \mathbb{E}(X))^2] \\ &= Var(X) \end{aligned}$$

Corollaire Variance d'une v.a. constante

Soit X une variable aléatoire réelle constante. Sa variance est nulle :

$$\text{Var}(X) = 0$$

Démonstration : Il suffit d'appliquer la formule d'invariance par translation ; pour tout réel μ :

$$\text{Var}(0 + \mu) = \text{Var}(0) = \mathbb{E}[(0 - \mathbb{E}(0))^2] = \mathbb{E}(0) = 0 \quad \blacksquare$$

► Variance nulle

Soit X une variable aléatoire réelle. Alors :

$$\text{Var}(X) = 0 \Leftrightarrow X = \mathbb{E}(X) \text{ avec une probabilité de } 1 \Leftrightarrow X \text{ est constante avec une probabilité de } 1$$

Démonstration : Il est clair que si X est une v.a. constante, alors son espérance est égale à elle-même ; sa variance est donc nulle.

Réciproquement, si X est une v.a. de variance nulle :

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}(X))^2] = \sum_{x_k \in X(\Omega)} (x_k - \mathbb{E}(X))^2 P(X = x_k) = 0$$

Tous les termes de la somme étant positifs ou nuls, il en résulte, pour tout x_k de $X(\Omega)$:

$$(x_k - \mathbb{E}(X))^2 P(X = x_k) = 0$$

Par suite, pour toutes les valeurs $x_k \neq \mathbb{E}(X)$:

$$P(X = x_k) = 0$$

Il reste :

$$P(X = \mathbb{E}(X)) = 1$$

qui est le résultat cherché. ■

► Formule de Koenig

Soit X une v.a. réelle admettant une espérance mathématique et un moment d'ordre 2. Alors :

$$\text{Var}(X) = \mathbb{E}[X^2] - [\mathbb{E}(X)]^2$$

Démonstration : On a :

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[(X - \mathbb{E}(X))^2] \\ &= \mathbb{E}[X^2 - 2X\mathbb{E}(X) + (\mathbb{E}(X))^2] \\ &= \mathbb{E}[X^2] - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}[(\mathbb{E}(X))^2] \quad (\text{linéarité de l'espérance}) \\ &= \mathbb{E}[X^2] - 2(\mathbb{E}(X))^2 + (\mathbb{E}(X))^2 \quad (\text{espérance de la constante } (\mathbb{E}(X))^2) \\ &= \mathbb{E}[X^2] - (\mathbb{E}(X))^2 \end{aligned} \quad \blacksquare$$



Le plus souvent, pour calculer des variances de variables aléatoires, la formule de Koenig s'avère donner des calculs légèrement plus simples que si on utilise la définition ; elle permet, en outre, d'identifier les deux premiers moments séparément.

3. Exemples de calcul de la variance

► Calcul de la variance dans le cas de la loi de Bernoulli

Considérons une variable aléatoire X suivant la loi de Bernoulli de paramètre p :

$$P(X = 1) = p, \quad P(X = 0) = 1 - p$$

Pour calculer sa variance, le plus simple est d'utiliser la formule de Koenig :

$$\mathbb{E}(X^2) = 0^2 \times P(X = 0) + 1^2 \times P(X = 1) = 0 + p = p$$

Il en résulte :

$$\text{Var}(X) = \mathbb{E}[X^2] - [\mathbb{E}(X)]^2 = p - p^2 = p(1 - p)$$



Dans la littérature anglo-saxonne, on rencontre souvent la notation $q = 1 - p$ (probabilité d'échec), qui conduit à la formule $\text{Var}(X) = pq$ ci-dessus, et des expressions (cf. infra) utilisant aussi la notation q .

► Calcul de la variance dans le cas de la loi uniforme

Considérons une variable aléatoire X suivant la loi uniforme sur l'ensemble fini $\{1, \dots, n\}$:

$$P(X = k) = \frac{1}{n}, \quad \mathbb{E}(X) = \frac{n+1}{2}$$

On calcule

$$\mathbb{E}(X^2) = \sum_{k=1}^n k^2 P(X = k) = \sum_{k=1}^n \frac{k^2}{n} = \frac{1}{n} \sum_{k=1}^n k^2 = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

Il en résulte, grâce à la formule de Koenig :

$$\text{Var}(X) = \mathbb{E}[X^2] - [\mathbb{E}(X)]^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{(n+1)(n-1)}{12} = \frac{n^2 - 1}{12}$$

► Calcul de la variance dans le cas de la loi géométrique de paramètre $p > 0$

Considérons une variable aléatoire X suivant la loi géométrique de paramètre $p \in]0, 1[$. On a alors $X(\Omega) = \mathbb{N}^*$, et, pour tout entier naturel non nul k :

$$P(X = k) = p(1 - p)^{k-1}$$

De plus :

$$\mathbb{E}(X) = \frac{1}{p}$$

On calcule :

$$\begin{aligned}
 \mathbb{E}(X^2) &= \sum_{k=1}^{+\infty} k^2 P(X = k) \\
 &= \sum_{k=1}^{+\infty} k^2 p (1-p)^{k-1} \\
 &= p \sum_{k=1}^{+\infty} k \left[\frac{d x^k}{dx} \right]_{x=1-p} \\
 &= p \frac{d}{dx} \left[\sum_{k=1}^{+\infty} (k x^k) \right]_{x=1-p} \\
 &= p \frac{d}{dx} \left[x \sum_{k=1}^{+\infty} k x^{k-1} \right]_{x=1-p} \\
 &= p \frac{d}{dx} \left[x \frac{d}{dx} \left(\sum_{k=0}^{+\infty} x^k \right) \right]_{x=1-p} \\
 &= p \frac{d}{dx} \left[x \frac{d}{dx} \left(\frac{1}{1-x} \right) \right]_{x=1-p} \\
 &= p \frac{d}{dx} \left[\frac{x}{(1-x)^2} \right]_{x=1-p} \\
 &= p \left[\frac{1}{(1-x)^2} + \frac{2x}{(1-x)^3} \right]_{x=1-p} \\
 &= p \left\{ \frac{1}{p^2} + \frac{2(1-p)}{p^3} \right\} \\
 &= \frac{1}{p} + \frac{2(1-p)}{p^2}
 \end{aligned}$$

(On applique le théorème de dérivation terme à terme des séries entières, puisque l'on est bien à l'intérieur du disque de convergence.)

Il en résulte, grâce à la formule de Koenig :

$$Var(X) = \mathbb{E}[X^2] - [\mathbb{E}(X)]^2 = \frac{1}{p} + \frac{2(1-p)}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

Il est à noter que ce résultat peut être également obtenu plus rapidement en calculant $\mathbb{E}(X(X-1))$.

► Calcul de la variance dans le cas de la loi de Poisson de paramètre $\lambda \in \mathbb{R}$

Considérons une variable aléatoire X suivant la loi de Poisson de paramètre $\lambda \in \mathbb{R}$. On a alors $X(\Omega) = \mathbb{N}$, et, pour tout entier naturel non nul k : $P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$.

On a :

$$\mathbb{E}(X) = \lambda$$

On calcule alors :

$$\begin{aligned}
 \mathbb{E}(X^2) &= \sum_{k=1}^{+\infty} k^2 P(X = k) \\
 &= \sum_{k=1}^{+\infty} k^2 \frac{e^{-\lambda} \lambda^k}{k!} \\
 &= \sum_{k=1}^{+\infty} k \frac{e^{-\lambda} \lambda^k}{(k-1)!} \\
 &= \sum_{k=1}^{+\infty} (k-1+1) \frac{e^{-\lambda} \lambda^k}{(k-1)!} \\
 &= \sum_{k=1}^{+\infty} (k-1) \frac{e^{-\lambda} \lambda^k}{(k-1)!} + \sum_{k=1}^{+\infty} \frac{e^{-\lambda} \lambda^k}{(k-1)!} \\
 &= \sum_{k=2}^{+\infty} (k-1) \frac{e^{-\lambda} \lambda^k}{(k-1)!} + \sum_{k=0}^{+\infty} \frac{e^{-\lambda} \lambda^{k+1}}{k!} \\
 &= \sum_{k=2}^{+\infty} \frac{e^{-\lambda} \lambda^k}{(k-2)!} + \lambda \sum_{k=0}^{+\infty} \frac{e^{-\lambda} \lambda^k}{k!} \\
 &= \sum_{k=0}^{+\infty} \frac{e^{-\lambda} \lambda^{k+2}}{k!} + \lambda \sum_{k=0}^{+\infty} \frac{e^{-\lambda} \lambda^k}{k!} \\
 &= \lambda^2 e^{-\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} + \lambda e^{-\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} \\
 &= \lambda^2 e^{-\lambda} e^{\lambda} + \lambda e^{-\lambda} e^{\lambda} \\
 &= \lambda^2 + \lambda
 \end{aligned}$$

Il en résulte, grâce à la formule de Koenig :

$$Var(X) = \mathbb{E}[X^2] - [\mathbb{E}(X)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$



Les quatre exemples ci-dessus utilisent la formule de Koenig. Dans de nombreux cas, on peut cependant recourir à des représentations de variables aléatoires comme sommes d'autres variables indépendantes, pour calculer beaucoup plus facilement leurs variances. C'est le cas, notamment, pour la loi binomiale et la loi binomiale négative.

À cet effet, il faut utiliser l'indépendance des termes intervenant dans ces sommes. Les résultats qui suivent apparaissent comme des cas particuliers faisant intervenir le concept de covariance et des lois jointes de couples de v.a., que nous rencontrerons un peu plus loin.

4. Espérance mathématique d'un produit de v.a. indépendantes

Théorème

Soient X et Y deux variables aléatoires indépendantes admettant, chacune, une espérance mathématique. Alors, la v.a. produit XY admet également une espérance mathématique, donnée par :

$$\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y).$$



Ce théorème peut être utilisé pour calculer la variance de la somme de deux v.a. X et Y indépendantes. Grâce à la formule de Koenig, et à la linéarité de l'espérance, on obtient, en développant les carrés :

$$\begin{aligned} \text{Var}(X + Y) &= \mathbb{E}(X^2 + Y^2 + 2XY) - (\mathbb{E}(X + Y))^2 \\ &= \mathbb{E}(X^2) + \mathbb{E}(Y^2) + 2\mathbb{E}(XY) - (\mathbb{E}(X))^2 - (\mathbb{E}(Y))^2 - 2\mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 + \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 \\ &= \text{Var}(X) + \text{Var}(Y). \end{aligned}$$

Nous avons donc démontré le résultat fondamental suivant :

► Additivité des variances pour les variables indépendantes

Corollaire

Soient X et Y deux variables indépendantes admettant des moments d'ordre 2. Alors :

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$



Ce corollaire implique immédiatement que si $X_1, \dots, X_n, n \in \mathbb{N}$, sont des variables indépendantes admettant des moments d'ordre 2, alors :

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i)$$

Le cas particulier suivant est particulièrement utile.

► Variance de la somme de variables indépendantes et identiquement distribuées (i.i.d.)

Corollaire

Soient $X_1, \dots, X_n, n \in \mathbb{N}$, des v.a. indépendantes et identiquement distribuées (i.i.d.), admettant, chacune, deux moments. Alors,

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = n \text{Var}(X_1)$$

Exemples

1. Calcul de la variance dans le cas de la loi binomiale $\mathcal{B}(n, p), n \in \mathbb{N}, p \in [0, 1]$

Considérons une variable aléatoire X suivant la loi binomiale $\mathcal{B}(n, p)$. Nous avons déjà vu que $\mathbb{E}(X) = np$, calcul qui a été grandement facilité par le fait qu'on peut considérer X comme la somme $\sum_{i=1}^n X_i$, où les $X_i, i = 1, \dots, n$, sont des variables i.i.d, de loi $\mathcal{B}(p)$, avec, pour tout entier i de $\{1, \dots, n\}$, $\mathbb{E}(X_i) = p$. Nous avons également vu que $\text{Var}(X_i) = p(1 - p)$. En utilisant le corollaire précédent, on obtient immédiatement :

$$\text{Var}(X) = np(1 - p)$$

Grâce à la formule de Koenig :

$$\begin{aligned} \mathbb{E}(X^2) &= np(1 - p) + n^2 p^2 \\ &= np(1 + (n - 1)p) \end{aligned}$$

Le calcul direct de cette variance (i.e. de ce second moment), à partir de la loi $P(X = k) = p^k (1 - p)^{n-k} \binom{n}{k}$, et de la formule :

$$\mathbb{E}(X^2) = \sum_{k=0}^n k^2 P(X = k)$$

est beaucoup plus ardu. C'est exactement la même chose pour le calcul qui suit.

2. Calcul de la variance dans le cas de la loi binomiale négative $\mathcal{BN}(n, p)$, $n \in \mathbb{N}$, $p \in]0, 1]$

Considérons une variable aléatoire X suivant la loi binomiale négative $\mathcal{BN}(n, p)$, $n \in \mathbb{N}$, $p \in [0, 1]$. Nous avons déjà vu que $\mathbb{E}(X) = \frac{n}{p}$, résultat qui vient immédiatement du

fait que X peut être considéré comme la somme $\sum_{i=1}^n X_i$, où les X_i , $i = 1, \dots, n$, sont des

variables i.i.d., de même loi $\mathcal{G}(p)$, telles que, pour tout entier i de $\{1, \dots, n\}$, $\mathbb{E}(X_i) = \frac{1}{p}$ et $\text{Var}(X_i) = (1 - p) p^{-2}$. En utilisant le corollaire précédent, on obtient immédiatement :

$$\text{Var}(X) = \frac{n(1 - p)}{p^2}$$

Grâce à la formule de Koenig :

$$\begin{aligned} \mathbb{E}(X^2) &= \frac{n(1 - p)}{p^2} + \frac{n^2}{p^2} \\ &= \frac{n(n + 1 - p)}{p^2} \end{aligned}$$

L'inégalité suivante est très utile pour évaluer les déviations d'une variable par rapport à sa moyenne (espérance).

5. Inégalité de Tchebychev¹

Soit X une v.a. admettant une variance $\text{Var}(X)$. Alors, pour tout réel $t \geq 0$:

$$P(|X - \mathbb{E}(X)| \geq t) \leq \frac{\text{Var}(X)}{t^2}$$

Démonstration : Il suffit d'appliquer l'inégalité de Markov :

$$P(|X - \mathbb{E}(X)| \geq t) = P(|X - \mathbb{E}(X)|^2 \geq t^2) \leq \frac{\mathbb{E}(|X - \mathbb{E}(X)|^2)}{t^2} \leq \frac{\text{Var}(X)}{t^2} \quad \blacksquare$$



Le phénomène qu'on vient d'observer peut être considéré comme « précurseur » de la concentration de la mesure : lorsqu'on répète des expériences indépendantes de nombreuses fois, les déviations autour des valeurs moyennes, toutes proportions gardées, ont tendance à s'estomper très rapidement. L'ultime manifestation de ce phénomène est celui du théorème central limite, qu'on verra ultérieurement.

Il est néanmoins possible de bien comprendre déjà ce phénomène, en regardant simplement les écarts-type des moyennes empiriques.

1. (ou Chebyshev, en transcription anglo-saxonne) Pafnouti Lvovitch Tchebychev (1821-1894), mathématicien russe, qui apporta de nombreuses contributions en probabilités et en statistiques.

► **Première manifestation de la concentration de la mesure**

Soit n un entier naturel non nul, et $(X_i)_{1 \leq i \leq n}$ une famille de variables i.i.d., d'espérance $\mathbb{E}(X_i) = \mu \in \mathbb{R}$, $i = 1, \dots, n$, de variance $\text{Var}(X_i) = \sigma^2 \in \mathbb{R}^+$, $i = 1, \dots, n$. On désigne par S_n leur « moyenne empirique » jusqu'au rang n :

$$S_n = \frac{1}{n} \sum_{i=1}^n X_i$$

Par linéarité de l'espérance, additivité et changement d'échelle de la variance, l'espérance et l'écart-type de la moyenne empirique sont donnés, pour tout entier $n \geq 1$, par :

$$\mathbb{E}(S_n) = \mu \quad , \quad \sqrt{\text{Var}(S_n)} = \frac{\sigma}{\sqrt{n}}$$

Ainsi, la dispersion de la moyenne empirique tend vers 0 à la vitesse $\frac{1}{n}$, alors que son espérance reste constante. On conçoit donc bien que S_n se concentre autour de μ . L'inégalité de Tchebychev donne alors, pour tout $x > 0$:

$$P(|S_n - \mu| > x) \leq \frac{\sigma^2}{nx^2}$$

La vitesse de décroissance vers 0 de la probabilité de déviation de S_n au-delà d'une distance x de μ , est, au moins de l'ordre de $\frac{1}{n}$ pour x fixé, et au moins de l'ordre de $\frac{1}{x^2}$ pour n fixé. C'est une concentration de type quadratique. Le théorème central limite montrera que la vitesse de concentration est typiquement beaucoup plus rapide, d'ordre exponentiel quadratique.



Revenons à l'exemple du temps d'attente X pour l'arrivée du premier client dans notre hypothétique magasin de luxe. Supposons à nouveau que l'on ne connaisse pas la loi de X . Par contre, on a mesuré, empiriquement, $\mathbb{E}(X) = 10$ et $\mathbb{E}(X^2) = 190$. Grâce à la formule de Koenig, on en déduit :

$$\text{Var}(X) = 190 - 100 = 90$$

On peut alors invoquer l'inégalité de Tchebychev pour estimer la probabilité que la déviation entre X et sa moyenne soit supérieure à cinq minutes :

$$P(|X - 10| > 5) \leq \frac{90}{25} = 3,6$$

Cette estimation n'a, évidemment, aucune valeur informative ! La difficulté vient du fait que la v.a. X est fortement non symétrique, et a une forte variance par rapport au carré de son espérance.

Considérons donc plutôt le temps d'attente Y jusqu'à l'arrivée du dixième client, et supposons que les durées entre deux clients successifs sont indépendantes les unes des autres. Il n'est alors pas mauvais de supposer que Y puisse être considéré comme somme de 10 variables i.i.d., de même loi que X . Par linéarité de l'espérance, additivité de la variance pour les variables indépendantes, on obtient :

$$\mathbb{E}(Y) = 100 \quad \text{et} \quad \text{Var}(Y) = 900$$

L'inégalité de Tchebychev conduit alors à :

$$P(|Y - 100| > 50) \leq \frac{900}{2500} = 0,36.$$

Si l'on changeait encore la question, et que l'on s'intéressait à la probabilité que le temps d'arrivée Z du 100^e client dépasse de plus de 500 minutes la durée moyenne d'attente de ce 100^e client, qui est de $\mathbb{E}(Z) = 1\,000$ minutes, on obtiendrait, par Tchebychev :

$$P(|Z - 1\,000| > 500) \leq \frac{9000}{250\,000} = 0,036.$$

1. Inégalité de Cauchy-Schwarz

Théorème

Soient X et Y deux variables aléatoires admettant, chacune, un moment d'ordre 2. Alors :

$$|\mathbb{E}(X Y)| \leq \sqrt{\mathbb{E}(X^2)} \sqrt{\mathbb{E}(Y^2)}$$

ou encore :

$$|\mathbb{E}(X Y)|^2 \leq \mathbb{E}(X^2) \mathbb{E}(Y^2)$$

Il y a égalité si et seulement si les v.a. X et Y sont colinéaires, i.e. s'il existe un réel t tel que $X = t Y$ presque sûrement (la probabilité associée est 1).

Démonstration : Pour tout réel t , par linéarité de l'espérance :

$$\mathbb{E}[(X + t Y)^2] = \mathbb{E}[X^2] + t^2 \mathbb{E}[Y^2] + 2 t \mathbb{E}[X Y]$$

X et Y admettant des moments d'ordre 2, l'existence de l'espérance $\mathbb{E}[X Y]$ est ainsi assurée.

De plus, le trinôme $\mathbb{E}[(X + t Y)^2]$ étant toujours positif, et ce quelle que soit la valeur du réel t , son discriminant réduit est donc négatif ou nul :

$$(\mathbb{E}[X Y])^2 - \mathbb{E}[X^2] \mathbb{E}[Y^2] \leq 0$$

ce qui conduit au résultat cherché.

Le cas d'égalité correspond à l'existence d'une racine double t_0 , qui est donc telle que :

$$\mathbb{E}[(X + t_0 Y)^2] = 0$$

ce qui équivaut à :

$$X + t_0 Y = 0 \text{ presque sûrement}$$

(i.e. la probabilité associée vaut 1).

X et Y sont donc presque sûrement colinéaires. La réciproque est évidente. ■

2. Espérance d'un produit de v.a. indépendantes

On rappelle, dans ce qui suit, le théorème sur l'espérance d'un produit de v.a. indépendantes ; sa démonstration est donnée ci-dessous.

Théorème

Soient X et Y deux variables aléatoires indépendantes admettant, chacune, une espérance mathématique. Alors, la v.a. produit $X Y$ a également une espérance mathématique, donnée par :

$$\mathbb{E}(X Y) = \mathbb{E}(X) \mathbb{E}(Y)$$

Démonstration : Par indépendance des v.a. X et Y , pour tout x_i de $X(\Omega)$, et tout y_j de $Y(\Omega)$:

$$|x_i y_j P(X = x_i, Y = y_j)| = |x_i y_j P(X = x_i) P(Y = y_j)| = |x_i| |y_j| P(X = x_i) P(Y = y_j)$$

X et Y admettant une espérance mathématique, $|x_i| P(X = x_i)$ et $|y_j| P(Y = y_j)$ sont les termes généraux de séries convergentes.

L'existence de l'espérance du produit XY est ainsi assurée. On a alors :

$$\begin{aligned} \mathbb{E}(XY) &= \sum_{x_i \in X(\Omega), y_j \in Y(\Omega)} x_i y_j P(X = x_i, Y = y_j) \\ &= \sum_{x_i \in X(\Omega), y_j \in Y(\Omega)} x_i y_j P(X = x_i) P(Y = y_j) \\ &= \left\{ \sum_{x_i \in X(\Omega)} x_i P(X = x_i) \right\} \left\{ \sum_{y_j \in Y(\Omega)} y_j P(Y = y_j) \right\} \\ &= \mathbb{E}(X) \mathbb{E}(Y) \end{aligned}$$



La réciproque est fautive ! Ce n'est pas parce que $\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y)$ que les v.a. X et Y sont indépendantes, loin de là ! Il y a néanmoins un type de variables aléatoires pour lesquelles cette réciproque est vraie : les variables de Bernoulli. Le lecteur pourra réfléchir à la propriété suivante : une variable de Bernoulli X est, en fait, équivalente à l'événement $A = \{X = 1\}$, puisque, si X n'est pas égale à 1, alors, nécessairement, $X = 0$. De fait, l'indépendance de deux variables de Bernoulli X et Y est équivalente à l'indépendance des événements $A = \{X = 1\}$ et $B = \{Y = 1\}$. On peut alors montrer que la relation $E[XY] = E[X] E[Y]$ est équivalente à $P[A \cap B] = P[A] P[B]$, qui implique, par définition, que A et B sont indépendants, et, par suite, que X et Y sont indépendants.

Exemple

Voici un contre-exemple assez général illustrant l'implication fautive de la remarque précédente. Soient X_1 et X_2 deux variables indépendantes de même loi, admettant, chacune, un moment non nul d'ordre 2. On considère les v.a. :

$$X = X_1 + X_2 \quad \text{et} \quad Y = X_1 - X_2$$

Par linéarité :

$$\mathbb{E}(Y) = 0$$

Ainsi,

$$\mathbb{E}(X) \mathbb{E}(Y) = 0$$

De plus,

$$\mathbb{E}(XY) = \mathbb{E}(X_1^2 - X_2^2) = 0$$

et donc $\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y)$.

Dans la plupart des exemples, X et Y ne sont pas indépendants. On peut noter, en particulier, que X et Y ne peuvent pas être de Bernoulli, sauf si une des deux variables X_1 ou X_2 a une variance nulle. Le lecteur vérifiera par exemple que, si X_1 et X_2 suivent la loi $\mathcal{B}(p)$, $p = 0,5$.

$$P(\{X = 1\} \cap \{Y = 0\}) = 0$$

alors que :

$$P(X = 1) P(Y = 0) = \frac{1}{2} \frac{1}{2} = \frac{1}{4} \neq 0$$

3. Covariance

Définition

Soient X et Y deux variables aléatoires admettant, chacune, un moment d'ordre 2. On appelle **covariance** du couple de v.a. (X, Y) la quantité :

$$Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))]$$

Soit X une variable aléatoire admettant un moment d'ordre 2. Alors :

$$Cov(X, X) = \mathbb{E}[(X - \mathbb{E}(X))^2] = Var(X)$$

► Covariance de deux v.a. indépendantes

Soient X et Y deux variables aléatoires indépendantes admettant, chacune, un moment d'ordre 2. Leur covariance est alors nulle :

$$Cov(X, Y) = 0$$



Attention ! La réciproque est fausse. Le contre-exemple précédent s'applique ici aussi (il suffit de soustraire les espérances)...

