

Math-F-112
Mathématiques 1
2023–2024
Prof. Dimitri LEEMANS

Informations au lecteur

Le présent document fait partie du syllabus pour le cours Math-F-112 pour l'année 2023–2024. Il reprend essentiellement les notes de cours de Nicolas Richard qui donnait le cours avant 2017.

Quand un terme est introduit, il est en général indiqué en italique et rappelé dans la marge. Par exemple : *l'italique* est une forme d'écriture qui consiste à pencher les caractères par rapport à leur forme normale. Ce sont ces termes qui forment l'index, en fin de document.

italique

Une version électronique de ce document est disponible (au format PDF), et présente deux avantages non-négligeables :

- * il est « cherchable », c'est-à-dire que par exemple grâce au raccourci clavier bien connu **Control-f**, il est possible de chercher dans le texte du document ; et
- * il est « cliquable », en ce sens que les références internes sont des liens hypertextes permettant de se déplacer dans le document par un clic de souris.

(Cette version contient accessoirement des passages en couleur.)

La version papier du document présente quant à elle tous les avantages du papier, comme par exemple :

- * il s'empporte facilement, partout,
- * il est simple à annoter, commenter, surligner, etc.

Enfin, une version électronique *partagée* est également disponible. Cette version présente l'intérêt d'être annotable de sorte que tout le monde puisse voir –si vous le souhaitez– les annotations que vous y ferez. C'est un outil potentiellement utile pour l'étude collective et partagée. Les détails sur cet outil seront donnés durant les premiers cours.

Notons enfin que le présent document souffre d'un problème. Chacun le sait : il suffit que des explications manquent pour que les mathématiques paraissent insurmontables. Il est intéressant de savoir que le même effet peut être obtenu en donnant *trop d'explications*. Noyer le poisson dans le flot des détails est particulièrement néfaste pour la compréhension. Le juste milieu, l'équilibre, dépend de chaque apprenant et est complexe à déterminer. Par ailleurs, un syllabus s'adressant à *tous* les étudiants, il est impossible de s'adapter à la situation de chacun. C'est pourquoi le lecteur est invité à avoir une lecture particulièrement active. Si des informations semblent manquer ou si des informations semblent contradictoires, les enseignants sont là pour répondre aux questions.

Table des matières

I	Préliminaires	11
1	Définitions, résultats et notations	11
1.1	Définitions	11
1.2	Résultats : théorèmes, lemmes, corollaires, axiomes,	11
1.3	Notations et variables	12
1.4	Alphabets	12
2	Logique, rigueur et notion de preuve	13
3	Manipuler le vrai et le faux	14
3.1	Connecteurs logiques	14
3.1.1	Et	14
3.1.2	Ou	15
3.1.3	Négation	15
3.1.4	Implication	15
3.1.5	Équivalence (ou bi-implication)	16
3.2	Quantificateurs	16
3.2.1	Pour tout	16
3.2.2	Il existe	16
3.2.3	Négation et quantificateurs	17
3.3	Tables de vérités	17
3.4	Raisonnements et preuves	17
3.4.1	Tiers-exclus	18
3.4.2	Démonstration par l'absurde	18
3.4.3	Contraposée	19
3.4.4	Démonstration par récurrence / induction	19
3.4.5	Modélisation	19
II	Les nombres	21
1	Une classification des nombres	21
2	Relations entre les nombres	23
2.1	Inégalités et notion d'ordre	23
2.2	Intervalles	24
2.3	Pourcentages	26
2.4	Principe de proportionnalité – Règle de trois	26
3	Manipulation et opération sur les nombres	27
3.1	Sommes	27
3.2	Moyenne arithmétique	28
3.3	Puissances	29
3.4	Valeur absolue	30
3.5	Écriture scientifique	31
3.6	Racines	31
3.7	Factorielle	32
3.8	Coefficients binomiaux, combinaisons	33
3.9	Formule du binôme de Newton	34
3.10	Triangle de Pascal	35
4	Équations et systèmes	36

4.1	Systèmes d'équations	36
5	Géométrie élémentaire	37
5.1	Longueurs	37
5.2	Aire	37
5.3	Volume	37
5.4	Surface latérale	38
III	Notations ensemblistes	39
1	L'ensemble vide	40
2	Union, intersection et différence	40
3	Diagrammes de Venn — Patatoïdes convexes	40
4	Couples et n-uples	41
IV	Fonctions	43
1	Domaine et image	44
2	Interprétation de la notion de fonction	44
2.1	Association	45
2.2	Transformation	45
3	Surjections, injections, bijections et réciproque	46
3.1	Injection	46
3.2	Surjection	47
3.3	Bijection	47
3.4	Inversibilité des fonctions	47
4	Graphes	47
5	Restrictions	47
6	Composées	47
7	Fonction identité	47
8	Antécédent	48
9	Parité	48
10	Quelques familles de fonctions	48
10.1	Fonctions linéaires, affines et polynomiales	48
10.1.1	Graphe	49
10.1.2	Fonctions polynomiales	50
10.2	Fonctions racines	50
10.3	Les fonctions exponentielles et logarithmes	50
10.3.1	Exponentielles	50
10.3.2	Logarithmes et logarithme naturel	50
10.3.3	Identités importantes	51
V	Systèmes de coordonnées	53
1	Sur une droite	53
1.1	Coordonnées cartésiennes	53
1.2	Coordonnées logarithmiques	53
2	Sur un plan	53
2.1	Coordonnées induites par des droites	54
2.2	Coordonnées cartésiennes du plan	54
2.2.1	Distance	54
2.3	Coordonnées polaires	54
2.4	Lien entre les systèmes de coordonnées	55
3	Graphes de fonctions	55
VI	Trigonométrie	57
1	La notion d'angle : généralités	57
2	Le radian	57
3	Fonctions trigonométriques dans le cercle	58
4	Valeurs importantes	58
4.1	Retrouver un angle à partir du sinus ou du cosinus	59

5	Relation fondamentale	60
6	Symétries	60
6.1	Symétrie par rapport à l'axe des abscisses	60
6.2	Symétrie par rapport à l'axe des ordonnées	60
6.3	Symétrie par rapport à l'origine	60
6.4	Symétrie par rapport à la première bissectrice	61
6.5	Angles décalés de 90°	61
7	D'autres identités remarquables : formulaire	62
8	Fonctions trigonométriques dans les triangles	62
8.1	Triangles rectangles	62
8.2	Triangles quelconques	63
VII	Géométrie analytique	65
1	Points	65
2	Vecteurs	65
2.1	Vecteurs libres vs vecteurs liés	66
3	Opérations de base	66
3.1	Centre de gravité	67
4	Produit scalaire	68
4.1	Définition	68
4.2	Lien avec les équations cartésiennes	69
4.2.1	Droites du plan	69
4.2.2	Plans dans l'espace	70
5	Produit vectoriel	70
5.1	Trièdres orientés	70
VIII	Géométrie analytique (suite)	71
1	Équations	71
1.1	Équations de droites	71
1.1.1	Dans le plan	72
1.1.2	Dans l'espace	72
1.2	Équations de plans	73
1.2.1	Dans l'espace	73
1.3	Équations de cercles	73
1.3.1	Cercle dans le plan	73
1.4	Équations d'autres objets dans l'espace	74
1.4.1	Cylindre	74
1.4.2	Sphère	74
2	Interprétations géométriques	74
2.1	Aires et volumes	74
2.2	Angles	75
2.3	Distances	75
IX	Fonctions réelles : bases	77
1	Opérations sur les fonctions	78
2	Fonctions élémentaires	78
3	Limites	82
3.1	Règles étendues	87
3.2	Notation de Landau	88
3.3	Une lemme pour les limites	90
4	Continuité	90
X	Fonctions réelles : dérivées	93
1	Dérivée en un point	93
1.1	Interprétation de la dérivée	95
1.1.1	Vitesse	95
2	Dérivées et opérations	96

2.1	Fonction réciproque	98
3	Fonction dérivée et dérivées d'ordre supérieur	99
4	Dérivées des fonctions élémentaires	101
4.1	Constantes, fonctions puissance et exponentielles	101
XI	Fonctions dérivables	103
1	Extrema et dérivées	103
2	Théorème des accroissements finis	104
3	Croissance, décroissance et monotonie.	106
4	Règle de l'Hospital	107
5	Approximations polynomiales de Taylor	108
6	Étude de fonctions	112
XII	Fonction réelles d'une variable réelle : primitives	117
1	Motivation	117
1.1	Équation aux dérivées pour une population	117
1.2	Équation aux dérivées pour le ressort	117
2	Définition	118
2.1	Notations	119
3	Intégration immédiate	120
4	Linéarité	122
5	Intégration par changement de variable	123
6	Intégration par parties	124
7	Décomposition en fractions simples	125
7.1	Irréductibilité	126
7.2	Division Euclidienne	126
7.3	Fraction simple	127
7.4	Un exemple	127
XIII	Fonction réelles d'une variable réelle : intégrales définies	129
1	Introduction	129
2	Définition et propriétés	130
3	Théorème fondamental	132
4	Intégration définie et indéfinie	133
5	Intégrales généralisées	134
5.1	Cas d'une fonction non bornée	134
5.2	Cas d'un domaine non borné	135
XIV	Fonctions d'une variable réelle à valeurs vectorielles	137
1	Régularité	137
2	Courbes	138
2.1	Re-paramétrisation	138
2.2	Vitesse et vecteur tangent	138
2.3	Accélération et vecteur normal	139
2.4	Vecteur binormal	139
2.5	Longueur d'une courbe	140
XV	Matrices et systèmes linéaires	141
1	Opérations sur les matrices	142
1.1	Somme	142
1.2	Produit	142
1.3	Produit par un scalaire	143
1.3.1	Déterminants	143
1.3.2	Transposition	145
1.3.3	Inverse	145
1.3.4	Trace	146
1.3.5	Propriétés des opérations sur les matrices	146

1.3.6	Propriétés sur les déterminants	147
2	Systèmes linéaires	148
2.1	Méthode du pivot de Gauß	149
2.2	Un exemple de la méthode	149
3	Inversion de matrices	151
XVI	Fonctions de plusieurs variables réelles à valeurs vectorielles	153
1	Une fonction comme une transformation	153
1.1	Motivation et transformation de la droite	153
1.1.1	Dilatation et dérivée	153
1.1.2	Dilatation et applications linéaires	153
1.2	Transformations générales	154
2	Graphes	154
2.1	Application du plan vers la droite	154
2.2	Applications de la droite vers le plan	156
3	Limites et continuité	156
4	Dérivabilité	158
5	Différentiabilité	159
6	Règles de dérivation	159
7	Dérivabilité et différentiabilité aux ordres supérieurs	160
8	Optimisation à plusieurs variables	160
XVII	Intégrales multiples	163
1	Définition et motivation	163
2	Propriétés	164
3	Coordonnées cartésiennes	165
4	Coordonnées polaires	165
4.1	Justification heuristique	166
4.2	Exemple d'intégration en coordonnées polaires	166
5	Coordonnées quelconques	167
6	Coordonnées cylindriques	168
7	Coordonnées sphériques	168
XVIII	Équations différentielles ordinaires	171
1	Problème de Cauchy	171
2	Équations du premier ordre	172
2.1	Interprétation géométrique	172
2.2	Équations à variables séparables	172
2.3	Équations du type homogène	174
2.4	Équations linéaires	175
3	Équations linéaires du second ordre à coefficients constants	177
3.1	Méthode de résolution	177
3.2	Exemples	179
XIX	Nombres complexes	183
1	Motivations	183
2	Définitions	184
2.1	Plan de Gauß	184
3	Opérations	184
3.1	Somme	184
3.2	Produit	185
3.3	Parties réelles et imaginaires	186
3.4	Conjugaison	186
3.5	Module	186
3.6	Inverse	187
3.7	Exponentiation	188
3.8	Racines carrées, cubiques,	188

	3.9	Argument	188
4		Forme polaire ou trigonométrique	189
	4.1	Fonctions trigonométriques	190
	4.2	Exponentielle complexe	190
5		Racine n^{e} d'un complexe	190
	5.1	Racines de l'unité	191
6		Interprétations géométriques	192
	6.1	Somme	192
	6.2	Produit	192
	6.3	Parties réelles et imaginaires	192
	6.4	Conjugaison	192
	6.5	Module	192
	6.6	Inverse	192
	6.7	Exponentiation	192
	6.8	Racines carrées, cubiques,	192
	6.9	Argument	192
7		Équations	193
	7.1	Premier degré	193
	7.2	Second degré	193
	7.3	Degrés supérieurs	193
8		Complexes et équations différentielles linéaires à coefficients constants	193
	8.1	Fonctions réelles à valeurs complexes	194
	8.2	Résolution d'une équation homogène avec les nombres complexes	194

Chapitre I

Préliminaires

Ce chapitre regroupe des notions et notations diverses et généralement basiques. Le contenu évolue avec l'expérience de l'auteur de ces notes dans l'enseignement des mathématiques, prenant en compte les difficultés des étudiants.

Les mathématiques se comprennent à la lumière de ce qui précède et de ce qui suit. Aussi le lecteur est invité à revenir régulièrement à ce chapitre afin que ce qui peut paraître abstrait au début puisse être éclairé par les concepts ultérieurs.

Les documents mathématiques sont bien souvent écrits à la première personne du pluriel, dans le but que le lecteur se sente impliqué lors de sa lecture. Nous nous conformerons donc à cet usage, et le lecteur est invité à ne pas s'en formaliser.

1 Définitions, résultats et notations

1.1 Définitions

Parfois, une notion ou un objet peut être long à décrire. Afin d'éviter les longues descriptions, on attribue alors un nom à la notion, ou à l'objet.

Définition I.1. Une *définition* est une manière d'attribuer un mot ou une notation à un concept.

définition

Exemple. La description du mot « définition » ci-dessus peut-être vue comme une définition.

Exemple. Un exemple plus classique et plus « mathématique » :

Définition I.2. Un *carré* est un quadrilatère plane ayant tous ses angles égaux et tous ses côtés de même longueur.

carré

Résultat I.3. La longueur de la diagonale d'un carré vaut la longueur d'un de ses côtés multipliée par la racine carrée de 2.

1.2 Résultats : théorèmes, lemmes, corollaires, axiomes, ...

Les textes mathématiques contiennent beaucoup d'affirmations considérées comme « vraies » (voir aussi section 3). Lorsque nous obtenons une affirmation qui est vraie, nous disons que nous avons obtenu un « résultat ». Dans le présent document, vous verrez beaucoup de ce qui suit :

Résultat I.4. Si on suppose telle ou telle chose, alors nous avons tel résultat.

Un théorème est généralement suivi d'une preuve, qui a pour but de convaincre le lecteur que le théorème est vrai.

Par exemple :

Résultat I.5. La longueur de la diagonale d'un carré vaut la longueur d'un de ses côtés multipliée par la racine carrée de 2.

Démonstration. La preuve, à l'aide du théorème de Pythagore, est laissée en exercice. (Eh oui, l'auteur est parfois fainéant et n'inclut pas toujours une vraie preuve!) □

Dans les textes mathématiques usuels, chaque résultat peut être

Un théorème C'est un résultat important, qui vaut la peine d'être mis en avant. Certains théorèmes portent le nom de quelqu'un ayant contribué à son élaboration ou à qui la communauté voulait rendre hommage, tel que le théorème de Pythagore. D'autres théorèmes ont un autre nom, tel que le « théorème de la valeur intermédiaire », ou le « théorème fondamental du calcul différentiel et intégral ».

Un lemme C'est généralement un résultat préliminaire à la démonstration d'un théorème important.

Une proposition Résultat vrai, n'ayant pas le statut de théorème.

Une propriété Généralement un résultat découlant (plus ou moins aisément) d'une définition.

Un corollaire C'est un résultat dont la démonstration découle normalement relativement aisément d'un théorème important.

Un axiome Résultat pris pour vrai sans démonstration.

(La liste est non-limitative et n'a pas force de loi : il y a des variations dans l'usage de ces termes !)

À l'exception de l'axiome (voir section 2 page ci-contre), qui a un statut particulier et n'est d'ailleurs pas vraiment un « résultat », les autres types de résultats ne diffèrent pas réellement entre eux. Il n'y en a pas de moins vrais que d'autres. Seul l'usage fait qu'on appelle un résultat donné tantôt un théorème, tantôt une proposition, etc. Dans le présent document, nous avons d'ailleurs pris le parti d'appeler tous les résultats « Résultat ».

Remarque (*). Le mot « proposition » est, dans le domaine de la logique, également utilisé pour nommer des affirmations dont on ne sait pas si elles sont vraies ou fausses. Nous ferons de notre mieux pour éviter cet usage.

1.3 Notations et variables

Lorsqu'un objet ou une notion (par exemple un nombre ou une construction particulière) est complexe à décrire, on peut lui attribuer une notation (une lettre ou un symbole) qui sera ré-utilisée plus tard.

Exemple. Voici comme nous pourrions définir le nombre π , en deux étapes. D'abord nous énonçons un théorème, puis nous formulons une définition.

Résultat I.6. *Le quotient entre la circonférence d'un cercle et son diamètre est une constante, indépendante du cercle considéré.*

Définition I.7. Ce rapport est noté π .

Nous pouvons ensuite énoncer des théorèmes faisant intervenir le nombre π .

Résultat I.8. *Le nombre π est irrationnel, et ses premières décimales sont 3.14159....*

On utilise beaucoup de symboles pour faire des notations, principalement des lettres. En particulier nous utilisons souvent des lettres grecques (dont π), raison pour laquelle nous rappelons les lettres grecques les plus utilisées dans la section suivante.

Chaque notation rencontrée est soit une notation bien connue, ou alors est une notation ad-hoc (définie le temps d'un énoncé, d'un paragraphe, d'un cours, ...).

Exemple. 1. Lorsqu'on écrit « L'aire d'un disque est πR^2 , où R est le rayon du cercle. », R est une notation ad-hoc tandis que π est la notation bien connue.

2. Lorsqu'on écrit « Si $x^2 - x - 2 = 0$, alors $x = 2$ ou $x = -1$. », on peut supposer que x est une notation ad-hoc et on comprend implicitement que x est un nombre dont il fallait trouver la valeur.

3. Si, quelques paragraphes plus loin, on parle de l'équation $x^2 - 1 = 0$, il faut se rendre à l'évidence : x n'a plus rien à voir avec le x de l'exemple précédent.

1.4 Alphabets

De l'alphabet hébreu, vous croiserez peut-être la lettre **ℵ** (« aleph »). Un autre symbole fréquemment rencontré est le symbole ∇ (« nabla ») (à ne pas confondre avec le « delta » majuscule grec : Δ).

TABLE II – L'alphabet grec pour mathématiciens.

Lettre grecque	variante	majuscule	« nom »
α			« alpha »
β			« beta »
χ			« chi »
δ		Δ	« delta »
ϵ	ε		« epsilon »
η			« eta »
γ		Γ	« gamma »
ι			« iota »
κ	κ		« kappa »
λ		Λ	« lambda »
μ			« mu »
ν			« nu »
ω		Ω	« omega »
ϕ	φ	Φ	« phi »
π	ω	Π	« pi »
ψ		Ψ	« psi »
ρ	ρ		« rho »
σ		Σ	« sigma »
τ			« tau »
θ	ϑ	Θ	« theta »
		Υ	« upsilon »
ξ		Ξ	« xi »
ζ			« zeta »

2 Logique, rigueur et notion de preuve

Attardons-nous sur la notion de *démonstration* (synonyme de *preuve*). Une démonstration, informelle-ment, est une succession d'affirmations qui découlent les unes des autres par applications de règles logiques à des résultats connus. Faire la preuve d'une affirmation, c'est écrire une preuve qui conduit à cette affirmation. Idéalement, une preuve doit pouvoir convaincre le lecteur de la véracité de l'énoncé.

démonstration

preuve

Exemple. Définissons « être divisible par » par « le quotient est un entier », et « pair » par « être divisible par 2 ». Nous pouvons alors affirmer : le carré de tout entier pair est divisible par 4.

Preuve/Démonstration.

Si p est un entier pair, cela veut dire qu'il est divisible par 2 (définition de « pair »); dès lors $p/2$ est un entier, notons-le q .

Nous avons donc $p/2 = q$

Dès lors, $(p/2)^2 = q^2$ (car appliquer la même opération à un objet donné amène au même résultat)

Or $(p/2)^2 = p^2/4$

De plus, q^2 est encore un entier, car le produit de deux nombres entiers est un nombre entier

Dès lors, p^2 est divisible par 4. □

Dans l'exemple, chaque étape justifiée par une définition (« être divisible par ») ou l'une des lignes précédentes combinée à une règle logique, par exemple « appliquer la même opération à un objet donné conduit au même résultat ». Une règle telle que celle-là est une règle qui ne se prouve pas : elle est simplement considérée comme vraie dans tous nos raisonnements, c'est un *axiome*.

axiome

Remarque (*). Demandons-nous ce que veut dire « $1 + 1 = 2$ ». Cela peut-il être prouvé? Quelle est la définition des symboles « 1 », « + » et « 2 »? Et même, le symbole d'égalité, « = »? Toutes ces questions, et tant d'autres, resteront malheureusement sans réponse dans le présent cours, mais retenons que sans supposer certaines choses communes, il est impossible de raisonner.

Remarque (*). Nous travaillons selon un mode appelé *hypothético-déductif*, c'est-à-dire que partant de certaines affirmations admises comme *hypothèses*, nous en *déduisons* d'autres affirmations.

Exemple. Considérons l'affirmation : si n est un entier divisible par 4, alors n est pair. Dans cette affirmation, nous ne savons pas ce qu'est n . Mais *sous l'hypothèse* que n est divisible par 4, *alors* n est pair.

Exercice. Pouvez-vous prouver l'affirmation précédente ?

Exemple. Nous constatons que dans une pièce tous les hommes portent la moustache. *Si nous savons* que l'homme recherché pour avoir mangé la mousse au chocolat se trouve dans la pièce, nous pouvons en déduire qu'il porte la moustache. Nous ne pouvons cependant pas en déduire que le coupable porte la moustache ni même que c'est un homme car rien ne dit que l'homme recherché est coupable.

3 Manipuler le vrai et le faux

En mathématiques, nous aimons discerner les *affirmations* (ou *énoncés*) qui sont vraies des autres. Les vraies seront appelées « théorème », « propriété », ou tout simplement « résultat ». Les autres sont généralement laissées dans l'oubli.

Exemple. Considérons les affirmations suivantes :

- * « x est positif » ;
- * « Si x est plus grand que 1 alors x est positif » ;
- * « Pour tout x entier, x est positif » ;
- * « Il existe x entier tel que x est positif » ;

Il se fait que la première est vraie ou fausse selon la valeur de x , la seconde est vraie quelle que soit la valeur de x , la troisième est fausse et la dernière est vraie.

Lorsque nous voulons noter une affirmation, nous les noterons généralement p , q ou r voire s dans cette section. Lorsque la valeur de vérité d'une affirmation peut (a priori) dépendre d'un paramètre, on peut expliciter cette dépendance en indiquant ce paramètre entre parenthèses. Le paramètre est alors appelé *variable libre* (en opposition aux « variables liées », voir 3.2.1 page 16).

Exemple. Par exemple nous pourrions appeler $p(x)$, $q(x)$, r et s les quatre affirmations ci-dessus. Dans ce cas nous pouvons écrire :

- * $p(x)$ peut être vraie ou fausse, selon la valeur de x ;
- * $q(x)$ est vraie quelque soit la valeur de x ;
- * r est fausse ;
- * s est vraie.

Remarque (*). La notion de « vérité » est ici purement formelle. Les mathématiques ne sont pas de la philosophie, et le mathématicien ne s'occupe pas de savoir s'il dit la « vraie vérité ». Nous dirons simplement qu'un énoncé est « vrai » s'il peut se déduire de ce qu'on suppose vrai dans notre modèle, c'est-à-dire s'il peut être prouvé.

3.1 Connecteurs logiques

3.1.1 Et

Si p désigne une affirmation, et q désigne une seconde affirmation, l'affirmation « p et q sont vraies » est une autre affirmation qui sera vraiment exactement dans le cas où p est vraie *et* q est vraie.

Cette nouvelle affirmation est notée $p \wedge q$, et la valeur de vérité dépend donc de celle de p et de celle de q de la manière suivante :

- * Si p et q sont vraies, alors $p \wedge q$ est vraie
- * si p est vraie mais q est fausse, alors $p \wedge q$ est fausse,
- * si p est fausse mais q est vraie, alors $p \wedge q$ est fausse,

affirmations

énoncés

variable libre

- * si p et q sont fausses, alors $p \wedge q$ est fausse.

Exemple. Si $p(x)$ est l'affirmation « x est un entier pair » et si $q(x)$ est l'affirmation « x est un entier divisible par 3 », alors l'affirmation $r(x) = p(x) \wedge q(x)$ est vraie pour les entiers multiples de 6, et uniquement ceux-là.

Exercice. Le lecteur est invité à relire l'exemple précédent, voire remplacer x par quelques valeurs différentes pour bien s'assurer qu'il a compris.

3.1.2 Ou

Si p désigne une affirmation, et q désigne une seconde affirmation, l'affirmation « p ou q est vraie » est une autre affirmation, notée $p \vee q$, dont la valeur de vérité dépend de celle de p et de celle de q de la manière suivante :

- * Si p et q sont vraies, alors $p \vee q$ est vraie
- * si p est vraie mais q est fausse, alors $p \vee q$ est vraie,
- * si p est fausse mais q est vraie, alors $p \vee q$ est vraie,
- * si p et q sont fausses, alors $p \vee q$ est fausse.

Exemple. Si $p(x)$ est l'affirmation « x est un entier pair » et si $q(x)$ est l'affirmation « x est un entier divisible par 3 », alors l'affirmation $r(x) = p(x) \vee q(x)$ est vraie pour les entiers multiples de 2 et pour les multiples entiers de 3, et uniquement ceux-là. En d'autres termes, $r(0), r(2), r(3), r(4), r(6)$ sont vraies, mais $p(1)$ et $p(5)$ ne le sont pas.

Remarque (*). En particulier, l'affirmation « 6 est pair ou multiple de 3 » est vraie. En d'autres termes, le « ou » décrit ci-dessus est *inclusif* : il suffit qu'une des affirmations soit vraie pour que $p \vee q$ le soit, mais les deux peuvent être vraies en même temps. Une autre possibilité serait un « ou exclusif » (parfois appelé *XOR*, pour « eXclusive OR » en anglais), mais il est essentiellement absent des raisonnements logiques en mathématiques.

3.1.3 Négation

Si p désigne une affirmation, l'affirmation « p n'est pas vraie » (ou encore « p est fausse ») est une autre affirmation, notée $\neg p$, dont la valeur de vérité dépend de celle de p de la manière suivante :

- * Si p est vraie, alors $\neg p$ est fausse,
- * si p est fausse, alors $\neg p$ est vraie.

3.1.4 Implication

Si p désigne une affirmation, et q désigne une seconde affirmation, l'affirmation « Si p est vraie alors q est vraie » (ou de manière équivalente « p implique q ») est une autre affirmation, notée $p \rightarrow q$, dont la valeur de vérité dépend de celle de p et de celle de q de la manière suivante :

- * Si p et q sont vraies, alors $p \rightarrow q$ est vraie,
- * si p est vraie mais q est fausse, alors $p \rightarrow q$ est fausse,
- * si p est fausse mais q est vraie, alors $p \rightarrow q$ est vraie,
- * si p et q sont fausses, alors $p \rightarrow q$ est vraie.

Il est capital de noter que ceci ne définit aucun lien de cause à effet ! La seule chose qui importe sont les valeurs de vérité des affirmations p et q .

Remarque (*). L'implication logique est une notion particulièrement difficile à cerner correctement pour certains. Notre esprit est pollué par la notion d'implication du langage naturel, qui sous-tend une notion de causalité.

Pour comprendre, ou au moins retenir, la notion mathématique, une façon de faire est de réfléchir à l'envers : l'implication sera « vraie » à condition de pouvoir prouver qu'elle n'est pas « fausse ». Comment prouver que $p \rightarrow q$ est faux ? Le seul moyen est d'avoir p qui soit vrai tout en ayant q qui soit faux. Dans tous les autres cas, elle sera donc « vraie ».

Exemple. Commençons par un exemple issu du langage naturel : notons r l'affirmation « S'il pleut je prends mon parapluie. » On peut décortiquer cette phrase en une affirmation p « Il pleut » et une affirmation q « je prends mon parapluie ». L'affirmation r est alors $p \rightarrow q$. La seule chose que cette affirmation dit, si elle est vraie, est que quand il pleut, je prends mon parapluie. S'il ne pleut pas, elle ne dit rien (et est donc « vraie » par défaut).

Exemple. Si $p(x)$ est l'affirmation « x est un entier pair » et si $q(x)$ est l'affirmation « x est un entier divisible par 3 », alors l'affirmation $r(x) = p(x) \rightarrow q(x)$ est vraie pour les entiers impairs et pour les entiers multiples de 3. Elle est fausse pour les autres.

3.1.5 Équivalence (ou bi-implication)

Si p désigne une affirmation, et q désigne une seconde affirmation, l'affirmation « p est vraie si et seulement si q est vraie » (ou de manière équivalente « p et q sont équivalentes ») est une autre affirmation, notée $p \leftrightarrow q$, dont la valeur de vérité dépend de celle de p et de celle de q de la manière suivante :

- * Si p et q sont vraies, alors $p \leftrightarrow q$ est vraie
- * si p est vraie mais q est fausse, alors $p \leftrightarrow q$ est fausse,
- * si p est fausse mais q est vraie, alors $p \leftrightarrow q$ est fausse,
- * si p et q sont fausses, alors $p \leftrightarrow q$ est vraie.

Il est capital de noter que ceci ne définit aucun lien de cause à effet ! La seule chose qui importe sont les valeurs de vérité des affirmations p et q .

3.2 Quantificateurs

3.2.1 Pour tout

Si $p(x)$ est une affirmation dont x est une variable libre (il pourrait y en avoir d'autres), alors l'affirmation « Pour tout x , $p(x)$ est vraie » est une autre affirmation, notée $\forall x, p(x)$, dont la valeur de vérité dépend de celle de $p(x)$ pour chaque valeur de x :

- * $\forall x, p(x)$ est vraie si pour tout x , $p(x)$ est vraie ;
- * $\forall x, p(x)$ est fausse si il y a au moins une valeur de x pour laquelle $p(x)$ est fausse.

Remarquons qu'il est en général sous-entendu que x prend les valeurs dans les nombres réels, sinon il faut préciser le domaine où x peut varier.

Exemple. L'affirmation $p(x)$ définie par $x^2 \geq 0$ est vraie pour tout nombre réel, donc l'affirmation $\forall x, p(x)$ est vraie.

Dans l'expression $\forall x, p(x)$, la variable x est dite *liée* ou *quantifiée*.

Remarque (*). Elle n'est pas libre car on ne peut pas lui donner *une* valeur : la construction doit être vraie pour toutes les valeurs ! Par contre, le nom de la variable n'a aucune importance : $\forall x, p(x)$ est exactement la même affirmation que $\forall u, p(u)$.

3.2.2 Il existe

Si $p(x)$ est une affirmation dont x est une variable libre (il pourrait y en avoir d'autres), alors l'affirmation « il existe x tel que $p(x)$ est vraie » est une autre affirmation, notée $\exists x : p(x)$, dont la valeur de vérité dépend de celle de $p(x)$ pour chaque valeur de x :

- * $\exists x p(x)$ est vraie si il existe une valeur de la variable x pour laquelle $p(x)$ est vraie ;
- * $\exists x p(x)$ est fausse si, quelle que soit la valeur de x considérée, $p(x)$ est fausse.

Exemple. * L'affirmation $p(x)$ définie par $x^2 \geq 0$ est vraie pour au moins un nombre réel, par exemple $x = 0$ (ou n'importe quel autre réel, en fait !), donc l'affirmation $\exists x : p(x)$ est vraie.

- * L'affirmation $p(x)$ définie par $x^2 \leq 0$ est vraie pour au moins un nombre réel, par exemple $x = 0$ (en fait c'est le seul !), donc l'affirmation $\exists x : p(x)$ est vraie.
- * L'affirmation $p(x)$ définie par $x^2 < 0$ n'est vraie pour aucun nombre réel, donc l'affirmation $\exists x : p(x)$ est fausse.

Comme précédemment, dans l'expression $\exists x : p(x)$, la variable x est dite *liée* ou *quantifiée*.

liée

quantifiée

liée

quantifiée

3.2.3 Négation et quantificateurs

Résultat I.9. Si $p(x)$ est une affirmation dont x est une variable libre, alors la négation de « $\forall x, p(x)$ » est « $\exists x : \neg(p(x))$ ».

Ceci est clair d'après les valeurs de vérités énoncées ci-dessus pour les quantificateurs « pour tout » et « il existe ».

3.3 Tables de vérités

En associant la valeur 1 à « vrai » et la valeur 0 à « faux », on peut résumer les connecteurs « ou, et, non, implique, équivalent à » dans des tables de la manière suivante (pour le « et ») :

p	q	$p \wedge q$	$p \vee q$	$p \leftrightarrow q$	$\neg p$	$p \rightarrow q$
0	0	0	0	1	1	1
0	1	0	1	0	1	1
1	0	0	1	0	0	0
1	1	1	1	1	0	1

Les colonnes de gauche sont remplies de sorte que chaque combinaison de vérité de p et q apparaisse une fois. Les colonnes à droite de la ligne verticale sont déduites des définitions des connecteurs logiques, à partir des colonnes de gauche.

Regardons la table de vérité de l'opération suivante : $p \vee \neg q$. Cela donne :

p	q	$\neg q$	$p \vee \neg q$
0	0	1	1
0	1	0	0
1	0	1	1
1	1	0	1

La colonne $\neg q$ sert de calcul intermédiaire : la colonne recherchée $p \vee \neg q$ s'obtient en appliquant la règle du « ou » sur la colonne p et la colonne $\neg q$. Une manière plus compacte de noter cela, sans ajouter de colonnes intermédiaires est la suivante :

$p \vee \neg q$
0 1 10
0 0 01
1 1 10
1 1 01

où la valeur de vérité s'inscrit directement sous la colonne du symbole correspondant. Pas à pas, la table ci-dessus a été construite comme suit :

$p \vee \neg q$		$p \vee \neg q$		$p \vee \neg q$
0	0	0	10	0 1 10
0	1	0	01	0 0 01
1	0	1	10	1 1 10
1	1	1	01	1 1 01

La colonne sous le symbole $\vee$ s'obtient en comparant la valeur dans la colonne sous p et la colonne sous $\neg$ (qui représente la valeur de $\neg q$).

On peut comparer les tables de vérité de $\neg(p \rightarrow q)$ et de $p \wedge (\neg q)$, et observer que la colonne finale est identique. Ceci montre que ces deux affirmations sont équivalentes.

3.4 Raisonnements et preuves

Raisonnement est un processus abstrait. Partant de certaines considérations (des hypothèses), nous en déduisons des choses qui doivent être vraies. Le raisonnement permet de décider d'actions à mener ou, en ce qui nous concerne, de prouver la véracité d'affirmations.

Exemple. J'ai 3 paires de chaussettes différentes. Si je choisis au hasard sans regarder, combien dois-je prendre de chaussettes dans ma valise pour être totalement certain d'avoir au moins deux chaussettes d'une même paire?

Exemple. Il y a cinq maisons de cinq couleurs différentes. Dans chacune de ces maisons, vit une personne de nationalité différente. Chacune de ces personnes boit une boisson différente, fume un cigare différent et a un animal domestique différent.

- * L'Anglais vit dans la maison rouge.
- * Le Suédois a des chiens.
- * Le Danois boit du thé.
- * La maison verte est à gauche de la maison blanche.
- * Le propriétaire de la maison verte boit du café.
- * La personne qui fume des Pall Mall a des oiseaux.
- * Le propriétaire de la maison jaune fume des Dunhill.
- * La personne qui vit dans la maison du centre boit du lait.
- * Le Norvégien habite dans la première maison.
- * L'homme qui fume des Blend vit à côté de celui qui a des chats.
- * L'homme qui a un cheval est le voisin de celui qui fume des Dunhill.
- * Le propriétaire qui fume des Blue Master boit de la bière.
- * L'Allemand fume des prince.
- * Le Norvégien vit juste à côté de la maison bleue.
- * L'homme qui fume des Blend a un voisin qui boit de l'eau.

Question : qui a le poisson ?

Cet énoncé est largement disponible sur l'Internet en plusieurs versions. La version ci-dessus est tiré de Wikipedia[[wiki:EnigmeEinstein](#)].

Nous allons maintenant explorer quelques types de raisonnements logiques courant en mathématique

3.4.1 Tiers-exclus

Tout affirmation peut être « vraie », ou « fausse ». Il n'y a pas d'autre possibilité : c'est le principe du « tiers¹-exclus ». Notons qu'il peut arriver qu'on ne soit pas capable de déterminer laquelle des possibilités est la bonne (par exemple parce qu'on y a pas réfléchi assez), et il peut même arriver qu'on puisse démontrer que personne ne saura déterminer laquelle des possibilités est la bonne, mais quoiqu'il arrive, l'énoncé sera l'un ou l'autre. S'il n'est pas l'un, il est l'autre.

3.4.2 Démonstration par l'absurde

Le principe du tiers-exclus a la conséquence suivante : pour montrer que quelque chose est vrai, on peut montrer qu'il ne peut pas être faux. Ce principe est appelé « démonstration par l'absurde » : supposons que l'énoncé soit faux, tirons-en quelque chose d'incompatible avec nos hypothèses, et nous en déduisons que l'énoncé devait être vrai.