Démonstration : Les v.a. X et Y étant indépendantes, il en est de même des v.a. $X - \mathbb{E}(X)$ et $Y - \mathbb{E}(Y)$. Par suite :

$$\begin{aligned} Cov(X, Y) &= \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] \\ &= \mathbb{E}[X - \mathbb{E}(X)] \mathbb{E}[Y - \mathbb{E}(Y)] \\ &= \{\mathbb{E}[X] - \mathbb{E}(X)\} \{\mathbb{E}[Y] - \mathbb{E}(Y)\} \\ &= 0 \end{aligned}$$

(On utilise, tout simplement, la linéarité de l'espérance par rapport aux constantes $\mathbb{E}(X)$ et $\mathbb{E}(Y)$.) ■

4. Propriétés de la covariance

Soient X et Y deux variables aléatoires admettant, chacune, un moment d'ordre 2. La covariance $Cov(X, Y)$ du couple de v.a. (X, Y) vérifie les propriétés suivantes :

- **Symétrie :**

$$Cov(Y, X) = Cov(X, Y)$$

- **Changement d'échelle :** pour tout couple de réels (λ, μ) :

$$Cov(\lambda X, \mu Y) = \lambda \mu Cov(X, Y)$$

- **Invariance par translation :** pour tout couple de réels (α, β) :

$$Cov(X + \alpha, Y + \beta) = Cov(X, Y)$$

- **Comparaison avec les écarts-types** : d'après l'inégalité de Cauchy-Schwartz :

$$|Cov(X, Y)| \leq \sigma(X) \sigma(Y)$$

Démonstration :

- **Symétrie** : De par la définition de la covariance :

$$Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] = \mathbb{E}[(Y - \mathbb{E}(Y))(X - \mathbb{E}(X))] = Cov(Y, X)$$

- **Changement d'échelle** : pour tout couple de réels (λ, μ) , par linéarité de l'espérance :

$$\begin{aligned} Cov(\lambda X, \mu Y) &= \mathbb{E}[(\lambda X - \mathbb{E}(\lambda X))(\mu Y - \mathbb{E}(\mu Y))] \\ &= \mathbb{E}[(\lambda X - \lambda \mathbb{E}(X))(\mu Y - \mu \mathbb{E}(Y))] \\ &= \mathbb{E}[\lambda \mu (X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] \\ &= \lambda \mu \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] \\ &= \lambda \mu Cov(X, Y) \end{aligned}$$

- **Invariance par translation** : pour tout couple de réels (α, β) , par linéarité de l'espérance :

$$\begin{aligned} Cov(X + \alpha, Y + \beta) &= \mathbb{E}[(X + \alpha - \mathbb{E}(X + \alpha))(Y + \beta - \mathbb{E}(Y + \beta))] \\ &= \mathbb{E}[(X + \alpha - \mathbb{E}(X) - \alpha)(Y + \beta - \mathbb{E}(Y) - \beta)] \\ &= \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] \\ &= Cov(X, Y) \end{aligned}$$

- **Comparaison avec les écarts-types** : d'après l'inégalité de Cauchy-Schwartz :

$$|Cov(X, Y)| \leq \sigma(X) \sigma(Y) \quad \blacksquare$$

► Coefficient de corrélation

Soient X et Y deux variables aléatoires non constantes admettant, chacune, un moment d'ordre 2. Le **coefficient de corrélation** du couple de v.a. (X, Y) est la quantité définie par :

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sigma(X) \sigma(Y)}$$

Propriété

Soient X et Y deux variables aléatoires non constantes admettant, chacune, un moment non nul d'ordre 2. Comme :

$$|Cov(X, Y)| \leq \sigma(X) \sigma(Y)$$

alors :

$$-1 \leq \rho(X, Y) \leq 1$$

► Formule de Koenig

Soient X et Y deux variables aléatoires telle que la covariance $Cov(X, Y)$ existe. Alors :

$$Cov(X, Y) = \mathbb{E}(X Y) - \mathbb{E}(X)\mathbb{E}(Y)$$

Démonstration : On a :

$$\begin{aligned} Cov(X, Y) &= \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] \\ &= \mathbb{E}[X Y + \mathbb{E}(X)\mathbb{E}(Y) - X\mathbb{E}(Y) - Y\mathbb{E}(X)] \\ &= \mathbb{E}[X Y] + \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(Y)\mathbb{E}(X) \quad (\text{linéarité de l'espérance}) \\ &= \mathbb{E}[X Y] - \mathbb{E}(X)\mathbb{E}(Y) \end{aligned}$$

■

Exemple

On considère deux variables X_1 et X_2 i.i.d., de loi $\mathcal{B}(p)$. On pose $X = X_1 + X_2$, et $Y = X_1 X_2$. (les v.a. X et Y correspondent, respectivement à la somme et au minimum des gains dans un jeu de pile ou face, où on lance, de façon indépendante, deux pièces, et où on gagne un euro chaque fois que tombe le côté « pile » (avec une probabilité de p) et rien du tout chaque fois que tombe le côté « face »).

Puisque :

$$\mathbb{E}(X_1) = \mathbb{E}((X_1)^2) = p$$

alors, par indépendance :

$$\mathbb{E}(X Y) = \mathbb{E}(X_1^2 X_2 + X_1 X_2^2) = p \times p + p \times p = 2p^2$$

et :

$$\mathbb{E}(X)\mathbb{E}(Y) = p \times p = p^2.$$

Grâce à la formule de Koenig :

$$Cov(X, Y) = 2p^2 - p^2 = p^2.$$

On a aussi :

$$Var(X_1) = Var(X_2) = p(1 - p)$$

Il en résulte :

$$\rho(X, Y) = \frac{p^2}{p(1 - p)} = \frac{p}{1 - p}$$

Il est intéressant de noter que le coefficient de corrélation entre X et Y est précisément égal à ce que l'on appelle, dans le langage des jeux d'argent, la « cote de la face pile ».

Nous avons déjà rencontré certaines propriétés des paires de variables aléatoires, particulièrement lorsqu'elles sont indépendantes, mais pas nécessairement, comme l'illustre le dernier exemple ci-dessus. Nous introduisons, dans ce qui suit, de nouveaux éléments permettant d'étudier, de façon plus systématique, ces paires de v.a.

1. Couple de variables aléatoires discrètes

Définition

Soit (Ω, \mathcal{A}, P) un espace probabilisé. On appelle **couple de variables aléatoires discrètes** sur (Ω, \mathcal{A}, P) toute application $Z : \Omega \rightarrow \mathbb{R}^2$, qui, à tout événement ω de Ω associe le couple $(X(\omega), Y(\omega))$, où X et Y sont deux variables aléatoires discrètes sur Ω .



Un couple de variables aléatoires peut être considéré comme une variable aléatoire à valeurs dans \mathbb{R}^2 .

Exemple

Considérons le lancer de deux dés. On désigne par X le plus grand des deux nombres obtenus, et par Y le plus petit. Alors, $Z = (X, Y)$ est un couple de variables aléatoires.

Théorème

Soient (Ω, \mathcal{A}, P) un espace probabilisé, et (X, Y) un couple de v.a. discrètes sur Ω . Supposons que les valeurs prises par chacune des deux composantes X et Y de ce couple sont :

$$X(\Omega) = \{x_1, \dots, x_m\} \quad , \quad Y(\Omega) = \{y_1, \dots, y_n\}$$

où $(m, n) \in (\mathbb{N}^* \cup \{+\infty\})^2$.

Alors, la famille d'événements :

$$(\{X = x_i\} \cap \{Y = y_j\})_{1 \leq i \leq m, 1 \leq j \leq n} = (\{X = x_i \text{ et } Y = y_j\})_{1 \leq i \leq m, 1 \leq j \leq n}$$

est un système complet d'événements de Ω .

On utilisera la notation :

$$\{X = x_i, Y = y_j\} = \{X = x_i\} \cap \{Y = y_j\}.$$

2. Lois

► Loi conjointe

Soient (Ω, \mathcal{A}, P) un espace probabilisé, et (X, Y) un couple de variables aléatoires discrètes sur (Ω, \mathcal{A}, P) . On appelle **loi conjointe** du couple de v.a. (X, Y) l'application, notée $P_{X,Y}$:

$$\begin{aligned} X(\Omega) \times Y(\Omega) &\rightarrow [0, 1] \\ (x, y) &\mapsto P(X = x, Y = y) \end{aligned}$$



► Lois marginales

Soient (Ω, \mathcal{A}, P) un espace probabilisé, et (X, Y) un couple de variables aléatoires discrètes sur (Ω, \mathcal{A}, P) . On appelle **lois marginales** les lois de probabilités respectives des v.a. X et Y .

Théorème

Soient (Ω, \mathcal{A}, P) un espace probabilisé, et (X, Y) un couple de v.a. discrètes sur Ω , de la forme :

$$X(\Omega) = \{x_1, \dots, x_m\} \quad , \quad Y(\Omega) = \{y_1, \dots, y_n\}$$

où $(m, n) \in \mathbb{N}^* \times \mathbb{N}^*$.

Alors, les lois des v.a. X et Y sont respectivement données par :

- pour tout entier i de $\{1, \dots, m\}$:

$$P(X = x_i) = \sum_{j=1}^n P(X = x_i, Y = y_j)$$

- pour tout entier j de $\{1, \dots, n\}$:

$$P(Y = y_j) = \sum_{i=1}^m P(X = x_i, Y = y_j)$$

► Lois conditionnelles

Soit (Ω, \mathcal{A}, P) un espace probabilisé, et (X, Y) un couple de variables aléatoires discrètes sur (Ω, \mathcal{A}, P) . Pour tout y de $Y(\Omega)$ tel que $P(Y = y) \neq 0$, l'application $P(Y = y)$:

$$\begin{aligned} X(\Omega) &\rightarrow [0, 1] \\ x &\mapsto P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} \end{aligned}$$

est appelée **loi conditionnelle de X sachant que $Y = y$** . On utilisera parfois la notation explicite $P(\cdot|Y = y)$, au lieu de $P(Y = y)$.

De même, pour tout x de $X(\Omega)$ tel que $P(X = x) \neq 0$, l'application $P(X = x)$:

$$\begin{aligned} Y(\Omega) &\rightarrow [0, 1] \\ y &\mapsto P(Y = y|X = x) = \frac{P(X = x, Y = y)}{P(X = x)} \end{aligned}$$

est appelée **loi conditionnelle de Y sachant que $X = x$** .



Nous avons déjà rencontré des variables aléatoires indépendantes, mais rappelons quand même la définition associée.

Définition

Soient (Ω, \mathcal{A}, P) un espace probabilisé, et (X, Y) un couple de variables aléatoires discrètes sur (Ω, \mathcal{A}, P) .

Les v.a. X et Y sont indépendantes si la loi conjointe est le produit des lois marginales, c'est-à-dire, pour tout (x, y) de $X(\Omega) \times Y(\Omega)$:

$$P(X = x, Y = y) = P(X = x) P(Y = y)$$

Un théorème limite pour les variables aléatoires discrètes : la loi des grands nombres

De façon très intuitive, il est clair que les caractéristiques d'un échantillon choisi au hasard se rapprocheront d'autant plus des caractéristiques réelles de la population que la taille de l'échantillon augmente. Curieusement, la taille de l'échantillon à considérer ne dépend que faiblement de la taille de la population. Ainsi, pour déterminer la préférence d'une population en chocolat noir corsé ou au lait, un sondage sur le même nombre d'individus dans un petit pays comme le Luxembourg ou, au contraire, dans un pays à plus forte population comme les États-Unis, suffit ! On peut se rappeler de l'exemple de l'approximation binomiale de la loi hypergéométrique, pour illustrer ce phénomène : l'approximation est bonne, que l'on ait ou non une population qui soit des centaines ou des centaines de milliers de fois plus grandes que l'échantillon.

En ce qui concerne le problème très précis de calculer une moyenne empirique, et son interprétation en tant qu'espérance de la population, on dispose d'un outil extrêmement général : la « loi des grands nombres », dont il existe au moins deux versions, une faible, une forte. Nous disposons déjà des outils pour en démontrer une version faible.

C'est le mathématicien suisse Jacob (Jacques, James) Bernoulli¹ qui en établit le premier, le modèle formel, en 1690. La plupart des sondages se basent, au minimum, sur ce résultat. Le cadre extrêmement vaste d'application de ce théorème a un prix : la loi des grands nombres, qui dit que la moyenne empirique de données i.i.d. converge vers leur espérance mathématique, ne donne pas d'ordre de grandeur de la vitesse de convergence, et ne permet donc pas d'obtenir ce que l'on appelle communément une « marge d'erreur », qui requiert le théorème central limite.

1. La loi faible des grands nombres

Théorème

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées, définies sur un même espace probabilisé (Ω, \mathcal{A}, P) , avec une espérance mathématique $\mathbb{E}(X)$. Alors, $\frac{X_1 + X_2 + \dots + X_n}{n}$ converge « en probabilité » vers $\mathbb{E}(X)$, ce qui se traduit mathématiquement par le fait que, pour tout $\varepsilon > 0$,

$$\lim_{n \rightarrow +\infty} P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - \mathbb{E}(X)\right| \geq \varepsilon\right) = 0$$

Démonstration : On peut présenter une démonstration, simple, et très complète, de la loi faible des grands nombres, en supposant que les variables X_n ont un moment d'ordre 2 fini. Cette hypothèse, beaucoup plus forte, permet même d'affaiblir les autres hypothèses du théorème, en supprimant en particulier l'hypothèse i.i.d.

1. (1654-1705), frère de Jean Bernoulli (1667-1748), lui aussi mathématicien, oncle de Daniel (1700-1782) et Nicolas Bernoulli (1695-1726).

Supposons plutôt que les X_n sont indépendants et possèdent tous la même espérance et la même variance. Soit alors :

$$M_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

On calcule alors, en utilisant la linéarité de l'espérance, l'additivité et le changement d'échelle de la variance :

$$\begin{aligned}\mathbb{E}[M_n] &= \frac{\mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}(X_n)}{n} = \mathbb{E}(X_1), \\ \text{Var}(M_n) &= \frac{\text{Var}(X_1) + \text{Var}(X_2) + \cdots + \text{Var}(X_n)}{n^2} = \frac{\text{Var}(X_1)}{n}\end{aligned}$$

Grâce à l'inégalité de Tchebychev, on obtient alors :

$$\begin{aligned}P\left(\left|\frac{X_1 + X_2 + \cdots + X_n}{n} - \mathbb{E}(X)\right| \geq \varepsilon\right) &= P(|M_n - \mathbb{E}(X)| \geq \varepsilon) \\ &\leq \frac{\text{Var}(M_n)}{\varepsilon^2} \\ &= \frac{\text{Var}(X_1)}{n \varepsilon^2}.\end{aligned}$$

Pour tout $\varepsilon > 0$, cette expression tend vers 0 lorsque n tend vers $+\infty$. Nous avons donc démontré le résultat suivant. ■

2. Une loi faible des grands nombres avec second moment, sans hypothèse i.i.d.

Théorème

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes définies sur un même espace probabilisé (Ω, \mathcal{A}, P) , toutes avec la même espérance mathématique $\mathbb{E}(X)$ et la même variance $\text{Var}(X)$ finie.

Alors, $\frac{X_1 + X_2 + \cdots + X_n}{n}$ converge en probabilité vers $\mathbb{E}(X)$.



En renforçant encore un peu les hypothèses concernant les moments, on peut démontrer facilement, grâce à un résultat connu sous le nom de lemme de Borel-Cantelli, la loi forte des grands nombres. Cette loi est néanmoins vraie en toute généralité, sans l'existence de moments d'ordres plus élevés que 1, mais sa démonstration sort largement du cadre de ce cours. Nous la mentionnons sans plus de commentaires.

3. La loi forte des grands nombres

Théorème

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires i.i.d. définies sur un même espace probabilisé (Ω, \mathcal{A}, P) , avec une espérance mathématique $\mathbb{E}(X)$ finie. Alors, avec une probabilité égale à 1, $\frac{X_1 + X_2 + \cdots + X_n}{n}$ converge vers $\mathbb{E}(X)$ lorsque l'entier n tend vers l'infini.

Jusqu'à présent, l'univers des possibles, Ω , était dénombrable. Il n'est pas toujours possible de modéliser les phénomènes de la vie réelle en se restreignant à de tels univers. Néanmoins, beaucoup des concepts développés dans le cadre discret restent encore valables lorsque Ω n'est pas nécessairement dénombrable. La théorie de la mesure, que nous ne traitons pas dans ce cours, permet de donner une définition précise des types d'espaces probabilisés (Ω, \mathcal{A}, P) qui sont utiles en modélisation probabiliste. Modulo ce point que nous ne détaillerons pas, il est trivial de redéfinir le concept de variable aléatoire et les quantités associées.

Définition

Soit (Ω, \mathcal{A}, P) un espace probabilisé. On appelle **variable aléatoire réelle** sur (Ω, \mathcal{A}, P) toute application $X : \Omega \rightarrow \mathbb{R}$.

La fonction de répartition F_X de X est la fonction de la variable réelle x définie par :

$$F_X(x) = P(X \leq x).$$

La loi, ou la distribution, de X , peut être identifiée à cette fonction F_X .

Propriété

Comme dans le cas discret, avec F_X définie comme ci-dessus, F_X est croissante,

$$\lim_{x \rightarrow -\infty} F_X(x) = 0 \text{ et } \lim_{x \rightarrow +\infty} F_X(x) = 1.$$

Dans ce qui suit, on se place, implicitement, dans un espace probabilisé (Ω, \mathcal{A}, P) , pour lequel l'univers Ω n'est pas dénombrable. Il s'agira donc simplement d'admettre sans démonstration que les variables aléatoires que nous serons amenés à utiliser sont définies sur de tels espaces.

1. Densité de probabilité

Définition

Une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$, à valeurs positives, intégrable (i.e. $\int_{\mathbb{R}} f(t) dt = \int_{-\infty}^{+\infty} f(t) dt$ existe) est appelée **densité de probabilité** si et seulement si :

$$\int_{-\infty}^{+\infty} f(t) dt = 1$$

Exemples

1. La fonction caractéristique de l'intervalle $[0, 1]$, $\mathbb{1}_{[0,1]}$, est une densité de probabilité, puisqu'elle est à valeurs positives, intégrable, et telle que :

$$\int_{-\infty}^{+\infty} \mathbb{1}_{[0,1]} dt = \int_0^1 1 dt = \int_0^1 dt = 1$$

2. La fonction f définie sur \mathbb{R} par :

$$f(x) = \begin{cases} 0 & \text{pour } x < 0 \\ e^{-x} & \text{pour } x \geq 0 \end{cases}$$

est une densité de probabilité pour les mêmes raisons, en particulier :

$$\int_{-\infty}^{+\infty} f(t) dt = \int_{-\infty}^0 0 dt + \int_0^{+\infty} e^{-t} dt = 0 + \left[(-e^{-t})\right]_0^{+\infty} = 0 - 0 + 1 = 1.$$

2. Variable aléatoire continue

Définition

Soit f une densité de probabilité. On dit qu'une variable aléatoire réelle X **suit la loi de densité** f si, pour tout couple de réels (a, b) tel que $a \leq b$:

$$P(X \in [a, b]) = \int_a^b f(t) dt$$

On dit aussi que la v.a. X est **continue**.



Pour toute v.a. continue,

$$P(X \in [a, b]) = P(X \in]a, b])$$

puisque l'intégrale de f sur l'intervalle ouvert à gauche $]a, b]$ est la même que celle sur l'intervalle fermé $[a, b]$, ou celle sur l'intervalle ouvert $]b, a[$.

Propriété

Soit X une v.a. réelle de densité f . Sa fonction de répartition F_X est telle que :

- Pour tout réel x :

$$F_X(x) = \int_{-\infty}^x f(t) dt$$

- F_X est continue sur \mathbb{R} .
- Si f est continue en $x \in \mathbb{R}$, alors F_X est dérivable en x , et :

$$F'_X(x) = f(x)$$

3. Espérance d'une variable aléatoire continue

Pour comprendre intuitivement ce que représente le concept d'espérance pour les v.a. continues, il est intéressant de considérer une variable aléatoire discrète prenant ses valeurs dans l'ensemble $\{1, \dots, n\}$.

Sa fonction de probabilité est donnée par :

$$p(k) = P(X = k)$$

pour $k = 1, 2, \dots, n$.

Si on change de notation et que l'on écrit :

$$f\left(\frac{k}{n}\right) = n p(k)$$

on est amené à s'intéresser à la v.a. $Y = \frac{X}{n}$, qui est toujours discrète, mais prend ses valeurs dans $\left\{\frac{1}{n}, \frac{2}{n}, \dots, 1\right\} \subset [0, 1]$.

On s'aperçoit, en particulier, que la fonction de probabilité de Y est donnée par :

$$P\left(Y = \frac{k}{n}\right) = n^{-1} p\left(\frac{k}{n}\right)$$

Par définition,

$$\mathbb{E}[X] = \sum_{k=1}^n k p(k)$$

Par linéarité :

$$\mathbb{E}[Y] = \frac{1}{n} \sum_{k=1}^n \frac{k}{n} p\left(\frac{k}{n}\right).$$

Enfin, s'il s'avère que la fonction p est telle que la suite $\left(f\left(\frac{k}{n}\right)\right)_{n \in \mathbb{N}}$, i.e. la suite $(n p(k))_{n \in \mathbb{N}}$, converge vers une limite $f(x)$, où f est une densité de probabilité, quand k/n converge vers $x \in [0, 1]$ (par exemple, en prenant $k = [nx]$, partie entière de nx), alors on reconnaît que la formule ci-dessus pour $\mathbb{E}[Y]$ est la somme de Riemann associée à l'intégrale $\int_0^1 t f(t) dt$. Si, de surcroît, la fonction f est continue par morceaux, nous avons vu que la somme de Riemann converge vers l'intégrale, i.e. :

$$\lim_{n \rightarrow \infty} \mathbb{E}[Y] = \int_0^1 t f(t) dt.$$

$np(k)$ peut aussi être interprété comme une approximation de la densité f d'une variable aléatoire Z , correspondant à la loi limite de la loi de Y . On divise $p(k)$ par la longueur de l'intervalle $\left[\frac{k-1}{n}, \frac{k}{n}\right]$, pour permettre à la probabilité de ne pas tendre vers une limite triviale. Plus précisément, toujours avec $k = [nx]$, on peut écrire, sous réserve d'existence des limites en jeu,

$$\begin{aligned} f(x) &= \frac{dF_Z(x)}{dx} \\ &= \lim_{n \rightarrow +\infty} \frac{P\left(Z \in \left[x - \frac{1}{n}, x\right]\right)}{\frac{1}{n}} \\ &= \lim_{n \rightarrow +\infty} n P\left[Y \in \left[\frac{k-1}{n}, \frac{k}{n}\right]\right] \\ &= \lim_{n \rightarrow +\infty} f\left(\frac{k}{n}\right) \\ &= \lim_{n \rightarrow +\infty} np(k) \end{aligned}$$

Et :

$$\mathbb{E}[Z] = \lim_{n \rightarrow +\infty} \mathbb{E}[Y] = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=1}^n \frac{k}{n} f\left(\frac{k}{n}\right) = \int_0^1 t f(t) dt.$$



Le lecteur assidu pourra vérifier « à la main » que la situation ci-dessus s'applique à des exemples que nous avons déjà rencontrés, y compris dans le cas où la v.a. X est de loi $\mathcal{B}(n, p)$. On trouve alors :

$$f(x) = (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2}\right)$$

Toutefois, même pour ce cas binomial, les calculs sont un peu ardu, et vont au-delà de ce que la théorie générale prédit, car il s'agit du cas particulier des variables binomiales, qui ont des propriétés très fortes. Le seul résultat que l'on peut obtenir par le théorème central limite est :

$$\lim_{n \rightarrow \infty} P\left[Y - \frac{1}{2} \leq x\right] = \int_{-\infty}^x f(z) dz$$

qui est celui du théorème de De Moivre-Laplace. Cette remarque permet non seulement d'expliquer la relation entre fonction de probabilité de variables discrètes et densités de variables continues, mais aussi de motiver la définition suivante.

Définition

Soit X une variable aléatoire réelle, de densité f , telle que l'intégrale $\int_{-\infty}^{+\infty} |x| f(x) dx$ converge. L'**espérance mathématique** de la v.a.r. X est la quantité, notée $\mathbb{E}(X)$, donnée par :

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx.$$

4. Propriétés

Les propriétés suivantes s'obtiennent trivialement à partir des propriétés correspondantes pour les intégrales. Les nouveaux concepts en jeu, les moments notamment, sont essentiellement identiques à ceux définis précédemment pour les variables discrètes. Nous omettons donc les démonstrations. Les interprétations « intuitives » sont les mêmes.

 Fiche 136

► Espérance et valeur absolue

Soit X une variable aléatoire admettant une espérance $\mathbb{E}(X)$. Alors :

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|) < +\infty$$

► Positivité de l'espérance

Soit X une variable aléatoire réelle à valeurs positives, admettant une espérance mathématique $\mathbb{E}(X)$. Alors :

$$\mathbb{E}(X) \geq 0.$$

► Croissance de l'espérance

Soient X et Y deux variables aléatoires réelles telles que $X \leq Y$, admettant chacune une espérance mathématique ($\mathbb{E}(X)$ et $\mathbb{E}(Y)$ respectivement). Alors :

$$\mathbb{E}(X) \leq \mathbb{E}(Y)$$

► Linéarité de l'espérance

Soient X et Y deux variables aléatoires réelles, admettant chacune une espérance mathématique ($\mathbb{E}(X)$ et $\mathbb{E}(Y)$ respectivement). Alors :

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

► Transfert de l'espérance

Soient X une variable aléatoire réelle, de densité f , et une fonction réelle g , telle que l'intégrale $\int_{-\infty}^{+\infty} |g(x)| f(x) dx$ converge.