Exemple. Prouvons que le nombre 0 n'a pas d'inverse, c'est-à-dire qu'il n'existe pas de nombre a tel que $0a = 1$. En effet, par l'absurde supposons qu'un tel nombre a existe. Alors $a0 = 0$ (car tout nombre multiplié par 0 vaut 0), dès lors nous aurions $0 = 1$, qui est une contradiction. Nous avons donc prouvé par l'absurde que 0 n'admet pas d'inverse.

Exemple. Prouvons qu'il n'existe pas de nombre strictement positif qui soit plus petit que tous les autres. En effet, supposons par l'absurde qu'un tel nombre existe et nommons-le x . Alors $x > 0$. Or la moitié de x est encore strictement positif, et par ailleurs strictement inférieure à x lui-même. Dès lors x n'est pas le plus petit nombre strictement positif, d'où la contradiction.

1. Rien à voir avec $\frac{1}{3} = 0.333\dots$, il s'agit d'un tiers comme on parlerait d'une « tierce personne ».

3.4.3 Contraposée

Un raisonnement par *contraposée* se base sur le fait que l'implication $p \rightarrow q$ est équivalente à l'implication $(\neg q) \rightarrow (\neg p)$.

contraposée

Exemple. Les affirmations « S'il pleut je prends mon parapluie » et « Si je ne prends pas mon parapluie, c'est qu'il ne pleut pas » sont parfaitement équivalentes du point de vue des valeurs de vérité. On peut également facilement passer de l'une à l'autre par un raisonnement par l'absurde.

Exemple. « Si x est divisible par 6, alors x est pair » est équivalente à « Si x n'est pas pair, alors x n'est pas divisible par 6 ».

3.4.4 Démonstration par récurrence / induction

Le raisonnement par *récurrence*, on dit également par *induction*, est un schéma de preuve particulier. Le principe est le suivant : supposons que nous voulions démontrer une proposition $p(n)$ pour tout entier naturel n (c'est-à-dire $n = 0, 1, 2, \dots$). Si nous pouvons prouver $p(0)$ (« le pas initial ») et $\forall k, p(k) \rightarrow p(k+1)$ (« l'étape d'induction »), alors le principe d'induction affirme que $\forall n, p(n)$.

récurrence

induction

Exemple. Montrons par récurrence que la somme des entiers de 0 à n vaut $\frac{n(n+1)}{2}$. C'est notre affirmation $p(n)$.

L'affirmation $p(0)$ est : « La somme des entiers de 0 à 0 vaut 0. » Ceci est clair. Le pas initial est donc prouvé.

Pour l'étape d'induction, fixons un nombre k quelconque, et supposons $p(k)$, c'est-à-dire supposons

$$0 + 1 + \dots + k = \frac{k(k+1)}{2}.$$

Nous voulons montrer $p(k+1)$, c'est-à-dire :

$$0 + 1 + \dots + (k+1) = \frac{(k+1)(k+2)}{2}.$$

Or la somme du membre de gauche « passe » par k , c'est-à-dire qu'on peut l'écrire sous la forme

$$\underbrace{0 + 1 + \dots + k}_{= \frac{k(k+1)}{2}} + (k+1) = \frac{k(k+1)}{2} + (k+1)$$

En mettant alors au même dénominateur, nous obtenons

$$\frac{k(k+1)}{2} + (k+1) = \frac{k(k+1) + 2(k+1)}{2}$$

et, en mettant $k+1$ en évidence, nous avons :

$$\frac{k(k+1) + 2(k+1)}{2} = \frac{(k+2)(k+1)}{2}$$

Mettant maintenant bout à bout nos dernières égalités, nous avons obtenu :

$$0 + 1 + \dots + (k+1) = \frac{(k+1)(k+2)}{2}$$

ce que nous voulions démontrer !

Par le principe d'induction, l'égalité $p(n)$ est donc vraie pour tout entier n .

3.4.5 Modélisation

Pour beaucoup, les mathématiques seront un outil pour étudier une certaine réalité. La réalité étant bien généralement trop complexe, il faudra la simplifier. Le passage d'une réalité complexe à une vision mathématisée moins complexe est la *modélisation* du problème.

modélisation

Exemple. « On lance une pièce de monnaie. » Il faut d'abord définir le problème : S'intéresse-t-on à la trajectoire de la pièce ? À la manière dont elle tourne ? Au déplacement d'air qu'elle provoque ? Aux processus qui font qu'on est capable ou pas de rattrapper la pièce avant qu'elle retombe ? Ou simplement au résultat « pile » ou « face » obtenu ?

Selon le problème, la mathématisation sera évidemment différente. Si c'est le résultat (« pile ou face ») qui nous intéresse, on pourra alors utiliser un modèle probabiliste dans lequel on a une chance sur deux d'obtenir pile, et autant d'obtenir face. Ceci ne prend pas en compte les éventuels déséquilibres de la pièce ou la probabilité qu'elle retombe sur la tranche, mais il vaut mieux simplifier en première approche.

Chapitre II

Les nombres

La notion de nombre est à la fois apprise dès la petite enfance et pourtant n'est jamais vraiment bien définie. Le lecteur sera déçu d'apprendre que nous ne définirons pas non plus la notion de nombre. Rappelons cependant ce qu'ils sont et comment les manipuler.

1 Une classification des nombres

Les entiers naturels Ce sont les nombres qui permettent de compter. L'ensemble des *entiers naturels* est noté $\mathbb{N}$. Il contient des nombres tels que 0, 1, 5, 42, 6 841 241 237 et tant d'autres.

entiers naturels

Le nombre 0 est donc un nombre entier naturel. La notation $\mathbb{N}_0$ désigne l'ensemble des naturels **sauf** ce 0.

Remarque (*). Notons que les entiers (comme les mathématiques en général) sont une **simplification** de la réalité : on compte, mais on ne dit pas ce qu'on compte. Cela a l'avantage de pouvoir faire des raisonnements généraux du type « $2 + 3 = 5$ », qu'il s'agisse de pommes, de demi-pommes ou de bombes à neutron. En transposant les résultats mathématiques au monde réel, il faut veiller à bien préciser ce qui est compté.

Remarque (*). L'ensemble des entiers naturels est « infini » mais ne contient aucun nombre appelé « l'infini ». Cela signifie seulement qu'on ne peut pas arriver au bout de tous les nombres entiers car il y a toujours un nombre entier « suivant » dans la liste. Nous reviendrons régulièrement sur cette notion « d'infini ».

Les entiers relatifs L'ensemble des *entiers relatifs*, noté $\mathbb{Z}$ et souvent simplement appelé l'ensemble des entiers, contient les entiers naturels et leur opposé : -42 , -5 , 4 , -9 , 2 , 6 et -38 sont tous des entiers.

entiers relatifs

Les rationnels Quand on compte « il y a une demi-pomme », on compte **une** (le nombre entier 1) demi-pomme. La notion de *moitié* est plus abstraite que le simple comptage, et demande d'introduire les nombres rationnels. Le nombre rationnel $\frac{1}{2}$ est un nombre caractérisé par le fait que son double vaut 1, ce qui n'existe pas si on se contente des nombres entiers.

Définition II.1. Les *nombres rationnels* sont les nombres qui peuvent s'écrire sous la forme $\frac{m}{n}$ avec m un entier et n un entier non-nul.

nombres
rationnels

Pour rappel, les fractions s'additionnent, se multiplient et se divisent comme suit :

Somme	$\frac{a}{b} + \frac{x}{y} = \frac{ay}{by} + \frac{xb}{by} = \frac{ay + xb}{by}$
Produit	$\frac{a}{b} \cdot \frac{x}{y} = \frac{ax}{by}$
Quotient	$\frac{\frac{a}{b}}{\frac{x}{y}} = \frac{a}{b} \cdot \frac{y}{x} = \frac{ay}{bx}$

Pour le quotient, notez la taille de la barre de fraction et la position verticale par rapport au reste de la ligne (en particulier le signe d'égalité).

Exemple. Voici un exemple où la taille de la barre de fraction et sa position relative ont leur importance.

$$\frac{\frac{1}{2}}{3} = \frac{\frac{1}{2}}{\frac{3}{1}} = \frac{1}{6} \qquad \frac{1}{\frac{2}{3}} = \frac{\frac{1}{1}}{\frac{2}{3}} = \frac{3}{2}$$

Remarque (*). Notons encore une fois la simplification du problème physique : en mathématiques, deux demis font un, alors qu'en pratique deux demi-pommes ne font pas une pomme. Si vous n'êtes pas convaincu, prenez dix milles dix-millièmes de pomme et tentez de la croquer.

Remarque (*). L'écriture d'un nombre rationnel sous forme de quotient de deux entiers n'est pas unique.

Exemple. Par exemple $\frac{2}{4}$ et $\frac{1}{2}$ représentent le même nombre rationnel : ils sont égaux. En d'autres termes, on peut écrire $\frac{2}{4} = \frac{1}{2}$.

Il existe une autre écriture des nombres rationnels : le développement décimal. Par exemple, on écrira $\frac{1}{2} = 0,5$. Si ce concept ne vous est pas familier, parlez-en à vos enseignants.

Le développement décimal de x est y :

x	y
1	1
$1/2$	0.5
$1/3$	0.333...3...
$1/11$	0.090909...09...
π	3.14159...

développement
décimal
périodique

Résultat II.2. *Un nombre rationnel admet un développement décimal périodique. Réciproquement, tout nombre admettant un développement périodique est rationnel.*

Exemple. Quelques développements décimal :

- * Celui de $\frac{1}{3}$ est 0.333...3.... C'est bien périodique, puisque 3 se répète.
- * Celui de $\frac{1}{2}$ est 0.5. C'est également périodique, en l'écrivant sous la forme 0.5000...0.... Le chiffre 0 se répète.
- * Pour $\frac{11}{13}$, cela donne 0,846 153 846 153 846 153 846 153...846 153....

nombres réels

Les nombres réels L'ensemble des *nombres réels* est plus complexe. Il contient notamment tous les ensembles précédents. Cet ensemble contient non-seulement les rationnels, mais également des nombres tels que $4 - \sqrt{2}$ ou $\pi^2 + \frac{2}{3}$. Décrire précisément l'ensemble des réels prendrait du temps, et ce n'est pas le but ici. Rappelons-nous simplement qu'on le note $\mathbb{R}$, et qu'on le représente le plus souvent par une droite « sans trou » (voir aussi la section « Sur une droite », page 53). Chaque nombre réel possède également un développement décimal.

Exemple. Le développement décimal de π n'est pas fini (il ne se termine jamais) et n'est pas non plus périodique, mais les premières décimales sont

$$\pi = 3,141\ 592\ 653\ 589\ 79...$$

Résultat II.3. *Un nombre réel admet un développement infini et non-périodique si et seulement si ce réel n'est pas un nombre rationnel.*

La précision et l'illusion de la précision En pratique, un moyen sûr de tomber sur des nombres est de faire des mesures. La température, la hauteur, la longueur, le nombre d'amis, la vitesse, le nombre de désintégrations par seconde, la fréquence cardiaque : tout cela se mesure et donne lieu à des nombres.

Le nombre d'amis que vous avez sur le premier réseau social venu est un nombre entier, déterminé sans équivoque : qu'il soit égal à 4 ou à 578, il peut être mesuré univoquement et précisément à un moment donné.

La température de la pièce, par contre, est sujette à l'imprécision : fait-il 20°C , 20.5°C ou en fait 20.42°C ? Tout dépend de l'appareil que vous utilisez pour mesurer. La mesure peut être erronée pour diverses raisons, que ce soit une imprécision à la lecture (graduations trop grossières), ou une erreur due à un appareil mal calibré.

D'autre part, à partir de certaines mesures, on peut calculer certaines valeurs. Par exemple on peut calculer qu'en roulant à vélo à 20 km h^{-1} pendant 10 minutes, la distance parcourue est de $3.333\dots 3\dots\text{km}$. La présence des décimales laisse croire à une précision importante dans le calcul, alors qu'en réalité la vitesse utilisée (20 à l'heure) était probablement une approximation grossière.

Retenons donc que le « nombre » en mathématiques est un objet très précis, alors que la mesure physique est généralement très floue. Il est par exemple tout à fait sans objet de se demander si la vitesse d'un cycliste est un nombre rationnel.

2 Relations entre les nombres

2.1 Inégalités et notion d'ordre

Les nombres peuvent être comparés entre eux.

Définition II.4. On dit que

- * « a est strictement inférieur à b », noté $a < b$, si a est plus petit et différent de b ;
- * « a est strictement supérieur à b », noté $a > b$, si a est plus grand et différent de b ;
- * « a est inférieur à b » (ou « inférieur ou égal »), noté $a \leq b$, si a est plus petit ou égal à b ;
- * « a est supérieur à b » (ou « supérieur ou égal »), noté $a \geq b$, si a est plus grand ou égal à b .

Exemple. Les affirmations suivantes sont vraies :

- * $2 \leq 3$
- * $2 < 3$
- * $-2 \leq 1$
- * $-2 \leq -1$
- * $2 \leq 2$
- * $2 \geq 2$
- * $5 > 3$

mais celles-ci sont fausses :

- * $2 \geq 3$
- * $-2 \geq 1$
- * $-2 \geq -1$
- * $2 < 2$
- * $2 > 2$
- * $5 < 3$

Remarque (*). L'inégalité $2 \leq 2$ peut perturber, mais c'est l'usage en mathématique : $\leq$ signifie « plus petit ou égal », donc en particulier puisque $2 = 2$, on a forcément $2 \leq 2$.

Nous supposons les deux résultats suivants connus.

Résultat II.5. Pour tous réels a, b, c , nous avons :

- * Si $a \leq b$ et $b \leq a$, alors $a = b$

* Si $a \leq b$ et $b \leq c$, alors $a \leq c$ (propriété de transitivité).

Résultat II.6. Soient a, b des nombres réels tels que $a \leq b$. Alors pour tout réel c nous avons

$$a + c \leq b + c.$$

De plus, si $c \geq 0$, on a

$$ac \leq bc$$

Nous pouvons alors prouver ceci :

Résultat II.7. Soient a, b, c, d des réels vérifiant $a \leq b$ et $c \leq d$. Alors

$$a + c \leq b + d.$$

De plus, si $0 \leq a$ et $0 \leq c$, alors

$$ac \leq bd.$$

Démonstration. Comme $a \leq b$, alors $a + c \leq b + c$ (résultat précédent). De même, comme $c \leq d$, nous avons $b + c \leq b + d$. Par transitivité, nous obtenons $a + c \leq b + d$ comme annoncé.

La preuve pour le produit (avec la condition supplémentaire) est laissée en exercice. $\square$

Exercice. Faire la preuve de la seconde partie du résultat II.7.

Exercice. Trouver des réels a, b, c, d tels que $a \leq b$, $c \leq d$ et $ac > bd$.

2.2 Intervalles

Définition II.8. Un *intervalle* s'entend comme un ensemble de nombres réels « sans trou » (on parle d'ensemble *connexe*).

Exemple. * Si on définit $[-2, \pi]$ comme l'ensemble des réels compris entre -2 et π , c'est un intervalle : il n'y a aucun trou entre ces deux nombres.

* L'ensemble des réels positifs (ou nuls), noté $\mathbb{R}^+$, est également un intervalle.

* L'ensemble des réels sauf 0, noté $\mathbb{R}_0$, n'est pas un intervalle car 0 manque : il y a un trou.

On peut classer les intervalles comme suit.

Résultat II.9. Si a et b sont des nombres réels, on définit ces notations :

- * $[a, b]$ désigne l'ensemble des réels de a (compris) à b (compris) ;
- * $]a, b]$ désigne l'ensemble des réels de a (non-compris) à b (compris) ;
- * $[a, b[$ désigne l'ensemble des réels de a (compris) à b (non-compris) ;
- * $]a, b[$ désigne l'ensemble des réels de a (non-compris) à b (non-compris) ;
- * $]-\infty, b[$ désigne l'ensemble des réels strictement inférieurs à b ;
- * $]-\infty, b]$ désigne l'ensemble des réels inférieurs ou égaux à b ;
- * $]a, \infty[$ désigne l'ensemble des réels strictement supérieurs à a ;
- * $[a, \infty[$ désigne l'ensemble des réels supérieurs ou égaux à a ;
- * $]-\infty, \infty[$ désigne l'ensemble des réels, également noté $\mathbb{R}$.

Dans les notations ci-dessus, a et $-\infty$ sont la borne inférieure des intervalles dans lesquels ils apparaissent, tandis que b et ∞ sont la borne supérieure des intervalles dans lesquels ils apparaissent.

Un intervalle est généralement « infini » car il contient une infinité de nombre. Cependant, les quatre premières notations définissent des intervalles *bornés* (c'est-à-dire que les deux bornes sont des nombres réels), tandis que les cinq suivantes forment des intervalles non-bornés (l'une des deux bornes au moins est infinie).

Exemple. * L'intervalle $[0, 1]$ contient une infinité de nombres, puisqu'il contient notamment $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}$, etc. Mais il est également borné, car aucun des nombres n'est inférieur à 0 ou supérieur à 1 (de par la définition de cet intervalle!).

intervalle

connexe

borne inférieure

borne supérieure

bornés

- * L'intervalle $[4, 4]$ contient un seul nombre : le nombre 4.
- * L'intervalle $]4, 4[$ ne contient aucun nombre, car rien aucun nombre n'est à la fois strictement supérieur et strictement inférieur à 4.

Un ensemble $A \subset \mathbb{R}$ est *ouvert* si, de tout point de A , on peut se déplacer un peu dans chaque direction tout en restant dans A . Avant de donner une définition rigoureuse, voyons deux exemples :

- Exemple.** 1. L'ensemble des réels strictement positifs $\mathbb{R}_0^+$ est un ouvert : pour tout point de $x > 0$, il est possible de s'éloigner un peu de x tout en restant strictement positif. Il suffit de s'éloigner d'une distance inférieure à x .
2. L'ensemble des réels positifs ou nuls $\mathbb{R}^+$ n'est *pas* un ouvert : si on part de 0, il est impossible de se déplacer « vers les négatifs » en restant dans $\mathbb{R}^+$. Quelle que soit la distance parcourue, on se retrouve hors de $\mathbb{R}^+$.
3. L'ensemble $\mathbb{N}$ n'est *pas* un ouvert de $\mathbb{R}$: partant d'un entier naturel quelconque, il suffit de se déplacer un petit peu pour ne plus être entier ! Bien sûr, en se déplaçant plus longtemps on retombe sur des entiers, mais ça « ne compte pas ». Pour le comprendre, il faut avoir la définition rigoureuse ci-dessous.

Définition II.10. Soit $A \subset \mathbb{R}$, un ensemble.

- * L'*intérieur* de A est composé des points $p \in \mathbb{R}$ tels qu'il existe $\epsilon > 0$ vérifiant $]p - \epsilon, p + \epsilon[\subset A$. (En particulier $p \in A$!) On note $\text{int } A$ l'intérieur de A . intérieur
- * L'ensemble A est *ouvert* si pour tout $p \in A$ il existe $\epsilon > 0$ tel que $]p - \epsilon, p + \epsilon[\subset A$. ouvert

Remarquons donc que l'intérieur est formé des points pour lesquels « si on se déplace un peu, on reste dans l'ensemble », et qu'un ensemble est ouvert si il est égal à son intérieur, c'est-à-dire si cette propriété est vérifiée par *tous* les points de A .

Définition II.11. Soit $A \subset \mathbb{R}$, un ensemble.

- * L'*adhérence* de A est composée des points $p \in \mathbb{R}$ tels que pour tout $\epsilon > 0$, il existe $a \in A$ avec $|a - p| < \epsilon$. On note $\text{adh } A$ l'adhérence de A . adhérence
- * L'ensemble A est *fermé* si tous les points qui adhèrent à A sont dans A . fermé

Remarque. Le mathématicien est facétieux : un ensemble ouvert peut également être fermé (mais c'est rare), et un ensemble qui n'est pas ouvert peut très bien ne pas être fermé (ça arrive souvent !).

Résultat II.12. Pour tout ensemble A , nous avons $\text{int } A \subset A \subset \text{adh } A$.

Démonstration. Le lecteur est invité à imaginer la preuve seul. Rappelons que le signe $\subset$ veut dire que tout élément de l'ensemble à la gauche du signe est aussi dans l'ensemble à la droite du signe. $\square$

- Exemple.** 1. Le nombre 0 est dans l'adhérence de $\mathbb{R}_0^+$: il existe des nombres strictement positifs qui sont très proches de 0. Aussi proche qu'on veut. Notons que, comme 0 n'est pas lui-même strictement positif, cela implique que $\mathbb{R}_0^+$ n'est pas fermé.
2. L'ensemble des réels positifs ou nuls $\mathbb{R}^+$ est fermé : aucun point hors de $\mathbb{R}^+$ n'est adhérent à $\mathbb{R}^+$.
3. L'ensemble $\mathbb{N}$ est fermé.
4. Soit $A =]0, 1]$. Le nombre 0 est adhérent à A . Le nombre 1 n'est pas dans l'intérieur de A . De ces deux affirmations, nous déduisons que A n'est ni fermé, ni ouvert.

Exercice. Parmi les différentes sortes d'intervalle, certains sont ouverts et pas d'autres. L'auteur de ces notes est un grand distrait et a oublié lesquels sont ouverts, lesquels sont fermés (certains sont les deux, et d'autres aucun des deux !). Indiquez les bonnes réponses ci-dessous :

- * $[a, b]$
- * $]a, b]$
- * $[a, b[$
- * $]a, b[$
- * $] -\infty, b[$

- * $]-\infty, b]$
- * $]a, \infty[$
- * $[a, \infty[$
- * $]-\infty, \infty[$

(On suppose que a et b sont des réels vérifiant $a < b$.)

2.3 Pourcentages

Un pourcentage est une façon commode de quantifier le rapport entre deux quantités.

Exemple. « 50% des humains sont des femmes. » Ceci signifie que dans la population, la moitié sont des femmes. En effet, $\frac{50}{100} = \frac{1}{2}$.

Le « pourcent » décrit une proportion « pour cent unités ». En d'autres termes, on décrit un nombre qui doit être divisé par cent.

Exemple. « Le taux de réussite des étudiants de bachelier de l'année passée était d'un tiers. » En d'autres termes, un étudiant sur trois avait réussi. Ceci est approximativement 33% car

$$\frac{1}{3} = \frac{\frac{1}{3}100}{100} = \frac{33.3...3...}{100} = 33.3...3...\%$$

Le pourcentage est une information globale, et non individuelle.

Exemple. Un taux de 33% ne veut pas dire qu'en prenant trois étudiants au hasard, un seul réussira.

Le pourcentage n'est pas de la voyance.

Exemple. Le taux de réussite peut très bien être différent cette année !

2.4 Principe de proportionnalité – Règle de trois

La règle de trois est un principe qui s'applique dans nombre de situations de la vie courante.

Exemple. Pour faire un quatre quarts pour quatre personnes, il faut 200 grammes de beurre¹. Combien de beurre faut-il pour cinq personnes ?

- * Pour 4 personnes, il faut 200 grammes.
- * Pour 1 personne, il faut donc $\frac{200}{4}$ grammes, soit 50 grammes.
- * Pour 5 personnes, il faut dès lors 5×50 grammes, soit 250 grammes.

Le principe est le suivant : ce qui compte n'est pas la quantité proprement dite, mais la proportion par rapport aux autres quantités.

Exemple (Quatre quarts, continué). On passe de 4 à 5 personnes, c'est-à-dire on multiplie par $\frac{5}{4}$. Dès lors toutes les quantités doivent augmenter de la même manière, y compris le beurre. On multiplie donc la quantité de beurre par $\frac{5}{4}$, cela fait bien 250 grammes.

On peut également considérer les choses en terme d'augmentation par rapport à la quantité de départ :

Exemple (Quatre quarts, continué). On passe de 4 à 5 personnes, c'est-à-dire une augmentation de $\frac{5-4}{4} = 25\%$. Dès lors toutes les quantités doivent augmenter de la même proportion. Y compris le beurre. Or 25% de 200 grammes, c'est 50 grammes. Il faut donc rajouter 50 grammes.

Attention cependant que toutes les grandeurs ne jouent pas forcément à ce jeu là. Par exemple les températures :

Exemple. Un quatre quarts pour quatre personnes se cuit au four à 180 °C. Un quatre quarts pour huit personnes ne se cuit *pas* à 360 °C !

1. Les données sont inventées. L'auteur décline toute responsabilité en cas d'accident culinaire ou d'incompatibilité entre l'attente et les résultats de ceux qui puiseraient une recette dans le présent syllabus.

3 Manipulation et opération sur les nombres

3.1 Sommes

Considérons les calculs suivants :

$$1 + 2 = 3, 1 + 2 + 3 = 6, 1 + 2 + 3 + 4 = 10, \dots$$

Comment exprimer un calcul pour sommer les cent nombres de 1 jusqu'à 100? Où jusqu'à une valeur n quelconque? Une solution classique est d'inventer une notation!

Nous définissons un *symbole de sommation* comme suit :

symbole de
sommation

Définition II.13. Si m et n sont des entiers, $m \leq n$, et si $a_m, a_{m+1}, \dots, a_n$ sont des nombres, on définit :

$$\sum_{k=m}^n a_k := a_m + a_{m+1} + a_{m+2} + \dots + a_{n-1} + a_n.$$

Exemple. La notion est très simple mais requiert quelques exemples. Il s'agit juste de donner une autre manière d'écrire les sommes.

$$\sum_{k=1}^{100} k = 1 + 2 + 3 + \dots + 100$$

$$\sum_{k=1}^{100} k^2 = 1 + 4 + 9 + \dots + 10000$$

$$\sum_{k=1}^{100} \frac{1}{k} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{100}$$

$$\sum_{k=1}^{100} 3 = 3 + 3 + 3 + \dots + 3 (= 300)$$

$$\sum_{k=5}^{10} (k-1) = 4 + 5 + 6 + 7 + 8 + 9$$

$$\sum_{k=5}^5 (k-1) = 4$$

$$\frac{1}{100} \sum_{k=1}^{100} k = \frac{1 + 2 + 3 + \dots + 100}{100}$$

$$\sum_{k=1}^{100} \frac{k}{100} = \frac{1}{100} + \frac{2}{100} + \frac{3}{100} + \dots + \frac{100}{100}$$

$$(n+1) + \sum_{k=1}^n k = (n+1) + (1 + 2 + 3 + \dots + n) = 1 + 2 + 3 + \dots + n + (n+1) = \sum_{k=1}^{n+1} k.$$

Résultat II.14. Le symbole somme vérifie les relations suivantes :

$$\sum_{k=m}^n (a_k + b_k) = \left(\sum_{k=m}^n a_k \right) + \left(\sum_{k=m}^n b_k \right)$$

$$\sum_{k=m}^n \lambda a_k = \lambda \sum_{k=m}^n a_k$$

$$\sum_{k=m}^n a_k = \sum_{k=m}^n a_{n+m-k}$$

Démonstration. Les preuves sont très simples. Donnons cell du dernier de ces résultats, puisqu'il peut surprendre au premier abord. Par définition, le membre de gauche est

$$\sum_{k=m}^n a_k = a_m + a_{m+1} + \cdots + a_n$$

et le membre de droite est

$$\begin{aligned} \sum_{k=m}^n a_{n+m-k} &= a_{n+m-(m)} + a_{n+m-(m+1)} + a_{n+m-(m+2)} + \cdots + a_{n+m-(n-1)} + a_{n+m-(n)} \\ &= a_n + a_{n-1} + a_{n-2} + \cdots + a_{m+1} + a_m \end{aligned}$$

Les deux sont donc égaux (l'un est simplement sommé dans un sens, l'autre l'est dans l'autre sens). $\square$

Donnons un premier résultat explicite faisant intervenir le signe somme. Il donne la solution au problème initial.

Résultat II.15.

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}$$

Remarque. Avant de passer à la preuve, notons que ce résultat n'a rien de magique. Il pourrait très bien s'écrire sans le signe somme :

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}.$$

Démonstration. Appliquons le résultat prouvé précédemment :

$$\sum_{k=1}^n k = \sum_{k=1}^n (n+1-k)$$

et calculons

$$2 \left(\sum_{k=1}^n k \right) = \left(\sum_{k=1}^n k \right) + \left(\sum_{k=1}^n k \right) = \left(\sum_{k=1}^n k \right) + \left(\sum_{k=1}^n (n+1-k) \right) = \sum_{k=1}^n (k + (n+1-k)) = \sum_{k=1}^n (n+1) = n(n+1)$$

d'où le résultat final. $\square$

Remarque. Avant de passer à la suite, remarquons que cette preuve n'a rien de magique. Elle pourrait très bien s'écrire sans le signe somme. Faisons le pour $n = 100$ afin d'illustrer :

$$\begin{aligned} (1 + 2 + 3 + \cdots + 100) + (1 + 2 + 3 + \cdots + 100) &= (1 + 2 + 3 + \cdots + 100) + (100 + 99 + \cdots + 3 + 2 + 1) \\ &= (1 + 100) + (2 + 99) + (3 + 98) + \cdots + (100 + 1) \\ &= 101 + 101 + 101 + \cdots + 101 \\ &= 100 \cdot 101 \\ &= 10100 \end{aligned}$$

dès lors $1 + 2 + 3 + \cdots + 100 = \frac{10100}{2} = 5050$.

3.2 Moyenne arithmétique

Le signe somme permet de décrire de façon générale une quantité bien connue : la moyenne. La moyenne est un estimateur numérique qui sera décrit plus précisément dans un cours de statistique, mais il est intuitivement connu des étudiants qui, depuis toujours, calculent « leur moyenne » !

Exemple. Si un étudiant a suivi trois cours, et a obtenu les notes x_1 , x_2 et x_3 , alors sa moyenne (arithmétique) est $\frac{x_1 + x_2 + x_3}{3}$.

Exemple. Si cinq étudiants ont suivi le cours de mathématiques, et ont eut des notes e_1, e_2, e_3, e_4, e_5 , alors on dira que la moyenne des notes pour ce cours est

$$\frac{e_1 + e_2 + e_3 + e_4 + e_5}{5}.$$

Définition II.16. La moyenne arithmétique de n nombres $x_1, \dots, x_n$ est la quantité

$$\bar{x} = \frac{\sum_{k=1}^n x_k}{n}$$

Résultat II.17. La moyenne arithmétique d'une séquence de nombres est toujours plus petite que le plus grands de ces nombres, et plus grande que le plus petit de ces nombres.

Démonstration. Considérons la moyenne $\bar{x}$ de n nombres $x_1, \dots, x_n$. L'un de ces nombres, « le plus petit », est inférieur ou égal à tous les autres. Il a donc un certain indice, disons j . De même, le plus grand a un certain indice disons J . C'est-à-dire qu'on a $x_j \leq x_k \leq x_J$ pour tout k entre 1 et n . Dès lors :

$$x_1 + x_2 + \dots + x_n \leq x_J + x_J + \dots + x_J = nx_J$$

(ceci est une application répétée de la proposition II.7) dont on tire

$$\bar{x} \leq x_J$$

c'est-à-dire : la moyenne d'une séquence de nombres est inférieure au plus grand de ces nombres.

La preuve se termine similairement avec x_j :

$$x_1 + x_2 + \dots + x_n \geq x_j + x_j + \dots + x_j = nx_j$$

d'où $\bar{x} \geq x_j$, ce qui est le résultat annoncé. $\square$

3.3 Puissances

Remarque (Sur la notation des produits). Généralement le produit entre deux quantités se note en juxtaposant ces quantités. Par exemple ab veut dire « a fois b », et $3x$ veut dire « 3 fois x ». Lorsque ces deux quantités sont des nombres tels que 3 et 5, cette notation amènerait à confondre le produit « 3 fois 5 » avec le nombre 35. Lorsqu'il y a ambiguïté, nous utilisons alors la notation $3 \cdot 5$ pour désigner le produit de 3 et de 5. Nous n'utilisons généralement *pas* la notation 3×5 , pour éviter de confondre le symbole $\times$ avec la variable x .

Pour tout réel x , la notation x^b (lire : « x exposant b », ou « x à la b^{e} puissance ») a le sens suivant, lorsque b est un entier naturel :

$$x^b = \underbrace{x \cdot x \cdot \dots \cdot x}_{b \text{ fois}}.$$

Avec comme cas particulier : $x^0 = 1$.

Résultat II.18. Pour tous réels x, y et pour tous naturels a, b , nous avons

$$x^a y^a = (xy)^a \qquad x^a x^b = x^{a+b} \qquad (3.1)$$

$$(x^a)^b = x^{ab} \qquad \frac{x^a}{x^b} = x^{a-b} \qquad (3.2)$$

(La dernière égalité suppose $a \geq b$, sinon $a - b$ n'est pas naturel.)

Exemple. Sans surprise, quelques lignes de calcul faisant intervenir des puissances :

$$\begin{aligned} 3^5 &= 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 = 243 \\ 3^4 3^2 &= (3 \cdot 3 \cdot 3 \cdot 3) \cdot (3 \cdot 3) = 3^6 = 729 \\ (3^2)^3 &= (3^2) \cdot (3^2) \cdot (3^2) = (3 \cdot 3) \cdot (3 \cdot 3) \cdot (3 \cdot 3) = 3^6 \\ \frac{2187}{81} &= \frac{3^7}{3^4} = \frac{3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3}{3 \cdot 3 \cdot 3 \cdot 3} = 3 \cdot 3 \cdot 3 = 3^3 = 27 \end{aligned}$$

Remarque (*). Sachant cela, à supposer que nous voulions donner du sens à x^{-4} , il serait agréable que $x^{-4}x^4 = x^{-4+4} = x^0 = 1$, c'est-à-dire $x^{-4} = \frac{1}{x^4}$. Notons que, pour que cela fonctionne, il faut supposer $x \neq 0$.

Remarque (*). Sachant cela, à supposer que nous voulions donner du sens à $x^{\frac{1}{4}}$, il serait agréable que $(x^{\frac{1}{4}})^4 = x^{\frac{4}{4}} = x^1 = x$. Pour que cela fonctionne, il faut supposer $x \geq 0$ car n'importe quoi mis à une puissance quatrième sera positif!

En fait, on peut donner un sens raisonnable à cette notation x^b pour tout $b \in \mathbb{R}$ à condition de se restreindre aux $x > 0$. Avec cette définition (que nous ne donnons pas), x^b est positif également. Il sort malheureusement du cadre de ce chapitre de faire la définition et la justification explicite. Néanmoins, ré-écrivons le résultat précédent dans sa forme plus générale :

Résultat II.19. Pour tous réels strictement positifs x, y , et pour tous réels a, b , nous avons

$$x^a y^a = (xy)^a \qquad x^a x^b = x^{a+b} \qquad (3.3)$$

$$(x^a)^b = x^{ab} \qquad \frac{x^a}{x^b} = x^{a-b} \qquad (3.4)$$

Les puissances permettent de répondre à certaines questions de comptage :

Question. Combien de séquences de 5 lettres peut-on réaliser avec les lettres A, C, G et T ?

Réponse. En première position, on a quatre lettres possibles. Pour chacun de ces quatre choix, nous avons à nouveau quatre choix possibles en deuxième position. Et ainsi de suite jusqu'à la cinquième position, nous avons donc au total le nombre de choix possibles suivant :

$$4 \times 4 \times 4 \times 4 \times 4 = 4^5 = 1024$$

Question. Combien de séquences de 8 chiffres peut-on réaliser avec les chiffres 0 et 1 ?

Réponse. Comme ci-dessus, avec cette fois seulement deux symboles pour huit positions :

$$2^8 = 256$$

3.4 Valeur absolue

La *valeur absolue* de x , notée $|x|$, est définie par :

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

(le symbole $:=$ indique une « égalité, par définition », de sorte que le membre de gauche est défini comme étant égal au membre de droite.)

Exemple. * Nous avons $|5| = 5$ car 5 est positif, donc on est dans le premier cas.

* Similairement, $|-5| = -(-5) = 5$ car -5 est négatif, donc on est dans le second cas.

* Le cas de 0 n'est pas spécial : $|0| = 0$, tout simplement (premier cas).

Résultat II.20. Pour tous x, y , nous avons

$$|x| = |-x| \qquad |x - y| = |y - x|$$

Démonstration. Pour la première égalité, on distingue trois cas :

* Si $x = 0$ est nul, alors $-x = 0$,

* Si $x > 0$, alors $-x < 0$ donc $x = -(-x)$,

* Si $x < 0$, alors $-x > 0$ donc $-x = -x$.

Pour la seconde égalité, notons simplement que $x - y = -(y - x)$. □

valeur absolue

Remarque. Si x et y sont des nombres réels, le nombre $|x - y|$ peut être vu comme la distance entre x et y . Le lecteur est invité à penser à cela seul avant de lire la suite de cette remarque.

Nous allons distinguer les cas et voir que ça marche dans chacun. Avant tout, remarquons que la distance entre x et y doit être la même que celle entre y et x . Or $|x - y| = |y - x|$, donc on peut supposer $x < y$ dans notre raisonnement, et il sera alors également valide pour $y < x$.