Alors $Y = g(X)$ est une variable aléatoire dont l'espérance est donnée par

$$\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(x) f(x) dx.$$

► Moment d'ordre n d'une v.a.r., $n \in \mathbb{N}$

Définition

Soient n un entier naturel non nul, et X une variable aléatoire réelle, de densité f , telle que l'intégrale $\int_{-\infty}^{+\infty} |x|^n f(x) dx$ converge. Le **moment (absolu) d'ordre n** de la v.a.r. X est la quantité donnée par :

$$\mathbb{E}[|X|^n] = \int_{-\infty}^{+\infty} |x|^n f(x) dx$$

► **Variance d'une v.a.r.**

Définition

Soit X une variable aléatoire réelle, de densité f , telle que l'intégrale $\int_{-\infty}^{+\infty} |x|^2 f(x) dx$ converge. La **variance** de la v.a.r. X est la quantité, notée $V(X)$, donnée par :

$$V(X) = \mathbb{E}[(X - \mathbb{E}(X))^2] = \int_{-\infty}^{+\infty} (x - \mathbb{E}(X))^2 f(x) dx$$

► **Formule de Koenig**

Soit X une v.a. réelle admettant une espérance mathématique et un moment d'ordre 2. Alors :

$$Var(X) = \mathbb{E}[X^2] - [\mathbb{E}(X)]^2 = \int_{-\infty}^{+\infty} x^2 f(x) dx - [\mathbb{E}(X)]^2.$$

► **Inégalité de Markov**

Soit X une variable aléatoire réelle, positive, admettant une espérance mathématique. Alors, pour tout réel $t > 0$:

$$P(X \geq t) \leq \frac{\mathbb{E}(X)}{t}$$

Démonstration : La démonstration qui suit s'avère plus simple que dans le cas discret. Il serait aussi possible d'unifier les cas discret et continu en utilisant l'intégrale de Stieltjes par rapport à une mesure donnée.

La v.a.r. X étant à valeurs positives, sa densité f est identiquement nulle sur $] -\infty, 0[$. On peut alors calculer, pour tout réel $t > 0$:

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{+\infty} x f(x) dx \\ &= \int_0^{+\infty} x f(x) dx \\ &= \int_0^t x f(x) dx + \int_t^{+\infty} x f(x) dx \\ &\geq \int_t^{+\infty} x f(x) dx \\ &\geq t \int_t^{+\infty} f(x) dx \\ &\geq t P(X \geq t) \end{aligned}$$

qui est le résultat cherché. ■

► **Inégalité de Tchebychev**

Soit X une v.a. admettant une variance $Var(X)$. Alors, pour tout réel $t \geq 0$:

$$P(|X - \mathbb{E}(X)| \geq t) \leq \frac{Var(X)}{t^2}$$

5. Fonction de survie d'une v.a.r. positive

Il est souvent utile de déterminer la probabilité qu'une variable positive dépasse un certain seuil ; si la variable peut être assimilée à une durée de vie, cette probabilité devient une chance de survie.

Définition

Soit X une v.a.r. positive. On appelle **fonction de survie**, la fonction G_X définie pour tout réel $x \geq 0$, par :

$$G_X(x) = P(X > x) = 1 - F_X(x).$$



G_X est décroissante et positive, avec :

$$\lim_{x \rightarrow +\infty} G_X(x) = 0$$

La positivité de X implique que $G_X(0) = 1$.

La notation anglo-saxonne pour la fonction de survie, y compris dans le domaine des assurances, est $S_X(x)$, ou, parfois ${}_x q_x$ dans le cas des assurances vie (la probabilité qu'un individu survive jusqu'à l'âge x).

6. Intervalle de confiance

En statistique, particulièrement dans l'étude des sondages, on a besoin du concept d'intervalle de confiance que nous définissons ci-dessous :

Définition

Soit α un réel de l'intervalle $[0, 1]$, et X une variable aléatoire réelle. On appelle **intervalle de confiance**, pour la v.a.r. X , tout intervalle I_α de \mathbb{R} tel que :

$$P(X \in I_\alpha) = 1 - \alpha$$

$1 - \alpha$, qui est, en général, exprimé sous forme de pourcentage, est le **degré de confiance** de l'intervalle considéré. On peut affirmer que l'on est certain à $100(1 - \alpha)$ pour cent que la variable X se trouve dans l'intervalle I_α .

En pratique, on parle souvent d'intervalle de confiance I_α de degré α si on est capable de démontrer que $P(X \in I_\alpha) \geq 1 - \alpha$. Ceci n'est évidemment pas une définition précise, mais permet d'affirmer qu'on est certain **au moins** à $100(1 - \alpha)$ % que la variable X se trouve dans l'intervalle I_α . Nous éviterons ici, lorsque cela est possible, ce type d'inégalité.

La probabilité α pouvant être répartie différemment de part et d'autre des bornes de l'intervalle de confiance, on écrit, en général, α sous la forme :

$$\alpha = \alpha_1 + \alpha_2 \quad , \quad (\alpha_1, \alpha_2) \in [0, 1]^2$$

Ainsi, $I_\alpha = [x_{\min}, x_{\max}]$, et :

$$P(X < x_{\min}) = \alpha_1$$

et :

$$P(X > x_{\max}) = \alpha_2$$

L'intervalle de confiance est dit :

- **bilatéral** si $\alpha_1 \neq 0$ et $\alpha_2 \neq 0$;
- **symétrique** si $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$;
- **disymétrique** si $\alpha_1 \neq \alpha_2$;
- **unilatéral** si $\alpha_1 = 0$ ou $\alpha_2 = 0$.



1. Par exemple, dans le cas où $\alpha_1 = \alpha$, $\alpha_2 = 0$, on a un intervalle de confiance unilatéral de la forme $[x_{\min}, +\infty[$, $x_{\min} \in \mathbb{R}$, ce qui veut dire que l'on est certain à $100(1 - \alpha)$ que $X \geq x_{\min}$.
2. Les valeurs du paramètre α les plus fréquentes sont 10 %, 5 %, 1 %, ce qui correspond donc respectivement, en termes de degré de confiance, à 90 %, 95 %, 99 %.
3. D'autre part, s'assurer que l'intervalle de confiance est le moins large possible permettra d'obtenir le maximum d'informations, en préférant les approximations précises du type $P(X \in I_\alpha) \approx 1 - \alpha$ à celles de la forme :

$$P(X \in I_\alpha) > 1 - \alpha$$

dont on ne contrôle pas la précision.

Les lois uniformes sont utiles pour modéliser des valeurs aléatoires qui sont bornées, mais pour lesquelles on ne dispose pas, a priori, d'autres informations.

Définition

Soient a et b deux réels tels que $a < b$. On dit qu'une v.a. réelle suit une **loi uniforme** sur l'intervalle $[a, b]$ si sa densité f est telle que, pour tout réel x :

$$f(x) = \frac{1}{b-a} \mathbb{I}_{[a,b]}(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{si } x \notin [a, b] \end{cases}$$

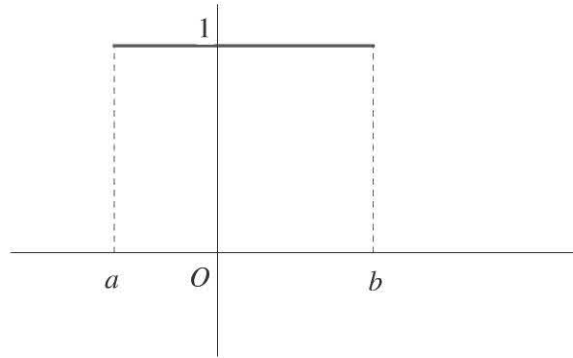


Figure 134.1 – La densité de la loi loi uniforme sur $[a, b]$.

1. Fonction de répartition pour la loi uniforme

Soient a et b deux réels tels que $a < b$. La fonction de répartition d'une v.a. suivant la loi uniforme sur l'intervalle $[a, b]$ est telle que, pour tout réel x :

$$F_X(x) = \int_{-\infty}^x f(t) dt = \frac{1}{b-a} \int_{-\infty}^x \mathbb{I}_{[a,b]}(t) dt = \begin{cases} 0 & \text{si } x \leq a \\ \frac{1}{b-a} \int_a^x \mathbb{I}_{[a,b]}(t) dt = \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ \frac{1}{b-a} \int_a^b \mathbb{I}_{[a,b]}(t) dt = 1 & \text{si } x \geq b \end{cases}$$

On peut le formuler différemment, en disant que le graphe de F_X est obtenu par interpolation linéaire de $(a, 0)$ à $(b, 1)$; ce graphe se doit d'être horizontal au-delà de (a, b) , par cohérence avec les limites en $-\infty$ et $+\infty$ et la croissance de F_X .

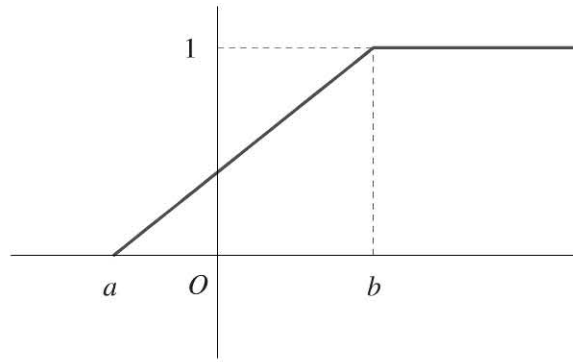


Figure 134.2 – La fonction de répartition de la loi uniforme sur $[a, b]$.

Propriété

Soit A un sous-intervalle de $[a, b]$ de longueur $|A|$. On a alors :

$$P(X \in A) = \frac{|A|}{b-a}.$$

En effet :

$$P(X \in A) = \int_a^b 1_A(t) \frac{1}{b-a} dt = \frac{1}{b-a} \int_a^b 1_A(t) dt = \frac{|A|}{b-a}.$$

2. Espérance d'une v.a.r. suivant la loi uniforme

Soient a et b deux réels tels que $a < b$, et X une v.a.r. suivant la loi uniforme sur l'intervalle $[a, b]$. L'espérance de X est :

$$\mathbb{E}(X) = \frac{a+b}{2}$$

Démonstration : Il est clair que $\int_{-\infty}^{+\infty} |x| f(x) dx = \frac{1}{b-a} \int_a^b |x| dx$ converge. On peut alors calculer :

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{+\infty} x f(x) dx \\ &= \frac{1}{b-a} \int_a^b x dx \\ &= \frac{1}{b-a} \frac{b^2 - a^2}{2} \\ &= \frac{a+b}{2} \end{aligned}$$

■

3. Variance d'une v.a.r. suivant la loi uniforme

Soient a et b deux réels tels que $a < b$, et X une v.a.r. suivant la loi uniforme sur l'intervalle $[a, b]$. La variance de X est :

$$V(X) = \frac{(b-a)^2}{12}$$

Démonstration : On calcule :

$$\begin{aligned}
 \mathbb{E}[X^2] &= \int_{-\infty}^{+\infty} x^2 f(x) dx \\
 &= \frac{1}{b-a} \int_a^b x^2 dx \\
 &= \frac{1}{b-a} \frac{b^3 - a^3}{3} \\
 &= \frac{1}{b-a} \frac{(b-a)(b^2 + ab + a^2)}{3} \\
 &= \frac{b^2 + ab + a^2}{3}
 \end{aligned}$$

La formule de Koenig permet alors d'en déduire :

$$\begin{aligned}
 V(X) &= \mathbb{E}[X^2] - [\mathbb{E}(X)]^2 \\
 &= \frac{b^2 + ab + a^2}{3} - \frac{(a+b)^2}{4} \\
 &= \frac{4(b^2 + ab + a^2) - 3(a+b)^2}{12} \\
 &= \frac{4(b^2 + ab + a^2) - 3a^2 - 3b^2 - 6ab}{12} \\
 &= \frac{a^2 + b^2 - 6ab}{12} \\
 &= \frac{(a-b)^2}{12} \\
 &= \frac{(b-a)^2}{12}
 \end{aligned}$$

■

4. La loi uniforme est stable par conditionnement

Soient a et b deux réels tels que $a < b$, et X une v.a.r. suivant la loi uniforme sur l'intervalle $[a, b]$. Soit $[c, d] \subset [a, b]$. Soit $P_{[c,d]}(\cdot)$ la loi conditionnelle de X sachant $\{X \in [c, d]\}$. Alors $P_{[c,d]}(\cdot)$ est la loi uniforme sur $[c, d]$.

Démonstration : On calcule, pour $x \in [c, d]$,

$$\begin{aligned}
 P_{[c,d]}(X \in]-\infty, x]) &= \frac{P(\{X \in [c, d]\} \cap \{X \in]-\infty, x])}{P(X \in [c, d])} \\
 &= \frac{P(X \in [c, x])}{P(X \in [c, d])} \\
 &= \frac{x - c}{b - a} \\
 &= \frac{b - a}{d - c} \\
 &= \frac{b - a}{x - c} \\
 &= \frac{b - a}{d - c}
 \end{aligned}$$

Cette fonction est la fonction de répartition de la loi uniforme sur $[c, d]$. Elle prend la valeur 1 pour $x = d$, et 0 pour $x = c$; ainsi, pour $x \notin [c, d]$, on récupère également les bonnes valeurs. ■



Le coût d'un déminage avec la loi uniforme

Un groupe terroriste a placé une mine sur une route de 100 km de long entre deux villages isolés dans le Sahara occidental ; un militant du groupe émet un appel téléphonique pour avertir les autorités de la présence de la mine, avant que personne n'ait emprunté la route, mais ne donne pas plus d'informations sur l'emplacement de celle-ci. Les autorités envoient une équipe de déminage pour désamorcer l'engin explosif. On désigne par X la distance à parcourir jusqu'à l'emplacement de la mine.

Il est prudent de supposer que X est une variable uniforme sur l'intervalle $[0, 100]$. On calcule, en particulier, l'espérance que l'emplacement de la mine soit :

$$\frac{100 + 0}{2} = 50$$

alors que l'écart-type de cet emplacement est :

$$\sqrt{\frac{(100 - 0)^2}{12}} = \frac{100}{2\sqrt{3}} \approx 28,87$$

La probabilité que les démineurs n'aient pas trouvé la mine au bout de 90 km est :

$$P(X > 90) = P(X \in [90, 100]) = \frac{100 - 90}{100 - 0} = 0,1.$$

Sachant que les démineurs n'ont toujours pas trouvé la mine au bout de 90 km, la probabilité qu'ils la trouvent avant le km 94 est celle qu'une variable Y uniforme sur $[90, 100]$ soit inférieure à 94 :

$$P(X < 94 | X > 90) = P(90 \leq Y < 94) = \frac{94 - 90}{100 - 90} = 0,4.$$

On peut aussi calculer le coût moyen associé au travail des démineurs. Les autorités estiment le que coût total Z est égal à 5 000 MAD (dirhams) de coût fixe, auquel il faut ajouter 100 MAD par kilomètre parcouru, ainsi qu'un terme égal à X^2 MAD pour prendre en compte le stress accumulé lors de missions longues, ce qui conduit donc à :

$$Z = 5000 + 100X + X^2$$

Par linéarité de l'espérance, et grâce à la formule de Koenig, on obtient :

$$\mathbb{E}(X) = 50$$

$$\mathbb{E}(X^2) = \text{Var}(X) + \mathbb{E}(X)^2 = \frac{10\,000}{12} + 50^2 = \frac{10\,000}{3},$$

Le coût moyen de l'opération de déminage est donc :

$$\begin{aligned} \mathbb{E}(Z) &= 5\,000 + 100\mathbb{E}(X) + (\mathbb{E}(X))^2 \\ &= 5\,000 + 100 \times 50 + \frac{10\,000}{3} \\ &= \frac{40\,000}{3} \\ &\approx 13\,333 \text{ MAD} \end{aligned}$$

i.e. environ 1 200 euros.

Les lois exponentielles sont très proches, en esprit, des lois géométriques. Elles sont utilisées pour modéliser des temps d'attente, avec la propriété d'absence de mémoire, i.e. dans le cas où la connaissance de la durée d'attente ne donne aucune information sur la durée restante.

Un exemple typique est la durée de vie d'un atome radioactif avant sa désintégration.

1. Loi exponentielle (de paramètre $\lambda \in \mathbb{R}_+^*$)

Définition

Soit λ un réel strictement positif. Une v.a. réelle X suit la **loi exponentielle de paramètre λ** si sa densité f est telle que, pour tout réel x :

$$f(x) = \lambda e^{-\lambda x} \mathbb{I}_{[0, +\infty[}(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$



On note : $X \sim \mathcal{E}(\lambda)$

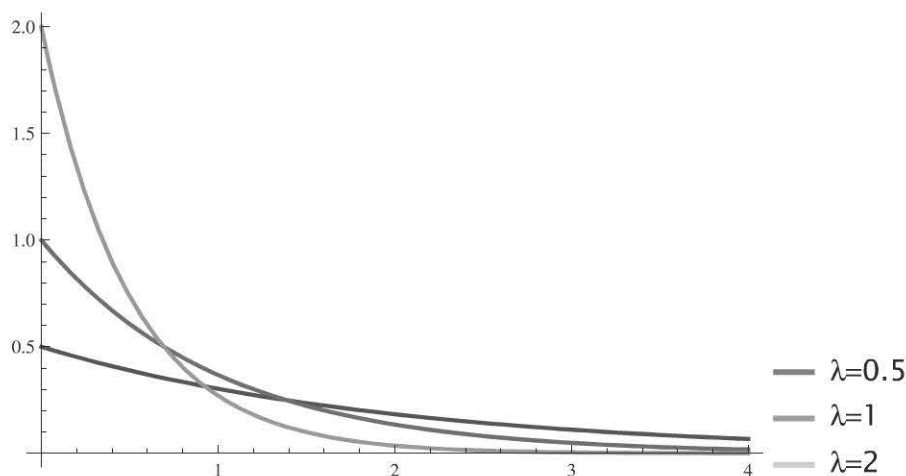


Figure 135.1 – La densité de la loi exponentielle, pour différentes valeurs du paramètre λ .

2. Fonction de répartition pour la loi exponentielle de paramètre $\lambda > 0$

Soit λ un réel strictement positif. La fonction de répartition d'une v.a. suivant la loi exponentielle de paramètre λ est telle que, pour tout réel x :

$$F_X(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x \lambda e^{-\lambda t} \mathbb{I}_{[0, +\infty[}(t) dt = \begin{cases} 0 & \text{si } x \leq 0 \\ \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$$

Exemple : Demi-vie d'un élément radioactif

Calculons, pour un élément radioactif de temps de vie $X \sim \mathcal{E}(\lambda)$, $\lambda \in \mathbb{R}_+^*$, la relation entre sa demi-vie et le paramètre λ .

Cette demi-vie est le temps nécessaire pour que l'intensité de la radioactivité diminue de moitié. Si on considère que la substance radioactive est formée d'un grand nombre d'atomes radioactifs dont les durées de vie sont indépendantes, on peut alors assimiler la demi-vie $t_{\frac{1}{2}}$ à la durée correspondant à 50 % des chances de vie, i.e.

$$P(X > t_{\frac{1}{2}}) = \frac{1}{2}$$

On obtient :

$$\frac{1}{2} = e^{-\lambda t_{\frac{1}{2}}}$$

ce qui conduit à :

$$t_{\frac{1}{2}} = \frac{\ln 2}{\lambda}$$

3. Fonction de survie

Soient λ un réel strictement positif et $X \sim \mathcal{E}(\lambda)$. La **fonction de survie** G_X de la loi de X vaut, pour tout réel $x \geq 0$:

$$G_X(x) = P(X > x) = 1 - F_X(x) = e^{-\lambda x}.$$

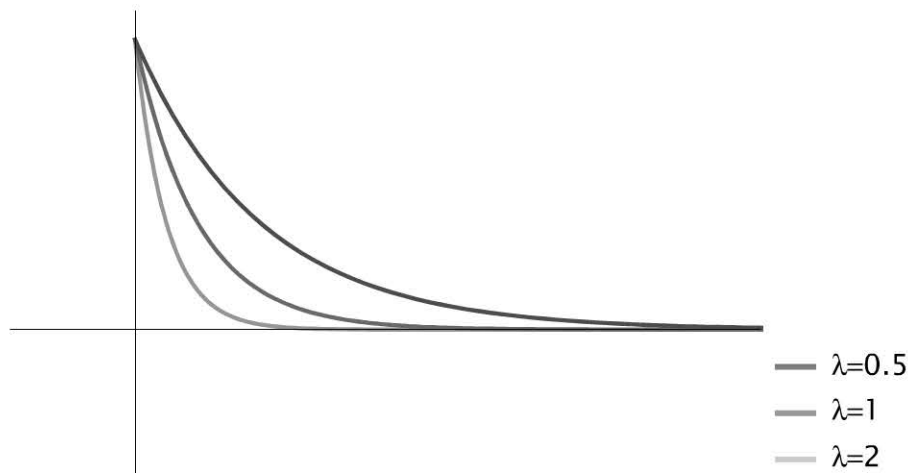


Figure 135.2 – La fonction de survie de la loi exponentielle, pour différentes valeurs du paramètre λ .

4. Absence de mémoire

Soient λ un réel strictement positif, et X une v.a. réelle positive à densité. Alors, X suit la loi exponentielle de paramètre λ si et seulement si elle vérifie la propriété d'absence de mémoire :

$$\forall s \geq 0, \forall t \geq 0 : P(X > t + s | X > t) = P(X > s)$$

Démonstration : On ne démontrera ici que la première implication. La réciproque est démontrée en remarque ci-dessous, en supposant de surcroît que la densité de la v.a.r. est continue.

Pour tout couple de réels positifs (s, t) :

$$P(X \geq t + s | X \geq t) = \frac{P(X > t + s \text{ et } X > t)}{P(X > t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s} = P(X > s) \quad \blacksquare$$



Pour démontrer la réciproque, considérons une v.a.r. positive X , de densité f_X continue sur \mathbf{R}_+ (f doit être nulle sur \mathbf{R}_-). On suppose que la propriété d'absence de mémoire est vérifiée. La fonction de survie G_X de X est de classe C^1 sur \mathbf{R}_+ , et :

$$\frac{dG_X}{dx} = f_X(x)$$

sur \mathbf{R}_+ . On a, pour tout $s \geq 0$ et tout $t \geq 0$:

$$\frac{P(\{X > t + s\} \cap \{X > t\})}{P(X > t)} = P(X > s)$$

ce qui s'écrit aussi :

$$G_X(t + s) = G_X(s) G_X(t)$$

En posant $s = \varepsilon > 0$, on en déduit :

$$\frac{G_X(t + \varepsilon) - G_X(t)}{\varepsilon} = \frac{G_X(t)(G_X(\varepsilon) - 1)}{\varepsilon}$$

Comme G_X est dérivable en 0 et en t , on en déduit :

$$\lim_{\varepsilon \rightarrow 0} \frac{G_X(\varepsilon) - 1}{\varepsilon} = f_X(0) \quad , \quad \lim_{\varepsilon \rightarrow 0} \frac{G_X(t + \varepsilon) - G_X(t)}{\varepsilon} = \frac{dG_X}{dx}(t)$$

On a donc montré que, pour tout $x \geq 0$,

$$\frac{dG_X}{dx}(x) = f_X(0) G_X(x)$$

On obtient ainsi une équation différentielle ordinaire qui admet pour unique solution, avec la condition initiale, $G_X(0) = 1$, la fonction G_X telle que, pour tout réel positif x :

$$G_X(x) = e^{-f_X(0)x}$$

En posant $\lambda = f_X(0)$, on obtient pour tout $x \geq 0$,

$$f_X(x) = \lambda e^{-\lambda x}$$

qui est le résultat cherché.

5. Espérance d'une v.a.r. suivant la loi exponentielle $\mathcal{E}(\lambda)$

Soient λ un réel strictement positif, et X une v.a.r. suivant la loi exponentielle $\mathcal{E}(\lambda)$. L'espérance de X est :

$$\mathbb{E}(X) = \frac{1}{\lambda}$$

Démonstration : Il est clair que $\int_{-\infty}^{+\infty} |x| f_{\lambda}(x) dx = \int_0^{+\infty} |x| \lambda e^{-\lambda x} dx$ converge. Une intégration par parties conduit à :

$$\begin{aligned} \int_{-\infty}^{+\infty} x f_{\lambda}(x) dx &= \int_0^{+\infty} x \lambda e^{-\lambda x} dx \\ &= \left[-x e^{-\lambda x} \right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} dx \\ &= \left[-\frac{e^{-\lambda x}}{\lambda} \right]_0^{+\infty} \\ &= \frac{1}{\lambda} \end{aligned}$$



Cette formule montre que le paramètre λ peut être interprété de façon similaire au paramètre p de la loi géométrique : c'est l'inverse du temps d'attente moyen, représenté par $\mathbb{E}(X)$. Ces deux paramètres peuvent, aussi, être assimilés à des fréquences d'arrivée.

6. Variance d'une v.a.r. suivant la loi exponentielle $\mathcal{E}(\lambda)$

Soient λ un réel strictement positif, et X une v.a.r. suivant la loi exponentielle $\mathcal{E}(\lambda)$. La variance de X est :

$$V(X) = \frac{1}{\lambda^2}$$

Démonstration : Il est clair que $\int_{-\infty}^{+\infty} x^2 f_{\lambda}(x) dx = \int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx$ converge. On peut alors utiliser deux intégrations par parties pour calculer :

$$\begin{aligned} \mathbb{E}[X^2] &= \int_{-\infty}^{+\infty} x^2 f_{\lambda}(x) dx \\ &= \int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx \\ &= \left[-x^2 e^{-\lambda x} \right]_0^{+\infty} + \int_0^{+\infty} 2x e^{-\lambda x} dx \\ &= \left[-\frac{2x e^{-\lambda x}}{\lambda} \right]_0^{+\infty} + \int_0^{+\infty} \frac{2 e^{-\lambda x}}{\lambda} dx \\ &= \int_0^{+\infty} \frac{2 e^{-\lambda x}}{\lambda} dx \\ &= \left[-\frac{2 e^{-\lambda x}}{\lambda^2} \right]_0^{+\infty} \\ &= \frac{2}{\lambda^2} \end{aligned}$$

La variance de la v.a.r. X s'obtient alors grâce à la formule de Koenig :

$$V(X) = \mathbb{E}[X^2] - (\mathbb{E}(X))^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Ces lois sont, peut-être, les plus importantes de toute la théorie des probabilités. Nous verrons, en particulier, le théorème central limite, qui prouve que, pour toute suite i.i.d. de variables $(X_n)_{n \in \mathbb{N}}$, d'espérance nulle et de variance σ^2 finie, la moyenne empirique $M_n = \frac{1}{n} \sum_{k=1}^n X_k$, qui tend vers 0 (cf. la loi des grands nombres), a une limite normale de variance σ^2 après qu'on l'ait multipliée par \sqrt{n} . Le caractère universel de ce théorème, ne dépendant de la loi des X_n qu'au travers de leur variance, est ce qui rend essentielles les lois normales, et explique pourquoi tant de phénomènes naturels et sociaux se modélisent bien avec des lois normales.