- * si $0 < y < x$, il est assez clair qu'on voudrait que $x - y$ soit la distance entre les deux. Par exemple la distance entre 2 et 5 serait $5 - 2 = 3$. Ça tombe bien : $|x - y| = x - y$ dans ce cas.
- * si $y < 0 < x$, on voudrait que la distance entre y et x soit égale à la distance entre y et 0 additionnée à la distance entre 0 et x . La première vaut $-y$ (car y est négatif), la seconde vaut x (car x est positif), donc $-y + x$ est la distance recherchée, ça tombe bien : c'est encore égal à $|x - y|$.
- * si $y < x < 0$, on a envie que la distance entre x et y soit la même que la distance entre $-x$ et $-y$. Or $y < x < 0$, donc $0 < -x < -y$, ce qui permet d'appliquer le premier cas : la distance entre $-x$ et $-y$ est $|-y - (-x)|$ c'est-à-dire $|x - y|$!

3.5 Écriture scientifique

Certains nombres sont trop grands ou trop petits pour être décrits dans les unités classiques. Par exemple la masse de Jupiter est d'environ 1 898 600 000 000 000 000 000 000 kg. Ridicule. Même en tonnes, ça n'enlève que trois zéros.

On prend alors l'habitude de décrire ces nombres en *notation scientifique* : $m10^e$. Le nombre m s'appelle la mantisse, le nombre e s'appelle l'exposant.

notation
scientifique

Exemple. Voici les masses de quelques planètes de notre système solaire

Planète	Masse (kg)
Jupiter	$1,898\,6 \cdot 10^{27}$
Saturne	$5,684\,6 \cdot 10^{26}$
Neptune	$1,024\,3 \cdot 10^{26}$
Uranus	$8,683\,2 \cdot 10^{25}$
Terre	$5,973\,6 \cdot 10^{24}$
Vénus	$4,868\,5 \cdot 10^{24}$
Mars	$6,418\,5 \cdot 10^{23}$
Mercure	$3,302 \cdot 10^{23}$

3.6 Racines

Regardons un nombre x , et considérons x^2 . Quelques valeurs :

x	x^2
0	0
0.10	0.01
0.50	0.25
0.90	0.81
0.99	0.9801
1	1
1.1	1.21
1.5	2.25
2	4
10	100

Il est clair que la colonne de droite, x^2 , prendra des valeurs depuis 0 jusqu'à autant que voulu, et les valeurs vont croissant sans jamais redescendre. Cette remarque montre que si t est un réel positif, il existe un unique x positif tel que $t = x^2$. Ce réel x est appelé racine carrée de t et se note $\sqrt{t}$.

Définition II.21. Si t est un réel positif ou nul, sa *racine carrée*, notée $\sqrt{t}$, est l'unique réel positif ou nul dont le carré vaut t .

racine carrée

Remarque. Si t est strictement positif, alors il existe exactement *deux* nombres dont le carré vaut t : $\sqrt{t}$ et $-\sqrt{t}$. Le premier est positif, le second est négatif.

On peut sans problème étendre ces raisonnements à x^3 , x^4 , etc.

racine n^e

Définition II.22. Si t est un réel positif ou nul, sa *racine n^e* , notée $\sqrt[n]{t}$, est l'unique réel positif ou nul dont la puissance n^e vaut t .

Remarque (*). Lorsque n est impair, cela garde un sens de parler de racine n^e d'un nombre t négatif. Par exemple, la racine cubique de -8 est $\sqrt[3]{-8} = -2$ car $(-2)^3 = -8$. Mais ceci ne fonctionne pas pour n pair car x^n est toujours positif si n est pair, donc il est impossible de trouver la racine n^e d'un nombre négatif dans ce cas.

Résultat II.23. Pour tout $x > 0$ et pour tout naturel n , nous avons :

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

Démonstration. D'après le résultat II.19, nous savons que le nombre $x^{1/n}$ mis à la puissance n vaut x . Nous savons également (d'après le commentaire juste au dessus du résultat II.19) qu'il est positif, donc c'est bien la racine n^e de x d'après notre définition. $\square$

3.7 Factorielle

factorielle

Définition II.24. Si $n \geq 1$ est un nombre naturel, on définit la *factorielle* de n , notée $n!$, le nombre obtenu en réalisant le produit de tous les naturels entre 1 et n . On définit à part la factorielle de 0 par $0! = 1$.

Exemple. Voici quelques valeurs :

$$0! = 1$$

$$1! = 1$$

$$2! = 2 \cdot 1 = 2$$

$$3! = 3 \cdot 2 \cdot 1 = 6$$

$$4! = 4 \cdot 3 \cdot 2 \cdot 1$$

$$5! = 120$$

Question. Combien de séquences de 4 lettres peut-on réaliser avec les lettres A, C, G et T en utilisant une seule fois chaque lettre ?

Réponse. Quatre lettres sont possibles en première position. Pour chaque tel choix, quel qu'il soit, il reste ensuite trois lettres possibles en deuxième position, puis deux, puis une. Au total cela fait le nombre de choix suivant :

$$4 \times 3 \times 2 \times 1 = 24 = 4!$$

Résultat II.25. Pour tout $n \geq 1$, $n! = n(n-1)!$.

Exemple. Nous avons $6! = 6 \cdot 5! = 6 \cdot 120 = 720$.

Résultat II.26. Pour tous naturels $1 \leq n < m$, nous avons :

$$\frac{m!}{n!} = m(m-1) \cdots (n+2)(n+1)$$

Démonstration. La preuve est intuitive : on divise le produit des m premiers entiers par le produit des n premiers entiers. Restent donc les entiers entre $n+1$ et m . $\square$

Exemple. Le nombre $40!$ contient 48 chiffres, mais il est facile de déterminer le quotient suivant :

$$\frac{40!}{37!} = 40 \cdot 39 \cdot 38 = 59280.$$

Question. Combien séquences de 2 lettres (différentes) peut-on former avec les lettres A, C, T, G ?

Réponse. $4 \times 3 = 12$.

Question. Combien de séquences de k lettres (différentes) peut-on former avec n lettres (différentes) données ?

Réponse.

$$n(n-1)(n-2) \cdots (n-(k-1)) = \frac{n!}{(n-k)!}$$

Remarque (*). Le mot « séquence » sous-entend que l'ordre a une importance.

3.8 Coefficients binomiaux, combinaisons

Ci-dessous, par « tas » ou « ensemble » nous voulons dire que l'ordre des lettres ne compte pas (et il n'y a pas de répétition possible dans un même tas).

Exemple. Avec les trois lettres ABC, quels tas de lettres peut-on former ?

A, B, C, AB, AC, BC, ABC.

Dans cet exemple il y a donc trois tas d'une lettre, trois de deux lettres, et un seul tas de trois lettres. L'ordre n'ayant pas d'importance, le tas « ABC » (par exemple) pourrait encore s'écrire ACB, BAC, BCA, CAB ou CBA.

Dans le cas général, supposons avoir n lettres à disposition, et considérons les tas de k lettres, avec $k \leq n$. Nous avons déjà qu'il y a $\frac{n!}{(n-k)!}$ mots de k lettres. Or chaque mot admet $k!$ permutations, donc nous avons « k factorielle fois » plus de mots que de tas. Nous devons donc diviser. En résumé nous avons :

Question. Combien d'ensembles de k objets (différents) peut-on former avec n objets (différents) donnés ?

Réponse. Il y a

$$\frac{n!}{(n-k)!k!}$$

ensembles de k objets. Ce nombre est également noté $\binom{n}{k}$ et est appelé coefficient binomial.

coefficient
binomial

Définition II.27. Le nombre défini par

$$\binom{n}{k} := \frac{n!}{(n-k)!k!}$$

est appelé un coefficient binomial. On le prononce « n au dessus de k » ou « n choose k » (avec votre plus bel English accent).

Nous avons compté le nombre d'ensembles de k lettres, mais combien d'ensembles y a-t-il en tout ? La réponse est à la fois simple et décevante, puisque chaque ensemble comporte entre 0 et n lettres, il suffit de compter toutes les possibilités pour chacune de ces valeurs de k , et faire la somme :

$$\sum_{k=0}^n \binom{n}{k}.$$

Attention : nous comptons aussi un ensemble vide, ne contenant aucune lettre ($k = 0$) de cette façon ! C'est la tradition, mais cela peut paraître surprenant.

Exemple. Reprenant l'exemple précédant, avec $n = 3$ (les lettres A, B et C), nous avons compté $3 + 3 + 1 = 7$ tas non-vides, soit 8 tas en tout.

Une autre façon de compter tous ces ensembles est la suivante : dans un ensemble donné, chacune des n lettres est ou n'est pas dans le ensemble. Il y a donc deux possibilités par lettre, soit au total 2^n possibilités. De la sorte, nous avons montré l'égalité suivante :

Résultat II.28.

$$\sum_{k=0}^n \binom{n}{k} = 2^n.$$

Nous pouvons, par d'autres arguments de comptage, montrer une autre égalité faisant intervenir les coefficients binomiaux : comptons le nombre d'ensembles de k objets qu'il est possible de faire avec $n + 1$ objets. Cela fait $\binom{n+1}{k}$. Parmi ces $n + 1$ objets, mettons-en un à part, et nommons-le ω . Chaque ensemble de k objets peut contenir ou ne pas contenir ω .

- * Combien d'ensembles de k objets ne contiennent pas ω ?
- * Combien d'ensembles de k objets contenant ω ?

(Le lecteur est invité à compléter ces réponses, en utilisant les coefficients binomiaux.) Grâce à ce raisonnement nous avons prouvé la proposition suivante.

Résultat II.29. *Pour tout naturel n , pour tout naturel $k \geq 1$:*

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$$

Démonstration. Une preuve combinatoire (faisant intervenir du comptage) a été faite ci-dessus, voici une preuve plus algébrique :

$$\begin{aligned} \frac{n!}{(n-k)!k!} + \frac{n!}{(n-(k-1))!(k-1)!} &= \frac{n!(n-k+1)}{(n-k)!k!(n-k+1)} + \frac{n!k}{(n-k+1)!k(k-1)!} \\ &= \frac{n!(n-k+1) + n!k}{k!(n-k+1)!} = \frac{n!(n-k+1+k)}{(n+1-k)!k!} = \binom{n+1}{k}. \end{aligned}$$

□

Notons également que faire un ensemble comportant k objets, c'est la même chose que de choisir les $n-k$ objets laissés hors de l'ensemble. Ceci prouve :

$$\binom{n}{k} = \binom{n}{n-k}$$

(Une preuve algébrique est très simple, et laissée en exercice.)

3.9 Formule du binôme de Newton

Résultat II.30 (Binôme de Newton). *Pour tous réels x, y et tout entier naturel n , la formule suivante est vérifiée :*

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}.$$

Démonstration. La preuve peut se faire par récurrence.

Pour $n = 0$, la relation s'écrit $(x+y)^0 = \binom{0}{0}x^0y^0$, c'est-à-dire $1 = 1$. La relation est donc prouvée.

Supposons la relation vraie pour $n = p$, où p est un entier naturel fixé. Prouvons la relation pour

$n = p + 1$.

$$\begin{aligned}
(x+y)^{p+1} &= (x+y)(x+y)^p \\
&= (x+y) \sum_{k=0}^p \binom{p}{k} x^k y^{p-k} \\
&= \sum_{k=0}^p \binom{p}{k} x^{k+1} y^{p-k} + \sum_{k=0}^p \binom{p}{k} x^k y^{p+1-k} \\
&= x^{p+1} + \sum_{k=0}^{p-1} \binom{p}{k} x^{k+1} y^{p-k} + \sum_{k=1}^p \binom{p}{k} x^k y^{p+1-k} + y^{p+1} \\
&= x^{p+1} + \sum_{k=1}^p \binom{p}{k-1} x^k y^{p-(k-1)} + \sum_{k=1}^p \binom{p}{k} x^k y^{p+1-k} + y^{p+1} \\
&= x^{p+1} + \sum_{k=1}^p \left(\binom{p}{k-1} + \binom{p}{k} \right) x^k y^{p-(k-1)} + y^{p+1} \\
&= x^{p+1} + \sum_{k=1}^p \binom{p+1}{k} x^k y^{p-(k-1)} + y^{p+1} \\
&= \sum_{k=0}^{p+1} \binom{p+1}{k} x^k y^{p-(k-1)}
\end{aligned}$$

où nous avons utilisé la relation :

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}.$$

□

Remarque (*). Cette formule du binôme est attribuée à Newton mais était connue bien avant la naissance de Sir Isaac Newton. La contribution de Newton est d'étendre cette formule au cas où l'exposant n n'est pas un entier, mais cette généralisation dépasse le cadre de ce chapitre.

3.10 Triangle de Pascal

Le triangle de Pascal permet de retenir facilement les coefficients binomiaux pour de petites valeurs de n . La première ligne commence à $n = 0$, puis $n = 1$, etc. Le nombre le plus à gauche de chaque ligne correspond à $k = 0$, puis $k = 1$, etc. C'est la relation

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}.$$

qui permet de reconstruire facilement ce tableau.

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

							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

4 Équations et systèmes

équation

solutions

Définition II.31. Une *équation* est une égalité faisant intervenir une ou plusieurs quantités inconnues. Résoudre une équation revient à déterminer l'ensemble des valeurs possibles pour la quantité inconnue de sorte que l'égalité soit vérifiée. Ces valeurs sont les *solutions* de l'équation.

Dans une de ses formes les plus simples, une équation fait intervenir une unique quantité inconnue : un nombre réel. La « quantité inconnue » (ou simplement « inconnue ») est souvent nommée x , mais ce nom n'a rien de magique.

Exemple. * L'équation $2x - 3 = 1$ a pour seule solution : $x = 2$.

- * L'équation $t^2 - 1 = 0$ (dont l'inconnue est t) a pour solutions : $t = 1$ et $t = -1$. On peut écrire que l'ensemble des solutions est $S = \{-1, 1\}$.
- * L'équation $\sin(x) = 0$ a pour solution $x = 0$, mais il y en a d'autres. Par exemple $x = \pi$. En fait, l'ensemble des solutions est formé de l'ensemble des multiples entiers de π . On peut noter $S = \{k\pi \text{ t.q. } k \in \mathbb{Z}\}$.
- * L'équation $x^2 + 1 = 0$ n'a pas de solution dans les nombres réels, car tout réel pris au carré est positif, donc $x^2 + 1$ ne peut jamais être égal à 0. L'ensemble des solutions est donc vide. On peut noter $S = \emptyset$.

Une équation peut faire intervenir plusieurs inconnues. Dans ce cas, *une* solution est la donnée d'une valeur pour chaque inconnue.

Exemple. L'équation $x^2 + y^2 = 0$ (dont les inconnues sont x et y) a une solution : $x = y = 0$. Il n'y en a pas d'autres. On peut noter $S = \{(0, 0)\}$.

Dans l'exemple ci-dessus, il y avait une équation, deux inconnues, mais une seule solution. C'est rare. Généralement une seule équation ne permet pas de déterminer les deux inconnues.

Exemple. L'équation $x + y^2 = 0$ possède une infinité de solutions : pour chaque nombre réel r , les valeurs $y = r$ et $x = -r^2$ fournissent une solution. On peut noter $S = \{(r, -r^2) \text{ t.q. } r \in \mathbb{R}\}$.

4.1 Systèmes d'équations

système
d'équations

Lorsqu'il y a plusieurs inconnues, il peut arriver qu'il y ait également plusieurs équations. On parle alors de *système d'équations*.

Exemple. Considérons le système d'équations suivant, dont les inconnues sont (x, y) :

$$\begin{cases} x + y = 1 \\ x - y = 3 \end{cases}$$

Les solutions sont $x = 2$ et $y = -1$. Il y a une unique solution.

Exemple. Considérons le système d'équations suivant, dont les inconnues sont (x, y) :

$$\begin{cases} x - y^2 = 1 \\ x + y^2 = 3 \end{cases}$$

En résolvant, on trouve $x = 2$ et $y^2 = 1$. Dès lors il y a deux possibilités :

- * Soit $x = 2$ et $y = 1$,
- * soit $x = 2$ et $y = -1$.

Il y a donc ici deux solutions !

5 Géométrie élémentaire

5.1 Longueurs

Chacun sait intuitivement ce qu'est la longueur. Plus un chemin est long, plus il faudra de temps à le parcourir. Pour mesurer une longueur, l'unité est le mètre ou ses dérivés (centimètre, kilomètres, etc.)

Exemple. Un carré est une figure du plan formée de quatre côtés de même longueur. Chacun des côtés a une longueur c . Le périmètre, c'est-à-dire la longueur du pourtour du carré, est alors $4c$.

Exemple. Un cercle est une figure du plan formée de l'ensemble des points à une distance $r > 0$ donnée d'un autre point (le centre du cercle). Le périmètre d'un tel cercle est $2\pi r$.

Le problème d'écrire mathématiquement la longueur d'une courbe quelconque (du plan ou de l'espace) sera résolu en section ?? page ??.

5.2 Aire

La notion d'aire est une notion plus complexe. Il s'agit ici d'estimer la place que prend une portion de plan. Lorsque la portion de plan est rectangulaire, la formule est simple : la hauteur multipliée par la largeur.

Exemple. L'aire de la surface délimitée par un carré est c^2 .

Exemple. L'aire d'un disque de rayon r (surface délimitée par le cercle de rayon r) est πr^2 .

Exercice. Ayant à sa disposition 300 mètres de clôture, une bergère veut délimiter un enclos pour ses blancs moutons avant qu'il pleuve. Afin que ses bêtes à quatre pattes aient le plus de place possible, quelle forme doit avoir l'enclos : un carré ou un cercle ?

On pourrait montrer, mais c'est hors sujet, la chose suivante :

Résultat II.32. *Le disque est la surface plane ayant l'aire la plus grande parmi toutes les figures dont le périmètre est fixé.*

L'écriture de l'aire d'une surface quelconque du plan sera décrit au chapitre XVII page 163.

5.3 Volume

La notion de volume est la généralisation logique de la notion d'aire. Pour un solide en parallélépipède rectangle (une brique), la formule est simple : surface de la base multipliée par la hauteur.

Exemple. Un cube de côté c a pour volume c^3 .

Exemple. Une boule de rayon r a pour volume $\frac{4}{3}\pi r^3$.

L'écriture du volume d'un solide quelconque de l'espace sera décrit au chapitre XVII page 163.

5.4 Surface latérale

De la même manière qu'une surface (par exemple un disque) est délimitée par une certaine courbe (dans l'exemple, un cercle), un solide (par exemple une boule) est généralement délimité par une surface de l'espace (dans l'exemple, une sphère). Cette surface, appelée *surface latérale*, possède une certaine aire, appelée *aire latérale*.

Exemple. La surface latérale d'un cube est composée de six faces carrées. L'aire latérale est donc égale à $6c^2$.

Résultat II.33. *L'aire latérale de la boule de rayon r est $4\pi r^2$.*

Le problème de décrire et calculer mathématiquement l'aire de la surface latérale d'un solide quelconque de l'espace sera approché au chapitre ?? page ??.

surface latérale

aire latérale

Chapitre III

Notations ensemblistes

De manière générale, un ensemble est une notion abstraite qui permet de rassembler des éléments sous une appellation commune. La définition qui suit est informelle, mais capture l'essentiel de ce qui est nécessaire à comprendre la notion et nous permettra de faire quelques petits raisonnements :

Définition III.1. Un *ensemble* est défini par les éléments qu'il contient.

ensemble

On note $x \in A$ le fait pour x d'appartenir à l'ensemble A . On dit aussi que x est un *élément* de A , ou que x *appartient* à A . À l'inverse, on note $x \notin A$ lorsque x n'est pas un élément de A .

élément

appartient

Exemple. * L'ensemble $\mathbb{N}$ de tous les entiers naturels est... un ensemble. Il contient 0 et 3 mais pas -3 : $0 \in \mathbb{N}$, $3 \in \mathbb{N}$, $-3 \notin \mathbb{N}$. (Notons que $-0 \in \mathbb{N}$ car $-0 = 0$.)

- * Si on considère l'ensemble E de tous les nombres entiers pairs, alors $2 \in E$, et $3 \notin E$. On peut aussi écrire : $E \ni 2$ et $E \not\ni 3$ (très rarement utilisé).
- * Considérons l'ensemble des étudiants inscrits à l'ULB. C'est un ensemble, mais bien sûr ce n'est pas un ensemble de nombres. On peut faire des ensemble de n'importe quoi.

Comme un ensemble est en fait défini par les éléments qu'il contient, pour décrire un ensemble on peut aussi décrire tous ses éléments explicitement. Pour cela, on écrit tous les éléments entre accolades.

Exemple. L'ensemble $E = \{2, \text{pomme}, \pi, -2\}$ est un ensemble de quatre éléments.

Remarque. Les ensembles $\{2, 3, \pi\}$ et $\{2, \pi, 3\}$ contiennent tous les deux exactement les mêmes éléments, dès lors il s'agit en fait d'un seul et même ensemble : *l'ordre des éléments n'importe pas*, les éléments ne sont pas a priori ordonnés.

Par ailleurs un élément ne peut pas appartenir « plusieurs fois » à un ensemble : soit il y est, soit pas. Dès lors $\{2, 3, \pi, 3, 2\}$ contient encore exactement les mêmes éléments que $\{2, 3, \pi\}$: c'est toujours le même ensemble. *Les répétitions n'importent pas*.

On peut également partir d'un ensemble donné et n'en sélectionner que ceux qui vérifient une condition donnée :

Exemple. Si on note $\mathbb{N}$ l'ensemble des nombres naturels, alors l'ensemble $\{x \in \mathbb{N} \text{ t.q. } x \text{ est pair}\}$ est simplement l'ensemble de tous les entiers pairs.

Définition III.2. On dit d'un ensemble A qu'il est *inclus* à B , noté $A \subset B$, si et seulement si pour tout $x \in A$, on a $x \in B$.

inclus

Corollaire. Pour deux ensembles A et B , si $A \subset B$ et $B \subset A$, alors $A = B$

Démonstration. L'affirmation $A \subset B$ indique que tout élément de A est dans B , et $B \subset A$ indique tout élément de B est dans A ; dès lors A et B ont les mêmes éléments, et sont donc le même ensemble. $\square$

1 L'ensemble vide

ensemble vide

Il y a un ensemble particulier : celui qui ne contient aucun élément. Il pourrait être noté $\{\}$ mais il est généralement noté $\emptyset$. On l'appelle l'*ensemble vide*. La propriété qui suit n'a aucun autre intérêt que de se familiariser avec la notion :

Résultat III.3. Pour tout ensemble E , nous avons $\emptyset \subset E$.

Démonstration. Il faut montrer que si $x \in \emptyset$, alors $x \in E$. Or $\emptyset$ ne contient aucun élément ! Donc il n'y a rien à vérifier, ce qui prouve que c'est vrai. $\square$

Remarque (*). Le raisonnement ci-dessus peut éventuellement perturber ou choquer. Cependant, qu'est-ce que cela voudrait dire si l'affirmation était fausse ? Prenons une analogie : tous les martiens sont verts. Comment prouver que c'est faux ? Il faut trouver un martien qui ne soit pas vert ! Or si les martiens n'existent pas, on n'y arrivera pas. Donc c'est vrai. (Voir aussi 3.1.4 page 15.)

2 Union, intersection et différence

intersection

union

différence

Si A et B sont deux ensembles, on note $A \cap B$ leur *intersection*, $A \cup B$ leur *union* et $A \setminus B$ leur *différence*. Les définitions sont les suivantes :

$$A \cap B = \{x \text{ t.q. } x \in A \text{ et } x \in B\}$$

$$A \cup B = \{x \text{ t.q. } x \in A \text{ ou } x \in B\}$$

$$A \setminus B = \{x \text{ t.q. } x \in A \text{ et } x \notin B\}$$

complémentaire

Lorsque B est inclus à A , on dit que $A \setminus B$ est le *complémentaire* de B dans A . Si l'ensemble A est évident dans le contexte, on dit simplement « le complémentaire de B » sans préciser.

Exemple. 1. L'ensemble des nombres entiers est l'union des nombres entiers pairs et des nombres entiers impairs.

2. Le complémentaire des entiers pairs dans $\mathbb{Z}$ est l'ensemble des entiers impairs.

3 Diagrammes de Venn — Patatoïdes convexes

patatoïde
convexe

Un diagramme de Venn est une façon de se représenter les opérations ensemblistes classiques (décrites ci-dessous). Chaque ensemble est représenté par une « patate ». Ce terme n'ayant pas l'air très scientifique, on utilisera parfois le terme *patatoïde convexe* afin d'avoir l'air sérieux.

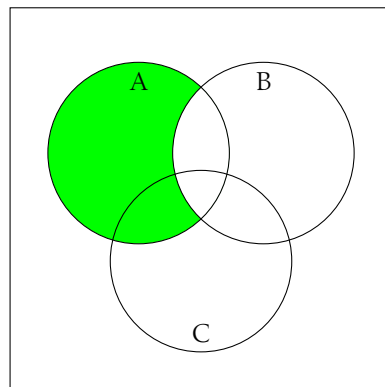
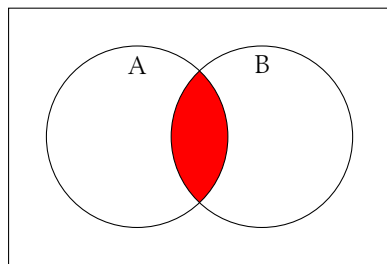
Si A et B sont des ensembles, on définit les trois opérations :

Intersection $A \cap B = \{x \text{ t.q. } x \in A \text{ et } x \in B\}$

Union $A \cup B = \{x \text{ t.q. } x \in A \text{ ou } x \in B\}$

Différence $A \setminus B = \{x \text{ t.q. } x \in A \text{ et } x \notin B\}$

Exercice. Voici deux diagrammes de Venn (en couleur, pour ceux qui ont la couleur) :



Que représentent-ils ?

Le lecteur est invité à dessiner lui-même quelques diagrammes de Venn :

$A \cup B$ (union)

$A \cap B$ (intersection)

$A \cap B \cap C$ (intersection triple)

$A \setminus B$ (différence)

4 Couples et n-uples

Un *couple de réels* est la donnée de deux réels, pour lesquels l'ordre importe et la répétition est possible! Si a et b sont deux réels, on peut former le couple (a, b) (et le couple (b, a)).

couple de réels

Deux couples de réels (a, b) et (c, d) sont *égaux* si et seulement si $a = c$ et $b = d$. Par exemple $(1, 3) \neq (1, 4)$ et $(1, 3) \neq (3, 1)$, mais $(3, 2) = (3, 2)$. L'ordre a une importance : $(a, b) = (b, a)$ si et seulement si $a = b$.

On définit de même un triple de réels : (a, b, c) . Deux triples sont égaux si leur premier élément est identique, ainsi que leur deuxième, ainsi que leur troisième éléments. On définit de la sorte des « uples » de plus en plus longs.

Notation : $\mathbb{R}$ est l'ensemble des réels, $\mathbb{R}^2$ l'ensemble des couples de réels, $\mathbb{R}^3$ l'ensemble des triples (ou triplets) de réels, etc. De manière générale, un élément de $\mathbb{R}^n$ est appelé un n -uple de réels.

On peut en fait définir la notion de couple pour n'importe quels ensembles A et B . On note $A \times B$ l'ensemble des couples (a, b) dont le premier élément a est dans A et le second élément b est dans B .

Exemple. L'ensemble des couples de réels est $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$. L'ensemble des triples de réels est $\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^3$.

Remarque (*). On pourrait se demander quel sens donner à la notation $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$: s'agit-il de $(\mathbb{R} \times \mathbb{R}) \times \mathbb{R}$ ou de $\mathbb{R} \times (\mathbb{R} \times \mathbb{R})$? C'est une bonne question, et je vous remercie de l'avoir posée.

La réponse est qu'il y a certes une différence formelle, mais pas de différence pratique. En pratique, donc, nous n'écrivons pas les parenthèses et nous nous contentons d'écrire des triples de réels sans nous soucier qu'il pourrait s'agir de $(a, (b, c))$ ou de $((a, b), c)$ sous le chapeau.

Chapitre IV

Fonctions

Définition IV.1. Se donner une *fonction* f s'est se donner un *ensemble de départ* A , un *ensemble d'arrivée* B et une règle permettant d'associer à chaque élément x de A un unique élément de B , l'*image* de x par f , noté $f(x)$. Inversement, on dit que x est un *antécédent* de $f(x)$.

La notation usuelle pour une fonction f de A dans B est

$$f : A \rightarrow B : x \mapsto f(x)$$

Exemple. Voici quelques exemples de fonctions :

- * Une fonction associant à chaque réel son carré

$$g : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^2;$$

- * la fonction valeur absolue

$$h : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto |x|;$$

- * la fonction *plancher*

$$i : \mathbb{R} \rightarrow \mathbb{Z} : x \mapsto \lfloor x \rfloor$$

où $\lfloor x \rfloor$ est défini comme le plus grand entier k tel que $k \leq x$;

- * la *fonction caractéristique* de l'ensemble $\mathbb{Q} \subset \mathbb{R}$

$$j : \mathbb{R} \rightarrow \{0, 1\} : x \mapsto \begin{cases} 1 & \text{si } x \in \mathbb{Q} \\ 0 & \text{si } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

(Ce sera probablement un de nos exemples les plus pathologiques — mais il en faut de temps en temps).

Dans tous nos exemples ci-dessus, les ensembles de départ et d'arrivée, A et B , sont des sous-ensembles de $\mathbb{R}$. On parle alors de *fonction d'une variable réelle à valeurs dans les réels*, ou simplement de *fonction réelle*. Au chapitre XIV, nous étudierons des fonctions dont l'ensemble d'arrivée est $\mathbb{R}^2$ voire $\mathbb{R}^3$. Dans le chapitre XVI, nous étudierons des fonctions de plusieurs variables, dont l'ensemble de départ est $\mathbb{R}^2$ ou $\mathbb{R}^3$.

Toutes les fonctions ne concernent pas forcément des nombres.

Exemple. Soit A l'ensemble des étudiants à l'ULB, soit B l'ensemble des nombres naturels. On associe, à chaque étudiant, son numéro matricule. Chaque étudiant à l'ULB possède un tel numéro matricule, cela définit donc une fonction !

Exemple. La notation

$$f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \frac{1}{x}$$

ne définit *pas* une fonction au sens précédent, car l'image de 0 n'a pas été définie. Par contre,

$$g : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{1}{x} & \text{si } x \neq 0 \\ 0 & \text{sinon.} \end{cases}$$

définit correctement une fonction.

fonction

ensemble de
départ

ensemble
d'arrivée

image

antécédent

plancher

fonction
caractéristique

fonction réelle

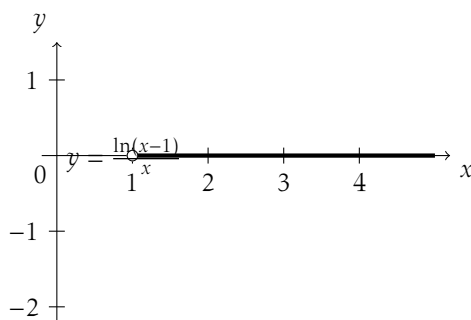


FIGURE IV1 – Le graphe de $(1, +\infty) \subset \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \frac{\ln(x-1)}{x}$.

Exercice. Comment choisir le domaine $A \subset \mathbb{R}$ pour que

$$f : A \rightarrow \mathbb{R} : x \mapsto \frac{1}{x^2 - 1}$$

définisse une fonction ?

Réponse : n'importe quel A ne contenant ni 1 ni -1 convient. Généralement, nous prenons « le plus grand possible », c'est-à-dire $A = \mathbb{R} \setminus \{-1, 1\}$.

Remarque (*). Cet exercice est exactement un problème de recherche des conditions d'existence, bien connu des étudiants sortant de l'enseignement secondaire.

Remarque (*). La règle qui donne $f(x)$ est souvent une simple formule (également appelée *expression algébrique*) contenant la variable x . Parfois c'est une recette plus complexe, comme dans le cas de la valeur absolue ou celui de la fonction g dans l'exemple ci-dessus. Parfois il n'y a pas du tout de formule, comme dans le cas de l'association de son matricule à chaque étudiant.

Remarque (*). Lorsqu'on demande d'exprimer une quantité « en fonction » d'une autre, la question est souvent ambiguë mais la notion de fonction mathématique n'est pas loin.

Exemple. * Exprimer $\log_6(5)$ en fonction de $\log_6 2$

1 Domaine et image

Pour une fonction $f : A \rightarrow B$, on appelle également *domaine* de f l'ensemble de départ A . Il est noté $\text{dom } f$.

Nous avons déjà défini l'image d'un élément x de A : c'est $f(x)$. Si on considère l'ensemble des images des points de A par f , l'ensemble obtenu est quant à lui appelé *ensemble image* (ou simplement *image*) de f , et il est noté $f(A)$ ou $\text{Im } f$. Donc

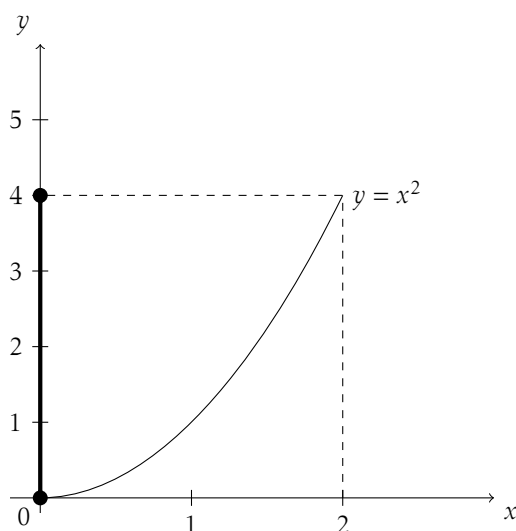
$$\text{Im } f = f(A) := \{f(x) \text{ t.q. } x \in A\}$$

On a toujours $f(A) \subset B$ mais pas forcément $f(A) = B$.

Exemple. L'image de la fonction $f : [0, 2] \rightarrow \mathbb{R} : x \mapsto x^2$ est $[0, 4]$; c'est bien un sous-ensemble de la droite réelle. Cet ensemble est représenté en trait gras sur la figure IV2.

2 Interprétation de la notion de fonction

La notion de fonction n'est pas introduite pour le plus pur plaisir du mathématicien, mais parce qu'elle correspond à une formalisation de diverses notions intuitives intéressantes. Relevons-en deux : les notions d'association (d'objets/valeurs à d'autres objets/valeurs) et de transformation (d'un ensemble géométrique).

FIGURE IV2 – Le graphe de $[0, 2] \rightarrow \mathbb{R} : x \mapsto x^2$.

2.1 Association

Dans cette vision de la notion de fonction, A représente un domaine comportant certains objets (physiques ou idéalisés), et B représente un ensemble de valeurs ou d'objets potentiellement associées aux objets du domaines. La fonction réalise l'association : à l'objet $x \in A$, la fonction f associe une valeur ou un objet noté $f(x)$.

Exemple. * Pour chaque point d'une pièce, on peut associer à la température (en degrés Celsius) de l'air en ce point. Dans ce cas, A est l'ensemble des points de la pièce, et B est l'ensemble des réels. La fonction prend un point p de la pièce et lui associe la température $f(p)$.

- * Pour chaque personne vivant sur Terre, on peut associer son âge (en nombre d'années) à un moment fixé. A est la population mondiale au moment fixé, B est l'ensemble des réels (on peut même se restreindre aux réels positifs.)
- * Pour chaque personne inscrite sur Facebook, on peut lui associer son nombre *d'amis facebook*. Ici le domaine A est un sous-ensemble de la population mondiale (à savoir uniquement les inscrits sur Facebook ; oublions les faux-comptes et les doublons), tandis que l'ensemble d'arrivée B est a priori $\mathbb{N}$, l'ensemble des entiers. (On pourrait dire $\mathbb{R}$, mais on sait bien que seuls les entiers seront utiles).

2.2 Transformation

Dans cette seconde vision des fonctions, on imagine que le domaine est « déplacé » dans l'ensemble d'arrivée.

Exemple.

La fonction $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto -x$ correspond à « retourner » la droite réelle. C'est la symétrie centrale de centre 0.

La fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 : (x, y) \mapsto (-x, y)$ correspond à « retourner » le plan par une symétrie orthogonale. Le « miroir » est la droite des ordonnées.

Pour la fonction $f : \mathbb{R} \rightarrow \mathbb{R}^2 : x \mapsto (x, x)$, on imagine que la droite réelle $\mathbb{R}$ est « envoyée » sur une droite du plan (à savoir la première bissectrice).

Pour la fonction $\mathbb{R} \rightarrow \mathbb{R} : x \mapsto |x|$, on imagine plutôt qu'on plie la droite en deux, en recollant la partie négative sur la partie positive.

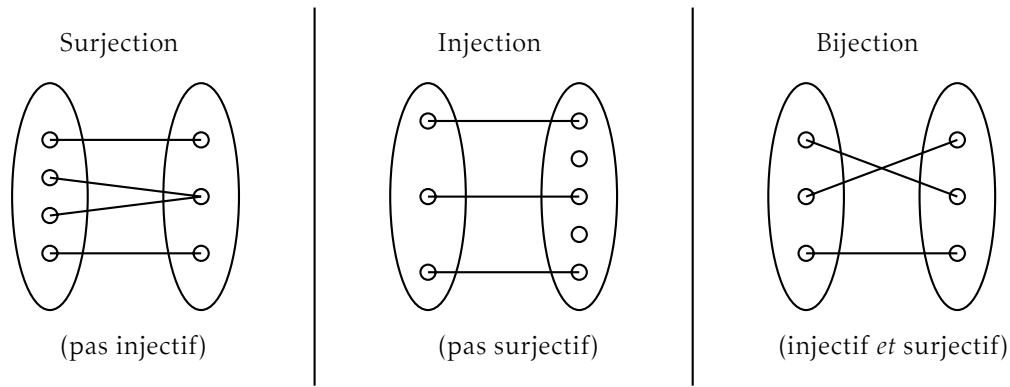


FIGURE IV3 – Injectivité, surjectivité.

3 Surjections, injections, bijections et réciproque

Comme rappelé plus haut, la fonction $f : A \rightarrow B$ attribue à tout élément x de A une unique image $f(x)$ dans B . Il se peut très bien qu'un élément de B ne soit image d'aucun élément de A , ou bien de plus d'un élément de A . Dans une certaine mesure, ces situations sont vues comme problématiques. On dit que

- * f est *surjective* si tout élément de B est l'image d'au moins un élément de A : pour tout $y \in B$ il existe $x \in A$ tel que $f(x) = y$, ce qui peut également s'exprimer $f(A) = B$;
- * f est *injective* si deux éléments distincts de A ont des images distinctes dans B : pour tout $x, x' \in A$ avec $x \neq x'$, on a $f(x) \neq f(x')$;
- * f est *bijjective* si f est à la fois injective et surjective.

Une fonction surjective s'appelle aussi une *surjection*. Très logiquement, une *injection* est une fonction injective et une *bijection* est une fonction bijective.

Quand $f : A \rightarrow B$ est bijective, chaque élément de B est l'image d'un unique élément de A et on peut définir la fonction *réciproque* $f^{-1} : B \rightarrow A$ par $f^{-1}(y) = x$, où x est l'antécédent (unique, par bijectivité) de y , c'est-à-dire tel que $f(x) = y$. On a donc

$$f^{-1}(y) = x \iff f(x) = y.$$

Exemple. Attention à ne pas confondre $f^{-1}(x)$ avec $f(x)^{-1} := 1/f(x)$! Par exemple, les fonctions

$$f : \mathbb{R}_0 \rightarrow \mathbb{R}_0 : x \mapsto x \quad \text{et} \quad g : \mathbb{R}_0 \rightarrow \mathbb{R}_0 : x \mapsto \frac{1}{x}$$

vérifient $f(x)^{-1} = g(x)$ pour tout $x \in \mathbb{R}_0$, par contre $f^{-1} = f$ et $g^{-1} = g$ (exercice facile).

Par exemple, la fonction

$$g : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^2$$

n'est ni surjective (car -1 est un élément de l'ensemble d'arrivée mais n'est le carré d'aucun réel), ni injective (car 4 est le carré de deux réels : 2 et -2). Mais si on change les ensembles de départ et d'arrivée, on obtient une fonction $\mathbb{R}_+ \rightarrow \mathbb{R}_+ : x \mapsto x^2$ qui est *bijjective*, et dont la fonction réciproque est $\mathbb{R}_+ \rightarrow \mathbb{R}_+ : x \mapsto \sqrt{x}$. Enfin, la fonction

$$i : \mathbb{R} \rightarrow \mathbb{Z} : x \mapsto \lfloor x \rfloor$$

est surjective (car tout entier k est son propre plancher) mais pas injective (car $0 = \lfloor 0 \rfloor = \lfloor 0,17 \rfloor = \lfloor 0,999 \rfloor$).

Soit une fonction $f : A \rightarrow B$. Nous définissons et illustrons (voir figure IV3) les notions d'injectivité, de surjectivité, de bijectivité et d'inversibilité.

3.1 Injection

La fonction f dite *injective* lorsque tout élément de l'ensemble d'arrivée de f a *au plus* (c'est-à-dire au maximum) un antécédent par f . Une telle fonction est appelée une *injection*.

C'est équivalent à dire que si $f(x) = f(y)$ pour un x et un y de A , alors forcément $x = y$.

3.2 Surjection

La fonction f est dite *surjective* lorsque tout élément de l'ensemble d'arrivée est image par f d'*au moins* un élément de l'ensemble de départ. En d'autres termes, f est surjective si et seulement si son image est l'ensemble d'arrivée tout entier. Une telle fonction est appelée une *surjection*.

surjective

C'est équivalent à dire que pour $a \in B$, l'équation $f(x) = a$ possède toujours une solution x .

3.3 Bijection

La fonction f est dite *bijjective* (ou est une *bijection*) si elle est à la fois injective et surjective.

bijjective

Ceci veut dire que les éléments du domaine et les éléments de l'ensemble d'arrivées se correspondent parfaitement par f . On pourrait dire qu'on a simplement renommé les éléments!

3.4 Inversibilité des fonctions

Une fonction $f : A \rightarrow B$ est *inversible* si il existe une fonction $g : B \rightarrow A$ telle que

inversible

$$g(f(x)) = x \quad \forall x \in A \quad \text{et} \quad f(g(y)) = y \quad \forall y \in B.$$

Cette définition peut en fait se raccrocher à la notion de fonction réciproque vue ci-dessus.

Résultat IV.2. 1. Une fonction est inversible si et seulement si c'est une bijection. Dans ce cas, l'inverse de f est f^{-1} .

2. Si g est l'inverse de f , alors f est l'inverse de g . En d'autres termes,

$$(f^{-1})^{-1} = f.$$

4 Graphes

Si $f : A \rightarrow B$ est une fonction, le *graphe* de f est le sous-ensemble de $A \times B$ défini par l'équation $y = f(x)$ (pour $x \in A$). En d'autres termes, le graphe de f est l'ensemble

graphe

$$\Gamma_f := \{(x, f(x)) \in A \times B \text{ t.q. } x \in \text{dom } f\}$$

Remarque (*). Cette définition, très formelle, permet de revenir à la notion de « graphe » telle que le lecteur la connaît sans doute : un dessin ! Nous verrons cela plus loin.

5 Restrictions

Si $f : A \rightarrow B$ est une fonction et $A' \subset A$, alors f définit une fonction sur A' appelée *restriction* de f à A' et notée $f|_{A'}$. On a donc

restriction

$$f|_{A'} : A' \rightarrow B : x \mapsto f(x)$$

La seule différence avec f est donc un domaine plus petit.

6 Composées

Si $f : A \rightarrow B$ et $g : C \rightarrow D$, avec $\text{Im } f \subset \text{dom } g$, alors on peut former $g \circ f : A \rightarrow D : x \mapsto g(f(x))$.

Remarque. Attention à l'ordre des opérations : $g \circ f$ se lit « g rond f » mais à x l'on applique d'abord f puis g .

7 Fonction identité

La fonction $A \rightarrow A : x \mapsto x$ est appelée l'*identité* sur A , et se note souvent Id_A . Donc $\text{Id}_A(x) = x$ pour tout $x \in A$.

identité

Remarque (*). Si $f : A \rightarrow B$ est une bijection, on a $f \circ f^{-1} = \text{Id}_B$ et $f^{-1} \circ f = \text{Id}_A$.

8 Antécédent

antécédent

Si x , élément de A , vérifie $f(x) = y$, on dit que x est un *antécédent* de y (pour la fonction f). Un élément y de B peut très bien avoir plusieurs antécédents ou n'en avoir aucun.

Exemple. Les antécédents de 4 par la fonction $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^2$ sont -2 et 2 . L'unique antécédent de 0 est 0. Par contre -4 n'a aucun antécédent pour cette fonction.

9 Parité

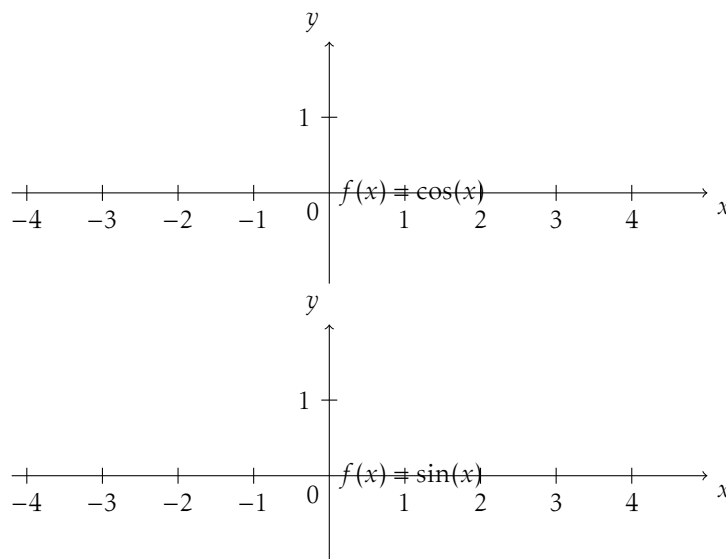
paire

Une fonction $f : A \rightarrow B$, avec $A \subset \mathbb{R}$ et $B \subset \mathbb{R}$ est :

impaire

- * *paire* si et seulement si pour tout x de A , on a $-x \in A$ et $f(x) = f(-x)$.
- * *impaire* si et seulement si pour tout x de A , on a $-x \in A$ et $f(-x) = -f(x)$.

Exemple. La fonction $\cos(x)$ est une fonction paire et la fonction $\sin(x)$ est une fonction impaire, mais $\ln(x-1)/x$ (voir figure IV1) n'est ni l'un ni l'autre.



Résultat IV.3. Si f est une fonction paire, son graphe est symétrique par rapport à l'axe des ordonnées (symétrie bilatère). Si f est impaire et si le graphe est dessiné dans les coordonnées cartésiennes d'un repère orthonormé, le graphe de f est symétrique par rapport à l'origine (symétrie centrale).

10 Quelques familles de fonctions

10.1 Fonctions linéaires, affines et polynomiales

fonction linéaire

Les fonctions linéaires et affines sont généralement bien connues car à la fois très simples et très importantes : les graphes de ces fonctions représentent des droites.

coefficient angulaire

Une *fonction linéaire* est, dans notre contexte, une fonction de la forme $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto ax$ pour un certain réel a . Ce nombre réel a s'appelle le *coefficient angulaire* ou la *pente*. On peut interpréter ce nombre comme suit : si (x_0, y_0) et (x_1, y_1) sont deux points distincts du graphe de f , alors $a = \frac{y_1 - y_0}{x_1 - x_0}$.

pente

Démonstration. Nous allons, dans un instant, prouver cette dernière égalité. Néanmoins il serait plus profitable au lecteur qu'il vérifie cette égalité par lui-même, et ne s'en réfère au calcul ci-dessous que si cela est nécessaire (ou pour vérification, par curiosité).

Puisque les points sont distincts on a $x_1 \neq x_0$, et par ailleurs puisqu'ils sont sur le graphe, on a $y_0 = f(x_0)$ et $y_1 = f(x_1)$, dès lors le quotient suivant a un sens et se calcule :

$$\begin{aligned} \frac{y_1 - y_0}{x_1 - x_0} &= \frac{f(x_1) - f(x_0)}{x_1 - x_0} = \frac{(ax_1 + b) - (ax_0 + b)}{x_1 - x_0} \\ &= \frac{ax_1 + b - ax_0 - b}{x_1 - x_0} = \frac{ax_1 - ax_0}{x_1 - x_0} = \frac{a(x_1 - x_0)}{x_1 - x_0} = a \quad (10.1) \end{aligned}$$

ce que nous avions annoncé. $\square$

Plus généralement, une *fonction affine* est de la forme $\mathbb{R} \rightarrow \mathbb{R} : x \mapsto ax + b$ pour certains réels a et b . Le nombre a s'appelle encore coefficient angulaire et possède la même interprétation que précédemment. Le nombre b s'appelle *ordonnée à l'origine* et est la valeur de la fonction lorsque la variable $x = 0$ (c'est évident, mais faites-le!).

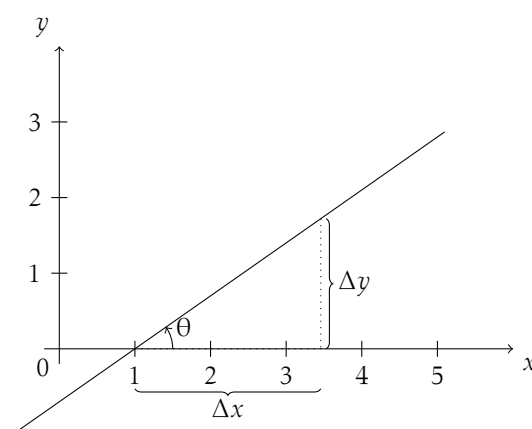
fonction affine

ordonnée à l'origine

10.1.1 Graphe

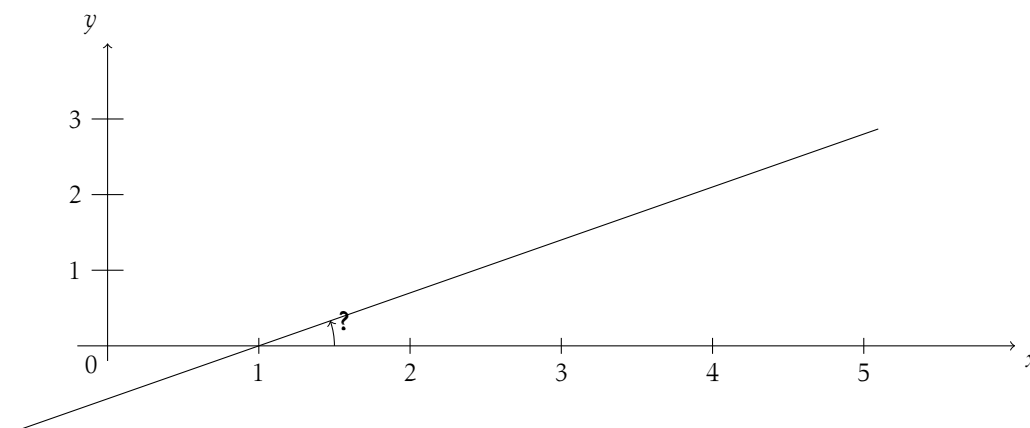
Le graphe de toute fonction affine est une droite, et si la fonction est en fait linéaire (c'est-à-dire $b = 0$) alors la droite passe par l'origine.

Lorsque le graphe d'une fonction affine est dessiné en coordonnées cartésiennes d'un repère orthonormé (c'est-à-dire que les axes des abscisses et des ordonnées sont orthogonaux et gradués de la même manière), alors le coefficient angulaire possède l'interprétation suivante : si θ est la mesure (avec son signe) de l'angle entre le demi-axe des abscisses positives et le graphe de la fonction au dessus de ce demi-axe, alors $a = \tan \theta$. Ceci est bien sûr incompréhensible sans un dessin :



Observez que les relations usuelles dans un triangle rectangle (voir aussi la section 8.1) impliquent en effet que $\tan \theta = \frac{\Delta y}{\Delta x}$, et que ce dernier quotient est égal au coefficient angulaire d'après l'équation (??).

Remarquons que cette interprétation ne tient plus si les axes ne sont pas gradués à l'identique :



Exercice. Déterminer l'angle approximatif dessiné, sachant que θ valait 35° (dans la figure précédente) et que la nouvelle figure a été obtenue en doublant la valeur de l'unité sur l'axe des abscisses par rapport au graphe précédent. (calculatrice autorisée).

10.1.2 Fonctions polynomiales

Les *fonctions polynomiales* sont des fonctions de la forme $x \mapsto a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ pour certaines constantes réelles $a_0, \dots, a_n$ et un certain entier n qu'on appelle *degré* de la fonction (si $a_n \neq 0$). Une fonction affine dont le coefficient angulaire est non-nul est donc une fonction polynomiale de degré 1.

Rappelons qu'une droite est caractérisée par deux points distincts. En particulier, si on choisit deux points dont les abscisses sont distinctes, on peut imaginer –et c'est vrai– qu'il existe une unique fonction affine dont le graphe passe par ces deux points. Il est nécessaire que les abscisses soient distinctes sinon la droite serait verticale, ce qui ne peut pas être le graphe d'une fonction.

Cette propriété se généralise aux polynômes comme suit : se donnant $n + 1$ points dont les abscisses sont distinctes deux-à-deux (c'est-à-dire qu'aucun point ne se trouve « au dessus » d'un autre), il existe une unique fonction polynomiale de degré au plus n (c'est-à-dire de degré n ou moins que n) dont le graphe passe par les $n + 1$ points.

10.2 Fonctions racines

Les fonctions *racine carrée* et *racine cubique* sont :

$$\mathbb{R}^+ \rightarrow \mathbb{R}^+ : x \mapsto \sqrt{x} \quad \text{et} \quad \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \sqrt[3]{x}$$

Plus généralement, pour n naturel pair, nous avons une fonction racine n^e :

$$\mathbb{R}^+ \rightarrow \mathbb{R}^+ : x \mapsto \sqrt[n]{x}$$

et pour n naturel impair, similairement avec un autre domaine :

$$\mathbb{R} \rightarrow \mathbb{R} : x \mapsto \sqrt[n]{x}.$$

Voir aussi la section 3.6 page 31 pour la définition de ces fonctions.

10.3 Les fonctions exponentielles et logarithmes

10.3.1 Exponentielles

Les fonction exponentielles sont les fonctions du type $f(x) = a^x$ pour un certain réel positif $a > 0$, appelé la *base* de l'exponentielle. Lorsqu'on parle de la fonction exponentielle, c'est la fonction exponentielle dont la base est $a = e$, où $e \simeq 2.7182818\dots$ est une constante appelée... « le nombre e . »

Le domaine des fonctions exponentielles est $\mathbb{R}$, leur image est l'ensemble $\mathbb{R}_0^+ =]0, \infty]$. Remarquons que lorsque base est inférieure à 1, l'exponentielle dans cette base est décroissante, alors que pour une base plus grande que 1, l'exponentielle est croissante.

La fonction exponentielle a pour propriété remarquable d'être sa propre dérivée :

$$\frac{d}{dx} \exp x = \exp x.$$

Cette propriété en fait une fonction omniprésente dans la description de nombreux phénomènes naturels, à chaque fois qu'une quantité varie proportionnellement à sa valeur instantanée. C'est une raison pour laquelle on parle de la fonction exponentielle.

10.3.2 Logarithmes et logarithme naturel

Les fonctions logarithmes sont les fonctions réciproques des fonctions exponentielles. Lorsque $a > 0$ et $a \neq 1$, la fonction exponentielle $x \mapsto a^x$ est une bijection. Sa fonction réciproque est le *logarithme en base a* , il se note $\log_a$.

Exemple. Puisque $10^2 = 100$, il est vrai que $\log_{10}(100) = 2$.

fonctions
polynomiales

degré

base

logarithme en
base a

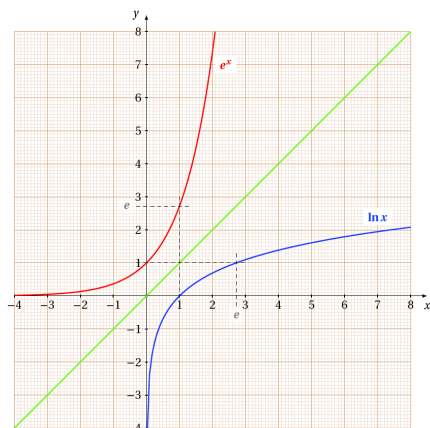


FIGURE IV4 – Graphe de $\exp x$ et de $\ln x$. Ces deux graphes sont symétriques l'un de l'autre par rapport à la droite d'équation $y = x$.

Il y a quelques fonctions logarithmes fort utilisées :

- * le logarithme décimal, en base 10, souvent noté simplement $\log$,
- * le logarithme en base e , souvent noté $\ln$, et
- * le logarithme en base 2, noté $\log_2$

sont les plus utilisés. C'est ce dernier qui est le plus souvent utilisé en mathématiques et en physique, il est appelé *logarithme naturel*, ou *logarithme népérien* (hommage à l'écossais JOHN NAPIER).

Les fonctions logarithmes sont définies sur le domaine $\mathbb{R}_0^+$ et ont pour ensemble image $\mathbb{R}$.

logarithme
naturel

logarithme
népérien

10.3.3 Identités importantes

Les identités suivantes sont essentielles. Ici, $a, b > 0$ et $x, y \in \mathbb{R}$.

- i. $\exp(\ln a) = a$ et $\ln(\exp x) = x$,
- ii. $\exp 0 = 1$ et $\ln 1 = 0$,
- iii. $\exp(x + y) = \exp(x)\exp(y)$ et $\ln(ab) = \ln(a) + \ln(b)$,
- iv. $\ln(a^x) = x \ln a$
- v. $\log_a(b) = \frac{\ln(b)}{\ln(a)}$.

Chapitre V

Systèmes de coordonnées

Définition V.1. Un *système de coordonnées* sur un ensemble est la donnée d'une bijection entre cet ensemble et une partie de $\mathbb{R}^n$ pour une certaine valeur de n .

système de coordonnées

Ne nous laissons pas impressionner par cette définition, et décrivons quelques systèmes de coordonnées utiles.

1 Sur une droite

Un *système de coordonnées* sur une droite est la donnée d'une bijection entre cette droite et $\mathbb{R}$, ou une partie de $\mathbb{R}$ (généralement un intervalle). L'unique réel associé à un point donné est appelé sa *coordonnée*¹ dans ce système de coordonnées.

système de coordonnées

coordonnée

Nous voyons maintenant deux systèmes de coordonnées classiques, mais il en existe tant qu'on veut!

1.1 Coordonnées cartésiennes

Pour introduire un système de coordonnées cartésiennes, il faut choisir une *origine* et un point *unité* sur la droite, correspondant aux réels 0 et 1 respectivement. Ensuite, tout autre réel x est placé de sorte que sa distance par rapport à l'origine soit égale à $|x|$ (la e x) multiplié par la distance entre 0 et 1, et du même côté de l'origine que 1 si x est positif, ou de l'autre côté si x est négatif. Cette description est probablement incompréhensible sans « l'essayer » : faites-le ! Ceci, bien que compliqué à décrire, est en fait très simple et naturel. C'est le système de coordonnées classique, et celui que nous utiliserons pour l'essentiel du temps.

origine

unité

1.2 Coordonnées logarithmiques

Soit a un réel strictement positif, différent de 1 ($a > 0$ et $a \neq 1$). Un système de *coordonnées logarithmiques* en base a se construit comme suit : il faut choisir une origine (qui représente 1) et un point placé à une certaine distance de l'origine (ce point représente a) et chaque déplacement d'une unité multiplie ou divise par a , selon la direction (il y a en effet deux possibilités : identique à la direction allant de 1 à a , ou opposée à celle-là).

coordonnées logarithmiques

Un système de coordonnées logarithmiques fournit une bijection entre la droite et $\mathbb{R}_0^+$: seuls les réels strictement positifs sont placés sur la droite de la sorte. Ce genre de système est pratique pour représenter des ordres de grandeurs en physique.

2 Sur un plan

Un système de coordonnées sur un plan est la donnée d'une bijection entre ce plan et (une partie de) $\mathbb{R}^2$.

1. Si le point P d'une droite a pour coordonnée x , nous dirons informellement qu'on a « placé » x en P .

2.1 Coordonnées induites par des droites

Une technique assez générale pour obtenir un système de coordonnées sur le plan est de choisir deux droites sécantes munies chacune d'un système de coordonnées (au sens de la section précédente), puis de procéder par projections sur ces deux droites pour obtenir deux coordonnées. Cependant nous aurons rarement besoin de tant de généralité, aussi nous illustrons ce principe sur le cas particulier des coordonnées cartésiennes du plan (qui généralisent les coordonnées cartésiennes de la droite).

2.2 Coordonnées cartésiennes du plan

Des coordonnées cartésiennes du plan sont définies à l'aide du choix de deux droites sécantes : la première est appelée « axe des abscisses », la seconde est « l'axe des ordonnées ». Ces droites sont sécantes en un point appelé *origine*. Sur chacune de ces droites, on choisit de plus un point « unité », qui sert à créer un système de coordonnées cartésiennes sur cette droite. La donnée de l'origine et des deux points « unité » forme un *repère cartésien*. Si les droites sont orthogonales, on dit que le repère est *orthogonal* ; si en plus les points « unité » sont à la même distance de l'origine l'un et l'autre, on parle de *repère orthonormé*.

Ayant un tel repère, nous pouvons maintenant *repérer* un point P en lui attribuant des coordonnées (x, y) par projections comme décrit ci-après. D'abord, on projette P sur l'axe des abscisses (parallèlement à l'axe des ordonnées). Sur cet axe, nous avons une origine et une unité, donc nous pouvons associer une coordonnée cartésienne à ce point. Cette coordonnée, nous l'appelons l'*abscisse* de P et on la note souvent P_1 , ou P_x ou simplement x (mais bien sûr, n'importe quel nom non-ambigu conviendra!). De même, on projette P sur l'axe des ordonnées (parallèlement à l'axe des abscisses), et le point obtenu possède une coordonnée appelée *ordonnée* de P , notée souvent P_2 ou P_y ou y . Le couple (P_1, P_2) sont les *coordonnées cartésiennes* de P .

Remarque. L'ensemble des points (x, y) vérifiant $x = 1$ est une droite verticale à distance 1 de l'axe des ordonnées. Bien sûr en remplaçant 1 par toute autre constante réelle a , l'effet est similaire.

2.2.1 Distance

Résultat V.2. Soient deux points $\mathbf{p}_1$ et $\mathbf{p}_2$ de coordonnées respectives (x_1, y_1) et (x_2, y_2) . La distance entre ces points est donnée par le théorème de Pythagore :

$$d(\mathbf{p}_1, \mathbf{p}_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

2.3 Coordonnées polaires

On va établir une bijection (enfin, presque) entre $\mathbb{R}^+ \times [0; 2\pi)$ et le plan. Notons O un origine fixé du plan, choisissons une demi-droite issue de O (en général on la prend horizontale partant vers la droite de O).

Au point P , on associe d'une part P_r la distance entre P et l'origine, et d'autre part P_θ la mesure (en radians) de l'angle entre la demi droite du repère et la demi-droite issue de l'origine O et passant par P . La mesure P_θ n'est autre que la longueur de l'arc de cercle intercepté par cet angle, divisée P_r .

Exercice. Pourquoi ceci n'est-il pas une bijection comme annoncé? (Indice : quelles sont les coordonnées de l'origine?) En fait, il s'agit d'une bijection entre $\mathbb{R}_0^+ \times [0; 2\pi[$ et le plan privé de son origine.

Exercice. Marquer l'ensemble des points vérifiant l'équation $r = 1$ en interprétant d'abord (r, t) comme des coordonnées polaires, puis comme des coordonnées cartésiennes.

origine

repère cartésien

orthogonal

repère orthonormé

abscisse

ordonnée

coordonnées cartésiennes

2.4 Lien entre les systèmes de coordonnées

Si P est un point, on peut lui associer un couple coordonnées dans plusieurs systèmes de coordonnées. Dès lors on peut créer une bijection : $f : A \subset \mathbb{R}^2 \rightarrow B \subset \mathbb{R}^2 : (x, y) \mapsto F(x, y) = (F_1(x, y), F_2(x, y))$ qui à des coordonnées (x, y) dans le premier système associe les coordonnées du même point dans le second système.

Cela donnera lieu à l'étude des fonctions de **plusieurs variables** à valeurs **vectérielles**, i.e. des fonctions $A \subset \mathbb{R}^n \rightarrow \mathbb{R}^k$.

3 Graphes de fonctions

On peut représenter une fonction réelle à condition d'avoir préalablement choisi un système de coordonnées du plan. En général, le système choisi est le système cartésien orthonormé.

Pour dessiner le graphe d'une fonction réelle f , on représente (noircit, colorie, ou tout autre moyen qui aura votre faveur) chaque point (dont les coordonnées sont) de la forme $(x, f(x))$. En d'autres termes, on représente tout point dont les coordonnées (x, y) vérifient $y = f(x)$. Par définition d'une fonction, pour x donné, il y a un seul tel y .

Par exemple, voici le graphe de la fonction valeur absolue, dans deux systèmes de coordonnées différents :

$$h : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto |x|$$

Le graphe est un outil capital pour appréhender les propriétés qualitatives (par exemple, continuité et dérivabilité) et quantitatives (par exemple, valeurs des limites et dérivées, asymptotes) d'une fonction réelle d'une variable réelle. Cela reste le cas pour les fonctions réelles de deux variables réelles. L'accès au graphe est moins aisée pour les autres types de fonction, en particulier pour les fonctions vectorielles et les fonctions de plus de deux variables que nous rencontrerons plus tard. C'est une bonne chose que de raisonner sans systématiquement faire appel au graphe.

Graphe de la réciproque Si $f : A \rightarrow B$ est bijective, alors le graphe de sa réciproque $f^{-1} : B \rightarrow A$ peut s'obtenir en prenant l'image du graphe de f par une symétrie bilatérale d'axe $x = y$. En effet,

$$\Gamma_{f^{-1}} = \{(y, f^{-1}(y)) \in \mathbb{R}^2 \text{ t.q. } y \in B\} = \{(f(x), x) \in \mathbb{R}^2 \text{ t.q. } x \in A\}$$

Chapitre VI

Trigonométrie

La trigonométrie c'est l'étude des triangles. Plus spécifiquement, l'étude des liens entre la mesure des angles et la mesure des côtés. Avant d'étudier les liens, il faut savoir mesurer, et si la mesure des côtés ne pose aucun problème particulier, mesurer les angles est déjà une activité plus abstraite.

1 La notion d'angle : généralités

Mesurer un angle, c'est mesurer l'idée intuitive qu'on peut faire un « tour complet », mais aussi un demi-tour, un quart de tour, ou toute autre fraction de tour. La première notion de mesure d'un angle pourrait donc être simplement un nombre réel, décrivant le nombre de tours. Ce serait tout à fait envisageable, mais ça n'est pas l'unité qui est la plus communément utilisée.

En fait, il y a deux sous-unités classiques : le degré et le radian. De la même manière qu'un centimètre est le centième d'un mètre, on définit le degré et le radian en divisant l'unité « tour complet » en parties égales : un tour vaut « 360 degrés », ou encore « 2π radians ». On a donc $360^\circ = 2\pi \text{ rad}$.

En résumé, une *mesure* d'un angle est un réel θ entre 0 et 360 (en degrés), ou un réel entre 0 et 2π (en radians). Un tel réel s'appelle la *détermination principale* de l'angle considéré.

Si cela est la détermination principale, c'est qu'il y en a d'autres. Et en effet : s'il est possible de faire un demi tour, il est également possible de faire un tour et demi, et se retrouver dans la même position. Dès lors, 180° est « équivalent » à 540° ($540 = 180 + 360$). La détermination principale, c'est simplement enlever assez de tours complets pour ramener la mesure de l'angle entre 0° et 360° , ou entre 0 rad et $2\pi \text{ rad}$.

Un dernier point est qu'il est possible de faire un quart de tour dans un sens, ou dans l'autre. Ceci introduit une notion d'*orientation* d'un angle. Dans le plan, un angle mesuré dans le sens horlogique est négatif, tandis que mesuré dans le sens anti-horlogique est positif. Afin de rajouter à la confusion, on appelle « sens trigonométrique » le sens anti-horlogique.

mesure

détermination
principale

orientation

2 Le radian

Le fait qu'il y ait 360 degrés pour un tour complet semble purement historique, par contre la valeur 2π radians peut s'interpréter géométriquement. Un *radian* (1 rad) correspond à l'angle au centre d'un cercle qui intercepte, sur la circonférence, un arc dont la longueur est égale au rayon du cercle (voir figure VII1, à droite).

radian

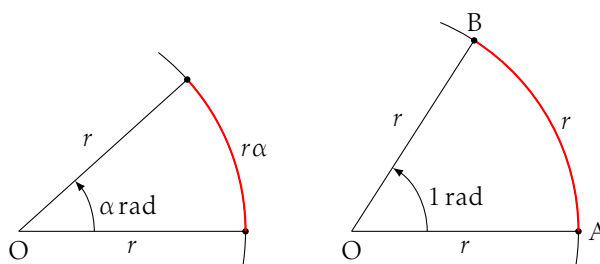


FIGURE VI1 – Angle de α radian (à gauche), et le cas particulier de l'angle de 1 radian (à droite). Dans le dessin de droite, les segments OA et OB sont de longueurs égales entre elles et égales à la longueur de l'arc de cercle AB.

En particulier, un angle de α rad intercepte un arc de longueur αr , si r est le rayon du cercle (à gauche, sur la figure); dès lors un angle de 2π intercepte un arc de longueur $2\pi r$, soit le cercle entier : cet angle correspond à un tour complet.

Le radian est donc une unité de choix dès qu'il faut faire intervenir des longueurs interceptées par des angles. Il s'avère également l'unité adéquate pour manipuler les fonctions trigonométriques (cosinus, sinus).

Pour passer du radian au degré et inversement, retenons que $2\pi \text{ rad} = 360^\circ$. L'habitude nous fera également retenir le tableau suivant :

en radians	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	π	2π
en degrés	30°	45°	60°	90°	180°	360°

3 Fonctions trigonométriques dans le cercle

Un *cercle trigonométrique* est un cercle de rayon 1, il permet de représenter géométriquement toutes les fonctions trigonométriques. Ces fonctions sont les fonctions sinus, cosinus, tangente, cotangente, sécante, cosécante et quelques autres plus rarement utilisées. Nous allons maintenant définir ces fonctions de manière géométrique, en mesurant des longueurs de côtés dans des triangles construits dans un cercle trigonométrique.

Pour fixer les idées, imaginons un système de coordonnées cartésiennes orthonormé dont l'origine est au centre du cercle. On considère un réel θ (un angle en radians, qu'on va supposer entre 0 et 2π pour simplifier la présentation), et un point P sur le cercle, tel que l'arc de cercle partant de (1,0) allant (dans le sens anti-horlogique) jusque P a pour longueur θ .

Le point P possède deux coordonnées cartésiennes, P_x et P_y . On définit le *cosinus* de θ par $\cos(\theta) := P_x$, et son *sinus* par $\sin(\theta) := P_y$. Représentons la situation :

FIGURE VI2 – Définition du sinus et du cosinus.

À partir de ces deux fonctions, sinus et cosinus, nous pouvons définir les fonctions trigonométriques usuelles, à savoir :

$$\begin{aligned} \tan(\theta) &= \frac{\sin(\theta)}{\cos(\theta)} & \sec(\theta) &= \frac{1}{\cos(\theta)} \\ \cot(\theta) &= \frac{\cos(\theta)}{\sin(\theta)} & \operatorname{cosec}(\theta) &= \frac{1}{\sin(\theta)} \end{aligned}$$

L'ensemble de ces fonctions peut se représenter dans le cercle trigonométrique (voir figure VI3).

4 Valeurs importantes

Récapitulons les domaines et ensembles images de quelques-unes de ces fonctions trigonométriques :

$$\begin{aligned} \sin: \mathbb{R} &\rightarrow [-1, 1] & \tan: \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi \text{ t.q. } k \in \mathbb{Z} \right\} &\rightarrow \mathbb{R} \\ \cos: \mathbb{R} &\rightarrow [-1, 1] & \cot: \mathbb{R} \setminus \{k\pi \text{ t.q. } k \in \mathbb{Z}\} &\rightarrow \mathbb{R}. \end{aligned}$$

cercle
trigonométrique

cosinus

sinus

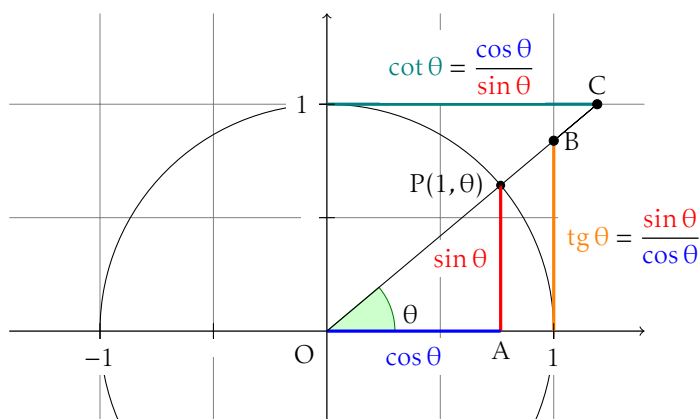


FIGURE VI3 – Illustration des fonctions sinus, cosinus, tangente et cotangente sur le cercle trigonométrique (ici, pour l'angle $\theta = 2\pi/9$ soit 40°).

et voici quelques valeurs importantes des fonctions sinus et cosinus :

angle	sin	cos	tan	cot
0	$\sqrt{0}/2 = 0$	1	0	$\nexists$
$\pi/6$	$\sqrt{1}/2 = 1/2$	$\sqrt{3}/2$	$\sqrt{3}/3$	$\sqrt{3}$
$\pi/4$	$\sqrt{2}/2$	$\sqrt{2}/2$	1	1
$\pi/3$	$\sqrt{3}/2$	1/2	$\sqrt{3}$	$\sqrt{3}/3$
$\pi/2$	$\sqrt{4}/2 = 1$	0	$\nexists$	0

(4.1)

(Il faut connaître ces valeurs et pouvoir reconstruire ce tableau de mémoire.)

4.1 Retrouver un angle à partir du sinus ou du cosinus

On considère un angle θ dont on suppose connaître le sinus, le cosinus, la tangente, etc. Que vaut θ ? Discutons du cas où c'est $\sin(\theta)$ qui est connu.

Ce problème se résout à l'aide du cercle trigonométrique (rappelez-vous de la figure VI2). Il est clair qu'en général deux angles θ sont possibles : l'un se trouve entre $-\frac{\pi}{2}$ et $\frac{\pi}{2}$, l'autre est son symétrique par rapport à l'axe vertical. Si en plus on connaît $\cos(\theta)$, alors il ne reste plus qu'un seul choix. (En fait seul le signe de $\cos(\theta)$ est alors nécessaire pour lever l'ambiguïté.)