1. Loi normale (de paramètres m, σ)

Définition

Soient μ et $\sigma > 0$ deux réels. Une v.a. réelle X suit la **loi normale (de paramètres μ, σ)**, $N(\mu, \sigma)$ (aussi appelée loi gaussienne de paramètres m, σ) si sa densité $f_{\mu, \sigma}$ est telle que, pour tout réel x :

$$f_{\mu, \sigma}(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



On note :

$$X \sim N(\mu, \sigma^2)$$

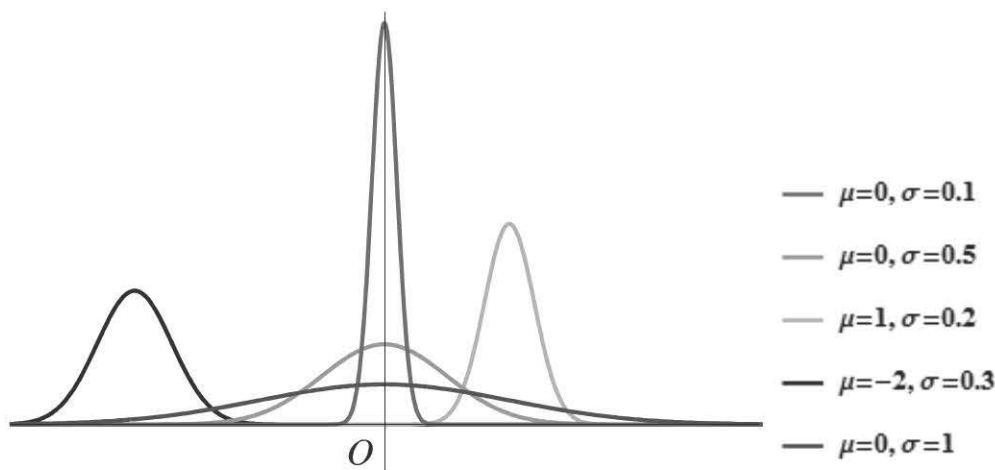


Figure 136.1 – La densité de la loi normale, pour différentes valeurs des paramètres μ et σ .

► Loi normale centrée réduite

Définition

On appelle **loi normale centrée réduite** la loi normale de paramètres $\mu = 0$, $\sigma = 1$, $\mathcal{N}(0, 1)$, dont la densité est donc définie, pour tout réel x , par :

$$f_{0,1}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

► Intégrale de Gauss

Définition

On appelle **intégrale de Gauss** l'intégrale convergente :

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$



Par parité :

$$\int_0^{+\infty} e^{-x^2} dx = \frac{1}{2} \int_{-\infty}^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

Proposition

Soient μ et $\sigma > 0$ deux réels, et X une v.a.r. suivant la loi normale $\mathcal{N}(\mu, \sigma)$. La v.a.r. $Y = \frac{X - \mu}{\sigma}$ suit la loi normale centrée réduite $\mathcal{N}(0, 1)$.

Démonstration : Soient a et b deux réels tels que $a \leq b$. Alors :

$$\begin{aligned} P(a < Y \leq b) &= P\left(a < \frac{X - \mu}{\sigma} \leq b\right) \\ &= P(\sigma a < X - \mu \leq \sigma b) \\ &= P(\sigma a + \mu < X \leq \sigma b + \mu) \\ &= \int_{\sigma a + \mu}^{\sigma b + \mu} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{y^2}{2}} dy \end{aligned}$$

(on effectue, tout simplement, le changement de variable $y = \frac{x - \mu}{\sigma}$.) ■

► Loi normale et transformation affine

Soient μ et $\sigma > 0$ deux réels, et X une v.a.r. suivant la loi normale de paramètres μ, σ , $\mathcal{N}(\mu, \sigma)$. Alors, pour tout couple de réels $(\alpha, \beta) \in \mathbb{R}^* \times \mathbb{R}$, la v.a.r. $\alpha X + \beta$ suit la loi normale de paramètres $\alpha\mu + \beta, \alpha^2\sigma^2$:

$$\alpha X + \beta \sim \mathcal{N}(\alpha\mu + \beta, \alpha^2\sigma^2)$$

Démonstration : Soient a et b deux réels tels que $a \leq b$. Alors :

$$\begin{aligned}
 P(a < \alpha X + \beta \leq b) &= P(a - \beta < \alpha X \leq b - \beta) \\
 &= P\left(\frac{a - \beta}{\alpha} < X \leq \frac{b - \beta}{\alpha}\right) \\
 &= \int_{\frac{a - \beta}{\alpha}}^{\frac{b - \beta}{\alpha}} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x - \mu)^2}{2\sigma^2}} dx \\
 &= \int_{a - \beta}^{b - \beta} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(\frac{x}{\alpha} - \mu)^2}{2\sigma^2}} \frac{1}{\alpha} dx \\
 &= \int_{a - \beta}^{b - \beta} \frac{1}{\alpha \sigma \sqrt{2\pi}} e^{-\frac{(x - \alpha\mu)^2}{2\alpha^2\sigma^2}} dx \\
 &= \int_a^b \frac{1}{\alpha \sigma \sqrt{2\pi}} e^{-\frac{(x - \alpha\mu - \beta)^2}{2\alpha^2\sigma^2}} dx
 \end{aligned}$$

■

Exemple : Loi normale et intervalle de confiance

Soient μ et $\sigma > 0$ deux réels, α un réel de l'intervalle $[0, 1]$, et X une v.a.r. suivant la loi normale $\mathcal{N}(\mu, \sigma^2)$. La v.a. centrée réduite $\frac{X - \mu}{\sigma}$ suit donc la loi normale standard $\mathcal{N}(0, 1)$. On cherche à déterminer l'intervalle de confiance symétrique I_α tel que :

$$P\left(\frac{X - \mu}{\sigma} \in I_\alpha\right) = 1 - \alpha$$

(c'est la symétrie, par rapport à l'axe des ordonnées, de la courbe représentative de la fonction $x \mapsto \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$ qui incite à rechercher un intervalle de confiance symétrique.)

À cet effet, on cherche I_α sous la forme :

$$I_\alpha = [-x_\alpha, x_\alpha] \quad , \quad x_\alpha \in \mathbb{R}_+^*$$

Pour que I_α soit symétrique, il faut donc :

$$P(X < -x_\alpha) = \frac{\alpha}{2} \quad , \quad P(X > x_\alpha) = \frac{\alpha}{2}$$

ce qui donne, pour les degrés de confiance usuels 90 %, 95 %, et 99 % :

$$x_{90\%} \simeq 1,645 \quad , \quad x_{95\%} \simeq 1,96 \quad , \quad x_{99\%} \simeq 2,58$$

On peut alors revenir à la v.a.r. initiale X :

$$\begin{aligned}
 P\left(\frac{X - \mu}{\sigma} \in I_\alpha\right) &= P\left(\frac{X - \mu}{\sigma} \in [-x_\alpha, x_\alpha]\right) \\
 &= P(X - \mu \in [-\sigma x_\alpha, \sigma x_\alpha]) \\
 &= P(X \in [\mu - \sigma x_\alpha, \mu + \sigma x_\alpha])
 \end{aligned}$$

L'intervalle de confiance pour la v.a.r. X est donc :

$$[\mu - \sigma x_\alpha, \mu + \sigma x_\alpha]$$

► **Espérance d'une v.a.r. suivant la loi normale $\mathcal{N}(\mu, \sigma^2)$**

Soient μ et $\sigma > 0$ deux réels, et X une v.a.r. suivant la loi normale $\mathcal{N}(\mu, \sigma^2)$. L'espérance de X est :

$$\mathbb{E}(X) = \mu$$

Démonstration : Il est clair que $\int_{-\infty}^{+\infty} |x| f_{\mu, \sigma}(x) dx = \int_{-\infty}^{+\infty} \frac{|x|}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$ converge. On peut alors calculer :

$$\begin{aligned} \int_{-\infty}^{+\infty} x f_{\mu, \sigma}(x) dx &= \int_{-\infty}^{+\infty} \frac{x}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \int_{-\infty}^{+\infty} \frac{x+\mu}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= \int_{-\infty}^{+\infty} \frac{x}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \\ &\quad + \int_{-\infty}^{+\infty} \frac{\mu}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= \int_{-\infty}^{+\infty} \frac{\sigma y}{\sigma \sqrt{2\pi}} e^{-\frac{y^2}{2}} \sigma dy + \int_{-\infty}^{+\infty} \frac{\mu}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= \sigma \int_{-\infty}^{+\infty} \frac{y}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy + \int_{-\infty}^{+\infty} \frac{\mu}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= \int_{-\infty}^{+\infty} \frac{\mu}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= \frac{\mu}{\sigma \sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= \frac{\mu}{\sigma \sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-x^2} \sqrt{2} \sigma dx \\ &= \frac{\mu}{\sigma \sqrt{2\pi}} \sqrt{2} \sigma \sqrt{\pi} \\ &= \mu \end{aligned}$$

puisque :

$$\int_{-\infty}^{+\infty} \frac{y}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \lim_{Y \rightarrow -\infty} \lim_{Z \rightarrow +\infty} \left[e^{-\frac{y^2}{2}} \right]_Y^Z = \lim_{Z \rightarrow +\infty} e^{-\frac{Z^2}{2}} - \lim_{Y \rightarrow -\infty} e^{-\frac{Y^2}{2}} = 1 - 1 = 0$$

■

► **Variance d'une v.a.r. suivant la loi normale $\mathcal{N}(\mu, \sigma^2)$**

Soient μ et $\sigma > 0$ deux réels, et X une v.a.r. suivant la loi normale $\mathcal{N}(\mu, \sigma^2)$. La variance de X est :

$$V(X) = \sigma^2$$

Démonstration : Il est clair que $\int_{-\infty}^{+\infty} x^2 f_{\mu,\sigma}(x) dx = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$ converge. On peut alors calculer :

$$\begin{aligned} \mathbb{E}[X^2] &= \int_{-\infty}^{+\infty} x^2 f_{\mu,\sigma}(x) dx \\ &= \int_{-\infty}^{+\infty} \frac{x^2}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \int_{-\infty}^{+\infty} \frac{(x+\mu)^2}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= \int_{-\infty}^{+\infty} \frac{x^2 + 2\mu x + \mu^2}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= \int_{-\infty}^{+\infty} \frac{x^2}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx + \int_{-\infty}^{+\infty} \frac{2\mu x}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx + \int_{-\infty}^{+\infty} \frac{\mu^2}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= \left[-\frac{\sigma^2 x}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \frac{\sigma^2}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \\ &+ \left[-\frac{\sigma^2 \mu}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \frac{\mu^2}{\sigma \sqrt{2\pi}} e^{-x^2} \sqrt{2}\sigma dx \\ &= \int_{-\infty}^{+\infty} \frac{(\sigma^2 + \mu^2)}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} dx \\ &= \int_{-\infty}^{+\infty} \frac{(\sigma^2 + \mu^2)}{\sigma \sqrt{2\pi}} e^{-x^2} \sqrt{2}\sigma dx \\ &= \int_{-\infty}^{+\infty} \frac{(\sigma^2 + \mu^2)}{\sqrt{\pi}} e^{-x^2} dx \\ &= \frac{(\sigma^2 + \mu^2)}{\sqrt{\pi}} \sqrt{\pi} \\ &= \sigma^2 + \mu^2 \end{aligned}$$

La variance de la v.a.r. X s'obtient alors grâce à la formule de Koenig :

$$V(X) = \mathbb{E}[X^2] - (\mathbb{E}(X))^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2 \quad \blacksquare$$

2. Théorème central limite¹

Théorème

Soient μ et $\sigma > 0$ deux réels, et $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires indépendantes, de même loi, de moyenne μ , de variance σ^2 .

Soit $(M_n)_{n \in \mathbb{N}^*}$ la suite des moyennes empiriques (arithmétiques), c'est-à-dire, pour tout entier naturel non nul n :

$$M_n = \frac{X_1 + \dots + X_n}{n} = \frac{1}{n} \sum_{k=1}^n X_k$$

1. C'est le mathématicien et physicien français Pierre-Simon de Laplace (1749-1827), qui publia, en 1809, la première démonstration du théorème. Mais le cas particulier où les variables suivent la loi de Bernoulli de paramètre $p = \frac{1}{2}$ avait déjà été étudié par Abraham De Moivre (1667-1754) en 1733.

et $(Z_n)_{n \in \mathbb{N}^*}$ la suite des v.a.r. centrées réduites définies pour tout entier naturel non nul n par :

$$Z_n = \frac{M_n - \mathbb{E}(M_n)}{\sqrt{\text{Var}(M_n)}} = \frac{M_n - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(M_n - \mu)}{\sigma}$$

Pour tout intervalle $[a, b]$ de \mathbb{R} ($a < b$), on a :

$$\lim_{n \rightarrow +\infty} P(a \leq Z_n \leq b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

On dit aussi que la suite de v.a.r. $(Z_n)_{n \in \mathbb{N}^*}$ converge en loi vers la loi normale standard $\mathcal{N}(0, 1)$, ou que la loi de la suite $(Z_n)_{n \in \mathbb{N}^*}$ converge vers $\mathcal{N}(0, 1)$ (en loi).



Le théorème central limite est un résultat essentiel en théorie des probabilités. Ce résultat est extrêmement puissant, il permet de dire que toute somme de variables aléatoires indépendantes et identiquement distribuées tend vers une variable aléatoire gaussienne, après normalisation. Il traduit aussi le fait que, dans la loi des grands nombres, les fluctuations autour de la limite de la moyenne sont de l'ordre de $\frac{1}{\sqrt{n}}$ (c'est une idée que nous avons déjà vue, dans une certaine mesure, dans le cadre de l'application de l'inégalité de Tchebyshev), mais, surtout, que la loi de ces fluctuations est universelle, dans la mesure où elle suit la loi normale, et ne dépend donc pas de la loi initialement suivie par les v.a. considérées.

Pour des variables de moyenne nulle, le seul élément qui subsiste de cette loi initiale est leur variance, et, du fait de la propriété de changement d'échelle de la variance, si on divise au préalable les variables par leur écart-type, on obtient un résultat universel, ne dépendant d'aucun paramètre.

Soit n un entier naturel très grand devant 1. Si on considère n v.a. X_1, \dots, X_n , de même loi, ayant une espérance μ et une variance σ^2 finie, la loi de leur moyenne empirique $\frac{X_1 + \dots + X_n}{n}$ peut être approximée par la loi normale $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$. On constate que la variance de cette approximation de la moyenne empirique est inversement petite d'autant que n est grand. C'est le phénomène de « concentration de la mesure », que nous avons observé en partie avec la loi des grands nombres, et que l'on retrouve ici, puisque la loi $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$ converge vers celle de la variable aléatoire constante égale à μ .

Pour s'assurer qu'on utilise le théorème central limite de manière efficace, il est utile de convertir le problème pratique à traiter pour qu'il dépende de la quantité :

$$Z_n = \frac{X_1 + \dots + X_n - n\mu}{\sigma \sqrt{n}}.$$

C'est, en effet, cette quantité Z_n qui, d'après le théorème central limite, a (approximativement) une loi universelle normale centrée réduite. L'opération qui consiste à passer de la suite $(X_i)_{i=1, \dots, n}$ à la variable Z_n est précisément la standardisation de sa somme partielle, i.e. une opération qui consiste à soustraire la moyenne de sa somme partielle, puis de la diviser par son écart-type.



Le lecteur assidu se demandera dans quelle mesure l'approximation du théorème central limite représente une bonne approximation. Il s'avère que, pour les variables qui ont aussi un troisième moment fini $\rho = \mathbf{E}[|X_1|^3] < +\infty$, le théorème dit « de Berry-Esseen », donne une réponse à cette question :

$$\left| P(a \leq Z_n \leq b) - \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \right| \leq \frac{C\rho}{\sigma^3 \sqrt{n}}$$

où la constante universelle C n'est pas connue de manière optimale, mais se situe entre 0,41 et 0,48. Ce résultat implique, par exemple, que si les probabilités calculées grâce au théorème central limite sont comprises entre 5 % et 95 %, pour que l'ordre de grandeur dans l'approximation soit au moins d'un ordre de grandeur inférieur à ces niveaux, i.e. pas plus de $\frac{1}{200}$, il est préférable de choisir n de l'ordre d'une dizaine de milliers. En pratique, on applique le théorème beaucoup plus fréquemment, souvent dès que $n \geq 100$, ce qui peut être abusif, puisque des probabilités de l'ordre de 5 % seraient alors nettement plus faibles que l'erreur commise si on appliquait le théorème central limite. Si on suppose que $C\rho\sigma^{-3}$ est de l'ordre de 1, pour s'assurer au moins que le niveau 5 % est au moins du même ordre que l'erreur dans le théorème de Berry-Esseen, alors il est préférable de prendre $n \geq 400$.

3. Théorème de De Moivre-Laplace^{2,3}

Théorème

Soit p un réel de l'intervalle $[0, 1]$, et X une variable aléatoire, de loi binomiale $\mathcal{B}(n, p)$. La loi de la v.a. :

$$Y = n^{-\frac{1}{2}} (X - np)$$

peut être approchée, lorsque l'entier n est très grand, par la loi normale $\mathcal{N}(0, p(1-p))$.

Démonstration : La démonstration de ce résultat est immédiate avec le théorème central limite, puisque X peut être considérée comme la somme de n variables indépendantes de loi $\mathcal{B}(p)$, d'espérance p , de variance $p(1-p)$. ■



La qualité de l'approximation proposée par le théorème de de Moivre-Laplace est nettement meilleure que celle du théorème de Berry-Esseen. En pratique, il est parfois légitime, pour obtenir des erreurs inférieures à 1 %, de s'assurer que l'entier n est plus grand que 30, mais il est plus sage de prendre $n \geq 100$, même si on est dans le cas de variables binomiales.

2. Abraham De Moivre (1667-1754). Après l'impulsion donnée par le suisse Jacob Bernoulli (1654-1705), De Moivre est un des premiers vrais « probabilistes ».

3. Pierre-Simon de Laplace (1749-1827). Il était non seulement mathématicien, spécialiste de calcul différentiel et intégral, et apporta de nombreuses contributions en théorie des probabilités, tout en étant aussi astronome et physicien.

Couples de variables aléatoires réelles

Avant de donner plus de résultats sur les variables gaussiennes, et d'introduire d'autres lois, il est important de présenter, pour le cas continu, les concepts rencontrés dans le cas discret.

1. Densité de probabilité sur \mathbb{R}^2

Définition

Une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, à valeurs positives, intégrable (i.e. $\iint_{\mathbb{R}^2} f(x, y) dx dy$ existe) est appelée **densité de probabilité sur \mathbb{R}^2** si et seulement si :

$$\iint_{\mathbb{R}^2} f(x, y) dx dy = 1$$

► Fonction de répartition d'un couple de v.a.r.

Soit (X, Y) un couple de variables aléatoires réelles. La **fonction de répartition** du couple de v.a. (X, Y) est la fonction, notée $F_{(X, Y)}$, telle que :

$$\begin{aligned} F_{(X, Y)} : \mathbb{R}^2 &\rightarrow [0, 1] \\ (x, y) &\mapsto P(X \leq x \text{ et } Y \leq y). \end{aligned}$$

Elle vérifie, pour tout couple de réels (a, b) :

$$\lim_{y \rightarrow +\infty} F_{(X, Y)}(a, y) = F_X(a) \quad , \quad \lim_{x \rightarrow +\infty} F_{(X, Y)}(x, b) = F_Y(b)$$

On dit que (X, Y) admet la **densité f** (ou **densité jointe f**) si f est une densité de probabilité, et si :

$$F_{X, Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x', y') dy' dx'.$$

► Densités marginales de deux v.a.r.

Soit (X, Y) un couple de variables aléatoires réelles, de densité de probabilité $f_{(X, Y)}$. Alors, les composantes X et Y sont des variables aléatoires à densité, dont les densités de probabilités respectives sont :

$$f_X(x) = \int_{-\infty}^{+\infty} f_{(X, Y)}(x, y) dy \quad , \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{(X, Y)}(x, y) dx$$

$f_X(x)$ est la densité marginale de la v.a.r. X , $f_Y(y)$ est la densité marginale de la v.a.r. Y .

Démonstration : Il faut montrer que les fonctions $x \mapsto f_X(x)$ et $y \mapsto f_Y(y)$ sont bien des densités de probabilité sur \mathbb{R} . À cet effet, on calcule :

$$P(X \leq x) = P((X, Y) \in]-\infty, x] \times \mathbb{R}) = \int_{-\infty}^x \left\{ \int_{-\infty}^{+\infty} f_{(X, Y)}(u, y) dy \right\} du = \int_{-\infty}^x f_X(u) du$$

et :

$$\int_{-\infty}^{+\infty} f_X(x) dx = \iint_{\mathbb{R}^2} f(x, y) dx dy = 1.$$

On procède de même pour f_Y . ■

► Densité conditionnelle de deux v.a.r., loi conditionnelle de deux v.a.r.

Soit (X, Y) un couple de variables aléatoires réelles, de densité de probabilité $f_{(X,Y)}$. On désigne par f_X et f_Y les densités marginales respectives de X et Y . Soit y un réel tel que $f_Y(y) > 0$. La **densité conditionnelle de X sachant que $Y = y$** est définie lorsque $f_X(x) > 0$, pour tout réel x , par :

$$f_{X|Y=y}(x) = \frac{f_{(X,Y)}(x, y)}{f_Y(y)}$$

Étant donnés deux réels a et b tels que $a < b$, la **loi conditionnelle de X sachant que $Y = y$** est définie par :

$$P(X \in [a, b] | Y = y) = \int_a^b f_{X|Y=y}(x) dx = \int_a^b \frac{f_{(X,Y)}(x, y)}{f_Y(y)} dx$$

De même, la **densité conditionnelle de Y sachant que $X = x$** est définie, pour tout réel y , par :

$$f_{Y|X=x}(y) = \frac{f_{(X,Y)}(x, y)}{f_X(x)}$$

Étant donnés deux réels a et b tels que $a < b$, la **loi conditionnelle de Y sachant que $X = x$** est définie par :

$$P(Y \in [a, b] | X = x) = \int_a^b f_{Y|X=x}(y) dy = \int_a^b \frac{f_{(X,Y)}(x, y)}{f_X(x)} dy$$



Le lecteur pourra se référer à la discussion précédant la définition de l'espérance d'une v.a.r. continue, qui permet de motiver la définition de la densité conditionnelle. En bref, pour un couple (U, V) de variables aléatoires discrètes prenant comme valeurs les paires d'entiers $((i, j); i, j = 1, \dots, n)$ et $(X, Y) = \left(\frac{U}{n}, \frac{V}{n}\right)$, l'expression $f_{(X,Y)}(x, y)$ représente la limite, sous réserve d'existence bien sûr, de :

$$n^2 P\left(X = \frac{[xn]}{n}, Y = \frac{[yn]}{n}\right)$$

alors que $\frac{f_{(X,Y)}(x, y)}{f_X(x)}$ est celle de :

$$n \frac{P\left(X = \frac{[xn]}{n}, Y = \frac{[yn]}{n}\right)}{P\left(X = \frac{[xn]}{n}\right)}.$$

2. Indépendance de deux v.a.r.

Définition

Deux v.a.r. X et Y sont dites indépendantes si, pour tout (x, y) de \mathbb{R}^2 :

$$P(X \leq x, Y \leq y) = P(X \leq x) P(Y \leq y)$$

► Caractérisation de l'indépendance de deux v.a.r. à densité

Deux v.a.r. X et Y , de densités de probabilités respectives f_X et f_Y sont indépendantes si et seulement si le couple (X, Y) a une densité $f_{(X,Y)}$ et que, pour tout (x, y) de \mathbb{R}^2 :

$$f_{(X,Y)}(x, y) = f_X(x) f_Y(y).$$

Exemples

1. On considère un couple de variables aléatoires (X, Y) , dont la densité jointe est donnée par :

$$f_{X,Y}(x, y) = \begin{cases} 2e^{-2x-y} & \text{si } x \geq 0 \text{ et } y \geq 0, \\ 0 & \text{si } x < 0 \text{ ou } y < 0. \end{cases}$$

On peut écrire :

$$f_{X,Y}(x, y) = f_X(x) f_Y(y)$$

où f_X et f_Y désignent, respectivement, les densités de la loi exponentielle $\mathcal{E}(2)$ et $\mathcal{E}(1)$. Ce sont donc les lois respectives des variables X et Y , qui sont indépendantes. On vérifie, en particulier, que ce produit est non nul si et seulement si x et y sont positifs.

2. On considère un couple de variables aléatoires (X, Y) dont la densité jointe est donnée par :

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{6} & \text{si } x \in [0, 2] \text{ et } y \in [0, 3], \\ 0 & \text{sinon} \end{cases}$$

On peut écrire :

$$f_{X,Y}(x, y) = f_X(x) f_Y(y)$$

où f_X et f_Y désignent, respectivement, les densités de la loi uniforme sur $[0, 2]$ et $[0, 3]$. Ce sont, encore, les lois respectives des variables X et Y , qui sont indépendantes. On vérifie, en particulier, que ce produit est non nul si et seulement si $x \in [0, 2]$ et $y \in [0, 3]$.

Les deux exemples ci-dessus mettent en lumière une propriété importante des couples de variables à densité indépendantes : le domaine sur lequel la densité jointe est non nulle (i.e. le **support** de la densité) est un rectangle, égal au produit cartésien des support des deux densités marginales.

3. On peut aisément construire un exemple mettant en jeu un couple de variables qui semblent, a priori, indépendantes, mais pour lesquelles la propriété concernant le support ne s'applique pas.

Soit donc (X, Y) un couple de variables aléatoires (X, Y) , dont la densité jointe est donnée par :

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{2} & \text{si } x \in [0, 2] \text{ et } y \in [0, 2] \text{ et } x \geq y, \\ 0 & \text{sinon} \end{cases}$$

Le support de ce couple de v.a. est un triangle rectangle isocèle, de côté de l'angle droit de longueur 2, donc de surface 2, ce qui explique la constante normalisatrice $\frac{1}{2}$. Cette densité est en fait uniforme sur une surface du plan d'aire finie. La dépendance entre X et Y est assez évidente, puisqu'avec une probabilité de 1, on dispose de la relation $0 \leq Y \leq X \leq 2$: par exemple, conditionnellement à $Y = 1,9$, la loi de X sera concentrée sur l'intervalle $[1,9, 2]$, alors que, conditionnellement à $Y = 0,1$, la loi de X sera concentrée sur $[0,1, 2]$. On peut aussi vérifier facilement par un calcul direct que ces lois conditionnelles sont uniformes sur ces intervalles. Cette dernière propriété se généralise à toutes les lois conditionnelles de lois jointes uniformes.

On constate également, dans cet exemple, et de manière générale pour les lois uniformes dans le plan, que les coordonnées X et Y ne sont pas indépendantes (sauf si le support est un rectangle), car la loi conditionnelle et la loi marginale sont distinctes.

Quelques résultats concernant la somme de variables aléatoires continues

On peut, de façon systématique, calculer la densité de la somme de deux v.a.r. indépendantes à densité.

1. Produit de convolution de deux fonctions

Définition

Soient f et g deux fonctions définies sur \mathbb{R} , à valeurs réelles, continues par morceaux, et intégrables sur \mathbb{R} . Leur **produit de convolution** est la fonction, notée $f * g$, définie, pour tout réel x , par :

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(x-y) g(y) dy$$

► Symétrie du produit de convolution de deux fonctions

Soient f et g deux fonctions définies sur \mathbb{R} , à valeurs réelles, continues par morceaux, et intégrables sur \mathbb{R} . Leur produit de convolution vérifie :

$$f * g = g * f$$

Démonstration : Ce résultat s'obtient immédiatement à l'aide d'un changement de variable. ■

2. Somme de deux variables aléatoires réelles indépendantes

Étant données deux variables aléatoires réelles indépendantes X et Y , de densités respectives f_X et f_Y , la densité f_{X+Y} de la v.a. somme $X + Y$ est donnée par :

$$f_{X+Y} = f_X * f_Y$$

Démonstration : Soient a et b deux réels tels que $a < b$. Alors :

$$\begin{aligned} P(a \leq X + Y \leq b) &= \iint_{x \in \mathbb{R}, y \in \mathbb{R}, a \leq x+y \leq b} f_X(x) f_Y(y) dx dy \\ &= \int_a^b \int_{-\infty}^{+\infty} f_X(s) f_Y(s-t) dt ds \\ &= \int_a^b (f_X * f_Y)(s) ds \end{aligned}$$

► Somme de deux variables aléatoires réelles indépendantes suivant respectivement les lois exponentielles de paramètres distincts λ et μ , $(\lambda, \mu) \in (\mathbb{R}_+^*)^2$, $\lambda \neq \mu$

Soient λ et μ deux réels strictement positifs et distincts, et X et Y deux variables aléatoires réelles indépendantes suivant respectivement les lois exponentielles de paramètres λ et μ . On désigne par f_X et f_Y les densités respectives des v.a. X et Y .

La densité de la v.a. somme $X + Y$ est définie, pour tout réel $x \geq 0$, par :

$$\begin{aligned}
 f_{X+Y}(x) &= (f_X * f_Y)(x) \\
 &= \int_{-\infty}^{+\infty} f_X(x-y) f_Y(y) dy \\
 &= \int_0^x \lambda e^{-\lambda(x-y)} \mu e^{-\mu y} dy \\
 &= e^{-\lambda x} \int_0^x \lambda \mu e^{(\lambda-\mu)y} dy \\
 &= e^{-\lambda x} \left[\frac{\lambda \mu e^{(\lambda-\mu)y}}{\lambda - \mu} \right]_0^x \\
 &= \lambda \mu e^{-\lambda x} \frac{e^{(\lambda-\mu)x} - 1}{\lambda - \mu} \\
 &= \lambda \mu \frac{e^{-\mu x} - e^{-\lambda x}}{\lambda - \mu}
 \end{aligned}$$

► **Somme de deux variables aléatoires réelles i.i.d. de loi exponentielle de paramètre $\lambda \in \mathbb{R}_+^*$**

La formule précédente peut être extrapolée au cas où $\lambda = \mu$, en remarquant que, pour tout $x > 0$,

$$\lim_{\mu \rightarrow \lambda} \frac{e^{-\mu x} - e^{-\lambda x}}{\lambda - \mu} = -x \lim_{\mu x \rightarrow \lambda x} \frac{e^{-\mu x} - e^{-\lambda x}}{\mu x - \lambda x} = -x \left[\frac{de^{-t}}{dt} \right]_{t=\lambda x} = x e^{-\lambda x}.$$

Lorsque X et Y sont i.i.d., de loi $\mathcal{E}(\lambda)$, la densité de $X + Y$ est :

$$f_{X+Y}(x) = \lambda^2 x e^{-\lambda x}.$$

Nous verrons, prochainement, que cette densité, i.e. celle de la somme de deux v.a. i.i.d. exponentielles de même paramètre, est un cas particulier de celle de la classe dite des lois Gamma.

Exemples

1. **Somme de deux variables aléatoires réelles indépendantes suivant respectivement les lois normales de paramètres respectifs μ_1, σ_1 et μ_2, σ_2 , $(\mu_1, \mu_2) \in \mathbb{R}^2$, $(\sigma_1, \sigma_2) \in (\mathbb{R}_+^*)^2$**

Soient μ_1, μ_2 deux réels, σ_1, σ_2 deux réels strictement positifs, et X et Y deux variables aléatoires réelles indépendantes suivant respectivement les lois normales $\mathcal{N}(\mu_1, \sigma_1^2)$ et $\mathcal{N}(\mu_2, \sigma_2^2)$. On désigne par f_X et f_Y les densités respectives des v.a. X et Y .

La densité de la v.a. somme $X + Y$ est définie, pour tout réel x , par :

$$\begin{aligned}
 f_{X+Y}(x) &= (f_X * f_Y)(x) \\
 &= \int_{-\infty}^{+\infty} f_X(x-y) f_Y(y) dy \\
 &= \int_{-\infty}^{+\infty} \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{(x-y-\mu_1)^2}{2\sigma_1^2}} \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}} dy \\
 &= \int_{-\infty}^{+\infty} \frac{1}{2\pi \sigma_1 \sigma_2} e^{-\frac{(x-y-\mu_1)^2}{2\sigma_1^2}} e^{-\frac{y^2}{2\sigma_2^2}} dy \\
 &= \int_{-\infty}^{+\infty} \frac{1}{2\pi \sigma_1 \sigma_2} e^{-\frac{(x-\mu_2-\mu_1)^2}{2\sigma_1^2}} e^{-y^2 \left(\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2} \right) - \frac{y(x-\mu_2-\mu_1)}{\sigma_1^2}} dy \\
 &= \int_{-\infty}^{+\infty} \frac{1}{2\pi \sigma_1 \sigma_2} e^{-\frac{(x-\mu_2-\mu_1)^2}{2\sigma_1^2}} e^{-\left(\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2} \right) \left\{ y^2 + \frac{y(x-\mu_2-\mu_1)}{\sigma_1^2 \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)} \right\}} dy \\
 &= \int_{-\infty}^{+\infty} \frac{1}{2\pi \sigma_1 \sigma_2} e^{-\frac{(x-\mu_2-\mu_1)^2}{2\sigma_1^2}} e^{-\left(\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2} \right) \left\{ \left(y + \frac{(x-\mu_2-\mu_1)}{\sigma_1^2 \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)} \right)^2 - \frac{(x-\mu_2-\mu_1)^2}{\sigma_1^4 \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^2} \right\}} dy \\
 &= \int_{-\infty}^{+\infty} \frac{1}{2\pi \sigma_1 \sigma_2} e^{-\frac{(x-\mu_2-\mu_1)^2}{2\sigma_1^2}} e^{-\left(\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2} \right) \left\{ y^2 - \frac{(x-\mu_2-\mu_1)^2}{4\sigma_1^4 \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^2} \right\}} dy \\
 &= \frac{1}{2\pi \sigma_1 \sigma_2} e^{-\frac{(x-\mu_2-\mu_1)^2}{2\sigma_1^2}} e^{-\left(\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2} \right) \frac{(x-\mu_2-\mu_1)^2}{\sigma_1^4 \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right)^2}} \int_{-\infty}^{+\infty} e^{-\left(\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2} \right) y^2} dy \\
 &= \frac{1}{2\pi \sigma_1 \sigma_2} e^{-\frac{(x-\mu_2-\mu_1)^2}{2\sigma_1^2}} e^{-\frac{(x-\mu_2-\mu_1)^2}{4\sigma_1^4 \left(\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2} \right)}} \frac{1}{\sqrt{\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2}}} \int_{-\infty}^{+\infty} e^{-y^2} dy \\
 &= \frac{\sqrt{\pi}}{2\pi \sigma_1 \sigma_2} \frac{1}{\sqrt{\frac{1}{2\sigma_1^2} + \frac{1}{2\sigma_2^2}}} e^{-\frac{(x-\mu_2-\mu_1)^2}{2\sigma_1^2} \left(1 - \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \right)} \\
 &= \frac{\sqrt{\pi}}{\sqrt{2\pi} \sigma_1 \sigma_2} \frac{\sigma_1 \sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} e^{-\frac{(x-\mu_2-\mu_1)^2}{2(\sigma_1^2 + \sigma_2^2)}} \\
 &= \frac{1}{\sqrt{2\pi} \sqrt{\sigma_1^2 + \sigma_2^2}} e^{-\frac{(x-\mu_2-\mu_1)^2}{2(\sigma_1^2 + \sigma_2^2)}}
 \end{aligned}$$

La v.a. somme $X + Y$ suit donc la loi normale $\mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

2. Somme de N variables aléatoires réelles indépendantes suivant respectivement les lois normales de paramètres respectifs $\mu_1, \sigma_1, \dots, \mu_N, \sigma_N, (\mu_1, \dots, \mu_N) \in \mathbb{R}^N, (\sigma_1, \dots, \sigma_N) \in (\mathbb{R}_+^*)^N, N \in \mathbb{N}^*$

Soit N un entier naturel non nul, μ_1, \dots, μ_N N réels, $\sigma_1, \dots, \sigma_N$ N réels strictement positifs, et X_1, \dots, X_N N variables aléatoires réelles indépendantes suivant respectivement les lois normales $\mathcal{N}(\mu_1, \sigma_1^2), \dots, \mathcal{N}(\mu_N, \sigma_N^2)$. Une récurrence immédiate à partir de l'exemple précédent montre que la v.a. somme $X_1 + \dots + X_N$ suit la loi normale $\mathcal{N}(\mu_1 + \dots + \mu_N, \sigma_1^2 + \dots + \sigma_N^2)$.



L'exemple précédent est d'importance, il met en lumière le fait que les variables normales sont stables par combinaisons linéaires affines de termes indépendants. Le seul autre exemple de ce type que nous avons rencontré est celui des variables de Poisson (bien que, dans ce cas, on n'ait que le droit d'ajouter les variables, sans les multiplier par des coefficients, ni leur ajouter des constantes). C'est, aussi, un cas très particulier du théorème central limite. En effet, si on considère une suite de variables $(X_n)_{n \in \mathbb{N}}$ i.i.d. de loi $\mathcal{N}(\mu, \sigma)$, alors, d'après ce même théorème central limite, la loi de :

$$\sqrt{n} M_n = \sqrt{n} \left(\frac{X_1 + \dots + X_n}{n} - \mu \right)$$

tend vers la loi normale $\mathcal{N}(0, \sigma^2)$.

Or, si on applique le résultat de l'exemple précédent, on montre facilement qu'en réalité la loi de $\sqrt{n} M_n$ est exactement égale à $\mathcal{N}(0, \sigma^2)$ (cf. l'exemple suivant). C'est le seul cas de suites i.i.d. où ait lieu cette égalité.

3. Une application aux intervalles de confiance

On donne ici un exemple d'illustration des intervalles de confiance, exploitant l'égalité entre la loi de $\sqrt{n} M_n$ et la loi normale $\mathcal{N}(0, \sigma^2)$. Il est important de comprendre que le théorème central limite permettra de reprendre cet argument dans le cas où le nombre n de variables est grand, ce, même si les lois individuelles ne sont pas normales (gaussiennes).

Soit N un entier naturel supérieur ou égal à 2, μ un réel, a priori inconnu, et X_1, \dots, X_N, N v.a.r. suivant chacune la loi normale $\mathcal{N}(\mu, 1)$. Les v.a.r. centrées $X_1 - \mu, \dots, X_N - \mu$, suivent donc chacune la loi normale centrée réduite $\mathcal{N}(0, 1)$. Les v.a.r. centrées réduites $\frac{X_1 - \mu}{N}, \dots, \frac{X_N - \mu}{N}$, suivent donc chacune la loi normale :

$$\mathcal{N}\left(\frac{\mu}{N} - \frac{\mu}{N}, \frac{1}{N^2}\right) = \mathcal{N}\left(0, \frac{1}{N^2}\right)$$

La v.a. somme :

$$\frac{X_1 + \dots + X_N - N\mu}{N} = \frac{X_1 + \dots + X_N}{N} - \mu$$

suit donc la loi normale :

$$\mathcal{N}\left(0, \frac{N}{N^2}\right) = \mathcal{N}\left(0, \frac{1}{N}\right)$$

On en déduit, finalement, que la v.a. $M_N = \sqrt{N} \left(\frac{X_1 + \dots + X_N}{N} - \mu \right)$ suit la loi normale :

$$\mathcal{N}\left(0, \frac{(\sqrt{N})^2}{N}\right) = \mathcal{N}(0, 1)$$

i.e. la loi normale centrée réduite.

On peut alors calculer :

$$P\left(\sqrt{N} \left| \frac{X_1 + \dots + X_N}{N} - \mu \right| \leq 1,96\right) = \int_{-1,96}^{1,96} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx \simeq 0,95$$

Ceci veut dire que l'on est à 95 % certain que la moyenne empirique $M_N = \frac{X_1 + \dots + X_N}{N}$ est dans l'intervalle $\left[\mu - \frac{1,96}{\sqrt{N}}; \mu + \frac{1,96}{\sqrt{N}}\right]$. On rencontre parfois la

notation abusive $M_N = \mu \pm \frac{1,96}{\sqrt{N}}$, avec, éventuellement, la mention marginale qu'il s'agit d'un intervalle de confiance à 95 %.

On peut encore réinterpréter ce résultat pour créer un intervalle de confiance pour l'espérance μ , si on suppose, comme c'est souvent le cas en pratique, que μ n'est pas observée, alors que M_N peut l'être, grâce à un sondage. C'est, aussi, une procédure d'estimation de μ . Dire que :

$$M_N \in \left[\mu - \frac{1,96}{\sqrt{N}}; \mu + \frac{1,96}{\sqrt{N}} \right]$$

équivalait à :

$$\mu \in \left[M_N - \frac{1,96}{\sqrt{N}}; M_N + \frac{1,96}{\sqrt{N}} \right] = I_{95}$$

Or :

$$P(\mu \in I_{95}) = P\left(M_N \in \left[\mu - 1,96/\sqrt{N}; \mu + 1,96/\sqrt{N}\right]\right) \approx 0,95$$

I_{95} est donc un intervalle de confiance à 95 % pour la valeur théorique non observée μ .

Exemple : Intervalles de confiance avec variance non unitaire

Examinons ce qui se produit lorsque $\sigma \neq 1$. Soient Y_1, Y_2, \dots, Y_N des variables i.i.d. de loi $\mathcal{N}(\mu, \sigma^2)$. En posant, pour tout entier k de $\{1, \dots, N\}$, $X_k = \frac{Y_k}{\sigma}$, on peut se ramener au cas précédent. À l'aide de la moyenne empirique :

$$\tilde{M}_N = \frac{Y_1 + \dots + Y_N}{N} = \sigma M_N$$

on obtient :

$$\sqrt{N}(\tilde{M}_N - \mu) = \sqrt{N}(\sigma M_N - \mu) \sim \mathcal{N}(0, \sigma^2)$$

et donc avec les mêmes calculs que précédemment,

$$P\left(\tilde{M}_N \in \left[\mu - \frac{1,96\sigma}{\sqrt{N}}; \mu + \frac{1,96\sigma}{\sqrt{N}}\right]\right) \approx 0,95.$$

Pour transformer ceci en un intervalle de confiance pour μ , nous butons sur une difficulté pratique qui vient du fait que, si l'on suppose ne pas connaître μ et qu'on veut se baser sur la moyenne empirique \tilde{M}_N pour l'estimer, notre intervalle de confiance dépendra de σ . Et il n'y a aucune raison de supposer que σ est plus facilement observable que μ .

Ce que font, en pratique, les statisticiens, consiste à se baser sur la notion dite du « *bootstrap* ». Ce terme anglo-saxon vient de l'expression colloquiale « *pull yourself up by your bootstraps* », dont le sens propre est « relève-toi en tirant sur les lacets (ou sangles) de tes propres bottes », qu'il faut interpréter par le fait qu'il faut se baser sur ses propres ressources pour effectuer un travail qui semble demander un aide extérieure.

Pour construire un intervalle de confiance pour μ , on utilise une procédure de ce type, qui fonctionnerait bien si on connaissait μ , pour estimer σ ; on utilise ainsi la valeur \tilde{M}_N comme estimateur de μ , ce qui introduit une petite erreur du fait de la circularité de l'argument, qui produit ensuite un intervalle de confiance pour μ . Nous n'entrerons pas dans les détails de l'estimation quantitative de cette erreur.

Toujours est-il que, si l'on a une estimation précise $\hat{\sigma}$ de la valeur de σ , il n'est pas faux d'affirmer que l'intervalle :

$$I_{95} = \left[\tilde{M}_N - \frac{1,96}{\sqrt{N}} \hat{\sigma}; \tilde{M}_N + \frac{1,96}{\sqrt{N}} \hat{\sigma} \right]$$

est un intervalle de confiance pour μ à 95 %.

Nous avons rencontré précédemment un exemple donnant la densité de la somme de deux variables X et Y i.i.d., de loi $\mathcal{E}(\lambda)$:

$$f_{X+Y}(x) = \begin{cases} \lambda^2 x e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Nous présentons, dans ce qui suit, la généralisation de ce résultat, avec les lois Gamma, extrêmement utiles pour modéliser les temps d'arrivée d'un $n^{\text{ième}}$ client, ou encore dans des études de fiabilité, où une défaillance technique est assimilée à un temps d'arrivée. Ces lois sont aussi fondamentales en statistique.

1. Fonction Gamma

Définition

La fonction Gamma est la fonction définie sur \mathbb{R}_+^* par :

$$\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{-t} dt$$

► Propriétés de la fonction Gamma

- Pour tout réel strictement positif α :

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$$

- Pour tout entier naturel n :

$$\Gamma(n + 1) = n !$$

-

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$



Comme, pour tout réel strictement positif α , $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$, la fonction Gamma peut être considérée comme une généralisation de la notion de factorielle aux réels positifs non entiers.

Démonstration :

- Pour tout réel strictement positif α , on obtient, en intégrant par parties :

$$\begin{aligned} \Gamma(\alpha + 1) &= \int_0^{+\infty} t^\alpha e^{-t} dt \\ &= [-t^\alpha e^{-t}]_0^{+\infty} + \int_0^{+\infty} \alpha t^{\alpha-1} e^{-t} dt \\ &= \int_0^{+\infty} \alpha t^{\alpha-1} e^{-t} dt \\ &= \alpha \Gamma(\alpha) \end{aligned}$$

- Pour tout entier naturel n :

$$\Gamma(n+1) = n\Gamma(n) = n \times (n-1) \times \Gamma(n-1) \times \dots \times 1 \times \Gamma(1) = n(n-1)\dots 1 = n!$$

puisque $\Gamma(1) = \int_0^{+\infty} e^{-t} dt = 1$.

-

$$\begin{aligned}\Gamma\left(\frac{1}{2}\right) &= \int_0^{+\infty} t^{-\frac{1}{2}} e^{-t} dt \\ &= \int_0^{+\infty} t^{-1} e^{-t^2} 2t dt \\ &= 2 \int_0^{+\infty} e^{-t^2} dt \\ &= \sqrt{\pi}\end{aligned}$$

■

2. Loi Gamma (de paramètres $\alpha > 0, \beta > 0$)

Soient α et β deux réels strictement positifs. Une v.a. réelle X suit la **loi Gamma (de paramètres α, β)**, $\Gamma(\alpha, \beta)$ (aussi appelée **loi eulérienne de paramètres α, β**) si sa densité $f_{\alpha, \beta}$ est telle que, pour tout réel x :

$$f_{\alpha, \beta}(x) = \begin{cases} \frac{x^{\alpha-1} \beta^\alpha e^{-\beta x}}{\Gamma(\alpha)} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

α est un paramètre de forme. β est un paramètre d'intensité, $\frac{1}{\beta}$ est le paramètre d'échelle associé.



1. On note :

$$X \sim \Gamma(\alpha, \beta)$$

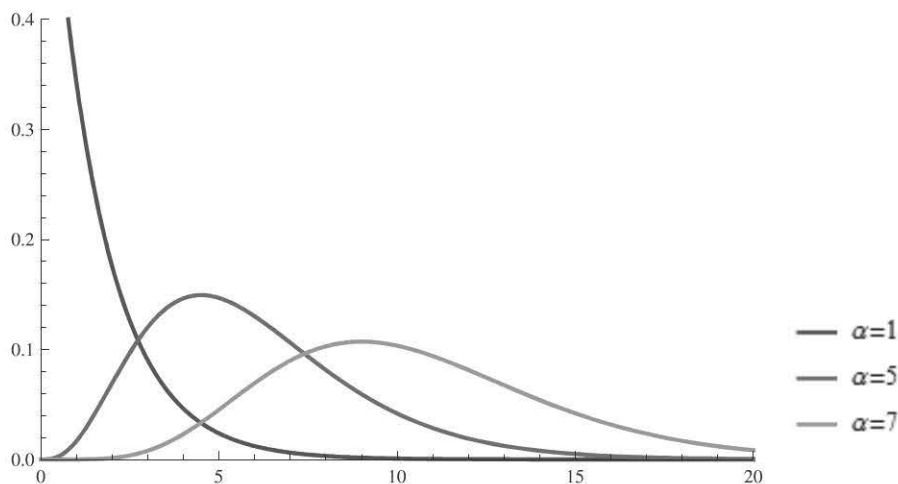


Figure 139.1 – La densité de la loi Gamma, pour différentes valeurs du paramètre α , dans le cas où $\beta = 1,5$.

2. Avec l'exemple rappelé au début de cette section, on voit que si X et Y sont i.i.d. $\mathcal{E}(\lambda)$, alors la loi de $X + Y$ est $\Gamma(2, \lambda)$, puisque $\Gamma(2) = 1! = 1$.
3. On voit aussi que la loi exponentielle $\mathcal{E}(\lambda)$ est tout simplement la loi $\Gamma(1, \lambda)$ puisque $\Gamma(1) = 0! = 1$.

► Loi d'Erlang (de paramètres $n \in \mathbb{N}^*$, $\beta > 0$)

Soient n un entier naturel non nul, et β un réel strictement positif. Une v.a. réelle X suit la **loi d'Erlang (de paramètres $n \in \mathbb{N}^*$)**, si elle suit une loi Gamma dont le paramètre de forme est entier ; sa densité $f_{n,\beta}$ est donc telle que, pour tout réel x :

$$f_{\alpha,\beta}(x) = \begin{cases} \frac{x^{n-1} \beta^n e^{-\beta x}}{(n-1)!} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

► Loi de distribution de Maxwell-Boltzmann (de paramètre $\alpha > 0$)

Soit a un réel strictement positif. Une v.a. réelle X suit la **loi de distribution de Maxwell-Boltzmann (de paramètre a)**, si elle suit une loi Gamma dont le paramètre de forme est $\frac{3}{2}$, et le paramètre d'échelle, $2a^2$; sa densité f_a est donc telle que, pour tout réel x :

$$f_a(x) = \begin{cases} \frac{x^{\frac{1}{2}} a^{-3} e^{-\frac{x}{2a^2}}}{\sqrt{2\pi}} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$



La loi de distribution de Maxwell-Boltzmann est utilisée en Physique statistique pour déterminer la répartition des particules entre les différents niveaux d'énergie. La théorie cinétique des gaz, qui suscite, en ce moment, beaucoup d'intérêt dans la communauté mathématique (voir, notamment, les travaux de Cédric Villani [41]), est basée sur cette distribution.