Cependant cette méthode ne fournit pas la valeur exacte de l'angle, seulement une idée approximative mesurée sur le dessin (sauf si on a la chance d'avoir un sinus ou un cosinus issu du tableau de la présente page – en pratique, dans nos exercices, ce sera souvent le cas!).

En fait si on donne une valeur r quelconque entre -1 et 1 (qui est l'ensemble des valeurs possibles pour le sinus d'un angle), nous avons maintenant compris qu'il existait un unique θ entre $-\frac{\pi}{2}$ et $\frac{\pi}{2}$ tel que $\sin(\theta) = r$. On note ce nombre $\arcsin(r)$, par définition. Et c'est là que la technologie entre en jeu : l'ordinateur peut calculer avec une redoutable précision la valeur de $\arcsin(r)$. C'est ensuite à l'humain de prendre le relais pour savoir si l'angle cherché est $\arcsin(r)$ ou son symétrique $\pi - \arcsin(r)$, selon les autres informations dont il dispose.

Les fonctions arccos et arctan permettent d'obtenir une information similaire si on connaît $\cos(\theta)$ ou $\tan(\theta)$.

Résumé. Si on connaît $r = \begin{cases} \sin(\theta) \\ \cos(\theta) \\ \tan(\theta) \end{cases}$ alors $\theta = \begin{cases} \arcsin(r) \text{ ou } \pi - \arcsin(r) \\ \arccos(r) \text{ ou } -\arccos(r) \\ \arctan(r) \text{ ou } \pi + \arctan(r) \end{cases}$

Bien entendu, plutôt que de retenir cette recette par cœur, il vaut mieux se rappeler comment elle est obtenue grâce au cercle trigonométrique.

5 Relation fondamentale

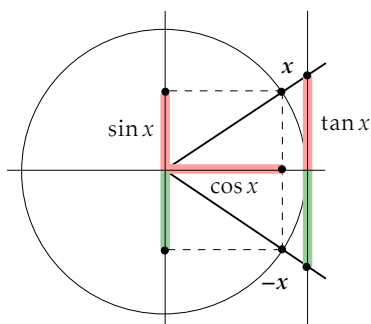
Le théorème de Pythagore appliqué dans le cercle trigonométrique implique l'importante relation $(\cos(\theta))^2 + (\sin(\theta))^2 = 1$, ce qu'on écrira plus souvent sous la forme

$$\cos^2 \theta + \sin^2 \theta = 1. \quad (5.2)$$

6 Symétries

Les fonctions trigonométriques vérifient certaines relations qu'il est indispensable de connaître. Elles découlent facilement de symétries géométriques.

6.1 Symétrie par rapport à l'axe des abscisses

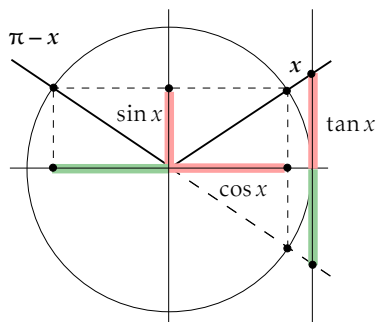


$$\sin(-x) = -\sin(x)$$

$$\cos(-x) = \cos(x)$$

$$\tan(-x) = -\tan(x)$$

6.2 Symétrie par rapport à l'axe des ordonnées

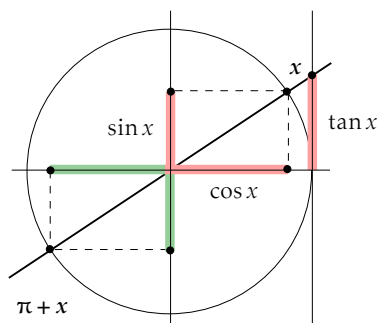


$$\sin(\pi - x) = \sin(x)$$

$$\cos(\pi - x) = -\cos(x)$$

$$\tan(\pi - x) = -\tan(x)$$

6.3 Symétrie par rapport à l'origine

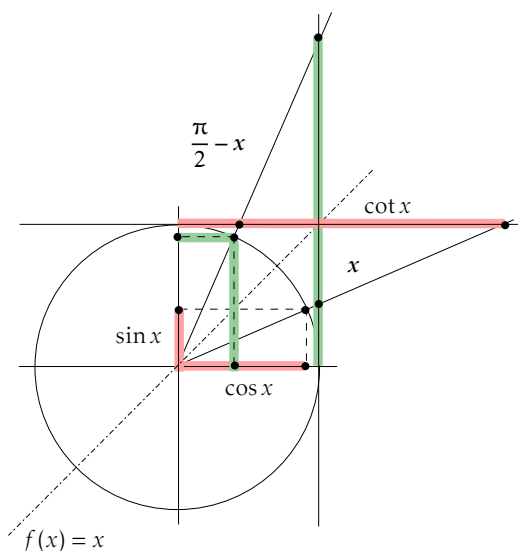


$$\sin(\pi + x) = -\sin(x)$$

$$\cos(\pi + x) = -\cos(x)$$

$$\tan(\pi + x) = \tan(x)$$

6.4 Symétrie par rapport à la première bissectrice

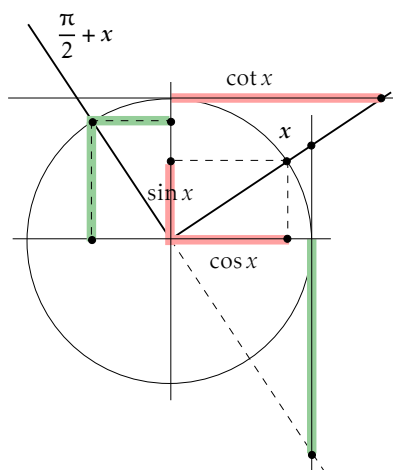


$$\sin\left(\frac{\pi}{2} - x\right) = \cos(x)$$

$$\cos\left(\frac{\pi}{2} - x\right) = \sin(x)$$

$$\tan\left(\frac{\pi}{2} - x\right) = \cot(x)$$

6.5 Angles décalés de 90°



$$\sin\left(\frac{\pi}{2} + x\right) = \cos(x)$$

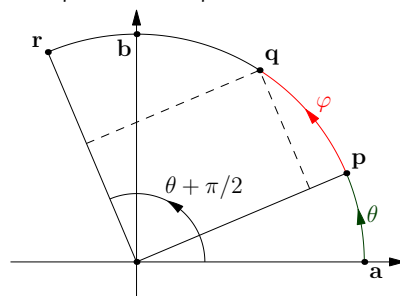
$$\cos\left(\frac{\pi}{2} + x\right) = -\sin(x)$$

$$\tan\left(\frac{\pi}{2} + x\right) = -\cot(x)$$

Preuve de la formule d'addition. Nous voulons prouver $\cos(\theta + \varphi) = \cos \theta \cos \varphi - \sin \theta \sin \varphi$.

Dans un repère cartésien, notons

- * $\mathbf{o} = (0, 0)$, $\mathbf{a} = (1, 0)$, $\mathbf{b} = (0, 1)$,
- * $\mathbf{p} = \cos(\theta)\mathbf{a} + \sin(\theta)\mathbf{b}$
- * $\mathbf{q} = (\cos(\theta + \varphi), \sin(\theta + \varphi))$
- * $\mathbf{r} = (\cos(\theta + \pi/2), \sin(\theta + \pi/2))$
 $= (-\sin(\theta), \cos(\theta))$



Dans le repère tourné d'un angle θ , on observe $\mathbf{q} = \cos(\varphi)\mathbf{p} + \sin(\varphi)\mathbf{r}$, et donc

$$\begin{aligned} (\cos(\theta + \varphi), \sin(\theta + \varphi)) &= \cos(\varphi)(\cos(\theta), \sin(\theta)) + \sin(\varphi)(-\sin(\theta), \cos(\theta)) \\ &= (\cos(\varphi)\cos(\theta) - \sin(\varphi)\sin(\theta), \cos(\varphi)\sin(\theta) + \sin(\varphi)\cos(\theta)) \end{aligned}$$

ce qui démontre la relation proposée, ainsi que la relation similaire concernant $\sin(\varphi + \theta)$. $\square$

7 D'autres identités remarquables : formulaire

Les identités suivantes sont souvent utiles :

$$\sin(A - B) = \sin A \cos B - \cos A \sin B$$

$$\sin(A + B) = \sin A \cos B + \cos A \sin B$$

$$\cos(A - B) = \cos A \cos B + \sin A \sin B$$

$$\cos(A + B) = \cos A \cos B - \sin A \sin B$$

$$\tan(A - B) = \frac{\tan A - \tan B}{1 + \tan A \tan B}$$

$$\tan(A + B) = \frac{\tan A + \tan B}{1 - \tan A \tan B}$$

$$\cos(2a) = \cos^2 a - \sin^2 a = 2 \cos^2 a - 1 = 1 - 2 \sin^2 a$$

$$\sin(2a) = 2 \sin a \cos a$$

$$\tan(2a) = \frac{2 \tan a}{1 - \tan^2 a}$$

$$\sin(3a) = 3 \sin a - 4 \sin^3 a$$

$$\cos(3a) = -3 \cos a + 4 \cos^3 a$$

$$\cos(A + B) + \cos(A - B) = 2 \cos A \cos B,$$

$$\cos(A + B) - \cos(A - B) = -2 \sin A \sin B,$$

$$\sin(A + B) + \sin(A - B) = 2 \sin A \cos B,$$

$$\sin(A + B) - \sin(A - B) = 2 \cos A \sin B.$$

$$\cos A \cos B = \frac{\cos(A - B) + \cos(A + B)}{2}, \text{ (en particulier } \cos^2 A = \frac{1 + \cos(2A)}{2} \text{)}$$

$$\sin A \sin B = \frac{\cos(A - B) - \cos(A + B)}{2}, \text{ (en particulier } \sin^2 A = \frac{1 - \cos(2A)}{2} \text{)}$$

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

$$\cos A \sin B = \frac{\sin(A + B) - \sin(A - B)}{2}$$

$$\cos p + \cos q = 2 \cos\left(\frac{p+q}{2}\right) \cos\left(\frac{p-q}{2}\right)$$

$$\cos p - \cos q = -2 \sin\left(\frac{p+q}{2}\right) \sin\left(\frac{p-q}{2}\right)$$

$$\sin p + \sin q = 2 \sin\left(\frac{p+q}{2}\right) \cos\left(\frac{p-q}{2}\right)$$

$$\sin p - \sin q = 2 \cos\left(\frac{p+q}{2}\right) \sin\left(\frac{p-q}{2}\right)$$

8 Fonctions trigonométriques dans les triangles

8.1 Triangles rectangles

Considérons un triangle rectangle ABC, rectangle en C, tel que dessiné à la figure VI4 (à gauche). Par une utilisation judicieuse du théorème de Thalès et du cercle trigonométrique (voir même figure, à droite), on prouve aisément les relations bien connues

$$\cos \theta = \frac{\text{côté adjacent}}{\text{hypothénuse}} = \frac{b}{c} \quad \sin \theta = \frac{\text{côté opposé}}{\text{hypothénuse}} = \frac{a}{c} \quad \tan \theta = \frac{\text{côté opposé}}{\text{côté adjacent}} = \frac{a}{b}$$

FIGURE VI4 – Un triangle rectangle (à gauche), et le même triangle avec un cercle dessiné autour d'un de ses sommets

8.2 Triangles quelconques

Soit un triangle de sommets A, B et C, avec la convention que $\hat{A}$, $\hat{B}$ et $\hat{C}$ désignent les angles (pris entre 0 et π) en chacun des sommets, et a , b et c les mesures des longueurs (positives) des segments BC, CA et AB respectivement. Ceci est illustré ci-dessous :

Loi des sinus Avec ces notations, la *loi des sinus* s'écrit

loi des sinus

$$\frac{\sin \hat{A}}{a} = \frac{\sin \hat{B}}{b} = \frac{\sin \hat{C}}{c}.$$

Démonstration. Prouvons l'une de ces égalités :

Notons h la hauteur issue de B. Alors d'une part $\sin \hat{C} = \frac{h}{a}$ et d'autre part $\sin \hat{A} = \frac{h}{c}$, ceci prouve

$$\frac{\sin \hat{A}}{a} = \frac{\sin \hat{C}}{c}$$

□

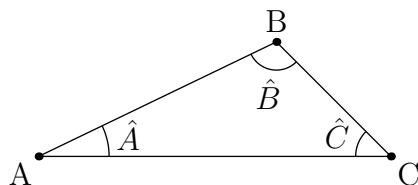
Loi des cosinus, ou théorème d'Al-Kashi Une généralisation du théorème de Pythagore est donnée par la *formule d'Al-Kashi*

formule
d'Al-Kashi

$$c^2 = a^2 + b^2 - 2ab \cos(\hat{C}).$$

On vérifiera qu'en effet, lorsque le triangle est rectangle en C ($\hat{C} = \pi/2$), alors $\cos(\hat{C}) = 0$ ce qui donne la formule de Pythagore :

$$c^2 = a^2 + b^2.$$



Démonstration. Notons $\mathbf{AC} = \mathbf{C} - \mathbf{A}$ pour le vecteur de A à C, etc. et développons c^2 grâce au produit scalaire :

$$\begin{aligned} c^2 &= \|\mathbf{AB}\|^2 = \|\mathbf{AC} + \mathbf{CB}\|^2 = \|\mathbf{AC}\|^2 + \|\mathbf{CB}\|^2 + 2 \mathbf{AC} \cdot \mathbf{CB} \\ &= \|\mathbf{AC}\|^2 + \|\mathbf{CB}\|^2 + 2\|\mathbf{AC}\|\|\mathbf{CB}\|\cos(\pi - \hat{C}) \\ &= \|\mathbf{AC}\|^2 + \|\mathbf{CB}\|^2 - 2\|\mathbf{AC}\|\|\mathbf{CB}\|\cos(\hat{C}) \end{aligned}$$

Attention, l'angle entre $\mathbf{AC}$ et $\mathbf{CB}$ est $\pi - \hat{C}$ car ils n'ont pas le même point base sur l'image.

□

Chapitre VII

Géométrie analytique

La géométrie analytique est une manière d'aborder des problèmes de géométrie grâce à un ou plusieurs systèmes de coordonnées. Le choix classique d'un système de coordonnées est un système cartésien orthonormé, ce qui fournit une bijection entre l'objet géométrique (nous concernant : une droite, un plan ou l'espace) et un objet analytique (dans le même ordre : $\mathbb{R}$, $\mathbb{R}^2$ ou $\mathbb{R}^3$). Ce choix est tellement standard que nous aurons coutume de parler de « la droite réelle » pour désigner $\mathbb{R}$, et de même pour le plan réel ($\mathbb{R}^2$) et l'espace réel ($\mathbb{R}^3$), prenant ainsi le modèle analytique pour le modèle géométrique lui-même. Nous étudierons bien sûr également les droites du plan, de l'espace, ou les plans de l'espace.

1 Points

Dans ce contexte, un élément de $\mathbb{R}^n$ est donc appelé un *point*, il s'écrit avec n nombres réels appelés ses *coordonnées*, que nous notons généralement $(p_1, p_2, \dots, p_n)$.

point

coordonnées

Remarque (*). On dit que $\mathbb{R}^n$ est un *modèle* ou *modélise* la droite ($n = 1$), le plan ($n = 2$) et l'espace ($n = 3$). Comme mentionné précédemment, nous aurons coutume de confondre l'objet géométrique avec le modèle.

modèle

modélise

Exemple. Le point $(1, -1)$ de $\mathbb{R}^2$ désigne l'élément du plan d'abscisse 1 et d'ordonnée -1 .

Remarque (*). Pour certaines valeurs de n fixées, d'autres coutumes existent : pour $n = 1$, il n'y a qu'une seule coordonnée donc on n'écrit généralement pas p_1 mais simplement p , ou toute autre lettre. Pour $n = 2$, il y a deux composantes qui sont souvent notées (x, y) (mais d'autres lettres s'utilisent). Similairement pour $n = 3$ en ajoutant z (ou autres). Et cætera. Toutes ces conventions d'écriture n'ont qu'un effet psychologique : on peut très bien en changer si le contexte le requiert.

2 Vecteurs

Au delà de l'étude des points de la droite, du plan, de l'espace, etc. nous pouvons également étudier les relations entre eux et, en particulier, les mouvements. Le mouvement le plus élémentaire est la translation, et une translation peut également se décrire par un élément de $\mathbb{R}^n$. Dans ce cas, les éléments de $\mathbb{R}^n$ sont appelés des *vecteurs*, et ses constituants sont les *composantes* du vecteur. (Les mouvements autres que les translations seront étudiés dans les chapitres suivants.)

vecteurs

composantes

Remarque (*). Ceci a la désagréable conséquence –pour l'étudiant– que le même objet $\mathbb{R}^n$ modélise plusieurs réalités (points et vecteurs). En pratique cela a peu d'importance, mais comprendre la différence conceptuelle permet de mieux apprécier le contexte et d'avancer plus efficacement dans l'étude.

Exemple. Le vecteur $(-1, 1)$ de $\mathbb{R}^2$ indique une translation dans la direction « Nord-Ouest » sur le plan.

La différence, fondamentale et intuitive, entre point et vecteur est que le premier représente un état, une position ; tandis que le second représente une évolution, un mouvement. En pratique, au moins en première approche, ces deux types d'objets sont représentés formellement par des éléments de $\mathbb{R}^n$, ce qui les rend indiscernables d'un point de vue de leur notation. Une différence de nomenclature aide à s'y retrouver : les constituants d'un point sont appelés ses coordonnées, tandis que les constituants d'un vecteur sont appelés ses composantes.

Remarque (*). Sur un dessin, un point (au sens ci-dessus) est généralement représenté par un point (dessiné); un vecteur est représenté le plus souvent par une flèche.

Exemple. L'élément $(1, 1)$ (de $\mathbb{R}^2$) peut représenter soit un point du plan, soit un déplacement dans le plan. En tant que point, il se trouve au « nord-est » du point $(0, 0)$; en tant que vecteur, il indique la direction « nord-est ».

Remarque (*). Dans ce cours, le mot *vecteur* sera utilisé dans des contextes où il serait également valide de parler de *point*. Dans la mesure où il n'y a formellement aucune différence cela ne doit pas porter à conséquence, mais le lecteur est invité à rester attentif.

2.1 Vecteurs libres vs vecteurs liés

Un vecteur se représente généralement par une flèche. Cette flèche a donc un début et une fin. Cependant, nous pourrions dessiner la même flèche à n'importe quel endroit! Son début est le « point-base » du vecteur, c'est son point d'attache ou encore son « origine ». Lorsque cela est important, on dit du vecteur $\mathbf{w}$ qu'il est basé (ou *lié*) en le point $\mathbf{p}$ s'il faut le représenter ou l'imaginer avec ce point pour origine. Si $\mathbf{p}$ et $\mathbf{q}$ sont deux points, on note parfois $\overrightarrow{\mathbf{qp}}$ le vecteur basé en $\mathbf{q}$ dont l'extrémité est en $\mathbf{p}$.

Exemple. Si on imagine un vecteur représentant la vitesse (direction et intensité) du vent en un point donné, cela n'a pas de sens que ce vecteur soit basé en un autre point que celui-là.

Si deux vecteurs liés ont même composantes, ils sont dits *équipollents*. Cela caractérise une situation dans laquelle les deux vecteurs sont « les mêmes » mais basés en des points éventuellement différents.

Lorsque le point-base n'a aucune importance on dit que le vecteur est un vecteur *libre*.

Exemple. La direction « Sud-Est » est donnée par un vecteur libre de composantes $(1, -1)$.

Remarque (*). Ceci n'a rien à voir avec la notion de partie « libre » et « liée » qui sera vue en la section ?? page ??.

3 Opérations de base

Les éléments de $\mathbb{R}^n$ (pour toute valeur de n fixée) peuvent être additionnés (« composante par composante ») entre eux, ou multipliés par un réel donné de la manière suivante : si $\mathbf{x}$ et $\mathbf{y}$ sont des éléments de $\mathbb{R}^n$ et $\lambda \in \mathbb{R}$, on définit

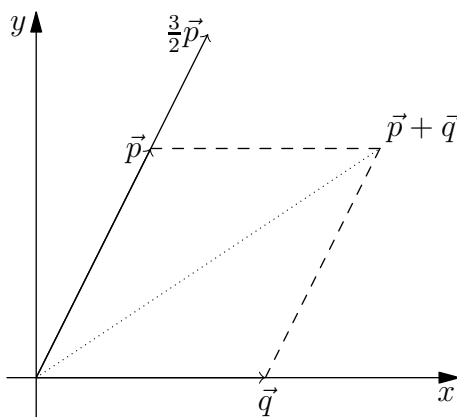
$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$

$$(x_1, \dots, x_n) - (y_1, \dots, y_n) = (x_1 - y_1, \dots, x_n - y_n)$$

$$\lambda(x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n)$$

(Il est également possible de multiplier deux éléments composante par composante, cependant c'est une opération qui se révèle largement moins utilisée. Nous n'en parlerons donc pas plus.) Ces opérations ont l'interprétation géométrique suivante :

- * la somme de deux vecteurs est donnée par la règle du parallélogramme ;
- * la multiplication par un réel (on dira aussi « par un *scalaire* ») multiplie la longueur.



Précisons notre pensée pour ce qui concerne la somme de points et de vecteurs :

- * la différence de deux *points* est un vecteur allant du second vers le premier : $\mathbf{p} - \mathbf{q}$ est le vecteur $\overrightarrow{qp}$;
- * la somme d'un *point* et d'un *vecteur* représente un déplacement à partir de ce point dans la direction du vecteur donné : le résultat est un point ;
- * la somme de deux ou plusieurs *vecteurs* représente la composée de déplacements successifs : le résultat est un vecteur.

Exemple. Si on considère $\mathbf{p} = (1, 2)$ comme un point, et $\mathbf{w} = (1, -1)$ comme un vecteur (donnant la direction « Nord-Ouest »), alors on conçoit leur somme $\mathbf{p} + \mathbf{w} = (1, 2) + (1, -1) = (2, 1)$ comme le point obtenu en translatant le point de départ, $\mathbf{p}$, dans la direction donnée $\mathbf{w}$.

Définition VII.1. L'élément particulier $(0, 0, \dots, 0)$ se note $\mathbf{0}$ et est appelé *vecteur nul*.

vecteur nul

Résultat VII.2. Quel que soit le vecteur $\mathbf{x} \in \mathbb{R}^n$, nous avons

$$\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}.$$

Avec ces définitions, nous attardant pour l'exemple sur le cas $n = 3$, si nous notons $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ les vecteurs suivants :

$$\mathbf{e}_1 = (1, 0, 0)$$

$$\mathbf{e}_2 = (0, 1, 0)$$

$$\mathbf{e}_3 = (0, 0, 1)$$

alors tout vecteur $\mathbf{x}$ de $\mathbb{R}^3$ peut s'écrire $x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$. Cette décomposition et ces vecteurs $\mathbf{e}_i$ peuvent évidemment se généraliser à toute valeur de n : ils contiennent simplement plus de composantes. Les vecteurs $\mathbf{e}_i$ sont appelés vecteurs de base de $\mathbb{R}^n$.

Résultat VII.3. Tout vecteur $\mathbf{v}$ de $\mathbb{R}^n$ s'écrit

$$v_1\mathbf{e}_1 + \dots + v_n\mathbf{e}_n$$

où $\mathbf{e}_i$ est un vecteur de $\mathbb{R}^n$ dont la j^{e} composante vaut 0 si $i \neq j$, et vaut 1 si $i = j$; en d'autres termes

$$(e_i)_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon.} \end{cases}$$

Remarque (*). Attention, ici $\mathbf{e}_1$ et ses compagnons ne sont pas les composante d'un hypothétique vecteur $\mathbf{e}$, mais des vecteurs eux-même ! Pour l'exemple, x_1 et $\mathbf{e}_1$ ont donc des rôles fondamentalement différents : le premier est un réel (ou « scalaire »), le second est un vecteur.

Définition VII.4. Si $\mathbf{u}_1, \dots, \mathbf{u}_k$ sont des vecteurs, une *combinaison linéaire* de ceux-ci est une somme de la forme $\lambda_1\mathbf{u}_1 + \dots + \lambda_k\mathbf{u}_k$.

combinaison linéaire

Remarque (*). On peut paraphraser la proposition précédente sous la forme : tout vecteur $\mathbf{v} \in \mathbb{R}^n$ est une combinaison linéaire des vecteurs $\mathbf{e}_1, \dots, \mathbf{e}_n$.

3.1 Centre de gravité

Définition VII.5. Si on considère des points $\mathbf{P}_1, \dots, \mathbf{P}_k$ dans $\mathbb{R}^n$, leur centre de gravité est :

$$\mathbf{O} + \frac{1}{k} \sum_i \mathbf{OP}_i$$

où $\mathbf{O}$ représente n'importe quel point de l'espace.

Remarque (*). Pouvez-vous démontrer que la quantité ci-dessus ne dépend pas du point $\mathbf{O}$ choisi ?

Exemple. Le centre de gravité de $(1, 1), (1, -1), (-1, 1), (-1, -1)$ est l'origine.

Le centre de gravité des points $(1, 2), (1, 3), (2, 1)$ est $(\frac{4}{3}, 2)$.

4 Produit scalaire

4.1 Définition

produit scalaire

Définition VII.6. Le *produit scalaire* entre deux vecteurs $\mathbf{v}$ et $\mathbf{w}$ de $\mathbb{R}^n$ est donné par

$$\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + \cdots + v_n w_n.$$

Résultat VII.7. Pour tous vecteurs $\mathbf{a}, \mathbf{a}', \mathbf{b} \in \mathbb{R}^n$ et $\lambda, \mu \in \mathbb{R}$, le produit scalaire vérifie les propriétés suivantes :

1. $(\lambda \mathbf{a} + \mu \mathbf{a}') \cdot \mathbf{b} = \lambda \mathbf{a} \cdot \mathbf{b} + \mu \mathbf{a}' \cdot \mathbf{b}$ (linéarité à gauche)
2. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$ (symétrique)
3. $\mathbf{a} \cdot \mathbf{a} \geq 0$ et de plus, $\mathbf{a} \cdot \mathbf{a} = 0 \iff \mathbf{a} = \mathbf{0}$ (défini positif).

Notons que les deux premières propriétés montrent qu'il y a également linéarité à droite ; on dit alors que le produit scalaire est *bilinéaire*.

On définit la *norme* associée à ce produit scalaire par

bilinéaire

norme

$$\|\mathbf{v}\| := \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + \cdots + v_n^2}$$

et cette norme vérifie donc la propriété $\|\mathbf{v}\| = 0 \iff \mathbf{v} = \mathbf{0}$.

Résultat VII.8 (Inégalité de Cauchy-Schwarz). Pour tout $\mathbf{a}, \mathbf{b}$ de $\mathbb{R}^n$, on a

$$|\mathbf{a} \cdot \mathbf{b}| \leq \|\mathbf{a}\| \|\mathbf{b}\|.$$

Démonstration. Considérons la quantité $(\mathbf{a} + t\mathbf{b}) \cdot (\mathbf{a} + t\mathbf{b})$, qui est positive. Nous pouvons la ré-écrire via la bilinéarité du produit scalaire sous la forme :

$$\begin{aligned} (\mathbf{a} + t\mathbf{b}) \cdot (\mathbf{a} + t\mathbf{b}) &= \mathbf{a} \cdot \mathbf{a} + 2t \mathbf{a} \cdot \mathbf{b} + t^2 \mathbf{b} \cdot \mathbf{b} \\ &= \|\mathbf{a}\|^2 + 2t \mathbf{a} \cdot \mathbf{b} + t^2 \|\mathbf{b}\|^2 \\ &= \left(t \|\mathbf{b}\| + \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|} \right)^2 - \left(\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|} \right)^2 + \|\mathbf{a}\|^2 \geq 0 \end{aligned}$$

Prenons maintenant $t = -\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|^2}$, ce qui annule la première parenthèse, et fournit donc

$$\|\mathbf{a}\|^2 \geq \left(\frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|} \right)^2$$

d'où (puisque $|\mathbf{a} \cdot \mathbf{b}| \geq 0$)

$$\|\mathbf{a}\| \geq \frac{|\mathbf{a} \cdot \mathbf{b}|}{\|\mathbf{b}\|}$$

qui est l'inégalité annoncée. □

Résultat VII.9. La norme $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ vérifie les propriétés suivantes :

- * $\|\mathbf{a}\| = 0 \iff \mathbf{a} = \mathbf{0}$;
- * $\|\lambda \mathbf{a}\| = |\lambda| \|\mathbf{a}\|$;
- * $\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$ (inégalité triangulaire).

Démonstration. Les deux premières propriétés sont évidentes de la définition de la norme et les propriétés du produit scalaire. La troisième se démontre en calculant

$$\begin{aligned} \|\mathbf{a} + \mathbf{b}\|^2 &= (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) \\ &= \mathbf{a} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} + 2 \mathbf{a} \cdot \mathbf{b} \leq \mathbf{a} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} + 2 \|\mathbf{a}\| \|\mathbf{b}\| \\ &= (\|\mathbf{a}\| + \|\mathbf{b}\|)^2 \end{aligned}$$

□

Résultat VII.10. Si $\mathbf{v}$ et $\mathbf{w}$ sont des vecteurs de $\mathbb{R}^n$, alors

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos(\theta)$$

où θ est l'angle formé par les vecteurs $\mathbf{v}$ et $\mathbf{w}$.

Preuve en dimension 2. Si $\mathbf{v} = (\|\mathbf{v}\| \cos(\phi), \|\mathbf{v}\| \sin(\phi))$ et $\mathbf{w} = (\|\mathbf{w}\| \cos(\psi), \|\mathbf{w}\| \sin(\psi))$, alors

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| (\cos(\phi) \cos(\psi) + \sin(\phi) \sin(\psi)) = \|\mathbf{v}\| \|\mathbf{w}\| \cos(\psi - \phi)$$

où $\psi - \phi$ représente bien l'angle entre les deux vecteurs. $\square$

Corollaire. Le produit scalaire s'annule si et seulement si les vecteurs sont orthogonaux (ou si l'un des deux vecteurs est nul).

Démonstration. Si les vecteurs ne sont pas nuls, alors leur norme n'est pas nulle, dès lors seul le cosinus peut s'annuler, ce qui n'arrive que pour des angles droits. $\square$

Résultat VII.11. Si $\mathbf{a}$ et $\mathbf{b}$ sont deux vecteurs de $\mathbb{R}^n$, alors

$$\mathbf{a}' = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}} \mathbf{b}$$

est le projeté de $\mathbf{a}$ sur la droite engendrée par $\mathbf{b}$. projeté

Démonstration. Le fait que $\mathbf{a}'$ est sur la droite engendrée par $\mathbf{b}$ est évident, puisqu'il en est un multiple par construction.

Quelle est la condition pour que $\mathbf{a}' = \lambda \mathbf{b}$ soit le projeté de $\mathbf{a}$? Il faut que le triangle formé par $\mathbf{a}$, $\mathbf{a}'$ et leur origine soit rectangle en $\mathbf{a}'$. C'est-à-dire il faut

$$\|\mathbf{a}'\|^2 + \|\mathbf{a} - \mathbf{a}'\|^2 = \|\mathbf{a}\|^2$$

ce qu'on ré-écrit

$$\lambda^2 \|\mathbf{b}\|^2 + \|\mathbf{a}\|^2 + \lambda^2 \|\mathbf{b}\|^2 - 2\lambda \mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\|^2$$

et est équivalent à

$$\lambda = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{b}\|^2}$$

ce que nous voulions observer. $\square$

Corollaire. Si $\mathbf{a}$ et $\mathbf{b}$ sont deux vecteurs, leur produit scalaire est donné par :

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{b}\| \|\mathbf{a}'\|$$

où $\mathbf{a}'$ est le projeté de $\mathbf{a}$ sur $\mathbf{b}$.

Démonstration. Il suffit de calculer le membre de droite grâce au résultat précédent. $\square$

4.2 Lien avec les équations cartésiennes

4.2.1 Droites du plan

Nous voulons justifier de manière géométrique l'équivalence des équations paramétriques et de l'équation cartésienne d'une droite du plan. Le passage de l'un à l'autre avait précédemment été effectué de manière algébrique.

Soit $E \subset \mathbb{R}^2$ une droite passant par $\mathbf{p}$ et est perpendiculaire au vecteur $\mathbf{w}$. Un point $\mathbf{r}$ appartient à la droite si et seulement si $\mathbf{r} - \mathbf{p} \perp \mathbf{w}$, c'est-à-dire si et seulement si $\mathbf{w} \cdot (\mathbf{r} - \mathbf{p}) = 0$. En d'autres termes (en notant $\mathbf{r} = (x, y)$, $\mathbf{p} = (p_1, p_2)$, $\mathbf{w} = (\alpha, \beta)$) si

$$\alpha x + \beta y = \alpha p_1 + \beta p_2.$$

Réciproquement, soit E l'ensemble d'équation $\alpha x + \beta y = \delta$. Notons $\mathbf{w} = (\alpha, \beta)$ et $\mathbf{r} = (x, y)$, et soit $p \in E$ de sorte que $\mathbf{w} \cdot \mathbf{p} = \delta$, ce qui permet de ré-écrire l'équation sous la forme

$$\mathbf{w} \cdot \mathbf{r} = \mathbf{w} \cdot \mathbf{p}$$

ce qui est l'équation obtenue précédemment pour une droite perpendiculaire à $\mathbf{w}$ et passant par $\mathbf{p}$.

4.2.2 Plans dans l'espace

Soit E un plan dans $\mathbb{R}^3$, $\mathbf{p}$ un point de E et soit $\mathbf{w}$ un vecteur perpendiculaire à ce plan. Un point $\mathbf{r}$ de $\mathbb{R}^3$ appartient à E si et seulement si $\mathbf{r} - \mathbf{p} \perp \mathbf{w}$, c'est-à-dire si $\mathbf{r} - \mathbf{p} \cdot \mathbf{w} = 0$. À nouveau, en écrivant $\mathbf{r} = (x, y, z)$, $\mathbf{p} = (p_1, p_2, p_3)$ et $\mathbf{w} = (\alpha, \beta, \gamma)$, nous ré-écrivons cette équation

$$\alpha x + \beta y + \gamma z = \alpha p_1 + \beta p_2 + \gamma p_3.$$

De même que pour les droites du plan, il est facile de constater que tout point vérifiant une équation du type

$$\alpha x + \beta y + \gamma z = \delta$$

se trouve sur un plan perpendiculaire à $\mathbf{w} := (\alpha, \beta, \gamma)$.

5 Produit vectoriel

Nous allons définir un produit entre deux vecteurs $\mathbf{a}$ et $\mathbf{b}$ de l'espace $\mathbb{R}^3$. Il est appelé « produit vectoriel » car son résultat est un vecteur. La définition que nous allons former n'aura pas de sens dans $\mathbb{R}^n$ pour des valeurs $n \neq 3$.

produit vectoriel

Définition VII.12. Si $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, on définit leur *produit vectoriel*

$$\mathbf{a} \times \mathbf{b} = (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1).$$

Le produit vectoriel est donc une application

$$\cdot \times \cdot : \mathbb{R}^3 \rightarrow \mathbb{R}^3 : (\mathbf{a}, \mathbf{b}) \mapsto \mathbf{a} \times \mathbf{b}.$$

Résultat VII.13. *Le produit vectoriel vérifie les identités suivantes :*

- * $\mathbf{a} \times \mathbf{b} \perp \mathbf{a}$ et $\mathbf{a} \times \mathbf{b} \perp \mathbf{b}$;
- * $\|\mathbf{a} \times \mathbf{b}\|^2 + |\mathbf{a} \cdot \mathbf{b}|^2 = \|\mathbf{a}\|^2 \|\mathbf{b}\|^2$;
- * $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\| \|\mathbf{b}\| \sin \theta$;
- * $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$ (*anti-symétrique*);
- * $(\alpha \mathbf{a} + \alpha' \mathbf{a}') \times \mathbf{b} = \alpha \mathbf{a} \times \mathbf{b} + \alpha' \mathbf{a}' \times \mathbf{b}$

pour tout $\alpha, \alpha' \in \mathbb{R}$ et $\mathbf{a}, \mathbf{a}', \mathbf{b} \in \mathbb{R}^3$, et où θ est l'angle formé par les vecteurs $\mathbf{a}$ et $\mathbf{b}$.

Démonstration. Tous ces points se prouvent à partir de la définition. Les calculs, simples mais fastidieux, sont laissés en exercice au lecteur.

Mentionnons tout de même que dans l'égalité $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\| \|\mathbf{b}\| \sin \theta$, le sinus est toujours positif car l'angle θ est, par définition, entre 0 et π . $\square$

5.1 Trièdres orientés

Un trièdre orienté est la donnée d'une séquence de trois vecteurs $\mathbf{u}, \mathbf{v}, \mathbf{w}$, tous issus d'un même point $\mathbf{p}$. L'ordre a ici son importance. Considérant votre main droite, si $\mathbf{p}$ représente votre poignet, $\mathbf{u}$ votre pouce et $\mathbf{v}$ votre index, on dira que le trièdre est droit (ou « orienté positivement ») si le vecteur $\mathbf{w}$ est du même côté du plan engendré par les vecteurs $\mathbf{u}$ et $\mathbf{v}$ que votre majeur. Dans le cas contraire, le trièdre n'est pas droit (parfois « gauche »).

Nous imposerons désormais toujours que les axes de coordonnées forment un trièdre droit, c'est-à-dire le trièdre basé en l'origine formé des vecteurs de base $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ (dans cet ordre). De la sorte, nous pouvons énoncer le résultat suivant :

Résultat VII.14. *Le trièdre formé de vecteurs $\mathbf{v}, \mathbf{w}$ et $\mathbf{v} \times \mathbf{w}$ est un trièdre droit.*

Chapitre VIII

Géométrie analytique (suite)

1 Équations

Dans $\mathbb{R}^n$, il existe deux manières classiques de décrire des sous-ensembles. Les équations paramétriques et les équations cartésiennes.

Une ou plusieurs *équations cartésiennes* donnent un ensemble de conditions sur le point (x, y) du plan ou (x, y, z) de l'espace (ou de manière générale sur le point $(p_1, \dots, p_n)$ de $\mathbb{R}^n$). Lorsque le point vérifie la ou les conditions, il est dans l'ensemble; sinon il n'y est pas. Les coordonnées du point sont appelées les *inconnues* de l'équation (ou des équations) cartésienne(s). On dit parfois que l'ensemble est donné sous forme *implicite*.

équations
cartésiennes

inconnues

Exemple. Dans le plan, l'équation $x^2 + y^2 = 4$ est une équation cartésienne pour le cercle centré en l'origine de rayon 2. Le point $(0, 2)$ et le point $(\sqrt{2}, \sqrt{2})$ sont dans cet ensemble, mais $(0, 0)$ ou $(1, 1)$ n'y sont pas.

Dans l'espace, la même équation $x^2 + y^2 = 4$ est une équation cartésienne pour le *cylindre* centré autour de l'axe des z et de rayon 2.

Contrairement aux équations cartésiennes, qui partent d'un point (x, y) inconnu et y imposent des conditions pour être dans l'ensemble, les *équations paramétriques* décrivent explicitement chacun des points de l'ensemble en fonction de certains paramètres (qui peuvent varier). Dans $\mathbb{R}^n$, il y a toujours n équations paramétriques : une pour chacune des coordonnées.

équations
paramétriques

Exemple. Par exemple $x = 2\cos(t)$ et $y = 2\sin(t)$ sont des équations paramétriques (de paramètre t) du cercle de rayon 2 centré en l'origine. Ici, le paramètre t doit varier au minimum dans $[0, 2\pi[$ pour pouvoir décrire tout le cercle, mais peut varier dans un ensemble plus grand : le cercle sera alors « parcouru » plusieurs fois, mais cela importe peu. Si t varie dans un ensemble plus petit, l'ensemble décrit sera alors simplement une partie du cercle.

Remarque (*). Formellement, pour le lecteur curieux, on pourrait dire qu'un ensemble $E \subset \mathbb{R}^n$ est donné sous forme cartésienne s'il existe une fonction $F : \mathbb{R}^n \rightarrow \mathbb{R}^k$ telle que

$$E = \{\mathbf{p} \text{ t.q. } F(\mathbf{p}) = 0\}$$

(pour une certaine valeur de k). On dirait au contraire que E est donné sous forme paramétrique s'il existe une fonction $F : U \rightarrow \mathbb{R}^n$ avec $E = \text{Im} F$ pour un certain sous-ensemble $U \subset \mathbb{R}^p$ et pour une certaine valeur de p .

1.1 Équations de droites

Une droite est caractérisée par deux points ou, de manière équivalente par un point et une direction. Une droite, dans $\mathbb{R}^n$, est donc un ensemble de la forme

$$\{\mathbf{p} + t\mathbf{v} \text{ t.q. } t \in \mathbb{R}\}$$

où le point $\mathbf{p}$ et le vecteur $\mathbf{v}$ sont des éléments de $\mathbb{R}^n$ donnés, avec $\mathbf{v} \neq 0$. $\mathbf{p}$ est un point par lequel passe la droite, $\mathbf{v}$ est une direction de cette droite. $\mathbf{v}$ est appelé un *vecteur directeur* de la droite.

vecteur directeur

Remarquons que $\mathbf{p}$ et $\mathbf{v}$ ne sont pas uniques : prendre $\mathbf{p}+\mathbf{v}$ au lieu de $\mathbf{p}$ et $3\mathbf{v}$ au lieu de $\mathbf{v}$ donneraient la même droite.

Définition VIII.1. Deux droites, de vecteurs directeurs $\mathbf{v}$ et $\mathbf{w}$ respectivement, sont parallèles s'il existe $\lambda \in \mathbb{R}$ tel que $\mathbf{v} = \lambda \mathbf{w}$.

1.1.1 Dans le plan

Dans le plan, nous parlons donc de l'ensemble

$$\{(a, b) + t(u, v) \text{ t.q. } t \in \mathbb{R}\}$$

où a, b, u, v sont des réels fixés. En d'autres termes, nous regardons tous les points (x, y) vérifiant

$$\begin{cases} x = a + tu \\ y = b + tv \end{cases}$$

pour au moins une valeur de t , c'est un système d'équations paramétriques de la droite (de paramètre t).

Cette forme paramétrique est très géométrique mais comme elle comporte deux équations, il est alors tentant de résoudre pour t (dont la valeur importe peu) ce qui donne

$$\frac{x-a}{u} = \frac{y-b}{v}$$

(ceci tient à condition que u et v ne soient pas nuls!) et qu'on ré-écrit souvent

$$xv - yu + (bu - av) = 0$$

Cette dernière équation est l'*équation cartésienne* de la droite. Cette équation est équivalente aux équations paramétriques : il n'y pas de conditions sur (u, v) en dehors de celle d'être différent du vecteur nul $(0, 0)$! Le lecteur s'en convaincra en vérifiant les cas $u = 0$ et $v = 0$ par lui-même. En d'autres termes :

Résultat VIII.2. Pour tout $(a, b) \in \mathbb{R}^2$ et $(u, v) \in \mathbb{R}^2$ avec $(u, v) \neq (0, 0)$, nous avons

$$\{(x, y) \text{ t.q. } xv - yu + (bu - av) = 0\} = \{(a, b) + t(u, v) \text{ t.q. } t \in \mathbb{R}\}.$$

Nous verrons plus loin que toute équation, dans le plan, de la forme $\alpha x + \beta y = \delta$ est l'équation d'une droite pour certaines valeurs de a, b, u, v .

Exercice. Si $p = (a, b)$ et $q = (c, d)$ sont deux points du plan, montrer que la droite

$$x = a + t(c - a) \quad y = b + t(d - b)$$

passse par ces deux points. Quel est un vecteur directeur $\mathbf{v}$?

1.1.2 Dans l'espace

Dans l'espace, les équations paramétriques (de paramètre t) sont données par

$$\begin{cases} x = a + tu \\ y = b + tv \\ z = c + tw \end{cases}$$

mais cette fois, en résolvant pour t , il n'est pas possible de se ramener à une seule équation : il restera toujours deux équations cartésiennes. On les écrit parfois sous la forme :

$$\frac{x-a}{u} = \frac{y-b}{v} = \frac{z-c}{w}$$

mais ceci rajoute des conditions : u, v, w doivent être tous les trois différent de 0 (alors qu'il suffit, du point de vue théorique, pour avoir une droite, que l'un d'entre eux soit non nul!) En pratique, la condition est souvent vérifiée donc ce sont souvent les équations cartésiennes qui sont données dans les exercices concrets.

Remarque (*). Les deux équations cartésiennes peuvent se comprendre comme donnant chacune un plan ; la droite est alors vue comme l'intersection de deux plans.

équation
cartésienne

1.2 Équations de plans

Un plan est donné par un point $\mathbf{p}$ et deux directions : $\mathbf{v}, \mathbf{w}$. Ses équations paramétriques sont alors

$$\{\mathbf{p} + \lambda\mathbf{v} + \mu\mathbf{w} \text{ t.q. } \lambda, \mu \in \mathbb{R}\}$$

avec la condition que $\mathbf{v}$ et $\mathbf{w}$ ne soient pas multiple l'un de l'autre (sinon on obtient une droite).

1.2.1 Dans l'espace

Les équations paramétriques sont explicitement données par

$$\begin{cases} x = p_1 + \lambda u_1 + \mu v_1 \\ y = p_2 + \lambda u_2 + \mu v_2 \\ z = p_3 + \lambda u_3 + \mu v_3 \end{cases}$$

où $\lambda, \mu \in \mathbb{R}$ sont les deux paramètres de ces équations paramétriques (en d'autres termes : lorsqu'ils varient dans $\mathbb{R}$, ils permettent de décrire tous les points du plan considéré.) et il est à nouveau possible de ramener ces équations paramétriques à une forme cartésienne :

$$\alpha x + \beta y + \gamma z = \delta$$

où les $\alpha, \beta, \gamma, \delta$ sont exprimables (mais pas de manière simple) en terme de $\mathbf{p}, \mathbf{u}$ et $\mathbf{v}$. Nous verrons la procédure générale plus tard. En pratique, il suffit de résoudre en éliminant les paramètres.

Exemple. Le plan d'équations $(x, y, z) = (1, \lambda, \lambda + \mu)$ se ré-écrit sous forme cartésienne : $x = 1$.

Exercice. Le plan d'équations $(x, y, z) = (\mu, \lambda, \lambda + \mu - 2)$ se ré-écrit sous forme cartésienne : $x + y - z = 2$.

Parallélisme Deux plans d'équations

$$\alpha x + \beta y + \gamma z = \delta$$

et

$$\alpha' x + \beta' y + \gamma' z = \delta'$$

sont parallèles si et seulement si (α, β, γ) et $(\alpha', \beta', \gamma')$ sont multiples l'un de l'autre. Pour prouver ces résultats et d'autres, nous introduisons des outils pratiques, en particulier la notion de produit scalaire !

1.3 Équations de cercles

Un cercle est l'ensemble des points d'un plan à distance donnée d'un centre.

1.3.1 Cercle dans le plan

Rappelons que la distance entre deux points $\mathbf{p} = (p_1, p_2)$ et $\mathbf{q} = (q_1, q_2)$ est donnée par

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2};$$

dès lors le cercle centré en (x_0, y_0) de rayon r du plan est donné par l'équation

$$\{(x, y) \text{ t.q. } (x - x_0)^2 + (y - y_0)^2 = r^2\}.$$

On peut également donner des équations paramétriques sous la forme :

$$\begin{aligned} x &= x_0 + r \cos(t) \\ y &= y_0 + r \sin(t) \end{aligned}$$

où t est le paramètre. Ici $t \in \mathbb{R}$ ou, plus simplement, $t \in [0, 2\pi[$

1.4 Équations d'autres objets dans l'espace

1.4.1 Cylindre

L'équation cartésienne d'un cylindre ressemble à celle d'un cercle dans le plan ; par exemple pour un cylindre dont l'axe est l'axe des « z » :

$$\{(x, y, z) \text{ t.q. } (x - x_0)^2 + (y - y_0)^2 = r^2\}$$

On peut également donner des équations paramétriques sous la forme :

$$\begin{aligned}x &= x_0 + r \cos(t) \\y &= y_0 + r \sin(t) \\z &= u\end{aligned}$$

où t, u sont les paramètres ; $t \in [0, 2\pi[$ et $u \in \mathbb{R}$.

1.4.2 Sphère

De même que pour un cercle dans le plan, la sphère est l'ensemble des points à distance r fixée d'un centre (x_0, y_0, z_0) fixé :

$$\{(x, y, z) \text{ t.q. } (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2\}.$$

Une version paramétrique classique est

$$\begin{aligned}x &= x_0 + r \cos \varphi \sin \theta \\y &= y_0 + r \sin \varphi \sin \theta \\z &= z_0 + r \cos \theta\end{aligned}$$

où φ est l'angle dans le plan des coordonnées xy (angle avec $(1, 0, 0)$) et θ est la colatitude (angle avec $(0, 0, 1)$). Les paramètres sont ici $\varphi \in [0, 2\pi[$ et $\theta \in [0, \pi]$.

2 Interprétations géométriques

2.1 Aires et volumes

Nous savons déjà que le produit scalaire ainsi que le produit vectoriel de deux vecteurs sont liés à la norme ainsi qu'à l'angle entre les deux vecteurs choisis.

Dans le résultat 4.1 page 69, nous avons noté que le produit scalaire d'un vecteur $\mathbf{v}$ avec un vecteur unitaire est égal à la longueur de la projection de $\mathbf{v}$ sur ce vecteur unitaire. Nous allons généraliser aux aires et volumes.

Résultat VIII.3. Dans $\mathbb{R}^3$, la norme $\|\mathbf{a} \times \mathbf{b}\|$ est égale à l'aire du parallélogramme construit sur $\mathbf{a}$ et $\mathbf{b}$.

Démonstration. En effet, projetons $\mathbf{a}$ sur $\mathbf{b}$, ce qui donne $\mathbf{a}' = \frac{\mathbf{a} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}} \mathbf{b}$ comme précédemment. Alors $\mathbf{a} - \mathbf{a}'$ représente la hauteur du parallélogramme considéré, de sorte que son aire est $\|\mathbf{a} - \mathbf{a}'\| \|\mathbf{b}\|$. Un calcul direct inspiré de la preuve précédente montre que le carré de cette norme vaut

$$\|\mathbf{a}\|^2 \|\mathbf{b}\|^2 - (\mathbf{a} \cdot \mathbf{b})^2$$

ce qui est le carré de la norme du produit vectoriel, comme nous l'avons vu auparavant. $\square$

Enfin, un résultat similaire peut encore se prouver pour les volumes de parallélépipèdes :

Résultat VIII.4. Dans $\mathbb{R}^3$, le produit mixte $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$ représente, en valeur absolue, le volume du parallélépipède construit sur les trois vecteurs $\mathbf{a}, \mathbf{b}, \mathbf{c}$.

produit mixte

2.2 Angles

Définition VIII.5. L'angle entre deux droites est donné par l'angle entre leurs vecteurs directeurs.

(Dans le plan, on peut également utiliser les vecteurs normaux.)

Définition VIII.6. L'angle entre deux plans de l'espace est donné par l'angle entre leurs vecteurs normaux.

Définition VIII.7. L'angle entre une droite et un plan de l'espace est donné par l'angle entre cette droite et sa projection orthogonale sur le plan.

2.3 Distances

Nous savons que la distance entre deux points $\mathbf{p}$ et $\mathbf{q}$ est la norme $\|\mathbf{q} - \mathbf{p}\|$. Notons $d(\mathbf{p}, \mathbf{q})$ cette quantité.

Définition VIII.8. La *distance* entre deux sous ensembles $E, F \subset \mathbb{R}^n$ est donnée par

distance

$$\min\{d(\mathbf{p}, \mathbf{q}) \text{ t.q. } \mathbf{p} \in E, \mathbf{q} \in F\}.$$

Résultat VIII.9. La distance entre le point $\mathbf{p}$ et la droite passant par $\mathbf{q}$ de vecteur directeur $\mathbf{v}$ est donnée par

$$\|\mathbf{q} - \mathbf{p}\|^2 - \frac{((\mathbf{p} - \mathbf{q}) \cdot \mathbf{v})^2}{\|\mathbf{v}\|^2}.$$

Le point de la droite réalisant ce minimum est donné par la projection $\mathbf{p}'$ de $\mathbf{p}$ sur la droite :

$$\mathbf{p}' = \mathbf{q} + \frac{(\mathbf{p} - \mathbf{q}) \cdot \mathbf{v}}{\|\mathbf{v}\|^2} \mathbf{v}.$$

Démonstration. Trouver le minimum de la distance revient à trouver le minimum du carré de la distance et puis à en prendre la racine carrée; donc il faut trouver le minimum de f définie par $f(t) = \|(\mathbf{q} + t\mathbf{v}) - \mathbf{p}\|^2$, où $\mathbf{q} + t\mathbf{v}$ est un point quelconque de la droite. Or

$$f(t) = \|\mathbf{q} - \mathbf{p}\|^2 + 2t(\mathbf{q} - \mathbf{p}) \cdot \mathbf{v} + t^2\|\mathbf{v}\|^2$$

est un polynôme du second degré en t dont le minimum s'obtient en général par $t = -\frac{(\mathbf{q} - \mathbf{p}) \cdot \mathbf{v}}{\|\mathbf{v}\|^2}$ ce qui donne la valeur annoncée

$$\|\mathbf{q} - \mathbf{p}\|^2 - \frac{((\mathbf{q} - \mathbf{p}) \cdot \mathbf{v})^2}{\|\mathbf{v}\|^2}.$$

Pour cette valeur de t , $\mathbf{q} + t\mathbf{v}$ est le point annoncé, ce qui correspond à la formule vue précédemment pour le projeté d'un point sur une droite. $\square$

Résultat VIII.10. La distance entre le point $\mathbf{p}$ et le plan de vecteur normal $\mathbf{w} = (\alpha, \beta, \gamma)$ passant par $\mathbf{q}$ est donnée par

$$\frac{|\mathbf{w} \cdot (\mathbf{p} - \mathbf{q})|}{\|\mathbf{w}\|}.$$

Démonstration. Le vecteur

$$\frac{\mathbf{w} \cdot (\mathbf{p} - \mathbf{q})}{\|\mathbf{w}\|^2} \mathbf{w}$$

est le projeté de $\mathbf{p} - \mathbf{q}$ sur la droite normale au plan, dès lors sa longueur est la distance recherchée. $\square$

Chapitre IX

Fonctions réelles : bases

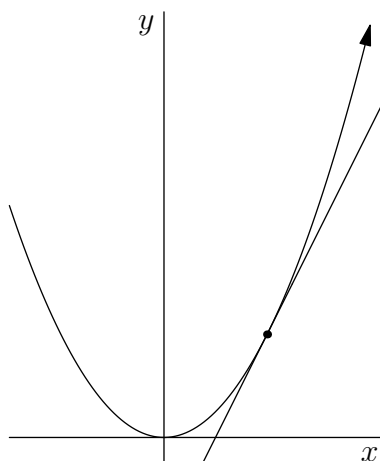
Le but de ce chapitre est d'énoncer différents concepts de base relatifs aux fonctions réelles d'une variable réelle, en particulier les concepts de

- * limite
- * continuité.

La notion de fonction est très générale : elle permet de représenter toute forme d'association ou de transformation (au moins tant que le hasard n'intervient pas). Ceci permet d'unir sous un même drapeau des champs d'intérêt très variés. En contrepartie, il est difficile d'établir des résultats concernant *toutes* les fonctions, car les « comportements » (quoique cela veuille dire) de ces fonctions sont à ce point variés. Nous devons alors nous restreindre à certaines classes de fonction.

Exemple. Nous savons généralement dessiner le graphe de certaines fonctions, par exemple la fonction qui à x associe x^2 se représente par une parabole car on dessine chaque point dont les coordonnées cartésiennes sont (x, x^2) pour une certaine valeur de x . Mais cette notion de graphe n'a aucun sens pour, par exemple, une fonction associant à chaque fruit sa couleur. Dessiner des graphes a du sens pour des fonctions dont le domaine est (un sous-ensemble de) $\mathbb{R}$ à valeurs dans $\mathbb{R}$, mais pas *n'importe quelle* fonction. Il faut donc se restreindre à cette classe de fonction si nous voulons dessiner des graphes.

Exemple. De même, parmi les fonctions dont nous savons dessiner le graphe, nous savons parfois dessiner la tangente à ce graphe en un certain point. Par exemple la droite dessinée ci-dessous est tangente à la parabole d'équation $y = x^2$.



Cependant, toutes les fonctions possédant un graphe n'admettent pas forcément des tangentes. Par exemple quelle serait la tangente au graphe de la fonction $x \mapsto |x|$ en $(0,0)$? Cela n'aurait pas de sens. Si des tangentes nous intéressent, il faudra donc nous restreindre un peu plus (en l'occurrence, aux fonctions dérivables).

Nous allons maintenant définir une telle classe de fonctions « agréables », à savoir les fonctions continues. Ces fonctions nous serviront de base pour introduire une classe de fonctions encore plus agréables au chapitre ?? page ?? : les fonctions dérivables.

1 Opérations sur les fonctions

restriction

Si $f : A \rightarrow B$ est une fonction et $A' \subseteq A$, alors f définit une fonction sur A' appelée *restriction* de f à A' et notée $f|_{A'}$. On a donc

$$f|_{A'} : A' \rightarrow B : x \mapsto f(x)$$

composée

La seule différence avec f est donc l'ensemble de départ.

Étant donné des fonctions $f : A \rightarrow B$ et $g : B \rightarrow C$, on peut définir la *composée* $g \circ f$ (lire « g rond f ») de ces fonctions par

$$g \circ f : A \rightarrow C : x \mapsto g(f(x))$$

Prenons maintenant $f : A \rightarrow B$, $g : A \rightarrow B$ deux fonctions réelles, c'est-à-dire telles que $B \subseteq \mathbb{R}$. Prenons également un nombre réel c . Alors on définit les fonctions réelles suivantes, toutes sur A (à l'exception du quotient, qui n'est pas défini quand le dénominateur est nul) :

somme

* La *somme* $f + g$ est définie par $(f + g)(x) := f(x) + g(x)$, pour $x \in A$.

produit par un scalaire

* Le *produit par un scalaire* cf est défini par $(cf)(x) := cf(x)$, pour $x \in A$.

produit

* Le *produit* fg est défini par $(fg)(x) := f(x)g(x)$, pour $x \in A$.

quotient

* Le *quotient* $\frac{f}{g}$ est défini par $\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$, pour $x \in A$ et $g(x) \neq 0$.

2 Fonctions élémentaires

élémentaire

Une fonction *élémentaire* est une fonction réelle d'une variable réelle construite en combinant des fonctions exponentielles, logarithmes, constantes ou racines n -èmes avec les opérations de composition, somme, produit et quotient (en restreignant le domaine à chaque fois que cela est nécessaire). L'ensemble des fonctions élémentaires est très vaste, et la fonction suivante en est un exemple parmi tant d'autres

$$]-\infty, 1[\rightarrow \mathbb{R} : x \mapsto \ln \sqrt{\frac{1-x}{1+x^2}} + 3e^{x^3+42}$$

Fonctions élémentaires « de base »

base

Soit b une *base*, c'est-à-dire un nombre réel positif et différent de 1. L'*exponentielle de base b* est la fonction

$$\mathbb{R} \rightarrow \mathbb{R}_0^+ : x \mapsto b^x$$

exponentielle de base b

Pour rappel, $\mathbb{R}^+ := \{x \in \mathbb{R} \text{ t.q. } x \geq 0\} = [0, \infty)$ est l'ensemble des réels positifs ou nuls et $\mathbb{R}_0^+ := \{x \in \mathbb{R} \text{ t.q. } x > 0\} =]0, \infty[$ est l'ensemble des réels strictement positifs. Définie ainsi, l'exponentielle de base b est une bijection. Sa fonction réciproque est le *logarithme de base b* :

$$\mathbb{R}_0^+ \rightarrow \mathbb{R} : x \mapsto \log_b x$$

logarithme de base b

Rappelons les identités les plus importantes relatives aux logarithmes. Pour tous réels positifs x, y , tout entier n (éventuellement négatif) et toute base c :

$$\log_b(b^x) = x$$

$$\log_b(xy) = \log_b x + \log_b y$$

$$\log_b(x^n) = n \log_b x$$

$$\log_b(x) = \frac{\log_c(x)}{\log_c(b)}$$

logarithme naturel

La base la plus utilisée est la base $e = 2.71828182\dots$. Le logarithme de base e est appelé *logarithme naturel* ou *logarithme népérien* (hommage à l'écossais John Napier) et noté $\ln$. En d'autres termes, le nombre $\ln(x)$ est l'unique nombre u tel que $e^u = x$.

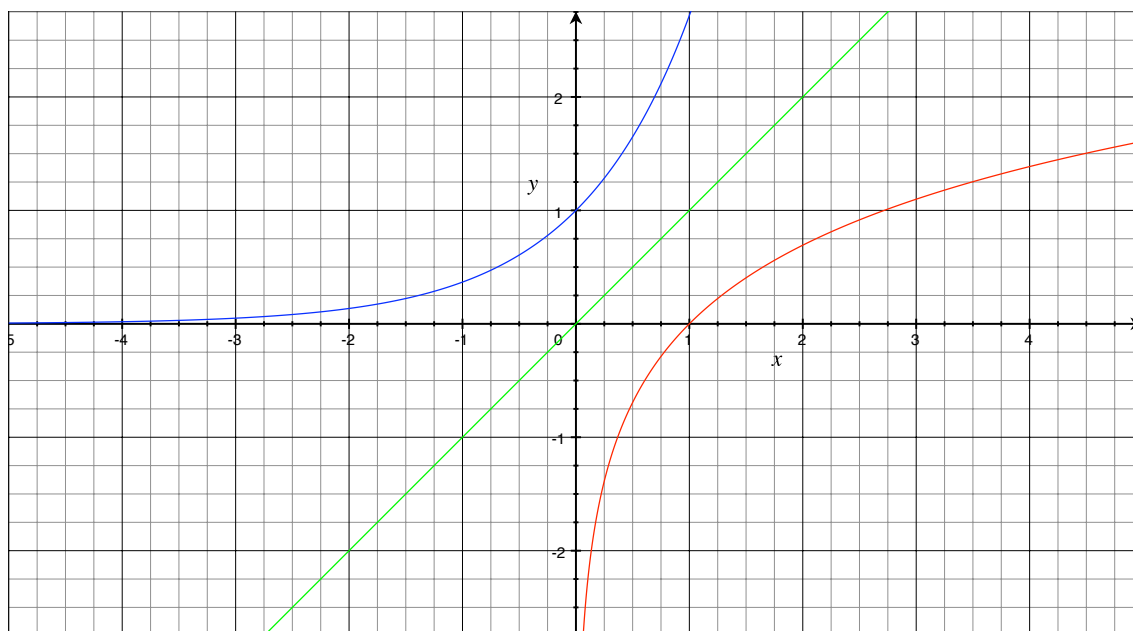
logarithme népérien

Une fonction *constante* est du type

$$\mathbb{R} \rightarrow \mathbb{R} : x \mapsto c$$

constante

avec $c \in \mathbb{R}$ (tous les points ont la même image).

FIGURE IX1 – Graphes de l'exponentielle et du logarithme en base e

Pour n un entier impair avec $n \geq 3$ et x un réel, on définit la *racine n -ème* de x , notée $\sqrt[n]{x}$ ou $x^{\frac{1}{n}}$ comme l'unique réel y tel que $y^n = x$. Si n est pair avec $n \geq 2$, on doit prendre $x \geq 0$ (car sinon il n'existe pas de racine n -ème) et $y \geq 0$ (car pour $x > 0$ il y a deux réels y tels que $y^n = x$, et il faut faire un choix). Si n est impair, on obtient une fonction racine n -ème

racine n -ème

$$\mathbb{R} \rightarrow \mathbb{R} : x \mapsto \sqrt[n]{x}$$

et si n est pair, on obtient une (autre) fonction racine n -ème

$$\mathbb{R}^+ \rightarrow \mathbb{R}^+ : x \mapsto \sqrt[n]{x}$$

Fonctions polynomiales

Une *fonction polynomiale* est une fonction élémentaire du type

fonction polynomiale

$$P : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \sum_{k=0}^n a_k x^k = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

où $a_0, a_1, a_2, \dots, a_n$ sont des constantes réelles appelées les *coefficients* de P . Le *degré* de P est n , à condition que $a_n \neq 0$. (Si P est identiquement nul, on définit parfois son degré comme $-\infty$.)

coefficients

degré

Une fonction polynomiale de degré 1 est aussi appelée fonction *linéaire* ou fonction *affine*. Le graphe d'une fonction linéaire est une droite (d'où le nom). Une fonction polynomiale de degré 2 est aussi appelée fonction *quadratique*. Le graphe d'une fonction quadratique est une parabole dont l'axe de symétrie est parallèle à l'axe des ordonnées.

linéaire

affine

quadratique

Fonctions rationnelles

Une *fonction rationnelle* est le quotient de deux fonctions polynomiales. C'est une fonction élémentaire définie sur $\mathbb{R}$ moins l'ensemble des racines du dénominateur. Si P et Q sont deux fonctions polynomiales, leur quotient R est la fonction rationnelle définie par

fonction rationnelle

$$R : A \rightarrow \mathbb{R} : x \mapsto \frac{P(x)}{Q(x)}$$

avec $A := \{x \in \mathbb{R} \text{ t.q. } Q(x) \neq 0\}$. Un exemple simple de fonction rationnelle est la fonction

$$\mathbb{R}_0 \rightarrow \mathbb{R} : x \mapsto \frac{1}{x}$$

Un autre exemple plus compliqué est la fonction

$$\mathbb{R} \rightarrow \mathbb{R} : x \mapsto \frac{x^3 - 5x + 7}{x^2 + 1}$$

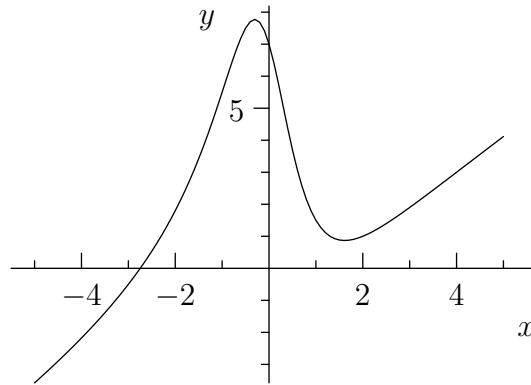


FIGURE IX2 – Graphe de $x \mapsto \frac{x^3 - 5x + 7}{x^2 + 1}$

Fonctions trigonométriques

cercle
trigonométrique)

Rappelons comment sont définies les fonctions trigonométriques $\sin$, $\cos$, tg et cotg . Dans le plan muni d'un repère cartésien, on considère le cercle de rayon 1 centré en l'origine, ou *cercle trigonométrique*. Soit $p = (x, y)$ un point du cercle trigonométrique et soit θ la mesure en radians de l'angle orienté entre les demi-droites ox et op (voir dessin).

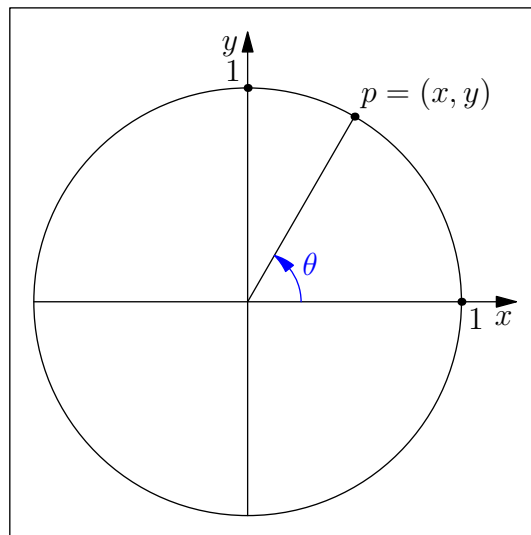


FIGURE IX3 – Le cercle trigonométrique

On définit :

$$\cos \theta := x \text{ et } \sin \theta := y$$

$$\text{tg } \theta := \frac{\sin \theta}{\cos \theta}, \quad \theta \neq n\pi + \frac{\pi}{2}, \quad n \in \mathbb{Z}$$

$$\cotg \theta := \frac{\cos \theta}{\sin \theta}, \quad \theta \neq n\pi, \quad n \in \mathbb{Z}$$

On obtient ainsi quatre fonctions trigonométriques :

$$\begin{aligned} \mathbb{R} &\rightarrow [-1, 1] : \theta \mapsto \cos \theta \\ \mathbb{R} &\rightarrow [-1, 1] : \theta \mapsto \sin \theta \\ \mathbb{R} \setminus \left\{ \dots, -\frac{3\pi}{2}, -\frac{\pi}{2}, \frac{\pi}{2}, \frac{3\pi}{2}, \dots \right\} &\rightarrow \mathbb{R} : \theta \mapsto \tg \theta \\ \mathbb{R} \setminus \{ \dots, -2\pi, -\pi, 0, \pi, 2\pi, \dots \} &\rightarrow \mathbb{R} : \theta \mapsto \cotg \theta \end{aligned}$$

Rappelons maintenant les propriétés les plus importantes de ces fonctions. En exploitant certaines des symétries du cercle trigonométrie, on trouve, pour tout $\theta \in \mathbb{R}$:

Réflexion d'axe ox	$\sin(-\theta) = -\sin \theta$	$\cos(-\theta) = \cos \theta$
Réflexion d'axe $x = y$	$\sin\left(\frac{\pi}{2} - \theta\right) = \cos \theta$	$\cos\left(\frac{\pi}{2} - \theta\right) = \sin \theta$
Réflexion d'axe oy	$\sin(\pi - \theta) = \sin \theta$	$\cos(\pi - \theta) = -\cos \theta$
Rotation de $\frac{\pi}{2}$	$\sin\left(\theta + \frac{\pi}{2}\right) = \cos \theta$	$\cos\left(\theta + \frac{\pi}{2}\right) = -\sin \theta$
Rotation de π	$\sin(\theta + \pi) = -\sin \theta$	$\cos(\theta + \pi) = -\cos \theta$
Rotation de 2π	$\sin(\theta + 2\pi) = \sin \theta$	$\cos(\theta + 2\pi) = \cos \theta$

De plus, on a les identités suivantes (pour tout $a, b \in \mathbb{R}$) :

$$\begin{aligned} \sin^2 a + \cos^2 a &= 1 \\ \cos(a - b) &= \cos a \cos b + \sin a \sin b \\ \cos(a + b) &= \cos a \cos b - \sin a \sin b \\ \sin(a + b) &= \sin a \cos b + \cos a \sin b \\ \sin(a - b) &= \sin a \cos b - \cos a \sin b \end{aligned}$$

et on peut en déduire d'autres identités dont par exemple les formules de Simpson (valides pour tout $p, q \in \mathbb{R}$) :

$$\begin{aligned} \cos p \cos q &= \frac{1}{2}(\cos(p + q) + \cos(p - q)) \\ \sin p \cos q &= \frac{1}{2}(\sin(p + q) + \sin(p - q)) \\ \sin p \sin q &= \frac{1}{2}(\cos(p - q) - \cos(p + q)) \end{aligned}$$

Fonctions trigonométriques réciproques

Arcsinus La fonction $\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow [-1, 1]$ a une fonction réciproque,

$$\arcsin : [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

Arccosinus La réciproque de $\cos : [0, \pi] \rightarrow [-1, 1]$ est

$$\arccos : [-1, 1] \rightarrow \left[0, \pi\right]$$

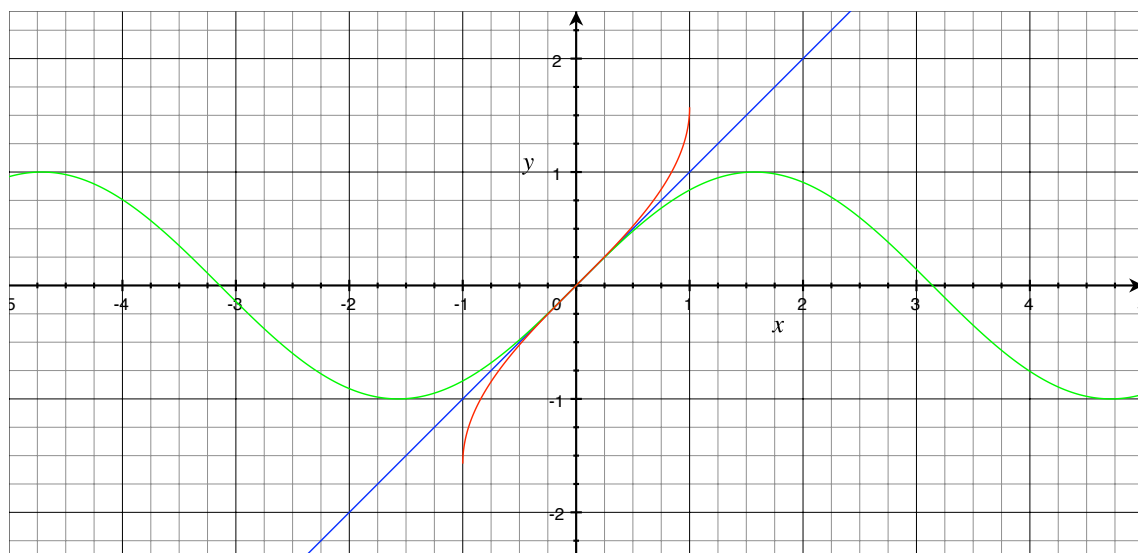


FIGURE IX4 – Graphes de sin et arcsin

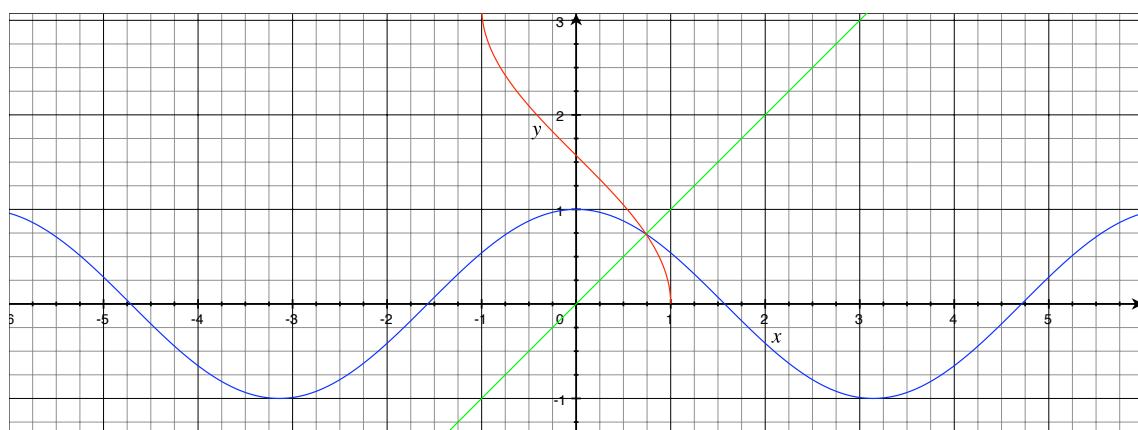


FIGURE IX5 – Graphes de cos et arccos

Arctangente La réciproque de $\text{tg} :]-\frac{\pi}{2}, \frac{\pi}{2}[\rightarrow]-\infty, \infty[$ est

$$\text{arctg} :]-\infty, \infty[\rightarrow]-\frac{\pi}{2}, \frac{\pi}{2}[$$

On peut aussi définir la réciproque de $\text{cotg} :]0, \pi[\rightarrow]-\infty, \infty[$.