► Loi du χ^2 (« Chi carré »)

Soit N un entier naturel non nul. Une v.a. réelle X suit la **loi du χ^2 à N degrés de liberté**, $\chi^2(N)$, si elle suit une loi Gamma $\Gamma\left(\frac{N}{2}, \frac{1}{2}\right)$; sa densité f_N est donc telle que, pour tout réel x :

$$f_N(x) = \begin{cases} \frac{x^{\frac{N}{2}-1} e^{-\frac{x}{2}}}{\sqrt{2^N} \Gamma\left(\frac{N}{2}\right)} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$



1. On note :

$$X \sim \chi^2(N)$$

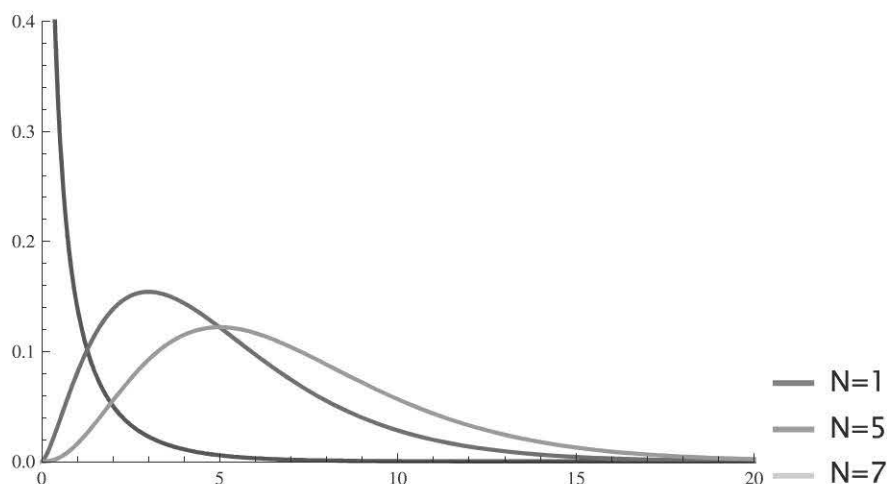


Figure 139.2 – La densité de la loi du χ^2 , pour différentes valeurs du paramètre N .

2. Il est intéressant d'étudier, dans un premier temps, la loi du χ^2 à un degré de liberté. Soit Z une variable aléatoire, de loi $\mathcal{N}(0, 1)$. Si on considère la v.a. positive $X = Z^2$, alors, pour tout $x > 0$,

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P(-\sqrt{x} \leq Z \leq \sqrt{x}) \\ &= 2 \int_0^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy. \end{aligned}$$

On peut alors en déduire, grâce à la formule de dérivation composée, et au théorème fondamental de l'analyse, l'expression de la densité de X , pour tout $x > 0$:

$$f_X(x) = \frac{1}{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}}.$$

Du fait de la positivité de X : $f_X(x) = 0$ pour $x < 0$; la valeur de f_X en 0 n'a pas besoin d'être définie, et peut, par convention, être considérée comme nulle. La densité f_X est donc aussi celle de la loi Gamma, de paramètres $\frac{1}{2}$ et $\frac{1}{2}$, ce qui constitue le résultat suivant.

3. Propriétés

Lemme

Si $Z \sim \mathcal{N}(0, 1)$ alors $Z^2 \sim \Gamma(1/2, 1/2)$.



Ce résultat est un cas particulier de la propriété suivante, lorsque $N = 1$.

Propriété

Soient N un entier naturel non nul, $\mu_1, \dots, \mu_N, \sigma_1 \geq 0, \dots, \sigma_N \geq 0$ des réels, et X_1, \dots, X_N des v.a.r. indépendantes suivant, respectivement, les lois normales $\mathcal{N}(\mu_1, \sigma_1), \dots, \mathcal{N}(\mu_N, \sigma_N)$.

Alors, la v.a.r. $\sum_{i=1}^N \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2$ suit la loi du χ^2 à N degrés de liberté :

$$\sum_{i=1}^N \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2 \sim \chi^2(N)$$

Du fait de la propriété de changement d'échelle de la loi normale, la loi du χ^2 à N degrés de liberté est, tout simplement, la loi de probabilité de la somme des carrés de N variables normales centrées réduites indépendantes entre elles.



La distribution $\chi^2(N)$ est utilisée en statistique inférentielle dans de nombreux contextes. La propriété ci-dessus est un cas particulier particulièrement simple. On peut s'en servir pour démontrer qu'on a aussi, avec notre amie la moyenne empirique $M_N = \frac{X_1 + \dots + X_N}{N}$,

$$S = \sum_{k=1}^N (X_k - M_N)^2 \sim \sigma^2 \chi^2(N-1)$$

Si jamais on a une idée extrêmement précise de la valeur de σ , alors, lorsque N n'est pas trop grand, on peut utiliser la statistique S pour déterminer si les données utilisées pour le calcul de S ont une bonne chance de provenir de données gaussiennes i.i.d., ou non.

Démonstration : Cette propriété est une conséquence immédiate du lemme ci-dessus, et de la proposition suivante. ■

Proposition

Soit N un entier naturel non nul, $\alpha_1, \dots, \alpha_N, \beta$ des réels strictement positifs, et X_1, \dots, X_N des v.a.r. indépendantes suivant, respectivement, les lois $\Gamma(\alpha_1, \beta), \dots, \Gamma(\alpha_N, \beta)$.

Alors, la v.a.r. $\sum_{i=1}^N X_i$ suit la loi Gamma de paramètres $\sum_{i=1}^N \alpha_i, \beta$:

$$\sum_{i=1}^N X_i \sim \Gamma\left(\sum_{i=1}^N \alpha_i, \beta\right)$$

Démonstration : Il suffit de démontrer cette propriété pour $N = 2$, le résultat pour $N \geq 2$ s'en déduisant par simple itération.

La densité de probabilité de la loi Gamma de paramètres $\alpha_1 + \alpha_2, \beta$, est donnée, pour tout réel $x > 0$, par :

$$\begin{aligned} f_{\alpha_1+\alpha_2, \beta}(x) &= \frac{x^{\alpha_1+\alpha_2-1} \beta^{\alpha_1+\alpha_2} e^{-\beta x}}{\Gamma(\alpha_1 + \alpha_2)} \\ &= \frac{x^{\alpha_1} x^{\alpha_2} \beta^{\alpha_1} \beta^{\alpha_2} e^{-\beta x}}{x \Gamma(\alpha_1 + \alpha_2)} \end{aligned}$$

On peut alors vérifier, grâce à des calculs assez pénibles que nous détaillerons pas ici, que :

$$f_{\alpha_1+\alpha_2, \beta} = f_{\alpha_1, \beta} * f_{\alpha_2, \beta} \quad \blacksquare$$



1. Les deux résultats ci-dessus montrent en outre, grâce au cas $N = 2$, que la loi exponentielle $\mathcal{E}(1)$ est aussi celle du χ^2 avec deux degrés de liberté, i.e. la loi de $Y^2 + Z^2$, où Y et Z sont i.i.d., de loi $\mathcal{N}(0, 1)$.
2. Dans le cas où, pour tout entier i , $\alpha_i = 1$, en choisissant $\lambda = \beta$, la proposition précédente implique que la loi $\Gamma(N, \lambda)$ est celle de la somme de N variables i.i.d de loi exponentielle $\mathcal{E}(\lambda)$.

► **Espérance d'une v.a.r. suivant la loi Gamma de paramètres $\alpha > 0, \beta > 0$**

Soient α et β deux réels strictement positifs, et X une v.a.r. suivant la loi Gamma $\Gamma(\alpha, \beta)$.
L'espérance de X est :

$$\mathbb{E}(X) = \frac{\alpha}{\beta}$$

Démonstration : Il est clair que $\int_{-\infty}^{+\infty} |x| f_{\alpha, \beta}(x) dx = \int_0^{+\infty} x \frac{x^{\alpha-1} \beta^\alpha e^{-\beta x}}{\Gamma(\alpha)} dx$ converge. On peut alors calculer :

$$\begin{aligned} \int_{-\infty}^{+\infty} x f_{\alpha, \beta}(x) dx &= \int_0^{+\infty} x \frac{x^{\alpha-1} \beta^\alpha e^{-\beta x}}{\Gamma(\alpha)} dx \\ &= \int_0^{+\infty} \frac{x^\alpha \beta^\alpha e^{-\beta x}}{\Gamma(\alpha)} dx \\ &= \int_0^{+\infty} \frac{x^\alpha \beta^{-\alpha} e^{-x}}{\Gamma(\alpha)} \frac{dx}{\beta} \\ &= \frac{1}{\beta \Gamma(\alpha)} \int_0^{+\infty} x^\alpha e^{-x} dx \\ &= \frac{1}{\beta \Gamma(\alpha)} \Gamma(\alpha + 1) \\ &= \frac{1}{\beta \Gamma(\alpha)} \alpha \Gamma(\alpha) \\ &= \frac{\alpha}{\beta} \quad \blacksquare \end{aligned}$$

► **Variance d'une v.a.r. suivant la loi Gamma de paramètres $\alpha > 0, \beta > 0$**

Soient α et β deux réels strictement positifs, et X une v.a.r. suivant la loi Gamma $\Gamma(\alpha, \beta)$.
La variance de X est :

$$V(X) = \frac{\alpha}{\beta^2}$$

Démonstration : Il est clair que $\int_{-\infty}^{+\infty} x^2 f(x) dx = \int_0^{+\infty} x^2 \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$ converge. On peut alors calculer :

$$\begin{aligned} \mathbb{E}[X^2] &= \int_{-\infty}^{+\infty} x^2 f_{\alpha, \beta}(x) dx \\ &= \int_0^{+\infty} x^2 \frac{x^{\alpha-1} \beta^\alpha e^{-\beta x}}{\Gamma(\alpha)} dx \\ &= \int_0^{+\infty} \frac{x^{\alpha+1} \beta^\alpha e^{-\beta x}}{\Gamma(\alpha)} dx \\ &= \int_0^{+\infty} \frac{x^{\alpha+1} \beta^{-\alpha-1} \beta^\alpha e^{-x}}{\Gamma(\alpha)} \frac{dx}{\beta} \\ &= \int_0^{+\infty} \frac{x^{\alpha+1} \beta^{-2} e^{-x}}{\Gamma(\alpha)} dx \\ &= \frac{1}{\beta^2 \Gamma(\alpha)} \Gamma(\alpha + 2) \\ &= \frac{1}{\beta^2 \Gamma(\alpha)} (\alpha + 1) \alpha \Gamma(\alpha) \\ &= \frac{\alpha(\alpha + 1)}{\beta^2} \end{aligned}$$

La variance de la v.a.r. X s'obtient alors grâce à la formule de Koenig :

$$V(X) = \mathbb{E}[X^2] - (\mathbb{E}(X))^2 = \frac{\alpha(\alpha+1)}{\beta^2} - \frac{\alpha^2}{\beta^2} = \frac{\alpha}{\beta^2} \quad \blacksquare$$



Nous avons vu, dans le cas où $\alpha = N$ est un entier naturel non nul, que $\Gamma(N, \lambda)$ est la loi de la somme $X = \sum_{k=1}^N X_k$ de N variables i.i.d. X_k , de loi $\mathcal{E}(\lambda)$. Les formules donnant les espérances et les variances conduisent alors à :

$$\mathbb{E}(X) = \frac{N}{\lambda} \quad \text{et} \quad \text{Var}(X) = \frac{N}{\lambda^2}$$

On aurait pu, aussi, obtenir ces formules directement par linéarité de l'espérance, et additivité de la variance, les v.a. en jeu étant indépendantes. Nous avons, d'ailleurs, déjà fait cette remarque à propos des lois du chi carré (χ^2).

Dans notre magasin de luxe de la rue de la Paix, la propriétaire a constaté que la durée d'attente avant l'arrivée d'un client ne donne pas d'information sur celles avant l'arrivée des autres clients. Elle pense que ces temps d'attentes sont uniformes au cours de la journée. En tenant compte de ces éléments, et de ses décisions précédentes, elle est en droit de modéliser la durée X_k d'attente entre le $(k-1)^{\text{ième}}$ et le $k^{\text{ième}}$ client comme une variable exponentielle, de paramètre $\lambda = 6$, et de considérer les X_k comme i.i.d.

Les deux employés et le gérant se sont mis d'accord pour que le premier prenne sa pause café du matin après l'arrivée du $10^{\text{ième}}$ client. Le temps d'attente pour cette pause est donc $X = \sum_{k=1}^{10} X_k$, de loi $\Gamma(10, 6)$.

La densité f_X de la v.a. X est nulle sur \mathbf{R}^- et, pour tout $x > 0$:

$$f_X(x) = \frac{x^9 6^{10} e^{-6x}}{9!} = \frac{5832}{35} x^9 e^{-6x}.$$

La durée moyenne de cette attente est :

$$\mathbb{E}(X) = \frac{10}{6} = \frac{5}{3} = 1 \text{ heure } 40 \text{ minutes}$$

Son écart-type est :

$$\sqrt{\text{Var}(X)} = \sqrt{\frac{10}{6^2}} = \frac{\sqrt{10}}{6} \approx 0,5271 \approx 32 \text{ minutes}$$

L'employé s'inquiète quand même un peu, car l'heure du déjeuner approche, et il n'a pas eu sa pause café ! Le magasin ouvre en effet à 10 heures, et le déjeuner est à 13 heures. La probabilité que la pause café n'arrive pas avant le déjeuner est donc :

$$P(X > 3) = \int_3^{\infty} f_X(x) dx = \int_3^{\infty} \frac{5832}{35} x^9 e^{-6x} dx.$$

Il est possible de calculer cette intégrale explicitement à l'aide de 9 intégrations par parties successives. On obtient :

$$P(X > 3) = \frac{35\,347\,283}{35} e^{-18} \approx 0,01538.$$

L'employé n'a donc pas trop à s'inquiéter !

Mais, étant de nature anxieuse, il se demande si une telle situation pourrait lui arriver au cours d'une année. Il décide donc de considérer que chaque jour de travail est indépendant des autres, et veut estimer le nombre Y de fois où il manquera sa pause café. Comme il travaille 250 jours par an, il assimile le nombre Y à une variable aléatoire binomiale de paramètres $n = 250$ et $p = 0,01538$, ce qui conduit à :

$$\mathbb{E}(Y) = np = 250 \times 0,01538 \approx 3,845$$

$$\sqrt{\text{Var}(Y)} = \sqrt{np(1-p)} = \sqrt{250 \times 0,01538 \times 0,9846} \approx 1,946$$

Il semble assez probable que l'événement redouté se produise une demi-douzaine de fois par an. Il souhaite donc connaître la probabilité $P(Y \geq 6)$. Comme n est très grand, il ne semble pas possible d'expliciter cette dernière expression. Comme p est assez proche de 0, cette situation semble intermédiaire entre celle où l'on utiliserait une approximation de Poisson, et une autre, où le théorème central limite s'appliquerait. L'employé décide d'essayer les deux calculs :

- si on utilise l'approximation de Poisson, on peut considérer que la loi de Y est proche de celle de la loi de Poisson de paramètre $\lambda = np = 3,845$, ce qui conduit alors à :

$$P(Y \geq 6) \approx e^{-3,845} \sum_{k=0}^5 \frac{3,845^k}{k!} \approx 0,8087$$

- si on utilise le théorème central limite, on peut considérer que $\frac{Y}{250}$ est la moyenne empirique de 250 variables de Bernoulli, de paramètre $p = 0,01538$. Il faudrait donc dire que $\frac{Y - \mathbb{E}(Y)}{\sqrt{\text{Var}(Y)}}$ est approximativement de loi $\mathcal{N}(0, 1)$, ce qui conduit à :

$$\begin{aligned} P(Y \geq 6) &= P\left(\frac{Y - 3,845}{1,946} \geq \frac{6 - 3,845}{1,946}\right) \\ &\approx P(\mathcal{N}(0,1) \geq 1,107) \\ &\approx 0,8659. \end{aligned}$$

On s'aperçoit que les deux approximations sont en léger désaccord. Mais, dans tous les cas, les nouvelles ne sont pas bonnes pour l'employé. À l'aide d'un ordinateur capable de calculer des factorielles d'un ordre élevé, on peut déterminer une valeur approximative de $P(Y \geq 6)$: 0,8102.

L'approximation de Poisson est donc la meilleure. Ce résultat était prévisible, puisqu'un bon critère pour pouvoir appliquer cette approximation est que n soit très grand devant 1, et que np soit très petit devant n . Le théorème central limite s'applique moins bien, car la moyenne empirique $\frac{Y}{N}$ ne se concentre pas assez autour de sa moyenne ; ainsi,

$$\sqrt{\text{Var}\left(\frac{Y}{N}\right)} \approx 1,946 \times 250^{-\frac{1}{2}} \approx 0,1231$$

n'est pas assez petit devant 1.

Exercices d'entraînement



Les corrigés sont disponibles en téléchargement sur le site dunod.com à partir de la page de présentation de l'ouvrage.

Espaces probabilisés

1 Tirage de boules

On considère deux urnes : l'une contient six boules rouges et trois noires, l'autre six noires et trois rouges. On choisit une urne au hasard puis on tire deux boules (simultanément). On obtient deux rouges. Quelle est la probabilité qu'on ait choisi la première urne ?

2 Généalogie

Soit n un entier naturel non nul.

On suppose connue la probabilité qu'une famille ait n enfants. On désigne par p_n ($n = 0, 1, \dots$) cette probabilité. On suppose que les 2^n compositions (filles/garçons) possibles aient la même probabilité conditionnelle (sachant que la famille est composée de n enfants).

2.a) Quelle est la probabilité qu'une famille n'ayant que des garçons soit composée de k enfants ?

2.b) On suppose que $p_n = \alpha p^n$ pour $n \geq 1$, où $\alpha > 0$, p est un réel de l'intervalle $]0, \alpha^{-1}[$, et

$$p_0 = 1 - \alpha p (1 + p + p^2 + \dots) = 1 - \frac{\alpha p}{1 - p}$$

Quelle est la probabilité qu'une famille ait exactement k garçons ?

2.c) Sachant qu'une famille a au moins un garçon, quelle est la probabilité (sous les mêmes hypothèses qu'à la question précédente) qu'elle en possède deux ou plus ? Généraliser à m ou plus, $m \geq 2$.

Probabilités discrètes

3 Jeux de dés (1)

Soit n un entier naturel non nul. On s'intéresse au lancer de $6n$ dés.

3.a) Définir l'univers Ω associé à cette expérience aléatoire.

3.b) Calculer la probabilité de l'événement « on obtient chacun des nombres 1, 2, ..., 6 exactement n fois ».

4 Jeux de dés (2)

Soit n un entier naturel non nul. On considère l'expérience aléatoire qui consiste à lancer successivement n fois un dé et à noter les résultats obtenus.

4.a) Définir l'univers Ω associé à cette expérience aléatoire.

4.b) Calculer la probabilité de l'événement « on obtient au moins un 2 ».

5 Tirages de boules

Soit n un entier naturel non nul. Dans une urne qui contient $2n$ boules numérotées de 1 à $2n$, on tire k boules successivement et sans remise ($1 \leq k \leq n$).

5.a) Définir l'univers Ω associé à cette expérience aléatoire.

5.b) Quelle est la probabilité que les k boules portent toutes des numéros impairs ?

6 Un peu de géométrie

Soit n un entier naturel non nul, strictement plus grand que 1. On marque n points sur un cercle, puis on en choisit 2 au hasard.

6.a) Quelle est la probabilité pour qu'ils soient voisins ?

6.b) Même question pour n points choisis sur une droite.

7 Rangements

Soit n un entier naturel non nul. n paires de chaussettes de couleur rouge sont placées au hasard dans n tiroirs.

- 7.a) Quelle est la probabilité pour que chaque tiroir soit occupé ?
- 7.b) Reprendre le même problème, en supposant les n paires de n couleurs différentes.

8 Sur la fonction de répartition

On s'intéresse au lancer, de manière indépendante, de trois pièces de monnaie. On suppose que la probabilité de tomber sur « pile » vaut $\frac{1}{2}$ pour chacune des pièces. On attribue la valeur 1 au côté « pile », et 0 au côté « face ». On définit alors la variable aléatoire X , qui correspond au nombre de fois où les 3 pièces tombent sur « pile ». Donner la loi de X , et sa fonction de répartition.

9 Légendes du baseball

Un batteur de baseball est considéré comme extrêmement bon s'il est capable de frapper la balle au moins 300 fois sur mille, en moyenne, au cours de sa carrière. Le meilleur batteur de tous les temps était le légendaire Ty Cobb¹. La probabilité qu'il avait de frapper la balle était $p = 0,3664$,

loin devant d'autres légendes comme Babe Ruth² (10^e à 0,3421) . Joe DiMaggio³ (41^e à 0,3246), Ichiro Suzuki⁴ (53^e à 0,3196), Jackie Robinson⁵ (98^e à 0,3113), ou Alex Rodriguez⁶ (204^e à 0,3000). En supposant que la performance de Ty Cobb était uniforme sur ses vingt-quatre années de jeu (une longévité extraordinaire également), calculer la probabilité pour que, sur cinq occasions de frapper la balle (cinq « at bat ») Ty Cobb la frappe au moins trois fois.

10 Une histoire de cerises chez les altermondialistes

Considérons un chargement de cerises en route pour le marché international de Rungis. Ces cerises sont destinées à être utilisées dans la préparation de desserts glacés pour une chaîne de restauration rapide McArnold. Des altermondialistes dressent un barrage routier et répandent le contenu du chargement sur la chaussée, sur la Nationale 7, dans la banlieue de Montélimard. L'un d'eux estime que le nombre total de cerises est de trente millions, alors qu'un autre remarque que dans

1. Prononcer « Taille Côôôb » ; au cours de sa carrière, au début du xx^e siècle, il fut le seul joueur ayant, pour la frappe de la balle, une probabilité de succès supérieure à 0,3 pendant 23 saisons consécutives, un record qui ne sera probablement jamais battu ; il a établi de nombreux autres records, qui ont duré plus de cinquante ans.
2. Considéré comme le meilleur joueur de tous les temps, et ce malgré un embonpoint évident ; il a détenu de nombreux records dans les années 1920, y compris le record du batteur le plus puissant – calculé grâce au « *slugging percentage* » – record qui tient encore. Il est donc toujours le batteur le plus redoutable de tous les temps, mais Ty Cobb était encore plus efficace.
3. Joe, fils d'immigrés italiens, jouait pour les New York Yankees ; il détient toujours le record, établi en 1941, du plus grand nombre de jeux consécutifs avec un hit (« *hitting streak* »). Il était toujours la coqueluche des Américains de nombreuses années après sa retraite, et épousa Marilyn Monroe en 1954. Sur son lit de mort en 1999, ses ultimes paroles furent « Je vais enfin pouvoir voir Marilyn ».
4. Un des rares joueurs de baseball avec une renommée internationale avant de jouer en Amérique du Nord, il signa aux Seattle Mariners en 2000, et seulement une année plus tard, devint simultanément le meilleur jeune de toutes les ligues nord-américaines (« *MLB Rookie of the Year* »), et le meilleur joueur de la Ligue américaine (« *American League MVP* »).
5. Quinze ans avant le mouvement des droits civils contre le régime d'apartheid américain (ou « Jim Crowe Laws », selon lesquelles les Afro-Américains ne pouvaient pas partager les mêmes lieux publics que les caucasiens), Jackie Robinson fut le premier joueur Afro-américain dans les ligues de première division (MLB). Il fut surnommé « *Rookie of the Year* » en 1947, et « MVP » (Most Valuable Player) de la National League en 1949. Son début de carrière est décrit en 2013 dans le film hollywoodien « 42 ».
6. D'origine dominicaine, surnommé « A-Rod », il finira une très longue carrière (1994-2017) dans l'équipe des New York Yankees, et restera probablement comme l'un des meilleurs joueurs de tous les temps – avec un talent prodigieux repéré dès sa plus tendre enfance – et ce, malgré de nombreuses controverses sur son recours à des stéroïdes dûment interdits. A-Rod fut le joueur le mieux payé de tous les temps (275 millions pour ses dix dernières années aux Yankees), et détient de nombreux records ; par exemple, en 2007, il a battu celui du joueur le plus jeune à atteindre 600 « home-runs », record détenu pendant près d'un siècle par Babe Ruth.

plusieurs échantillons successifs, il semble que la proportion de cerises non moisies soit de 80 %. Le premier compère en conclut qu'il doit y avoir environ 24 millions de bonnes cerises répandues sur la chaussée. Les compagnons ayant décidé de conserver le barrage routier pendant toute la journée, ils s'aperçoivent qu'ils n'ont pas prévu de casse-croûte. En accord avec le transporteur, qui ferme les yeux car il a de la sympathie pour les altermondialistes et n'aime pas la « malbouffe », ils prennent des poignées de cerises, par cinq, pour apaiser leur faim et leur soif. Quel est la loi du nombre de bonnes cerises dans la première poignée ? Calculer la probabilité qu'on puisse en consommer au moins quatre avec plaisir dans une poignée donnée.

11 Les clients de la boutique de l'Opéra Garnier

Le nombre de clients qui pénètrent dans la boutique de l'Opéra Garnier à Paris au mois de février est, en moyenne, de dix par heure en semaine. En l'absence de toute autre information, on peut (et même on doit) modéliser le nombre de clients qui rentrent dans ce magasin au cours d'une période de t heures comme une variable de Poisson de paramètre $\lambda = 10t$. Quelle est la probabilité qu'au moins cinq clients pénètrent dans la boutique entre dix heures et dix heures quinze ? Quelle est la probabilité qu'il n'arrive pas plus de trois clients pendant la période allant de dix heures à dix heures dix ou de onze heures à onze heures cinq ?