3 Limites

Distances

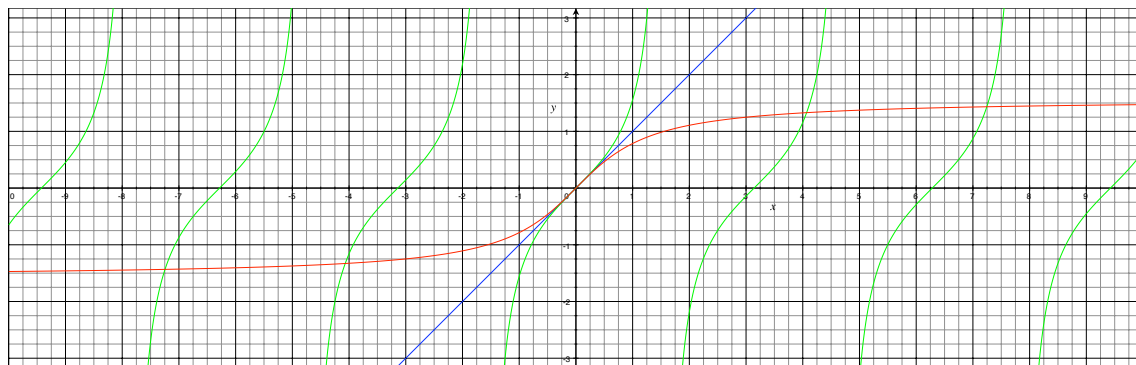
La distance entre deux réels quelconques x et y est $|x - y|$, la valeur absolue de leur différence. Pour un nombre $\delta > 0$, demander $|x - y| < \delta$ revient donc à demander que la distance entre x et y soit inférieure à δ . On a les équivalences suivantes :

$$|x - y| < \delta \iff x \in]y - \delta, y + \delta[\iff y - \delta < x < y + \delta$$

Points adhérents

Définition IX.1. Un réel a est dit *adhérent* à un sous-ensemble A de $\mathbb{R}$ si pour tout $\delta > 0$ il existe un point x de A tel que $|x - a| < \delta$. On note $\text{adh } A$ l'ensemble des points adhérents à A .

adhérent

FIGURE IX6 – Graphes de tg et arctg

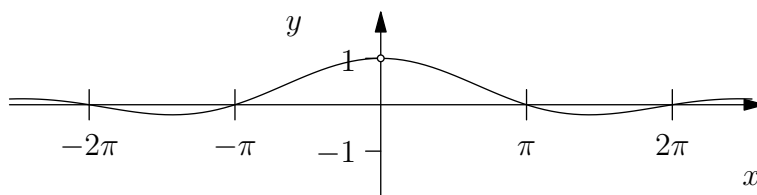
Ceci est équivalent à exiger que tout intervalle ouvert non-vidé centré en a contienne au moins un point de A . Ceci équivaut également à demander que pour tout rayon $\delta > 0$ il existe un point x de A situé à une distance inférieure à δ de a .

Exemple. Les points adhérents à l'ensemble ... forment l'ensemble ... :

$]0, 1]$	$[0, 1]$
$[0, 1[$	$[0, 1]$
$[0, 1]$	$[0, 1]$
$]0, 1[$	$[0, 1]$
$\mathbb{R}$	$\mathbb{R}$
$]-\infty, 2[$	$]-\infty, 2]$
$\mathbb{R}_0$	$\mathbb{R}$
$\mathbb{Q}$	$\mathbb{R}$
$\mathbb{N}$	$\mathbb{N}$

Définition de limite

Exemple. Considérons la fonction $f : \mathbb{R}_0 \rightarrow \mathbb{R} : x \mapsto \frac{\sin x}{x}$. Son graphe :

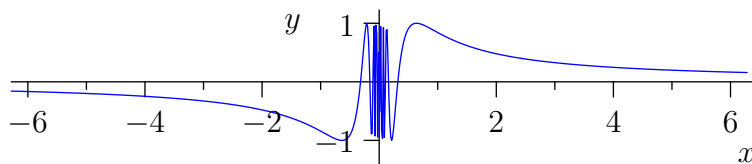


La valeur $f(0)$ n'a aucun sens car 0 n'est pas dans le domaine, car $\frac{0}{0}$ n'a aucun sens.

Mais on a envie de dire $f(0) = 1$.

À la place, on dira $\lim_{x \rightarrow 0} f(x) = 1$, et on va définir cela rigoureusement.

Exemple. Considérons la fonction $f : \mathbb{R}_0 \rightarrow \mathbb{R} : x \mapsto \sin(1/x)$. Son graphe :



La valeur $f(0)$ n'a aucun sens car 0 n'est pas dans le domaine, car $\frac{1}{0}$ n'a aucun sens.

Quelle valeur a-t-on envie que $f(0)$ prenne? Aucune, mais à la place, on dira $\lim_{x \rightarrow 0} f(x)$ n'existe pas.

Définition IX.2. Soit $f : A \rightarrow B$ où $A, B \subseteq \mathbb{R}$ et soit a un point adhérent à A . Un nombre réel L est la limite de f en a si

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x \in A, |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon.$$

Si L est la limite de f en a , on écrit $\lim_{x \rightarrow a} f(x) = L$. Remarquons que cette notation n'est vraiment raisonnable que si la limite est unique. Il se fait que c'est le cas (si celle-ci existe).

Intuitivement, L est la limite de f en a si $f(x) \approx L$ dès que $x \approx a$ (le signe $\approx$ correspond, ici, à l'idée intuitive d'approximer). Dans l'énoncé formel donné ci-dessus, $f(x) \approx L$ correspond à $|f(x) - L| < \epsilon$ et $x \approx a$ correspond à $|x - a| < \delta$. On peut ainsi se faire une bonne idée de ce que signifie la définition : ϵ est une précision à garantir et δ un rayon « de sécurité ». Dès que x est à une distance inférieure à δ de a , alors l'erreur absolue commise en remplaçant $f(x)$ par la limite L est inférieure à ϵ .

L'ordre des quantificateurs « pour tout » et « il existe » est importante. On choisit d'abord une précision ϵ et puis ensuite un rayon δ adapté à cette précision. Typiquement, au plus la précision ϵ est petite, au plus il faudra choisir le rayon δ petit.

Exemple. Montrons que

$$\lim_{x \rightarrow 0} \sqrt{x} = 0$$

en utilisant la définition. Comment doit-on choisir $\delta = \delta(\epsilon)$ pour que $|x - 0| < \delta$ garantisse $|\sqrt{x} - 0| < \epsilon$? (Evidemment, on doit prendre x dans le domaine, donc $x \geq 0$ ici.) Réponse : il suffit de prendre $\delta = \epsilon^2$, car

$$|x - 0| < \epsilon^2 \Rightarrow x < \epsilon^2 \Rightarrow \sqrt{x} < \epsilon \Rightarrow |\sqrt{x} - 0| < \epsilon$$

pour $x \geq 0$. Pour garantir une précision de $\epsilon = 0,01 = 10^{-2}$ on peut prendre $\delta = \epsilon^2 = 10^{-4} = 0,0001$.

Remarques

Définition IX.3. S'il n'existe pas de $L \in \mathbb{R}$ tel que L est la limite de f en a est L , on dit que la limite de f en a n'existe pas (dans $\mathbb{R}$).

Résultat IX.4. Si L_1 et L_2 sont deux nombres vérifiant la définition de « limite de f en a , alors $L_1 = L_2$

Démonstration. Supposons $L_1 > L_2$ (le cas inverse est semblable). Dès lors $\epsilon = (L_1 - L_2)/2 > 0$.

Alors il existe $\delta_1 > 0$ tel que

$$\forall x \in A, |x - a| < \delta_1 \Rightarrow |f(x) - L_1| < \epsilon.$$

et il existe $\delta_2 > 0$ tel que

$$\forall x \in A, |x - a| < \delta_2 \Rightarrow |f(x) - L_2| < \epsilon.$$

et il existe $x \in A$ avec $|x - a| < \min\{\delta_1, \delta_2\}$.

Pour ce x , on a :

$$-(L_1 - L_2)/2 < f(x) - L_1 \quad f(x) - L_2 < (L_1 - L_2)/2$$

c'est-à-dire :

$$(L_1 + L_2)/2 < f(x) < (L_1 + L_2)/2$$

d'où une contradiction! □

Remarque (*). Si $a \in A$, et la limite de f en a existe, alors nécessairement elle est égale à $f(a)$. Cela est dû au fait qu'on peut prendre $x = a$ dans la définition. Plus précisément nous avons le résultat suivant.

Résultat IX.5. Soit $f : A \rightarrow \mathbb{R}$ et $a \in A$. Si $\lim_{x \rightarrow a} f(x)$ existe, alors elle vaut $f(a)$.

Démonstration. Supposons que la limite vaut L :

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x \in A, |x - a| < \delta \Rightarrow |f(x) - L| < \epsilon.$$

Comme $a \in A$, on peut prendre $x = a$. Comme $|0 - 0| < \delta$ est toujours vrai, on a forcément $|f(a) - L| < \epsilon$. Pour tout $\epsilon > 0$! Donc $f(a) = L$. □

Limite épointée, limites latérales

Les limites

$$\lim_{\substack{x \rightarrow a \\ x \neq a}} f(x) \quad , \quad \lim_{\substack{x \rightarrow a \\ x > a}} f(x) \quad \text{et} \quad \lim_{\substack{x \rightarrow a \\ x < a}} f(x)$$

sont définies de manière similaire à

$$\lim_{x \rightarrow a} f(x)$$

Par exemple, pour $L \in \mathbb{R}$, on dit que la *limite épointée de f en a* est L si

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x \in A, \begin{cases} |x - a| < \delta, \text{ et} \\ x \neq a \end{cases} \implies |f(x) - L| < \epsilon.$$

limite épointée
de f en a

ce qui s'écrit

$$\lim_{\substack{x \rightarrow a \\ x \neq a}} f(x) = L$$

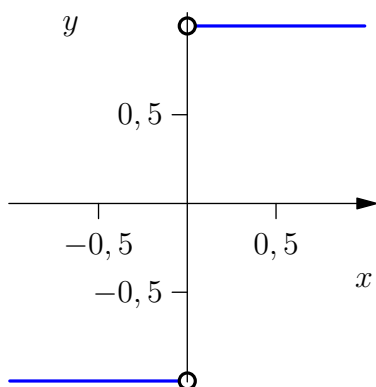
Les *limites latérales de f en a* $\lim_{\substack{x \rightarrow a \\ x > a}} f(x)$ et $\lim_{\substack{x \rightarrow a \\ x < a}} f(x)$ se définissent de la même manière : remplacer $x \neq a$ dans la définition ci-dessus par $x < a$ ou $x > a$. Ces limites se notent également $\lim_{x \rightarrow a^+} f(x)$ et $\lim_{x \rightarrow a^-} f(x)$.

limites latérales
de f en a

Exemple. La limite lorsque $x \rightarrow 0$ de $|x|/x$ n'existe pas, mais

$$\lim_{\substack{x \rightarrow 0 \\ x < 0}} \frac{|x|}{x} = -1$$

$$\lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{|x|}{x} = 1$$



Résultat IX.6. $\lim_{\substack{x \rightarrow a \\ x \neq a}} f(x)$ existe si et seulement si $\lim_{\substack{x \rightarrow a \\ x > a}} f(x)$ et $\lim_{\substack{x \rightarrow a \\ x < a}} f(x)$ existent et sont égales.

Résultat IX.7. Lorsque $a \in A$, $\lim_{x \rightarrow a} f(x)$ existe si et seulement si la limite épointée existe et vaut $f(a)$.

Limites « restreintes »

De manière générale, on notera

$$\lim_{\substack{x \rightarrow a \\ \text{condition sur } x}} f(x) = L$$

lorsque

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x \in A, \begin{cases} |x - a| < \delta, \text{ et} \\ \text{condition sur } x \end{cases} \implies |f(x) - L| < \epsilon.$$

Limites infinies, limites à l'infini

Nous allons maintenant nous intéresser aux cas où $L = \pm\infty$ ou $a = \pm\infty$, c'est-à-dire qu'on s'intéresse à décrire la situation où la variable ou la fonction « tendent vers l'infini ». Il n'y a pas de difficulté particulière à adapter notre définition de l'égalité

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

dans ces cas où $L = \pm\infty$ ou $a = \pm\infty$. Il suffit de traduire $f(x) \approx L$ et $x \approx a$ par les conditions adéquates sur $f(x)$ et x , comme ci-dessous. Il est important, cependant, de noter que chaque utilisation du symbole ∞ s'accompagne d'une définition ! L'infini n'intervient pas comme un cas particulier de ce qu'on a vu : ce n'est pas un nombre, c'est un concept intuitif auquel il faut donner consistance par une définition appropriée.

Par exemple lorsque $L = \infty$, la condition $|f(x) - L| < \epsilon$ devient $f(x) > M$ (où M est un nombre qu'on imagine grand), et $|x - a| < \delta$ devient $x < -N$ si $a = -\infty$ (où N est un nombre qu'on imagine grand). Lorsque a est infini (positif ou négatif), il faut également demander un analogue d'être « adhérent » pour l'infini : cela revient à dire qu'il existe des nombres aussi grands (si $a = +\infty$) ou aussi petits (si $a = -\infty$) que voulu. Pour l'exemple, écrivons l'un de ces nouvelles définitions :

Définition IX.8. Soit $f : A \rightarrow \mathbb{R}$ une fonction réelle dont le domaine n'est pas minoré. On dit que $f(x)$ tend vers ∞ pour x tendant vers $-\infty$ si l'affirmation suivante est vraie :

$$\forall M > 0, \exists N > 0 : \forall x \in A : x < -N \Rightarrow f(x) > M$$

on note cette situation $\lim_{x \rightarrow -\infty} f(x) = \infty$.

Les autres cas de limite infinie / à l'infini s'approchent de la même façon.

Limites et opérations

Si les deux limites, $\lim_{x \rightarrow a} f(x)$ et $\lim_{x \rightarrow a} g(x)$ existent dans $\mathbb{R}$, alors $\lim_{x \rightarrow a} [f(x) + g(x)]$, $\lim_{x \rightarrow a} cf(x)$, $\lim_{x \rightarrow a} [f(x) - g(x)]$ et $\lim_{x \rightarrow a} f(x)g(x)$ existent aussi et on a

$$\begin{aligned} \lim_{x \rightarrow a} [f(x) + g(x)] &= \lim_{x \rightarrow a} f(x) + \lim_{x \rightarrow a} g(x) \\ \lim_{x \rightarrow a} cf(x) &= c \lim_{x \rightarrow a} f(x), \text{ pour n'importe quelle constante } c \in \mathbb{R} \\ \lim_{x \rightarrow a} [f(x) - g(x)] &= \lim_{x \rightarrow a} f(x) - \lim_{x \rightarrow a} g(x) \\ \lim_{x \rightarrow a} f(x)g(x) &= \lim_{x \rightarrow a} f(x) \lim_{x \rightarrow a} g(x) \end{aligned}$$

Si, de plus $\lim_{x \rightarrow a} g(x) \neq 0$, alors $\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$ existe et on a

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

Enfin, si on a une fonction composée $g \circ f$ et si on sait que $\lim_{x \rightarrow a} f(x) = L$ et que $\lim_{t \rightarrow L} g(t)$ existe alors $\lim_{x \rightarrow a} g(f(x))$ existe et on a

$$\lim_{x \rightarrow a} g(f(x)) = \lim_{t \rightarrow L} g(t)$$

Remarque. En d'autres termes, on peut poser $t =$ fonction de x , et remplacer tous les x par des t , et considérer la limite pour t .

Exemple. Supposons savoir que $\sin x/x$ tend vers 1 en 0.

Alors

$$\lim_{x \rightarrow 0} \frac{\sin(3x)}{x} = 3 \lim_{x \rightarrow 0} \frac{\sin(3x)}{3x} = 3 \lim_{t \rightarrow 0} \frac{\sin t}{t} = 3$$

Exemple. Calculons :

$$\lim_{x \rightarrow 1} \frac{\sin(x^2 - 1)}{x - 1} = \lim_{x \rightarrow 1} \frac{(x + 1) \sin(x^2 - 1)}{(x - 1)(x + 1)} = \lim_{x \rightarrow 1} (x + 1) \lim_{x \rightarrow 1} \frac{\sin(x^2 - 1)}{x^2 - 1} = \lim_{x \rightarrow 1} (x + 1) \lim_{t \rightarrow 0} \frac{\sin t}{t} = 2$$

Un résultat naturel est le suivant :

Résultat IX.9 (Théorème du sandwich). *si $f, g, h : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ sont trois fonctions à valeurs réelles telles que*

$$\forall x \in A, f(x) \leq g(x) \leq h(x)$$

et pour lesquelles les limites $\lim_{x \rightarrow a} f(x)$ et $\lim_{x \rightarrow a} h(x)$ sont identiques. Alors la limite de g vaut cette limite commune :

$$\lim_{x \rightarrow a} g(x) = \lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} h(x).$$

Démonstration. On sait :

$$\forall \epsilon > 0, \exists \delta_1 > 0 : \forall x (|x - a| < \delta_1 \rightarrow |f(x) - L| < \epsilon)$$

et similairement pour h avec un δ_2 .

Soit $\epsilon > 0$. On prend δ_1 et δ_2 tels que ci-dessus.

On définit alors $\delta = \min(\delta_1, \delta_2)$. Alors pour $|x - a| < \delta$ on a :

$$|f(x) - L| < \epsilon \text{ et } |h(x) - L| < \epsilon$$

c'est-à-dire

$$-\epsilon < f(x) - L < \epsilon \text{ et } -\epsilon < h(x) - L < \epsilon$$

donc on a :

$$-\epsilon < f(x) - L \leq g(x) - L \leq h(x) - L < \epsilon$$

dès lors $|g(x) - L| < \epsilon$ comme attendu. □

Notons que ce résultat vaut encore lorsque a est infini et lorsque la limite commune de g et h est infinie. Pouvez-vous en faire la preuve ?

Un cas particulier du résultat est le suivant : si $|f(x)| \leq g(x)$ pour tout $x \in A$, et si g tend vers 0, alors f tend également vers 0.

3.1 Règles étendues

Les règles de calcul précédentes restent valables, sans modification, si on prend $a = +\infty$ ou $a = -\infty$.

Dans une certaine mesure, il est également possible de conserver ces règles lorsque certaines des limites sont $\pm\infty$. Il faut cependant prendre des *conventions*. Ces conventions sont une manière compacte d'énoncer des résultats sur les limites, qui peuvent se démontrer — elles ne sont en aucun cas des égalités ayant a priori du sens :

$$\begin{aligned} \infty + \infty &= \infty \\ (-\infty) + (-\infty) &= -\infty \\ c \cdot (\pm\infty) &= \begin{cases} \pm\infty & \text{si } c > 0 \\ \mp\infty & \text{si } c < 0 \end{cases} \text{ (remarquez le changement de signe.)} \\ \infty \cdot \infty &= \infty \\ \infty \cdot (-\infty) &= -\infty \\ (-\infty) \cdot (-\infty) &= \infty \end{aligned}$$

Pour le quotient, on pourra aussi étendre en prenant les nouvelles conventions suivantes :

$$\frac{L}{0^+} = \begin{cases} +\infty & \text{si } L > 0 \\ -\infty & \text{si } L < 0 \end{cases}$$

$$\frac{L}{0^-} = \begin{cases} -\infty & \text{si } L > 0 \\ +\infty & \text{si } L < 0 \end{cases}$$

$$\frac{\pm\infty}{0^+} = \pm\infty$$

$$\frac{\pm\infty}{0^-} = \mp\infty \text{ (remarquez le changement de signe.)}$$

Ci-dessus, les notations 0^+ et 0^- se comprennent comme suit : une limite qui vaut 0^+ signifie que $f(x) \rightarrow 0$ pour $x \rightarrow a$ et $f(x) \geq 0$ pour $x \approx a$. Formellement, la condition $f(x) \geq 0$ pour $x \approx a$ se définit par : il existe $\epsilon > 0$ tel que pour tout $x \in A$ vérifiant $|x - a| < \epsilon$ on a $f(x) \geq 0$. En d'autres termes « il existe un intervalle ouvert autour de a sur lequel f ne prend que des valeurs positives. » La notation 0^- se définit de manière analogue.

Exemple. La fonction $x \mapsto x$ a une limite nulle en 0. On peut cependant être plus précis et écrire :

$$\lim_{\substack{x \rightarrow 0 \\ <}} x = 0^- \qquad \lim_{\substack{x \rightarrow 0 \\ >}} x = 0^+$$

(ceci ne contredit pas la proposition IX.6, car la limite est bien 0 dans chaque cas ; la différence entre gauche et droite est qu'on décore ce 0 d'un signe qui donne une information supplémentaire.)

Les limites non couvertes par ce qui précède doivent être étudiées au cas-par-cas. On dit qu'elles relèvent de cas d'*indétermination* ou de *forme indéterminée*. Les cas d'indétermination les plus fréquents quand on veut appliquer les règles ci-dessus sont les suivants :

indétermination

forme
indéterminée

$$\infty - \infty, 0 \cdot \infty, 1^\infty, \infty^0, \frac{0}{0} \text{ et } \frac{\infty}{\infty}.$$

Exemple.

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

On ne peut pas appliquer la règle de calcul du quotient ! Il faut écrire :

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

Exemple.

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

On ne peut pas appliquer la règle de calcul du quotient ! Il faut écrire :

$$\lim_{x \rightarrow \infty} \frac{1 + 1/x}{1 - 1/x} = \lim_{x \rightarrow \infty} \frac{1 + 0}{1 - 0} = 1$$

3.2 Notation de Landau

Il arrive souvent qu'on s'intéresse à comparer des fonctions « sur le long terme ».

Exemple. Deux populations de bactéries peuvent avoir le même nombre d'individus au début, mais ce nombre va-t-il rester comparable tout le temps ?

Exemple. Si deux algorithmes pour factoriser le nombre n prennent respectivement $f(n)$ et $g(n)$ secondes, comment savoir lequel est le plus rapide lorsque n devient grand ?

Nous n'allons pas répondre à ces (vagues) questions, mais nous présentons une définition utilisable dans ce cadre.

Définition IX.10. Soient f et g deux fonctions. Supposons qu'il existe $r \in \mathbb{R}$ et $c \in \mathbb{R}_0^+$ tels que pour tout $x \in]r, +\infty[$ on a

$$x \in \text{dom } f, x \in \text{dom } g \text{ et } |f(x)| \leq c \cdot |g(x)|$$

On dit alors « f est un $O(g)$ » (prononcer « f est un grand O de g »).

Remarque. Cette notion indique essentiellement que f/g reste borné. Généralement on compare une fonction f intéressante à une fonction g bien connue. On dira aussi que l'ordre de grandeur de f est inférieur à celui de g .

Exemple. * $10x$ est $O(x)$. En effet, si $x > 0$, on a bien $10x \leq cx$ (par exemple $c = 10$ fonctionne).

- * Inversement, x est $O(10x)$ car $x \leq c10x$ pour un certain c .
- * $\sin(x) + x^2$ est en $O(x^2)$ car $\sin(x) + x^2 \leq x^2 + x^2 = 2x^2$ si $x > 2$.
- * x^2 est $O(x)$ mais x n'est pas $O(x^2)$.

Résultat IX.11. Considérons la limite $\lim_{x \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right|$.

- * Si elle existe dans $\mathbb{R}$, alors $f(x)$ est $O(g(x))$.
- * Si elle est infinie, alors $f(x)$ n'est pas $O(g(x))$.

Démonstration. Si la limite existe et vaut $L \in \mathbb{R}$, on prend $\epsilon = 1$ dans la définition de limite. Ceci assure alors l'existence d'un $N > 0$ tel que pour tout $x \geq N$ on ait

$$\left| \left| \frac{f(x)}{g(x)} \right| - |L| \right| < 1$$

Dès lors $\left| \frac{f(x)}{g(x)} \right| < |L| + 1$, ce qui prouve l'affirmation en prenant $c = |L| + 1$ (et $r = N$).

Dans le second cas, quel que soit $M > 0$, on sait qu'il existe $N > 0$ tel que pour $x > N$, on a $\left| \frac{f(x)}{g(x)} \right| > M$. En particulier il ne peut pas exister de c vérifiant la définition de « grand O ». $\square$

Exemple. Pour tout réel A : Ax^n est $O(x^k)$ si et seulement si $n \leq k$. En d'autres termes : x^n est d'un ordre de grandeur inférieur à x^k si et seulement si $n \leq k$.

En effet le quotient de la proposition précédente vaut $|Ax^{n-k}|$:

- * Si $n \leq k$ alors l'exposant $n - k$ est négatif ou nul, donc le quotient tend vers 0 ou $|A|$.
- * Si $n > k$, alors l'exposant est strictement positif et donc le quotient tend vers $+\infty$.

Une définition similaire est la suivante :

Définition IX.12. Lorsque $\lim_{x \rightarrow \infty} \left| \frac{f(x)}{g(x)} \right| = 0$, on dit alors que $f(x)$ est $o(g(x))$.

Remarque. * En particulier $f(x)$ est en $O(g(x))$ mais ceci est plus fort.

- * Si $g(x)$ tend vers 0 et $f(x)$ est un $o(g(x))$, alors f tend vers 0 plus vite que g !
- * Les définitions de o et O tiennent encore lorsque x tend vers $a \in \mathbb{R}$ et vers $-\infty$. Il faut alors préciser quelle est la limite pour x !

Exemple. * x^2 est un $o(x^3)$ (pour $x \rightarrow \infty$),

- * x^3 n'est pas un $o(x^2)$ (pour $x \rightarrow \infty$),
- * x^3 est un $o(x^2)$ (pour $x \rightarrow 0$),
- * x^2 n'est pas un $o(x^3)$ (pour $x \rightarrow 0$).

3.3 Une lemme pour les limites

Nous aurons ultérieurement besoin de ce résultat sur les limites. Il peut être passé en première lecture.

Résultat IX.13. Soit g une fonction réelle. Si la limite $\lim_{x \rightarrow a} g(x)$ existe et vaut $L > 0$, alors il existe δ tel que pour tout $x \in]a - \delta, a + \delta[$, l'inégalité $g(x) > 0$ est vérifiée.

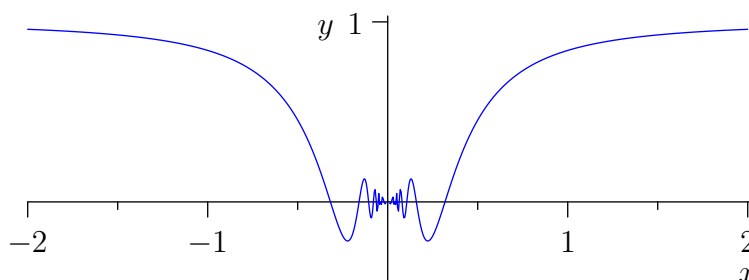
Démonstration. On prend $\epsilon = L$ dans la définition de limite, et on obtient $\delta > 0$ tel que pour $x \in]a - \delta, a + \delta[$, $|g(x) - L| < \epsilon = L$ dont on tire

$$-L < g(x) - L < L$$

Et en particulier $g(x) > 0$ (inégalité de gauche). $\square$

Remarque (*). Soit g une fonction réelle pour laquelle $\lim_{x \rightarrow a} g(x) \geq 0$, il est possible que $g(x) < 0$ pour des valeurs de x arbitrairement proches de 0!

Exemple. Pour $g(x) = x \sin(1/x)$, on a $\lim_{x \rightarrow 0} g(x) = 0$.



Néanmoins, il n'existe pas d'intervalle autour de 0 dans lequel g serait positive.

4 Continuité

Définition IX.14. Soit $f : A \rightarrow B$ une fonction réelle et $a \in A$. La fonction f est *continue au point* a si

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

Définition IX.15. f est *discontinue au point* a si f n'est pas continue au point a .

Définition IX.16. f est *continue (sur son domaine)* si elle est continue en chaque point a de son domaine. Par exemple, la fonction

$$f : \mathbb{R}_0 \rightarrow \mathbb{R} : x \mapsto \frac{1}{x}$$

est continue car f est continue en chaque $a \in \mathbb{R}_0$.

Définition IX.17. Une fonction f est *discontinue* si elle n'est pas continue, c'est-à-dire s'il existe au moins un point a de son domaine en lequel f est discontinue.

Exemple. $g : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^2$ et $h : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto |x|$ sont continues, $i : \mathbb{R} \rightarrow \mathbb{Z} : x \mapsto [x]$ est discontinue en a pour $a \in \mathbb{Z}$ et continue en a pour $a \in \mathbb{R} \setminus \mathbb{Z}$. Enfin, la fonction caractéristique des rationnels est discontinue en chaque point $a \in \mathbb{R}$.

Continuité et opérations

Soient deux fonctions f et g continues en un point a . Soit $c \in \mathbb{R}$ une constante. Alors

$f + g$ est continue en a

cf est continue en a

fg est continue en a

continue au point a

discontinue au point a

continue (sur son domaine)

discontinue

Si, de plus, $g(a) \neq 0$, alors

$$\frac{f}{g} \text{ est continue en } a.$$

Ensuite, supposons $f : A \rightarrow B$ et $g : B \rightarrow C$, avec $A, B, C \subseteq \mathbb{R}$. Donc la composée $g \circ f$ est bien définie comme fonction de A dans C . Si f est continue en a et g est continue en $f(a)$ alors

$$g \circ f \text{ est continue en } a.$$

Intervalles

Définition IX.18. Un *intervalle* réel est un sous-ensemble de $\mathbb{R}$ « sans trou », plus précisément d'une intervalle des formes suivantes :

$$]a, b[, \quad]a, b], \quad [a, b[, \quad [a, b], \quad]-\infty, b[, \quad]-\infty, b], \quad]a, \infty[, \quad [a, \infty[, \quad \mathbb{R}$$

pour a, b des réels. Un intervalle d'un des quatre premiers types est dit borné; les autres sont non-bornés.

Propriétés importantes des fonctions continues

Définition IX.19. Soit $E \subset \mathbb{R}$ un ensemble. On dit que $m \in \mathbb{R}$ est un majorant de E si $m \geq e$ pour tout $e \in E$.

Résultat IX.20 (Propriété de la borne supérieure). Si $E \subset \mathbb{R}$ est majoré, alors il existe un plus petit majorant noté $\sup E$. Similairement si E est minoré, il possède un plus grand minorant noté $\inf E$.

Cette propriété est une propriété fondamentale de l'ensemble des nombres réels, que nous prenons comme axiome (admise sans preuve).

Exemple. Si $A = [0, 1[\cup [2, 5]$, alors :

$$\inf A = 0 \qquad \qquad \qquad \sup A = 5$$

Résultat IX.21 (Théorème des bornes atteintes). Si $f : A \rightarrow \mathbb{R}$ est continue alors pour tout $a < b \in A$ il existe u, v deux réels dans $[a, b]$ tels que $f([a, b]) = [f(u), f(v)]$.

Idee de preuve du théorème de la borne atteinte. Nous n'allons nous intéresser qu'au cas de la borne supérieure. Le cas de la borne inférieure s'en déduit.

Montrons d'abord que $F := f([a, b])$ est majoré. Pour $N > 0$, considérons l'ensemble E_N des éléments $x \in [a, b]$ vérifiant $f(x) \geq N$.

Notons que les ensembles E_N diminuent lorsque N augmente. Si E_N est vide pour un certain N , alors N majore F . Sinon, E_N n'est jamais vide, et toujours minoré par a et majoré par b .

Considérons alors l'ensemble E des valeurs $\sup E_N$ pour $N > 0$. Cet ensemble est minoré par a . Considérons donc $c = \inf E$.

Par construction, c est arbitrairement proche de nombres dont l'image est au delà de N , pour tout N . D'un autre côté par continuité, les nombres proches de c ont des images proches de $f(c)$. Ceci fournit une contradiction.

L'ensemble $F = f([a, b])$ étant majoré, il possède un supremum, notons le s .

Pour $\epsilon > 0$, on s'intéresse maintenant aux ensembles E_ϵ des x tels que $f(x) > s - \epsilon$. Il existe toujours de tels x puisque $s - \epsilon$ n'est pas un majorant de F . Cet ensemble est toujours minoré par a et majoré par b .

On s'intéresse comme précédemment à l'ensemble des nombres $\sup E_\epsilon$, et plus précisément à leur infimum. Ce nombre est le nombre v recherché.

Ceci montre que $f(v)$ est la plus grande valeur prise par f sur $[a, b]$. Similairement on peut démontrer l'existence de u tel que $f(u)$ est la plus petite valeur. Le fait que $f([a, b]) = [f(u), f(v)]$ suivra alors du théorème suivant (de la valeur intermédiaire). □

Exemple. L'image de $[0, 2\pi]$ par la fonction sinus est $[-1, 1]$, qu'on peut effectivement ré-écrire $[\sin(3\pi/2), \sin(\pi/2)]$.

Il est intuitivement clair que si le graphe d'une fonction continue « démarre » en dessous d'une droite horizontale et se termine au dessus de cette droite, il doit croiser la droite quelque part. C'est l'objet du résultat suivant :

Résultat IX.22 (Théorème de la valeur intermédiaire.). Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue. Pour tout $\gamma \in \mathbb{R}$ strictement compris entre $f(a)$ et $f(b)$, il existe $c \in]a, b[$ tel que $f(c) = \gamma$.

Idée de preuve. Nous supposons que $f(a) < \gamma < f(b)$: le cas $f(b) < f(a)$ se déduit d'arguments similaires. Considérons

$$S = \{x : x \in [a, b] \text{ et } f(x) < \gamma\}.$$

Cet ensemble S est non-vidé puisque $a \in S$ et S est majoré par b . Donc S possède un supremum. Écrivons $c = \sup S$. Évidemment $c \in [a, b]$.

Nous démontrons que $f(c) = \gamma$ en raisonnant par l'absurde :

- * Si $f(c) < \gamma$, alors il existe un point x légèrement plus grand que c tel que $f(x)$ est encore plus petit que γ , ce qui contredit le fait que c majore S .
- * Si $f(c) > \gamma$, nous voyons qu'en fait $f(x) > \gamma$ à partir d'un point légèrement plus petit que c , ce qui contredit le fait que c est le plus petit majorant de S .

□

Résultat IX.23. Soit f une fonction réelle bijective définie sur un intervalle. Si f est continue en a , alors sa réciproque est continue en $f(a)$.

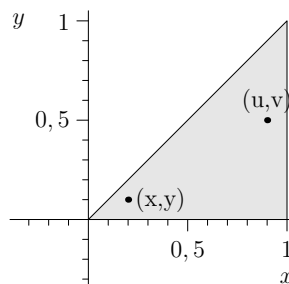
Exemple. Voyons un exemple dont le domaine n'est pas un intervalle. Soit

$$f : [0, 1[\cup [2, 3[: x \mapsto \begin{cases} x & \text{si } x < 1 \\ x - 1 & \text{sinon.} \end{cases}$$

Alors f est continue partout, mais f^{-1} n'est pas continue au point 1. Ci-dessous, la fonction f et la fonction f^{-1} sont représentées :

Résultat IX.24. Si $f : I \rightarrow \mathbb{R}$ est une injection continue définie sur un intervalle, alors f est strictement monotone (c'est-à-dire strictement croissante ou strictement décroissante).

Démonstration. Considérons le domaine T des points $(x, y) \in I \times I$ avec $x > y$. Ceci forme un triangle.



Supposons qu'il existe (x, y) dans T et (u, v) dans T avec $f(x) - f(y)$ positif et $f(u) - f(v)$ négatif. Considérons alors le point

$$\mathbf{p}(t) = (p_1(t), p_2(t)) = (x, y) + t((u, v) - (x, y)).$$

Lorsque t varie, ce point va du point (x, y) au point (u, v) en ligne droite, et donc en restant dans le triangle T . La quantité $f(p_1(t)) - f(p_2(t))$ est

- * positive en $t = 0$ (correspondant au point (x, y)) et
- * négative en $t = 1$.