12 Cléo et Chloé jouent à pile ou face (1)

Cléo et Chloé jouent à pile ou face avec une pièce non truquée. Cléo joue quatre fois. Chloé joue trois fois. Calculer la probabilité que les deux amies ont de tirer au total au moins 5 fois le côté « pile ».

13 Cléo et Chloé jouent à pile ou face (2)

Chloé et Cléo changent maintenant légèrement de jeu. Cléo lance la pièce jusqu'à obtenir trois fois le côté « pile » ; Chloé lance la pièce jusqu'à obtenir deux fois le côté

« pile ». Calculer la probabilité qu'il leur faille au moins huit tirages.

Variables à densité

14 Modélisation continue des temps d'attente

On s'intéresse, ici, aux temps d'attente des clients dans un magasin de luxe, situé rue de la Paix, à Paris. On suppose que la propriétaire est âgée de 58 ans, et que la clientèle a plus l'habitude de fréquenter le Café des Deux Magots, à Saint-Germain-des-Prés, plutôt que des cafés proches de cette boutique de luxe, comme, par exemple, le Café de la Paix. Supposons que la propriétaire ait calculé, comme fréquence d'arrivée des clients, 6 par heure, en moyenne. Elle décide de modéliser l'arrivée du premier client sans se restreindre aux moments d'arrivée mesurés en minutes entières. Elle a remarqué, contrairement à l'avis de son gérant, que la durée d'attente jusqu'à l'arrivée du premier client ne semble pas être liée au temps passé à attendre sans que le premier client n'arrive ; en réalité, elle pense que son gérant est superstitieux de considérer qu'une attente longue est de mauvais augure pour le reste de la journée.

Les seules informations utiles à retenir sont donc les suivantes :

- on modélise le temps d'attente X comme une variable qui peut prendre n'importe quelle valeur réelle positive ; on supposera, à cet effet, que X a une densité f_X .
- La propriétaire a noté la propriété d'absence de mémoire. on peut donc, légitimement, supposer que X est de loi exponentielle.
- La propriétaire a calculé la fréquence moyenne du nombre de clients, qui vaut 6 : on peut interpréter ce résultat en disant que l'on doit prendre $\lambda = 6$ (l'unité de X est l'heure). On peut aussi se baser sur le fait que si l'on a, en moyenne, six clients par heure, on peut répéter l'expérience en cherchant plutôt la durée moyenne d'attente du premier client (voire la durée moyenne d'attente entre l'arrivée de deux clients successifs, ces durées étant supposées indépendantes). On trouvera alors

vraisemblablement $\mathbb{E}[X] = \frac{1}{6}$, on attend le client dix minutes en moyenne. Comme $\mathbb{E}[X] = \frac{1}{\lambda}$, ceci conduit précisément à $\lambda = 6$.

(Cette dernière opération est un cas très particulier de ce que l'on appelle en statistique la « méthode des moments » pour l'estimation de paramètres.)

Calculer :

$$P(X > 10 \text{ min}) \quad , \quad P(X > 20 \text{ min}) \quad , \quad P(X > 60 \text{ min})$$

15 Retour au magasin de luxe

Supposons qu'un employé de notre magasin de luxe de la rue de la Paix souhaite fumer une cigarette sans incommoder les clients. Lorsqu'il est pressé, cette activité lui prend 60 secondes. Il est intéressant de calculer la probabilité qu'il parvienne à finir sa cigarette avant que n'arrive le prochain client.

15.a) Supposons, d'abord, que l'employé ne soit pas seul dans le magasin, et qu'il ait promis à son coéquipier de servir le second client, et qu'il parvienne à allumer discrètement sa cigarette et à sortir du magasin au moment même où entre le premier client. Quelle est sa probabilité de succès ?

15.b) Supposons maintenant que l'employé se retrouve seul dans le magasin, et que le premier client rentre et reste trois minutes avant de ressortir (il avait passé une commande réglée sur internet, et n'a eu qu'à récupérer celle-ci).

La durée d'attente X entre le premier et le second client est donc supérieure à trois minutes : $X > \frac{3}{60}$.

L'employé se précipite dehors, une minute après le départ du client (il ne veut pas faire mauvaise impression au cas où celui-ci serait toujours dans la rue).

Ainsi, $X > \frac{4}{60}$.

Quelle est la probabilité qu'il arrive à fumer sa cigarette avant l'arrivée du second client ?

16 Une application du théorème central limite

On s'amuse à lancer un dé 400 fois de suite, et on désigne par X la v.a. donnant le nombre d'apparitions d'un multiple de 2 (i.e. deux, quatre ou six). X suit la loi binomiale $\mathcal{B}\left(400, \frac{1}{2}\right)$. Calculer son espérance et son écart-type, puis la probabilité que X soit compris entre 220 et 230.

17 Les élections présidentielles en Floride (Fiction)

Lors de l'élection présidentielle américaine de novembre 2000, la chaîne de télévision CNN a effectué un sondage à la sortie des urnes (« exit poll ») des bureaux de vote de l'Etat de Floride, sur un échantillon de 1 000 personnes. Le candidat « W » recueille a pour cent des suffrages des 1 000 personnes interrogées. On suppose que la valeur a est proche de 50.

17.a) À l'aide du théorème central limite, donner un intervalle de confiance à 95% pour la véritable proportion μ de votants qui ont choisi le candidat « W » en Floride. On assimile cette proportion au score S_W réalisé par « W ».

17.b) À partir de quelle valeur de $|a - 50|$ peut-on se baser sur le sondage pour connaître le résultat de l'élection avec 95 % de certitude ?

17.c) Sur un effectif total de 6 000 000 de votants floridiens, la différence de votes entre le score de « W » et celui de l'autre candidat est de 4 000 voix. Combien de personnes aurait-il fallu interroger dans l'« exit poll » de CNN pour pouvoir déclarer le vainqueur avec 95 % de certitude ? On notera N le nombre recherché.

Formulaires

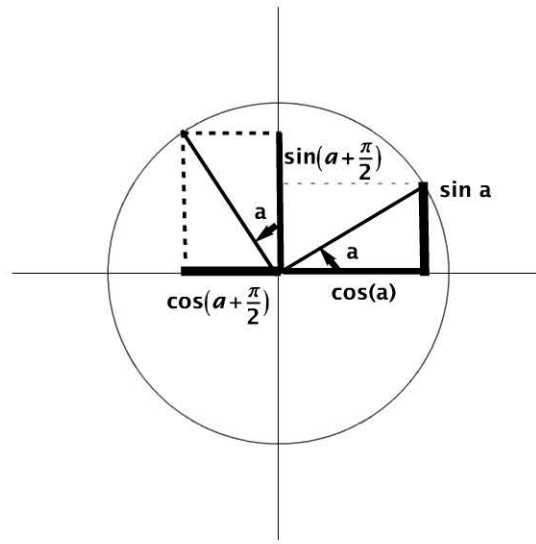
Formulaire de trigonométrie

1. Formules d'addition

Pour tout couple de réels (a, b) :

$$\begin{cases} \cos(a+b) = \cos a \cos b - \sin a \sin b \\ \cos(a-b) = \cos a \cos b + \sin a \sin b \\ \sin(a+b) = \sin a \cos b + \cos a \sin b \\ \sin(a-b) = \sin a \cos b - \cos a \sin b \end{cases}$$

$$\begin{cases} \cos\left(a + \frac{\pi}{2}\right) = -\sin a \\ \cos\left(\frac{\pi}{2} - a\right) = \sin a \\ \sin\left(a + \frac{\pi}{2}\right) = \cos a \\ \sin\left(\frac{\pi}{2} - a\right) = \cos a \end{cases}$$



Formule d'addition avec un angle droit.

et, lorsque $(a, b) \in \left(\mathbb{R} \setminus \left\{\frac{\pi}{2} + k\pi, k \in \mathbb{Z}\right\}\right)^2$:

$$\begin{cases} \tan(a+b) = \frac{\tan a + \tan b}{1 - \tan a \tan b} \text{ si } a+b \in \mathbb{R} \setminus \left\{\frac{\pi}{2} + k\pi, k \in \mathbb{Z}\right\} \\ \tan(a-b) = \frac{\tan a - \tan b}{1 + \tan a \tan b} \text{ si } a-b \in \mathbb{R} \setminus \left\{\frac{\pi}{2} + k\pi, k \in \mathbb{Z}\right\} \end{cases}$$

2. Formules de duplication

Pour tout réel a :

$$\begin{cases} \cos(2a) = \cos^2 a - \sin^2 a = 2 \cos^2 a - 1 = 1 - 2 \sin^2 a \\ \sin(2a) = 2 \sin a \cos a \end{cases}$$

et, lorsque $a \neq \frac{\pi}{2} [\pi]$ et $2a \neq \frac{\pi}{2} [\pi]$:

$$\tan(2a) = \frac{2 \tan a}{1 - \tan^2 a}$$

3. Formules de linéarisation

Pour tout réel a :

$$\begin{cases} \cos^2 a = \frac{1 + \cos(2a)}{2} \\ \sin^2 a = \frac{1 - \cos(2a)}{2} \end{cases}$$

4. Formules de Simpson¹ pour la transformation de produits en sommes

Pour tout couple de réels (a, b) :

$$\begin{cases} \cos a \cos b = \frac{1}{2} \{ \cos(a+b) + \cos(a-b) \} \\ \sin a \sin b = \frac{1}{2} \{ \cos(a-b) - \cos(a+b) \} \\ \sin a \cos b = \frac{1}{2} \{ \sin(a+b) + \sin(a-b) \} \end{cases}$$

5. Formules de Simpson pour la transformation de sommes en produits

Pour tout couple de réels (p, q) :

$$\begin{cases} \cos p + \cos q = 2 \cos\left(\frac{p+q}{2}\right) \cos\left(\frac{p-q}{2}\right) \\ \cos p - \cos q = -2 \sin\left(\frac{p+q}{2}\right) \sin\left(\frac{p-q}{2}\right) \\ \sin p + \sin q = 2 \sin\left(\frac{p+q}{2}\right) \cos\left(\frac{p-q}{2}\right) \\ \sin p - \sin q = 2 \sin\left(\frac{p-q}{2}\right) \cos\left(\frac{p+q}{2}\right) \end{cases}$$

6. Expression en fonction de la tangente de l'angle moitié

Pour tout réel $a \neq \pi [2\pi]$, on pose

$$t = \tan\left(\frac{a}{2}\right)$$

On a alors :

$$\begin{cases} \cos a = \frac{1 - t^2}{1 + t^2} \\ \sin a = \frac{2t}{1 + t^2} \end{cases}$$

Si, de plus, $a \neq \frac{\pi}{2} [\pi]$, on a :

$$\tan a = \frac{2t}{1 - t^2}$$

1. Thomas Simpson (1710-1761), mathématicien anglais, essentiellement connu pour ses travaux sur le calcul infinitésimal (la méthode de Simpson, qui permet un calcul approché de l'aire sous une courbe), mais qui fut aussi l'auteur d'un important traité de trigonométrie.

Dérivées usuelles

Fonction	Dérivée	Domaine de validité
$x \mapsto x^\alpha, \alpha \in \mathbb{R}^* \setminus \{-1\}$	$x \mapsto \alpha x^{\alpha-1}$	\mathbb{R}_+^*
$x \mapsto \frac{1}{x}$	$x \mapsto -\frac{1}{x^2}$	\mathbb{R}^*
$x \mapsto \ln x$	$x \mapsto \frac{1}{x}$	\mathbb{R}_+^*
$x \mapsto e^x$	$x \mapsto e^x$	\mathbb{R}
$x \mapsto \cos x$	$x \mapsto -\sin x$	\mathbb{R}
$x \mapsto \sin x$	$x \mapsto \cos x$	\mathbb{R}
$x \mapsto \tan x$	$x \mapsto 1 + \tan^2 x = \frac{1}{\cos^2 x}$	$\mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi, k \in \mathbb{Z} \right\}$
$x \mapsto \cotan x = \frac{1}{\tan x}$	$x \mapsto -\frac{1}{\sin^2 x}$	$\mathbb{R} \setminus \pi\mathbb{Z}$
$x \mapsto \operatorname{ch} x = \frac{e^x + e^{-x}}{2}$	$x \mapsto \operatorname{sh} x$	\mathbb{R}
$x \mapsto \operatorname{sh} x = \frac{e^x - e^{-x}}{2}$	$x \mapsto \operatorname{ch} x$	\mathbb{R}
$x \mapsto \operatorname{th} x = \frac{\operatorname{sh} x}{\operatorname{ch} x}$	$x \mapsto 1 - \operatorname{th}^2 x$	\mathbb{R}
$x \mapsto \operatorname{coth} x = \frac{1}{\operatorname{th} x}$	$x \mapsto -\frac{1}{\operatorname{sh}^2 x}$	\mathbb{R}^*

Dérivées des fonctions réciproques usuelles

Fonction	Dérivée	Domaine de validité
$x \mapsto \arccos x$	$x \mapsto -\frac{1}{\sqrt{1-x^2}}$	$] -1, 1[$
$x \mapsto \arcsin x$	$x \mapsto \frac{1}{\sqrt{1-x^2}}$	$] -1, 1[$
$x \mapsto \arctan x$	$x \mapsto \frac{1}{1+x^2}$	\mathbb{R}
$x \mapsto \operatorname{argsh} x$	$x \mapsto \frac{1}{\sqrt{x^2+1}}$	\mathbb{R}
$x \mapsto \operatorname{argch} x$	$\frac{1}{\sqrt{x^2-1}}$	$]1, +\infty[$
$x \mapsto \operatorname{argth} x$	$x \mapsto \frac{1}{1-x^2}$	$] -1, 1[$

Primitives usuelles

Fonction	Primitive	Domaine de validité
$x \mapsto \frac{1}{x}$	$x \mapsto \ln x $	\mathbb{R}^*
$x \mapsto e^x$	$x \mapsto e^x$	\mathbb{R}
$x \mapsto \cos x$	$x \mapsto \sin x$	\mathbb{R}
$x \mapsto \sin x$	$x \mapsto -\cos x$	\mathbb{R}
$x \mapsto \tan x$	$x \mapsto -\ln \cos x $	$\mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi, k \in \mathbb{Z} \right\}$
$x \mapsto \frac{1}{\cos^2 x}$	$x \mapsto \tan x$	$\mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi, k \in \mathbb{Z} \right\}$
$x \mapsto \frac{1}{\sin^2 x}$	$x \mapsto -\cotan x = -\frac{1}{\tan x}$	$\mathbb{R} \setminus \pi \mathbb{Z}$
$x \mapsto -\frac{1}{\sqrt{1-x^2}}$	$x \mapsto \arccos x$	$] -1, 1[$
$x \mapsto \frac{1}{\sqrt{1-x^2}}$	$x \mapsto \arcsin x$	$] -1, 1[$
$x \mapsto \frac{1}{1+x^2}$	$x \mapsto \arctan x$	\mathbb{R}
$x \mapsto \operatorname{ch} x = \frac{e^x + e^{-x}}{2}$	$x \mapsto \operatorname{sh} x$	\mathbb{R}
$x \mapsto \operatorname{sh} x = \frac{e^x - e^{-x}}{2}$	$x \mapsto \operatorname{ch} x$	\mathbb{R}
$x \mapsto \operatorname{th} x = \frac{\operatorname{sh} x}{\operatorname{ch} x}$	$x \mapsto \ln(\operatorname{ch} x)$	\mathbb{R}
$x \mapsto \frac{1}{\operatorname{ch}^2 x}$	$x \mapsto \operatorname{th} x$	$\mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi, k \in \mathbb{Z} \right\}$
$x \mapsto \frac{1}{\operatorname{sh}^2 x}$	$x \mapsto -\operatorname{coth} x$	\mathbb{R}^*
$x \mapsto \frac{1}{\sqrt{x^2+1}}$	$x \mapsto \operatorname{argsh} x$	\mathbb{R}
$x \mapsto \frac{1}{\sqrt{x^2-1}}$	$x \mapsto \operatorname{argch} x$	\mathbb{R}
$x \mapsto \frac{1}{1-x^2}$	$x \mapsto \operatorname{argth} x$	\mathbb{R}
$x \mapsto \ln x$	$x \mapsto x \ln x - x$	\mathbb{R}_+^*

Limites usuelles des fonctions puissances

Soit n un entier naturel non nul. Alors :

$\lim_{x \rightarrow -\infty} x^{2n} = +\infty$	$\lim_{x \rightarrow +\infty} x^{2n} = +\infty$
$\lim_{x \rightarrow 0^+} x^{2n} = 0^+$	$\lim_{x \rightarrow 0^-} x^{2n} = 0^+$
$\lim_{x \rightarrow -\infty} x^{2n+1} = -\infty$	$\lim_{x \rightarrow +\infty} x^{2n+1} = +\infty$
$\lim_{x \rightarrow 0^+} x^{2n+1} = 0^+$	$\lim_{x \rightarrow 0^-} x^{2n+1} = 0^-$
$\lim_{x \rightarrow -\infty} x^{-2n} = 0^+$	$\lim_{x \rightarrow +\infty} x^{-2n} = 0^+$
$\lim_{x \rightarrow 0^+} x^{-2n} = +\infty$	$\lim_{x \rightarrow 0^-} x^{-2n} = +\infty$
$\lim_{x \rightarrow -\infty} x^{-2n-1} = 0^-$	$\lim_{x \rightarrow +\infty} x^{-2n-1} = 0^+$
$\lim_{x \rightarrow 0^+} x^{-2n-1} = +\infty$	$\lim_{x \rightarrow 0^-} x^{-2n-1} = -\infty$

Rang d'une matrice

Définition

Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$.

On appelle **rang de la matrice** A , et on note $rg A$, la taille maximale d'une matrice carrée inversible extraite² de A .

Exemples

1.

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

La matrice A étant de taille 3×4 , son rang sera inférieur ou égal à 3.

La matrice $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ extraite de A ayant un déterminant non nul, elle est inversible :

$$rg A = 3$$

2.

$$B = \begin{pmatrix} 0 & 1 \\ 0 & 2 \end{pmatrix}$$

La matrice B étant de taille 2×2 , son rang sera inférieur ou égal à 2.

Seules les matrices $\begin{pmatrix} 1 \end{pmatrix}$ et $\begin{pmatrix} 2 \end{pmatrix}$ extraites de B ont un déterminant non nul :

$$rg B = 1$$

2. c'est-à-dire obtenue en enlevant des lignes et/ou des colonnes à A .

Bibliographie

Partie 1

Ouvrages ou références

- [1] *Histoires de logarithmes*, Commission Inter-IREM d'épistémologie et d'histoire des mathématiques, Éllipses, 2006.
- [2] *Der Briefwechsel von Gottfried Wilhelm Leibniz mit Mathematikern*, Hrsg. von C. J. Gerhardt. 1. Bd, Published 1899 by Mayer & Müller in Berlin, Library of Congress QA29 L45 A4 1899.
- [3] Mauricio Garay, *La géométrie des lumières à la Belle Époque*, Quadrature, numéro 87, 2013.
- [4] Nicolas Lerner, *Lecture Notes on Real Analysis*, UPMC, 2011, <http://www.math.jussieu.fr/~lerner/>.
- [5] Claire David et Pierre Gosselet, *Équations aux dérivées partielles*, Dunod, Paris, 2012.
- [6] Claire David, *Calcul Vectoriel*, Dunod, Paris, 2012.
- [7] *L'ami de la religion et du roi*, Journal ecclésiastique, Volume 46, 1825.
- [8] Jean-Robert Argand, *Essai sur une manière de représenter des quantités imaginaires dans les constructions géométriques*, Gauthier-Villars, Paris, 1874, <http://www.bibnum.education.fr/mathematiques/geometrie/essai-sur-une-maniere-de-representer-des-quantites-imaginaires-dans-les-cons>
- [9] Michèle Audin, *Géométrie*, EDP Sciences, Les Ullis, 2006.
- [10] Jos Leys, Étienne Ghys, Aurélien Alvarez, *Dimensions, une promenade mathématique*, http://www.dimensions-math.org/Dim_fr.htm.
- [11] Adrien Douady et John H. Hubbard, *Étude dynamique des polynômes complexes*, Prépublications mathématiques d'Orsay 2/4 (1984 / 1985).
- [12] Benoît Mandelbrot, *Les objets fractals : Forme, hasard et dimension*, Flammarion, Paris, 2010.
- [13] <http://www.maths.anu.edu.au/~barnsley>.
- [14] Michael Barnsley, *SuperFractals*, Cambridge University Press, 2006.
- [15] <http://www.superfractals.com>.
- [16] Daniel Lignon et Roger Beslan, *Les maths, cent théorèmes*, Le polygraphe, Angers, 2008.
- [17] Arthur Cayley, *À memoir on the Theory of Matrices*, Philosophical Transactions, Londres, 1858.
- [18] <http://www.math93.com/theoreme/matrices.html>.
- [19] Denis Serre, *Matrices : Theory and Applications*, Springer, New-York, 2002.
- [20] David Lay, *Linear Algebra and its Applications*, 4th edition, Addison-Wesley, New-York, 2012.
- [21] Michèle Audin, *Géométrie*, EDP Sciences, Les Ullis, 2006.
- [22] Raymond Queneau, *Bâtons, chiffres et lettres*, Gallimard, Paris, 1965.
- [23] Henri Lebesgue, *L'œuvre mathématique de Vandermonde*, in *Notices d'Histoire des Mathématiques*, Université de Genève 1958, pp. 18-39.

- [24] Ilia Itenberg, *Leçons de Mathématiques d'aujourd'hui*, Volume 4, Cassini, Paris, 2010.
- [25] Erwan Brugalle, *Un peu de géométrie tropicale*, Quadrature, n° 74, p. 10-22, Paris, 2009.
- [26] *Algèbres tropicales et plus court chemin*, Quadrature, n° 72, pp. 22-28, avril 2009.
- [27] Stella Brassesco, Miguel A. Méndez, The asymptotic expansion for $n!$ and Lagrange inversion formula, <http://arxiv.org/pdf/1002.3894v1.pdf>.
- [28] Maryse Beguin, *Théorie de la mesure et de l'intégration pour les probabilités*, Ellipses, Paris, 2013.
- [29] Jean-Marie Arnaudiès, Henri Fraysse, *Cours de Mathématiques*, tomes 1 à 3, Dunod, Paris, 1997.
- [30] Xavier Gourdon, *Les Maths en tête, Algèbre*, Éllipses, Paris, 2009.
- [31] Xavier Gourdon, *Les Maths en tête, Analyse*, Ellipses, Paris, 2009.
- [32] Jean-Pierre Ramis, André Warusfel (dir.), *Mathématiques Tout-en-un pour la Licence - Niveau L1 - 2^e édition*, Dunod, Paris, 2013.
- [33] George B. Thomas Jr., Maurice D. Weir, Joel Hass, Thomas' *Calculus : Global Edition*, Pearson Education, Edinbourg, 2009.

Sites web

- [34] <http://www.ljll.math.upmc.fr/~ledret>.
- [35] Jacques Fejoz, Calcul vectoriel et matriciel de première année, <http://www.ceremade.dauphine.fr/~fejoz/>.
- [36] <http://demonstrations.wolfram.com>.
© 2012 Wolfram Demonstrations Project www.wolfram.com
- [37] Wolfram Alpha LLC. 2011. <http://www.wolframalpha.com>.

Partie 3

- [38] Marquis de Condorcet, *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*, 1785.
- [39] <http://www.sciencepresse.qc.ca/blogue/2011/04/28/democratie-loeil-mathematiciens>.
- [40] George B. Thomas Jr., Maurice D. Weir, Joel Hass, Thomas' *Calculus: Global Edition*, Pearson Education, Edinbourg, 2009.
- [41] A review of mathematical topics in collisional kinetic theory. In *Handbook of Mathematical Fluid Dynamics*, S. Friedlander and D. Serre, Eds, Elsevier Science, 2002.
- [42] <http://demonstrations.wolfram.com>.
© 2012 Wolfram Demonstrations Project www.wolfram.com
- [43] Wolfram Alpha LLC. 2011. <http://www.wolframalpha.com>.
- [44] Claire David et Pierre Gosselet, *Équations aux dérivées partielles*, Dunod, Paris, 2012.
- [45] Claire David, *Calcul Vectoriel*, Dunod, Paris, 2012.
- [46] Werner Heisenberg, *Les principes physiques de la théorie des quanta*, Gauthier-Villars (1932), Réédition par Jacques Gabay (1989).
- [47] William Feller, *An introduction to Probability theory and its applications*, third edition, Wiley, 1968.

Index

A

Abbasides 50
Abbé Buée 210
Abélien 215
Abraham De Moivre 217, 469, 471
Absence de mémoire 402, 462
Abu Al-Wafa 50
Additivité
 de l'intégrale double 194
 des variances 436
Adhérence
 (d'un intervalle) 7
 de \mathbb{R} 3
ADN 370
Adrien Douady 210
Affixe 211
Aire d'un parallélogramme 300, 301
Al-Kashi 50
Al-Khwarizmi 50
Algébriquement clos 225
Algorithme 50
 de Gauss-Jordan 274
 Fang-Shen 274
Angle entre deux vecteurs de \mathbb{R}^3 298
Anneau 215
Anneau commutatif 214
Antécédent 5
Application 5
Application identité 72, 354
 de \mathbb{R}^n 335
 de l'espace \mathbb{R}^3 319
 du plan \mathbb{R}^2 305
Application linéaire 354
 de l'espace \mathbb{R}^3 318
 de l'espace \mathbb{R}^n 335
 du plan 304
Arccosinus 76
Arcsinus 75
Arctangente 77
Argand Jean-Robert 206, 210
Argument 211
 cosinus hyperbolique 79
 principal 211
 sinus hyperbolique 79
 tangente hyperbolique 80
Arrangements 370

avec répétition 371

Aryabhata 49
Associativité du barycentre 288
Astrolabe 49
Asymptote verticale 19
Auto-indéfinie-divisibilité de la loi de Poisson 417
Automorphisme 338, 307, 321, 354
Axe
 d'une rotation 328
 de symétrie vertical 31
Axiome de la borne supérieure 21, 123

B

Bachman 16, 144
Barycentre
 d'un système pondéré de n points 287, 288
 d'un système pondéré de deux points 286, 287
Base directe de \mathbb{R}^3 298
Baseball 490
Bernoulli Daniel 196
Bernoulli Jacques 44
Bernoulli Nicolas 44
Bijectivité 72
Bombelli 209
Bootstrap 479
Borel-Cantelli (Lemme de) 448
Branche parabolique d'axe
 d'équation $y = mx$, $m \in \mathbb{R}$ 20
 horizontal 20
 vertical 19
Briggs Henry 38
Buée 210

C

Caractérisation
 de l'indépendance de deux v.a.r. à densité 474
 des fonctions constantes dérivables 63
 des fonctions croissantes dérivables 64
 des fonctions décroissantes dérivables 64
 des fonctions strictement croissantes dérivables 64
 des fonctions strictement décroissantes dérivables 64

- des projecteurs 358
 - des symétries 358
 - séquentielle de la continuité 53
 - de Weierstrass 51
 - Cardan 209
 - Cauchy Augustin 210, 250
 - Causalité 380
 - Centre de gravité 288
 - Cesàro Ernesto 134
 - Changement
 - d'indices 182
 - de base 308, 322, 339
 - de variable 171
 - Chebyshev 437
 - Coefficient de corrélation 443
 - Coefficients binomiaux 265
 - Cofacteur d'indice (i, j) 259, 270
 - Comatrice 259, 270
 - Combinaisons 372
 - linéaire 282, 331, 351
 - Comparaison avec une intégrale
 - convergente (sur un intervalle de longueur finie) 191
 - convergente (sur un intervalle de longueur infinie) 189
 - divergente (sur un intervalle de longueur finie) 191
 - divergente (sur un intervalle de longueur infinie) 190
 - Comparaison des fonctions puissances non entières 43
 - Complémentaire d'un ensemble 4
 - Composition (de fonctions) 22
 - Concentration de la mesure 470
 - Condition
 - initiale (pour une équation différentielle linéaire du premier ordre homogène) 102
 - nécessaire de convergence d'une série 150
 - Condition nécessaire et suffisante de convergence 188
 - d'une intégrale de Riemann (sur un intervalle de longueur infinie) 187
 - d'une série de Riemann 152
 - d'une série géométrique 151
 - Conditions initiales 117
 - Condorcet 398
 - Conjugaison 311, 325, 342
 - Conjugué 208
 - Continuité
 - des fonctions composées 54
 - des fonctions usuelles 54
 - à droite 52
 - à gauche 51
 - d'une fonction en un point 51
 - des fonctions composées 55
 - des fonctions usuelles 55
 - et opérations algébriques 55
 - Convergence d'une série numérique 149
 - Convexité et dérivabilité 99
 - Corde 58
 - sous-tendue 49
 - Corps de base 350, 351
 - Cosécante 50
 - Cosinus 45
 - hyperbolique 47
 - Couple de variables aléatoires discrètes 445
 - Coupures de Dedekind 21
 - Courbe représentative (d'une fonction) 18
 - Covariance 442
 - de deux v.a. indépendantes 442
 - Cramer 250
 - Critère
 - de Cauchy 142
 - de d'Alembert 154
 - Croissance 26
 - (au sens strict) 26
 - (pour une intégrale) 159, 165
 - de l'espérance 421, 453
 - comparées 43
 - Curie Pierre 44
- ## D
- D'Alembert Jean le Rond 210
 - Décomposition en éléments simples 231
 - sur $\mathbb{R}(X)$ d'une fraction rationnelle impaire 238
 - sur $\mathbb{R}(X)$ d'une fraction rationnelle paire 236
 - Décroissance 26
 - Dedekind Richard 21
 - Degré de confiance 455
 - De l'Hôpital Guillaume 196
 - De Gennes Pierre-Gilles 44
 - Demi-tangente
 - à droite 62
 - à gauche 63
 - De Nicée Hipparque 49
 - Densité
 - conditionnelle de deux v.a.r. 473
 - de probabilité 450
 - de probabilité sur \mathbb{R}^2 472
 - jointe 472

marginales de deux v.a.r. 472
 Dérivabilité
 des fonctions composées 60, 61
 des fonctions définies par morceaux 62
 en un point 58
 en un point et opérations algébriques 59
 et opérations algébriques 61
 sur un intervalle 61
 Décroissance (au sens strict) 26
 Dérivée 58
 à gauche d'une fonction 63
 Dérivée $n^{\text{ième}}$ 65
 d'un quotient de fonctions, $n \in \mathbb{N}$ 66
 d'une combinaison linéaire de fonctions,
 $n \in \mathbb{N}$ 65
 de la composée de deux fonctions, $n \in \mathbb{N}$
 66
 Déterminant
 d'une application linéaire de \mathbb{R}^2 311
 d'une application linéaire de \mathbb{R}^3 325
 d'une application linéaire de \mathbb{R}^n 342
 d'une matrice de taille 2×2 254
 d'une matrice de taille 3×3 259
 d'une matrice diagonale 271
 d'une matrice triangulaire 271
 de deux vecteurs du plan 254
 de trois vecteurs de l'espace 261
 du produit de deux matrices d'ordre n 271
 Déterminisme 380
 Développement
 asymptotique 95, 146
 limité au voisinage d'un point 81
 limité de la fonction exponentielle au voi-
 sinage de zéro 85
 limité de la fonction logarithme népérien
 au voisinage de 1 84
 limité des fonctions *sinus* et *cosinus* au
 voisinage de zéro 85
 limités usuels 89
 Diagonalisabilité 344
 Di Bruno Faà 66
 Différence de deux ensembles 4
 Dilatation 253
 Discriminant réduit 226
 Divergence d'une série numérique 149
 Division euclidienne (de deux polynômes)
 230
 Domaine de définition (d'une fonction) 18
 Domination 16, 144
 Droite
 affine 290
 asymptote à une courbe 19

réelle achevée 3
 vectorielle 290

E

Écart-type 431
 Écriture cartésienne 207
 Ensemble 2
 Ensemble vide 4, 6
 Épreuve 376
 Équation
 caractéristique pour une suite récurrente
 linéaire d'ordre deux 118
 différentielle linéaire du premier ordre
 avec second membre 103
 homogène associée à une équation diffé-
 rentielle linéaire du premier ordre
 avec second membre 103
 Équiprobabilité 379
 Équivalence
 de deux fonctions 17
 de deux suites 145
 Espace
 caractéristique 344
 complet 142
 probabilisé 377
 propre 344
 vectoriel 347
 vectoriel de dimension finie 352
 vectoriel de dimension infinie 352
 vectoriel des suites 112
 vectoriel produit 353
 Espérance 419
 et valeur absolue 453
 mathématique d'un produit de v.a. indé-
 pendantes 435
 Espérance d'une v.a.r.
 suivant la loi Gamma de paramètres $\alpha >$
 $0, \beta > 0$ 485
 suivant la loi normale $\mathcal{N}(\mu, \sigma^2)$ 468
 suivant la loi uniforme (sur un intervalle
 $[a, b]$) 458
 suivant la loi exponentielle $\mathcal{E}(\lambda)$ 463
 Euler 44, 210
 Événement 376
 élémentaire 376
 indépendants 388
 négligeable 379
 presque sûr 378
 Expérience aléatoire 376
 Exponentielle complexe 213

Extremum

- d'une fonction sur un intervalle 67
- local d'une fonction 68

F

Factorielle 368

Famille d'événements observables) 377

Fermat 380

Fincke Thomas 50

Formule de Taylor Lagrange 71

Fonction 5

- n fois dérivable, $n \in \mathbb{N}$ 65
 - concave 96
 - continue par morceaux 162
 - continue partout, mais nulle part dérivable 59
 - contractante 131
 - convexe 96
 - cosinus 45
 - d'une variable aléatoire discrète 390
 - de Dirichlet 184
 - de probabilité 390
 - de répartition d'un couple de v.a.r. 472
 - de répartition pour la loi exponentielle de paramètre $\lambda > 0$ 461
 - de répartition pour la loi uniforme 457
 - de répartition d'une variable aléatoire discrète 392
 - de survie 462
 - de survie d'une v.a.r. positive 455
 - en escalier 157
 - exponentielle 41
 - Fonction de classe C^∞ 65
 - Fonction de classe C^n , $n \in \mathbb{N}$ 65
 - Gamma 480
 - indéfiniment dérivable 65
 - logarithme népérien 39
 - polynomiale 35
 - puissance 33, 498
 - puissance non entière 43
 - réciproque 72
 - sinus 45
 - tangente 46
 - valeur absolue 37
 - continues sur un intervalle 55
 - définies par morceaux 23
 - monotones 26
 - périodiques 32
 - symétriques des racines 228
- Forme
- canonique 175

indéterminée 13

Formule

- de Bayes 386, 387
- de changement de base pour un vecteur 308, 322, 339
- de Grassmann 334
- de Koenig 432, 444, 454
- de la moyenne 167
- de Leibniz 66, 234
- de Moivre 217
- de Stirling 179, 369
- de Taylor-Young 84, 87
- de Vandermonde 374, 418
- des probabilités composées généralisée 384
- des probabilités totales 385
- des probabilités composées 384
- du binôme de Newton 373
- du binôme de Newton 265, 337
- du rang dans \mathbb{R}^2 307
- du rang dans \mathbb{R}^3 321
- du rang dans \mathbb{R}^n 338
- d'addition 50, 493
- d'Euler 217
- de Cramer 279
- de duplication 493
- de linéarisation 494
- de Simpson pour la transformation de produits en sommes 494
- de Simpson pour la transformation de sommes en produits 494

G

Gauss Carl Friedrich 250

Géométrie tropicale 303

Girard Albert 50

Gottfried Leibniz 44

Grand O 16, 17, 144

Graphe

- d'une fonction 18
- d'une fonction réciproque 74

Grassmann Hermann 334

Grauert Hans 210

Grothendieck Alexandre 210

Groupe

- commutatif 215, 350
- linéaire 355
- linéaire d'ordre n 272
- linéaire de \mathbb{R}^n 338
- orthogonal 312, 327
- spécial orthogonal 313, 316, 327, 329, 347

Guillaume François Antoine de l'Hôpital
196

H

Heisenberg Werner 380

Hilbert David 210

Homographie 126

Homothétie 240

I

Idempotent 358

Identité de polarisation 313, 327

Image 5

d'un intervalle fermé et borné par une
fonction continue 57

d'un intervalle par une fonction continue
strictement monotone 73

d'une application linéaire 307, 320, 337,
355

Imaginaire pur 207

Imparité 28

Imre Simon 303

Inégalité

de Cauchy-Schwarz 294, 440

de convexité 97

de Jensen 97, 425

de la moyenne 166

de Markov 426, 437, 454

de Tchebychev 437, 454

des pentes croissantes 98

triangulaire 36, 294

Indépendance de deux v.a.r. 473

Indice

de colonne 262

de ligne 262

Injectivité 72, 307, 321, 337, 338, 354

des fonctions continues 73

Intégrale

de Gauss 466

d'une fonction continue par morceaux 163

d'une fonction en escalier 158

de Lebesgue 184

de Riemann 158, 187

de Stieltjes 454

double sur un pavé (du plan \mathbb{R}^2) et valeur
absolue 194

double sur un pavé (du plan \mathbb{R}^2) 193

impropre 185, 186

de Wallis 202, 369

Intégration

des développements limités 84

par parties 170

Interprétation géométrique de la valeur abso-
lue 36

Intersection d'ensembles 4

Intervalle 7

de confiance 455, 478

de confiance bilatéral 456

de confiance disymétrique 456

de confiance symétrique 456

de confiance unilatéral 456

fermé borné 6

ouvert 6

ouvert et borné 6

semi-ouvert et borné 6

de confiance avec variance non unitaire
479

Invariance par transvection 271

Inverse

d'un produit de matrices 275

d'une suite 133

de la transposée d'une matrice 274

Involutif 358

Isobarycentre

d'un ensemble de 3 points, A, B, C 288

d'un ensemble de n points 288

Isométrie 313, 326

J

Jensen 97

Julia Gaston Maurice 210

K

\mathbb{K} -espace vectoriel 351

Kronecker Leopold 266

L

Lagrange 67, 148, 250

L'Almageste 49

Landau Edmund 16, 144

Laplace Simon 471

Lebesgue Henri-Léon 183

Leibniz Gottfried 66, 250

Limite(s)

d'un produit de fonctions 15

d'un quotient de fonctions 15

d'une fonction en $+\infty$ ou $-\infty$ 12

d'une fonction polynomiale en $+\infty$ ou $-\infty$
35

d'une somme de fonctions 15

- finie à droite (ou par valeurs supérieures) 10
- finie à gauche (ou par valeurs inférieures) 10
- finie d'une fonction en un point 8
- infinie à droite (ou par valeurs supérieures) 11
- infinie à gauche (ou par valeurs inférieures) 11
- infinie d'une fonction en un point 10
- uniforme 163
- usuelles des fonctions puissances non entières 43
- Linéarité
 - de l'espérance 420, 453
 - (pour une intégrale) 158, 164
 - de l'intégrale double sur un pavé (du plan \mathbb{R}^2) 193
- Logarithme
 - de base a , $a \in \mathbb{R}_+^*$ 40
- Loi(s)
 - binomiale 397
 - binomiale négative 404
 - conditionnelle 446
 - conditionnelle de deux v.a.r. 473
 - conjointe 445
 - d'Erlang (de paramètres $\alpha \in \mathbb{N}^*$, $\beta > 0$) 482
 - d'une variable aléatoire discrète 390
 - de Bernoulli 394
 - de distribution de Maxwell-Boltzmann (de paramètre $a > 0$) 482
 - de Pólya 404
 - de Pascal 404
 - de Poisson 412
 - des cosinus 50
 - du χ^2 (à N degrés de liberté, $N \in \mathbb{N}^*$) 482
 - exponentielle (de paramètre $a \in \mathbb{R}_+^*$) 461
 - faible des grands nombres 447, 448
 - forte des grands nombres 448
 - géométrique 400
 - Gamma (de paramètres $\alpha > 0$, $\beta > 0$) 481
 - hypergéométrique 409
 - marginales 446
 - normale (de paramètres μ , σ) 465
 - normale centrée réduite 466
 - normale et transformation affine 466
 - uniforme 396, 457
- Majorant 123
- Mandelbrot Benoît 210
- Matrice
 - 2×2 252
 - 3×3 256
 - colonne 263
 - de dilatation 268
 - de passage 308, 310, 322, 324, 339, 341
 - de permutation 267
 - de taille $m \times n$ 262
 - de transposition 267
 - de transvection 269
 - diagonale 252, 256, 261
 - identité d'ordre n 266
 - inverse 272
 - invertible 272
 - ligne 262
 - orthogonale 312, 313, 326
 - singulière 272
 - triangulaire 255, 261
 - triangulaire inférieure 252, 256
 - triangulaire supérieure 252, 256
 - semblables 311, 325, 342
- Maximum
 - d'une fonction sur un intervalle 67
 - local d'une fonction 67
- Médianes 288
- Méridiens 49
- Méthode
 - de variation de la constante 109
 - de Simpson 179
 - des moments 492
 - des rectangles 176
 - des trapèzes 178
 - du Pivot de Gauss 274
- Miche Rolle 69
- Mineur d'indice (i, j) 259, 270
- Minimum local d'une fonction 67
- Minorant 123
- Module 211
- Modulo 2π 212
- Moment d'ordre n d'une v.a.r., $n \in \mathbb{N}$ 453
- Monotonie 26
- Multiplication
 - d'un vecteur par un réel 282, 331
 - d'une suite par un scalaire 112, 133
- Multiplicité 225
 - d'une valeur propre 343

N

- Napier John 38
- Nasir Al-Din Al-Tusi 50

Négligeabilité 16, 144

Neper John 38

Newton Isaac 44

Nombre d'or 302

Nombres de Bernoulli 90, 91

Norme 294

Norme euclidienne 293

Notation

« + » 3

« - » 3

« * » 3

« + » 3

« - » 3

de Landau 16, 144

Noyau

d'une application linéaire 306, 320, 337, 355

d'une matrice 306, 320, 337

O

Opérateur orthogonal 312, 326

Opération élémentaire 253, 257, 258

Opérations algébriques sur les fonctions continues 54

Oughtred William 50

P

Parallèles 49

Paramètre de forme 481

Parité 28

Partie

entière 52

entière (d'une fraction rationnelle) 231

imaginaire 207

polaire (d'une fraction rationnelle) 231

principale d'ordre $n \in \mathbb{N}$ d'un développement limité 82

réelle 207

Partition 379

Pas d'une subdivision 157

Pascal Blaise 380

Pavé (du plan \mathbb{R}^2) 193

Période fondamentale 32

Périodicité 32

Petit o 16, 144

Pôle (d'une fraction rationnelle) 230

Pôle d'ordre p (d'une fraction rationnelle) 230

Pivot de Gauss 274

Plan

affine 291

vectériel 291

Poincaré Henri 95

Polynôme(s)

caractéristique 343

de Tchebychev 122

formels 229

Position de la courbe représentative

d'une fonction convexe

par rapport à ses tangentes 99

par rapport à ses cordes 97

Positivité

(pour une intégrale) 158, 164

de l'espérance 421, 453

de l'intégrale double sur un pavé (du plan \mathbb{R}^2) 194

Primitives 62, 168

usuelles 497

Principe

d'incertitude de Heisenberg 380

de superposition 103

Probabilité 377

conditionnelle 384

sur un univers fini 378

Problème des N corps 95

des partis 380

des trois corps 95

Processus stochastique

à saut 412

continu 412

Produit

cartésien de n ensembles, $n \in \mathbb{N}$, $n \geq 2$ 5

cartésien de deux ensembles 4

cartésien de trois ensembles 5

de convolution de deux fonctions 475

de suites 113, 132

scalaire 293

vectériel 298

Projecteur 357

Projection 357

Prolongement par continuité en un point 53

Propriétés

de la covariance 442

de la fonction Gamma 480

Ptolémée 49

Puissance (quelconque) d'un réel strictement positif 42

R

Racine

$n^{\text{ème}}$ d'un réel strictement positif, $n \in \mathbb{N}^*$ 42

carrée complexe de -1 207

- d'ordre n 225
 - d'un polynôme 224
 - double 225
 - triple 225
 - $n^{\text{ièmes}}$ 221
 - Raison (d'une série géométrique) 151
 - Rang
 - d'une application linéaire 307, 321, 338
 - d'une matrice 499
 - \mathbb{R} -espace vectoriel 350
 - \mathbb{R} 3
 - Récurrence
 - double 121
 - forte 121
 - Règle
 - de l'Hôpital 196
 - du bonhomme d'Ampère 298
 - Régularité d'une fonction convexe 98
 - Relation
 - de Chasles (pour une intégrale) 160, 165
 - de récurrence d'ordre p 118
 - de récurrence linéaire d'ordre 2 116
 - coefficients-racines pour un polynôme de degré n 227
 - coefficients-racines pour un polynôme de degré 2 227
 - de Viète 228
 - Repère de l'espace 285
 - Représentation paramétrique
 - d'un plan affine 292
 - d'une droite 290
 - Reste d'ordre n , $n \in \mathbb{N}$, d'une série convergente 150
 - Restriction d'une fonction à un intervalle 22
 - Riemann Bernhard 156
 - Rotation 242
 - vectorielle 314, 328
 - Ruban de Möbius 302
- S**
- Scipione dal Ferro 209
 - Sécante 50, 58
 - Second membre
 - de type « exponentielle \times polynôme » 105
 - de type « exponentielle \times sinus » 108
 - de type « exponentielle \times cosinus » 108
 - Segment 6
 - Sens de variation des fonctions puissances 33
 - Séquençage de l'ADN 370
 - Série 149
 - absolument convergente 150
 - de Riemann 152
 - entière 152
 - géométrique 151
 - harmonique 152
 - semi-convergente 150
 - Sextique de Barth 302
 - Sigma-additivité 378
 - σ -algèbre 377
 - Similitude
 - directe 245
 - indirecte 247
 - Simpson Thomas 179, 494
 - Singleton 6
 - Sinus 45
 - Sinus hyperbolique 47
 - Solution d'une équation différentielle linéaire
 - du premier ordre
 - avec second membre 104
 - homogène 101
 - homogène avec une condition initiale 102
 - Somme
 - d'une série convergente 149
 - de deux séries convergentes 150
 - de deux v.a. indépendantes 416, 417, 475
 - de deux variables aléatoires discrètes indépendantes 416
 - de deux variables aléatoires réelles indépendantes 476
 - de deux vecteurs 282, 330
 - de sous-espaces vectoriels de \mathbb{R}^n 333
 - de suites 112, 132
 - des n premiers carrés, $n \in \mathbb{N}$ 198
 - des n premiers cubes, $n \in \mathbb{N}$ 198
 - des n premiers entiers naturels, $n \in \mathbb{N}$ 198
 - directe de sous-espaces vectoriels 334, 356
 - directe de sous-espaces vectoriels de \mathbb{R}^n 334, 356
 - de Riemann 164, 176
 - partielles 149
 - Sous-espace vectoriel 333, 334, 354, 356
 - Sous-suite 140
 - Spectre 343
 - Stabilité par addition 412
 - Stirling 179
 - Stricte positivité (pour une intégrale) 165
 - Structure
 - de corps 214
 - de groupe 347
 - Subdivision adaptée à une fonction
 - continue par morceaux 162
 - escalier 157

Subdivision
 d'un intervalle 157
 régulière 157
 Suite
 adjacentes 146
 arithmétique 114
 arithmético-géométrique 115
 bornée 124
 constante 124
 convergente 127
 croissante 124
 décroissante 124
 de Cauchy 142
 définies explicitement 114
 divergente 127
 extraite 140
 géométrique 115
 homographique 126, 135
 majorée 123
 minorée 123
 monotone 124
 négative 124
 positive 124
 stationnaire 124
 Support (d'une densité) 474
 Surjectivité 72, 307, 321, 338, 354
 Sylvester 250
 Symétrie 357
 du produit de convolution de deux
 fonctions 475
 Symboles de Kronecker 266
 Système
 complet d'événements 379
 de Cramer 279
 homogène 278
 linéaire 278

T

Tableau de variations 27
 Tangente 46
 hyperbolique 48
 Tartaglia Nicolas 209
 Tchebychev 122, 218, 437
 Termes diagonaux 262
 Théorème
 central limite 469
 de d'Alembert-Gauss 225
 de Berry-Esseen 471
 de Bolzano-Weierstrass 141
 de Cesàro 134
 de De Moivre-Laplace 452, 471

 de Fubini 194
 de la base incomplète 333
 de Rolle 69
 de transfert 422
 de Weierstrass 56, 163
 des accroissements finis 70
 des bornes atteintes 56
 des fonctions réciproques 73
 des gendarmes 15, 147
 des valeurs intermédiaires 56
 du jury de Condorcet 398
 du point fixe 131
 du rang dans \mathbb{R}^2 307
 du rang dans \mathbb{R}^3 321
 du rang dans \mathbb{R}^n 338
 fondamental de l'analyse 168
 Trace d'une matrice
 de taille 2×2 253
 d'une matrice de taille 3×3 257
 d'une matrice de taille $n \times n$ 262
 Transformation linéaire 354
 de l'espace \mathbb{R}^3 318
 de l'espace \mathbb{R}^n 335
 du plan 304
 Transformation orthogonale 312, 313, 326
 Translation 240
 Transposée 252–255, 257–259, 261, 263, 270,
 271
 transposition 253
 Transvection 253
 Triangle de Pascal 66, 373
 Tribu 377
 Trigonalisabilité 345
 Ty Cobb 490

U

Unicité
 de la limite (d'une fonction) 14
 du développement limité 82
 Union d'ensembles 4
 Univers 376

V

Valeur
 absolue d'un réel 36
 absolue d'une intégrale 159, 165
 propre 343
 remarquables des fonctions sinus, cosinus
 et tangente 46
 Vandermonde 250, 374, 418
 Varahamihira 49

- Variable aléatoire (v.a.) 389
 - continue 450
 - discrète 389
 - discrètes indépendantes 391
 - Variance 431
 - d'une v.a. constante 432
 - d'une v.a.r. 454
 - d'une v.a.r. suivant la loi exponentielle $\mathcal{E}(\lambda)$ 464
 - d'une v.a.r. suivant la loi Gamma de paramètres $\alpha > 0, \beta > 0$ 485
 - d'une v.a.r. suivant la loi normale $\mathcal{N}(\mu, \sigma^2)$ 468
 - d'une v.a.r. suivant la loi uniforme (sur un intervalle $[a, b]$) 458
 - de la somme de variables indépendantes et identiquement distribuées (i.i.d.) 436
 - Vecteur normal
 - à un plan affine 295
 - à un plan vectoriel 295
 - libres 283, 331, 352
 - liés 283, 331, 352
 - propre 343
 - unitaire 293
 - Viète 228
 - Voisinage d'un point 7
 - Volume orienté d'un parallélépipède 301
- W**
- Wallis John 202, 369
 - Weierstrass 51, 59
 - Wessel Caspar 210
 - Weyl Hermann 210
- Y**
- Young William Henry 87