Dès lors, par continuité, elle s'annule pour une certaine valeur de t , c'est-à-dire il existe t tel que

$$f(p_1(t)) = f(p_2(t))$$

Comme f est injective, cela implique $p_1(t) = p_2(t)$. Mais ceci n'est pas possible car le point $\mathbf{p}(t)$ est dans T et donc $p_1(t) > p_2(t)$! □

Chapitre X

Fonctions réelles : dérivées

Dans ce chapitre nous étudierons la notion de dérivée, ainsi que les propriétés de la dérivée. L'opération de dérivation est si importante qu'elle est présente virtuellement partout à travers les sciences.

Concrètement, la dérivée donne une formalisation de la variation d'une fonction « en un point ». Si cette fonction représente la position, la dérivée représente une vitesse.

1 Dérivée en un point

Points intérieurs

Intuitivement, un point $a \in A$ est intérieur à A si on peut « bouger » tout en restant dans le domaine. Formellement, un point est *intérieur* à un sous-ensemble A de $\mathbb{R}$ s'il existe un intervalle contenant ce point et entièrement contenu dans A . En d'autres termes, un réel a est intérieur à A , il existe un nombre $\delta > 0$ tel que $]a - \delta, a + \delta[\subseteq A$.

intérieur

Pentes

Considérons la droite D du plan passant par les points (x, y) et $(x + \Delta x, y + \Delta y)$. Supposons que $\Delta x \neq 0$. En particulier, les deux points sont distincts. Alors la *pente* de D est le nombre

pente

$$m := \frac{\Delta y}{\Delta x}$$

c'est-à-dire le quotient de la différence des ordonnées par la différence des abscisses. Imaginons maintenant un point mobile p se déplaçant le long de la droite D . La pente mesure la quantité dont p monte (ou descend, si la pente est négative) verticalement quand p avance d'une unité horizontalement.

Définition de dérivée

Soit $f : A \rightarrow B$ une fonction réelle ($A, B \subseteq \mathbb{R}$), et soit a un point intérieur à A . Si la limite

$$\lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}$$

existe, on appelle cette limite le *nombre dérivé de f en a* et on le note $f'(a)$. Donc

nombre dérivé
de f en a

$$f'(a) = \lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}$$

Quand cette limite existe, on dit que f est *dérivable* au point a , ou encore que $f'(a)$ existe.

dérivable

Résultat X.1. Si f est dérivable en u , alors f est continue en u .

Démonstration. On suppose que $\lim_{x \rightarrow u} \frac{f(x) - f(u)}{x - u}$ existe et vaut un nombre réel ($f'(u)$). Alors pour tout $x \neq u$:

$$f(x) = \frac{f(x) - f(u)}{x - u}(x - u) + f(u)$$

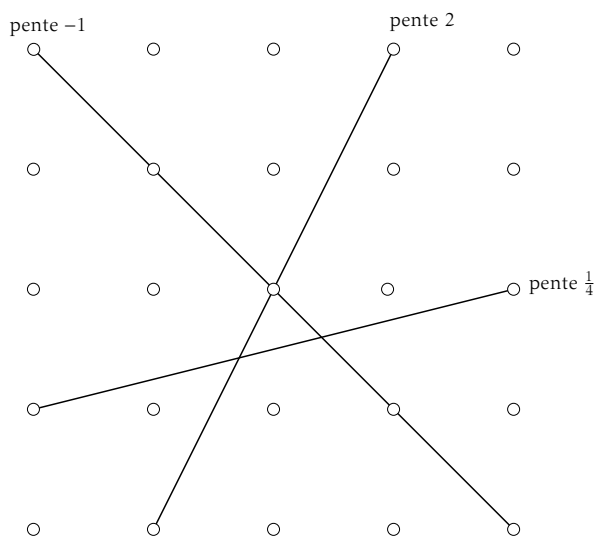


FIGURE X1 – Différentes pentes

et donc en passant à la limite :

$$\lim_{x \rightarrow u} f(x) = \lim_{x \rightarrow u} \frac{f(x) - f(u)}{x - u} (x - u) + f(u) = f'(u)0 + f(u) = f(u).$$

□

Approximation de Taylor d'ordre 1

Une autre approche très visuelle du même concept est celui de « différentielle » ou de « tangente ». Soit $f : A \rightarrow B$ une fonction réelle et $a \in A$ un point dans l'intérieur de son domaine. Nous dirons informellement qu'une fonction réelle g approxime bien f près de a si $g(x)$ s'approche de $f(x)$ plus vite que x ne s'approche de a . Pour formaliser ce concept, introduisons une notion :

Définition X.2. Soit u une fonction réelle et a un point adhérent à son domaine. On note $u(x) = o(x-a)$ (et on dit « u est un petit o de $x-a$ ») si

$$\lim_{x \rightarrow a} \frac{u(x)}{x-a} = 0$$

c'est-à-dire si $u(x)$ tend plus vite¹ vers 0 que $x-a$ (lorsque x tend vers a).

Nous cherchons donc une fonction g telle que $f(x) = g(x) + o(x-a)$. Cependant, si on s'autorise n'importe quel g , il suffit de prendre $g = f$ ce qui donne une excellente « approximation » mais ne simplifie rien du tout. Afin d'obtenir une fonction g très simple, nous n'allons considérer dans un premier temps que des polynômes de degré au plus égal à 1 (c'est-à-dire inférieur ou égal à 1). C'est-à-dire qu'on cherche, géométriquement, à trouver la droite qui approxime le mieux la fonction près de a .

Définition X.3. Le polynôme $g(x) = kx + m$ est l'*approximation de Taylor d'ordre 1 en a de f* si $f(x) = kx + m + o(x-a)$ pour $x \rightarrow a$.

Résultat X.4. Une fonction réelle admet une approximation de Taylor d'ordre 1 en a si et seulement si elle est dérivable en a ; de plus l'approximation de Taylor est alors donnée par $x \mapsto f'(a)(x-a) + f(a)$. En particulier, l'approximation de Taylor est unique, ce qui justifie la terminologie.

Ce polynôme de degré 1 s'appelle aussi *polynôme de Taylor*, ou *approximation linéaire*, ou encore *approximation affine*, voire parfois *linéarisation* de f en a . On dit également que la fonction L définie par $L(x) = f'(a)(x-a) + f(a)$ est la meilleure approximation linéaire de f près du point a .

1. Cette notion de « plus vite » ou « moins vite » est informelle. Par exemple, rappelons-nous que le graphe des fonctions $x \mapsto x^n$ « s'écrase » près de 0 lorsque n augmente : dès lors x^2 tend plus vite vers 0 que x , mais moins vite que x^3 ou x^4 .

approximation
de Taylorpolynôme de
Taylorapproximation
linéaireapproximation
affine

linéarisation

Résultat X.5. Si une fonction f est dérivable au point a , alors f est continue au point a .

Démonstration. Notons que pour tout x différent de a ,

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

donc, en prenant la limite pour $x \rightarrow a$, on a (en appliquant les règles de calcul)

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

ce qui prouve que la limite vaut $f(a)$, c'est-à-dire f est continue en a . □

Notations

D'autres notations sont utilisées pour la dérivée d'une fonction f au point a : on note $f'(a)$ également

$$\frac{df}{dx}(a)$$

ce qui se signifie littéralement « la dérivée de f par rapport à x en a ». Si on pose $y = f(x)$ alors on note $f'(a)$ également

$$y'(a) \quad \text{ou encore} \quad \frac{dy}{dx}(a)$$

Par ailleurs, nous pouvons reformuler la formule du résultat X.4 en écrivant :

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

ou encore

$$f(x + \delta x) = f(x) + f'(x)\delta x + o(\delta x)$$

Cette dernière variation, strictement identique à la précédente, sera notamment utilisée dans le cadre des cours de physique.

1.1 Interprétation de la dérivée

Droite tangente

Si f est dérivable au point a , la droite d'équation

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

est la *droite tangente au graphe de f au point $(a, f(a))$* ou *droite tangente de f en a* .

droite tangente
au graphe de f
au point $(a, f(a))$

1.1.1 Vitesse

Considérons une droite munie d'un repère cartésien, c'est-à-dire une droite graduée.

Soit $f : \mathbb{R} \rightarrow \mathbb{R} : t \mapsto f(t)$ une fonction dont la valeur $f(t)$ représente la coordonnée d'un mobile à l'instant t le long de cette droite.

Exemple. Si on note $f(t)$ la hauteur d'un yoyo au cours du temps, on pourrait modéliser ceci par $f(t) = \cos(t)$.

- * À l'instant $t = 0$, le mobile est à hauteur 1,
- * puis descend pour arriver à hauteur -1 ,
- * remonte ensuite pour retourner à hauteur 1,
- * et ainsi de suite...

droite tangente
de f en a

Dans cette situation, la quantité

$$\frac{f(b) - f(a)}{b - a}$$

représente la distance parcourue rapportée au temps passé à la parcourir : c'est la *vitesse moyenne* du mobile entre les instants a et b .

Remarque. Notons que le signe de la vitesse moyenne indique le sens du déplacement !

Lorsque a et b sont de plus en proches, la vitesse moyenne tend vers la quantité

$$\lim_{b \rightarrow a} \frac{f(b) - f(a)}{b - a} = f'(a).$$

C'est la *vitesse instantanée* du mobile à l'instant a .

La vitesse instantanée est la dérivée de la position par rapport au temps.

Remarque (*). Connaissant la vitesse instantanée en chaque instant, comment retrouver la position ? Notons qu'il faut au minimum connaître la position initiale ! La réponse à cette question sera donnée par la notion de primitive, section ?? page ??.

2 Dérivées et opérations

Somme, produit, quotient

Si f et g sont des fonctions dérivables au point a , et $c \in \mathbb{R}$ est une constante, alors $f + g$, cf et fg sont des fonctions² dérivables en a et

$$\begin{aligned}(f + g)'(a) &= f'(a) + g'(a) \\ (cf)'(a) &= cf'(a) \\ (fg)'(a) &= f'(a)g(a) + f(a)g'(a) \quad (\text{Règle de Leibniz})\end{aligned}$$

Si $g(a) \neq 0$, alors $\frac{f}{g}$ est dérivable au point a et

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

Détaillons quelque preuves :

Résultat X.6. Si a est une constante et f une fonction, alors $(af)' = af'$

Démonstration. Il faut montrer que pour tout u dans le domaine de f' , $(af)'(u) = af'(u)$. On calcule simplement :

$$\begin{aligned}\lim_{x \rightarrow u} \frac{(af)(x) - (af)(u)}{x - u} &= \lim_{x \rightarrow u} \frac{af(x) - af(u)}{x - u} \\ &= \lim_{x \rightarrow u} a \frac{f(x) - f(u)}{x - u} \\ &= a \lim_{x \rightarrow u} \frac{f(x) - f(u)}{x - u} \\ &= af'(u)\end{aligned}$$

□

Résultat X.7. Si f, g sont deux fonctions, alors $(f + g)' = f' + g'$

2. Ces fonctions sont définies à la section 1

Démonstration. Vérifions l'égalité en un point u où f et g sont toutes les deux dérivables :

$$\begin{aligned}\lim_{x \rightarrow u} \frac{(f+g)(x) - (f+g)(u)}{x-u} &= \lim_{x \rightarrow u} \frac{f(x) + g(x) - f(u) - g(u)}{x-u} \\ &= \lim_{x \rightarrow u} \frac{f(x) - f(u)}{x-u} + \frac{g(x) - g(u)}{x-u} \\ &= f'(u) + g'(u)\end{aligned}$$

d'où le résultat. □

Résultat X.8. Si f et g sont des fonctions dérivables en u , alors $(fg)'(u) = f'(u)g(u) + f(u)g'(u)$

Démonstration.

$$\begin{aligned}\lim_{x \rightarrow u} \frac{(fg)(x) - (fg)(u)}{x-u} &= \lim_{x \rightarrow u} \frac{f(x)g(x) - f(u)g(u)}{x-u} \\ &= \lim_{x \rightarrow u} \frac{f(x)g(x) - f(x)g(u) + f(x)g(u) - f(u)g(u)}{x-u} \\ &= \lim_{x \rightarrow u} \frac{f(x)(g(x) - g(u)) + (f(x) - f(u))g(u)}{x-u} \\ &= \lim_{x \rightarrow u} \frac{f(x)(g(x) - g(u))}{x-u} + \frac{(f(x) - f(u))g(u)}{x-u} \\ &= \lim_{x \rightarrow u} f(x) \frac{g(x) - g(u)}{x-u} + \lim_{x \rightarrow u} \frac{(f(x) - f(u))}{x-u} g(u) \\ &= f(u)g'(u) + f'(u)g(u)\end{aligned}$$

□

Résultat X.9. La dérivée $1/f$ est $-f'/f^2$ en tout point où f ne s'annule pas.

Démonstration.

$$\begin{aligned}\lim_{x \rightarrow u} \frac{\frac{1}{f(x)} - \frac{1}{f(u)}}{x-u} &= \lim_{x \rightarrow u} \frac{\frac{f(u) - f(x)}{f(x)f(u)}}{x-u} \\ &= \lim_{x \rightarrow u} \frac{1}{f(x)f(u)} \frac{f(u) - f(x)}{x-u} \\ &= -\frac{f'(u)}{f(u)^2}\end{aligned}$$

□

Exercice. Démontrer $(f/g)' = (f'g - fg')/g^2$ en combinant les deux dernières propositions.

Composition

Si a est un point intérieur au domaine de $g \circ f$, si g est dérivable en $f(a)$, et f dérivable en a , alors $g \circ f$ est dérivable en a et

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

Une autre manière d'appréhender la composée de fonctions est de penser à des variables réelles, disons x , y et t , qui sont liées les unes aux autres. Par exemple, il se pourrait que y soit fonction de x et que x soit fonction de t . Notons cela $y = y(x)$ et $x = x(t)$. Alors, à travers x , la variable y est une fonction de t . Donc cela a un sens de dériver y par rapport à t . Si t_0 est le point où on dérive et $x_0 = x(t_0)$ alors la formule donnée ci-dessus se réécrit

$$\frac{dy}{dt}(t_0) = \frac{dy}{dx}(x_0) \frac{dx}{dt}(t_0)$$

ou encore, plus simplement (mais attention aux points où on dérive!)

$$\frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt}$$

Notez que dans l'écriture ci-dessus, les dx se « simplifient ». Ceci est à prendre avec un grain de sel : on a en aucun cas défini les quantités dx , dy et dt isolément ! Par contre, cette manière de voir les choses donne un moyen infaillible pour retenir la règle de dérivation d'une fonction composée.

Idée de preuve. On calcule la limite :

$$\begin{aligned}\lim_{x \rightarrow a} \frac{f(g(x)) - f(g(a))}{x - a} &= \lim_{x \rightarrow a} \left(\frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \right) \left(\frac{g(x) - g(a)}{x - a} \right) \\ &= \lim_{x \rightarrow a} \frac{f(g(x)) - f(g(a))}{g(x) - g(a)} \lim_{x \rightarrow a} \frac{g(x) - g(a)}{x - a} \\ &= \lim_{t \rightarrow g(a)} \frac{f(t) - f(g(a))}{t - g(a)} \lim_{x \rightarrow a} \frac{g(x) - g(a)}{x - a} \\ &= f'(g(a))g'(a) \\ &\text{on a utilisé la continuité de } g \text{ en } a!\end{aligned}$$

□

Remarque (*). Nous avons passé sous silence le fait que $g(x)$ et $g(a)$ pourraient être égaux.

2.1 Fonction réciproque

Supposons que $f : A \rightarrow B$ soit une fonction réelle bijective dérivable dont la réciproque est dérivable, de sorte qu'on puisse appliquer la formule de dérivation de fonctions composées à l'égalité

$$f^{-1}(f(x)) = x.$$

On a donc

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

c'est-à-dire $(f^{-1})'(f(a)) = \frac{1}{f'(a)}$.

Cette formule d'apparence compliquée permet en fait de calculer la dérivée de la réciproque à partir de la dérivée de la fonction de départ. En réalité, nous avons un résultat un peu plus fort puisqu'il n'est pas nécessaire de supposer la dérivabilité de la réciproque :

Résultat X.10. Si $f : A \rightarrow B$ est une bijection dérivable en a avec $f'(a) \neq 0$, alors sa réciproque f^{-1} est dérivable en $f(a)$ et

$$(f^{-1})'(f(a)) = \frac{1}{f'(a)}.$$

Démonstration. Comme f est une bijection dérivable en a , elle est également continue en a et son inverse est donc continue en $f(a)$. Dès lors nous avons $\lim_{t \rightarrow f(a)} f^{-1}(t) = f^{-1}(f(a)) = a$. On a donc successivement

$$\begin{aligned}\frac{1}{f'(a)} &= \lim_{x \rightarrow a} \frac{x - a}{f(x) - f(a)} = \lim_{x \rightarrow a} \frac{f^{-1}(f(x)) - f^{-1}(f(a))}{f(x) - f(a)} \\ &= \lim_{t \rightarrow f(a)} \frac{f^{-1}(t) - f^{-1}(f(a))}{t - f(a)}\end{aligned}$$

où la dernière égalité s'obtient par la règle de composition des limites. Ceci est la définition du nombre dérivé de f^{-1} en $f(a)$. □

Justification intuitive des règles de dérivation

Attention, ce qui suit n'est pas une justification formelle, mais offre un autre point de vue sur les preuves et idées de preuves vues précédemment.

Remarque (*). Avant tout, notons que le fait que la dérivée de f en a est $f'(a)$ peut s'interpréter par³

$$f(x) \approx f'(a)(x - a) + f(a) \quad \text{pour } x \approx a.$$

3. Le signe $\approx$ signifie « approche ».

Partant de cette remarque, considérons deux fonctions f et g et écrivons :

$$\begin{cases} f(x) \approx f'(a)(x-a) + f(a) \\ g(x) \approx g'(a)(x-a) + g(a) \end{cases} \quad \text{pour } x \approx a$$

En additionnant les deux équations, on est amené à conclure

$$f(x) + g(x) \approx (f'(a) + g'(a))(x-a) + f(a) + g(a) \quad \text{pour } x \approx a$$

et donc que $(f+g)'(a) = f'(a) + g'(a)$ car si $h(x) \approx m(x-a) + h(a)$ pour $x \approx a$ pour une fonction h et un nombre m , alors nécessairement $m = h'(a)$.

On peut aussi multiplier les deux équations pour trouver

$$f(x)g(x) \approx (f'(a)g(a) + g'(a)f(a))(x-a) + f(a)g(a) + f'(a)g'(a)(x-a)^2 \quad \text{pour } x \approx a$$

ce qui donne

$$f(x)g(x) \approx (f'(a)g(a) + g'(a)f(a))(x-a) + f(a)g(a) \quad \text{pour } x \approx a$$

car $(x-a)^2$ est négligeable par rapport aux deux premiers termes quand $x \approx a$. (On cherche une approximation de $f(x)g(x)$ par une fonction de degré 1.) En regardant le coefficient de $(x-a)$, on trouve la dérivée du produit fg en a . On peut aussi faire le même genre de raisonnement pour le quotient f/g (voyez-vous comment?).

Voyons ce qu'on obtient pour la fonctions composée $g \circ f$. Dans ce cas, on a

$$\begin{aligned} f(x) &\approx f'(a)(x-a) + f(a) \quad \text{pour } x \approx a \quad \text{et} \\ g(y) &\approx g'(f(a))(y-f(a)) + f(g(a)) \quad \text{pour } y \approx f(a) \end{aligned}$$

En remplaçant y par $f(x)$ dans le membre de gauche, et par l'approximation de $f(x)$ dans le membre de droite on trouve

$$g(f(x)) \approx g'(f(a))(f'(a)(x-a) + f(a) - f(a)) + f(g(a)) \quad \text{pour } x \approx a$$

=0

Le coefficient de $(x-a)$ est $g'(f(a))f'(a)$, ce qui explique la formule donnant la dérivée d'une composée : en composant les approximations de f autour de a et de g autour de $f(a)$ par des fonctions linéaires on trouve une approximation de $g \circ f$ autour de a par une fonction linéaire. Le coefficient de $(x-a)$ dans cette fonction linéaire est la dérivée de $g \circ f$ en a : $(g \circ f)'(a) = g'(f(a))f'(a)$.

3 Fonction dérivée et dérivées d'ordre supérieur

Fonction dérivée

Soit $f : A \rightarrow B$ une fonction réelle d'une variable réelle ($A, B \subseteq \mathbb{R}$). Appelons C l'ensemble des points intérieurs de A où f est dérivable. Alors $C \subseteq A$ est le domaine de définition de la fonction

$$f' : C \rightarrow \mathbb{R} : x \mapsto f'(x)$$

appelée *fonction dérivée* ou simplement la *dérivée* de f . Remarquons que C est un sous-ensemble de A (éventuellement égal à celui-ci), mais l'ensemble image de la dérivée f' n'a a priori aucun lien avec B .

fonction dérivée

dérivée

Remarque (*). Attention, si f désigne une fonction, f' désigne sa fonction dérivée. Si $f(x)$ est l'expression algébrique d'une fonction, on peut noter indifféremment $f'(x)$ ou $f(x)'$ la dérivée par rapport à x de cette expression. Cependant, si l'expression est composée, $f(g(x))'$ et $f'(g(x))$ sont en général des expressions différentes ! La première représente la dérivée de la fonction $t \mapsto f(g(t))$, calculée au point x , tandis que $f'(g(x))$ désigne le nombre dérivé de f (c'est-à-dire la fonction $t \mapsto f'(t)$) au point $g(x)$. Il faut bien distinguer les deux, sans quoi la règle

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

n'aurait aucun sens.

Dérivée seconde, troisième, etc...

Si f est dérivable dans un intervalle ouvert $]a, b[$, et si sa dérivée f' est aussi dérivable dans $]a, b[$, alors on définit la *dérivée seconde*, notée f'' par

$$f''(x) = (f')'(x); \quad x \in]a, b[$$

Si f'' admet une dérivée dans $]a, b[$, on l'appelle *dérivée troisième*, notée f''' ou f^3 et ainsi de suite, c'est-à-dire

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

si f^n est dérivable dans $]a, b[$, avec $n \in \mathbb{N}$.

Interprétation de la dérivée seconde

La dérivée seconde d'une fonction f en un point a est liée à la concavité que possède le graphe de f en a , c'est-à-dire la tendance qu'ont les tangentes de f d'être sous le graphe de f (concavité tournée vers le haut) ou au dessus du graphe de f (concavité tournée vers le bas).

Nous avons vu que $L : x \mapsto f(a) + f'(a)(x - a)$ est la meilleure approximation linéaire de $f(x)$ pour x proche de a . La dérivée seconde de f en a permet, de la même manière, d'approximer f par une fonction quadratique (polynôme de degré au plus 2) autour de a . La meilleure approximation quadratique de $f(x)$ pour x proche de a est :

$$Q(x) = f(a) + f'(a)(x - a) + f''(a) \frac{(x - a)^2}{2}$$

La fonction Q est appelée *l'approximation de Taylor d'ordre 2* de f en a . Son graphe est une parabole (dont l'axe de symétrie est parallèle à oy). Cette parabole est tournée vers le haut si $f''(a) > 0$ et vers le bas si $f''(a) < 0$ (c'est le coefficient devant x^2).

Cette remarque sur Q s'étend à f elle-même :

- * Si $f''(a) > 0$, la dérivée est croissante, donc la pente de la tangente a tendance à remonter près de a , donc la concavité en a est tournée vers le haut !
- * Inversement, $f''(a) < 0$ indique une concavité tournée vers le bas.

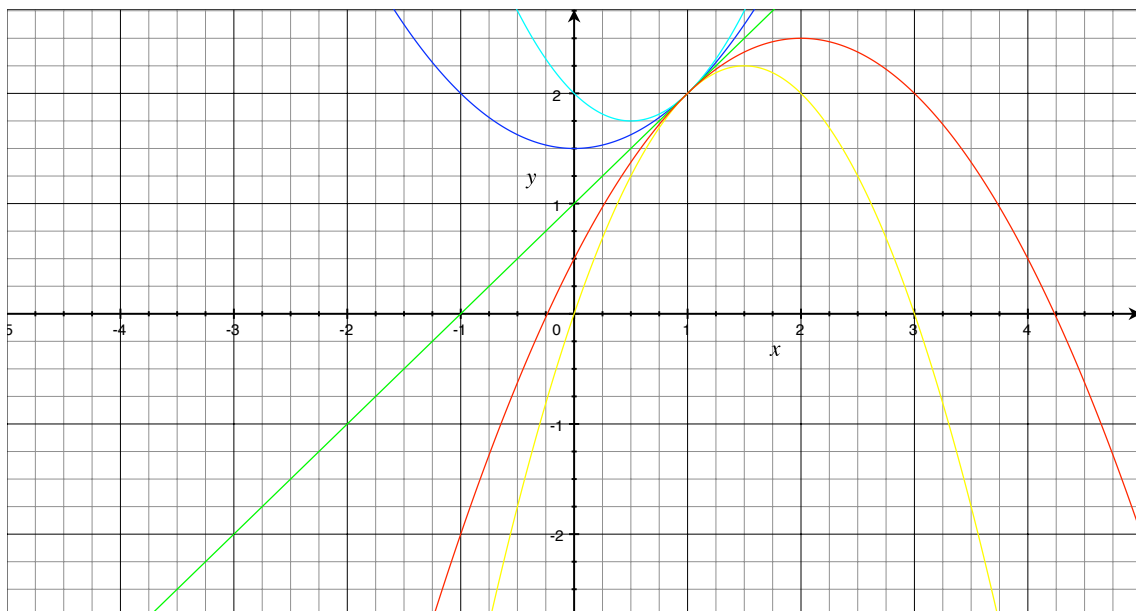
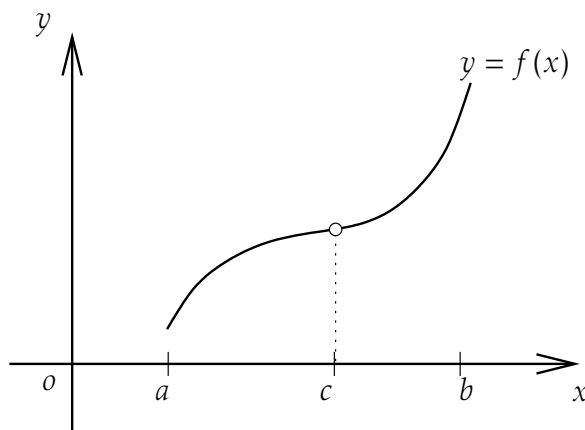


FIGURE X2 – L'effet d'une modification de $f''(a)$ (sans changer $f(a)$ ni $f'(a)$) dans l'expression de Q fait varier la concavité en a .

Définition X.11. Lorsque f'' change de signe en c (et souvent s'y annule), alors la concavité du graphe de f change en cpaue et on dit que c est un *point d'inflexion* de f .



4 Dérivées des fonctions élémentaires

La dérivée de toute fonction élémentaire est encore une fonction élémentaire et peut être calculée de manière automatique, une fois qu'on connaît les dérivées des fonctions élémentaires « de base », en utilisant les règles de dérivation données ci-dessus pour la somme, le produit, le quotient et la composée de fonctions.

4.1 Constantes, fonctions puissance et exponentielles

La dérivée d'une fonction constante $f(x) = c$ ($c \in \mathbb{R}$) est nulle :

$$f'(a) = \lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{f(x) - f(a)}{x - a} = \lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{c - c}{x - a} = 0.$$

(Remarquons que le numérateur $c - c$ vaut 0 indépendamment de x , il ne s'agit donc pas d'un cas d'indétermination.)

La dérivée de f définie par $f(x) = x$ se calcule également aisément à partir de la définition :

$$f'(a) = \lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{f(x) - f(a)}{x - a} = \lim_{\substack{x \rightarrow a \\ x \neq a}} \frac{x - a}{x - a} = 1$$

Pour dériver une fonction monôme $f(x) = x^n$ (avec $n \in \mathbb{N}$), on applique la règle de dérivation d'un produit :

$$\begin{aligned} (x^0)' &= 0 \\ (x^1)' &= 1 \\ (x^2)' &= (x \cdot x)' = 1 \cdot x + x \cdot 1 = 2x \\ (x^3)' &= (x^2 \cdot x)' = 2x \cdot x + x^2 \cdot 1 = 3x^2 \\ &\vdots \\ (x^n)' &= nx^{n-1} \end{aligned}$$

Sans démonstration, mentionnons que la dérivée de x^α pour α une constante réel (et $x > 0$) prend la même forme :

$$(x^\alpha)' = \alpha x^{\alpha-1}$$

Nous n'avons pas défini la constante e de l'exponentielle e^x avec assez de rigueur pour le montrer, cependant le fait que la dérivée de e^x est $(e^x)' = e^x$ est particulièrement important. En particulier, on en déduit la formule

Résultat X.12. $(\ln(x))' = \frac{1}{x}$

Démonstration. Ceci suit de la dérivation de fonction composée : $\ln$ est l'inverse de l'exponentielle, dès lors nous avons vu l'égalité

$$\ln'(e^x) = \frac{1}{e^x}$$

pour tout $x \in \mathbb{R}$. En particulier, si $y \in \mathbb{R}_0^+$, on peut prendre $x = \ln(y)$

$$\ln'(y) = \frac{1}{y}$$

ce qui est la relation annoncée. □

De ces deux résultats nous déduisons ;

$$(b^x)' = ((e^{\ln b})^x)' = (e^{\ln b \cdot x})' = e^{\ln b \cdot x} \cdot (\ln b \cdot x)' = b^x \cdot \ln b$$

et

$$\log'_b(x) = \left(\frac{\ln(x)}{\ln(b)} \right)' = \frac{1}{x \ln(b)}$$

Fonctions trigonométriques

Sans démonstration, rappelons simplement les dérivées des fonctions trigonométriques usuelles (les deux dernières s'obtiennent à partir des deux premières via les règles de calcul précédentes.)

$f(x)$	$f'(x)$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\operatorname{tg} x$	$\frac{1}{\cos^2 x}$
$\operatorname{cotg} x$	$-\frac{1}{\sin^2 x}$

Fonctions réciproques : fonctions trigonométriques réciproques

Les fonctions arcsin, arccos et arctg sont les réciproques de fonctions trigonométriques, dès lors leur dérivée s'obtiennent avec la formule vue plus haut : pour tout angle $\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ on a $\cos \theta = \sqrt{1 - \sin^2 \theta}$ et donc

$$(\arcsin x)' = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - x^2}}$$

Pour $\theta \in [0, \pi]$ on a $\sin \theta = \sqrt{1 - \cos^2 \theta}$ et donc

$$(\arccos x)' = \frac{1}{-\sin(\arccos x)} = -\frac{1}{\sqrt{1 - x^2}}$$

Enfin, la dérivée de la fonction tan peut s'écrire $1 + \operatorname{tg}^2 \theta$, donc

$$(\operatorname{arctg} x)' = \frac{1}{1 + \operatorname{tg}^2(\operatorname{arctg} x)} = \frac{1}{1 + x^2}.$$

Chapitre XI

Fonctions dérivables

Les fonctions dérivables sur leur domaine jouissent de propriétés les rendant plus simples à étudier.

1 Extrema et dérivées

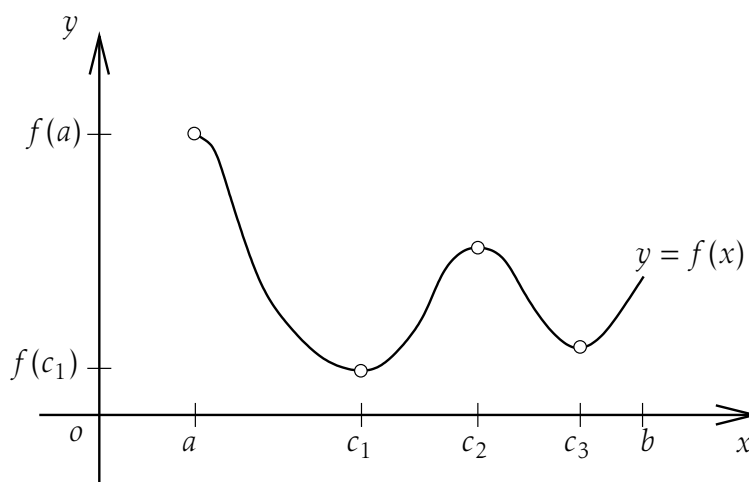
Définition XI.1. Soit f une fonction réelle et a un point de son domaine. Ce point est

- * un *minimum* (global) de f si $f(a) \leq f(x)$ pour tout $x \in \text{dom } f$;
- * un *maximum* (global) de f si $f(a) \geq f(x)$ pour tout $x \in \text{dom } f$;
- * un *minimum local* de f s'il existe $\delta > 0$ tel que $f(a) \leq f(x)$ pour tout $x \in]a - \delta, a + \delta[\cap \text{dom } f$;
- * un *maximum local* de f s'il existe $\delta > 0$ tel que $f(a) \geq f(x)$ pour tout $x \in]a - \delta, a + \delta[\cap \text{dom } f$.

Un *extremum*, qui peut être local ou global, est un point qui est soit un maximum, soit un minimum.

Bien sûr, un extremum global est aussi, d'après ces définitions, également un extremum local.

Exemple. La fonction suivante, définie sur l'intervalle $[a, b[$ a un minimum global en c_1 , un maximum global en a , un minimum local en c_3 et un maximum local en c_2 :



Il n'y a pas de maximum local en b puisque $f(b)$ n'est même pas défini !

Trouver les extrema d'une fonction en général n'est pas chose aisée, mais pour les fonctions dérivables nous avons des théorèmes agréables permettant de cibler la recherche.

Résultat XI.2 (Extrema et dérivée). Soit f une fonction réelle. Si f est dérivable en a et a est un extremum local, alors $f'(a) = 0$.

minimum

maximum

minimum local

maximum local

extremum

Démonstration. Supposons par l'absurde $f'(a) > 0$ (le cas $f'(a) < 0$ se traite similairement.) Dans ce cas, si on prend $\epsilon > 0$ (et nous prenons $\epsilon = \frac{f'(a)}{2}$), on obtient un $\delta > 0$ par la définition de limite pour lequel

$$|x - a| < \delta \Rightarrow -\epsilon < \frac{f(x) - f(a)}{x - a} - f'(a) < \epsilon$$

ce qui, pour notre choix de ϵ , donne

$$|x - a| < \delta \Rightarrow \frac{f'(a)}{2} < \frac{f(x) - f(a)}{x - a} < \frac{3f'(a)}{2}.$$

ce qui montre que $\frac{f(x) - f(a)}{x - a} > 0$, c'est-à-dire le numérateur et le dénominateur ont même signe. Dès lors pour **tout** x (strictement) entre a et $a + \delta$, il faut avoir

$$f(x) - f(a) > 0$$

ce qui empêche a d'être un minimum local, tandis que pour **tout** x (strictement) entre $a - \delta$ et a , nous avons forcément

$$f(x) - f(a) < 0$$

ce qui empêche a d'être un maximum local. Contradiction!

Similairement, le cas $f'(a) < 0$ conduit à une contradiction également. Il faut donc forcément $f'(a) = 0$. $\square$

point critique

Un point c (intérieur au domaine de la fonction) où $f'(c) = 0$ est un *point critique*, et on dit en plus que $f(c)$ est une *valeur critique*. Une fonction réelle f étant donnée, quels sont les points c où f peut avoir un extremum?

valeur critique

1. les points critiques, où $f'(c) = 0$;
2. les points non-intérieurs au domaine (le « bord »);
3. les points où f n'est pas dérivable.

Tous ces points sont des « candidats extrema ». Nous dans les pages qui suivent les outils classiques utilisant le signe de la dérivée pour déterminer les extrema avec précision.

Exemple. * Si $f : \mathbb{R}^+ \rightarrow \mathbb{R} : x \mapsto \sqrt{x}$, alors f admet un minimum en $a = 0$, mais a n'est pas intérieur au domaine!

- * Si $f : \mathbb{R}^+ \rightarrow \mathbb{R} : x \mapsto x$, alors f admet un minimum en $a = 0$, mais a n'est pas intérieur au domaine!
- * Si $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto |x|$, alors f admet un minimum en $a = 0$, mais f n'est pas dérivable en a !
- * Si $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^3$, alors $f'(0) = 0$, mais 0 n'est pas un extrémum!

2 Théorème des accroissements finis

Résultat XI.3 (Théorème de Rolle). Soit f une fonction réelle continue sur un intervalle $[a, b]$ (avec $a < b$) et dérivable dans $]a, b[$. Si $f(a) = f(b)$, alors il existe un nombre $c \in]a, b[$ tel que $f'(c) = 0$.

Démonstration. f est continue sur un intervalle fermé, donc admet un minimum $f(u)$ et un maximum $f(v)$ pour certains $u, v \in [a, b]$.

Si le minimum et le maximum sont tous les deux sur les bords (a et b), alors f est forcément constante, puisqu'on suppose $f(a) = f(b)$!

Supposons donc que u ou v est dans $]a, b[$. Disons u (ça serait pareil si c'était v). Mais alors f atteint un minimum en un point intérieur à son domaine (où elle est dérivable), et donc $f'(u) = 0$! $\square$

Résultat XI.4 (Théorème des accroissements finis). Soit f continue sur $[a, b]$ et dérivable dans $]a, b[$. Alors il existe un nombre $c \in]a, b[$ tel que

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

Démonstration. On applique le théorème de Rolle à la fonction

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

Notons que g est continue sur $[a, b]$ car somme de fonctions continues sur $[a, b]$, et dérivable sur $]a, b[$ car somme de fonctions dérivables sur $]a, b[$. De plus,

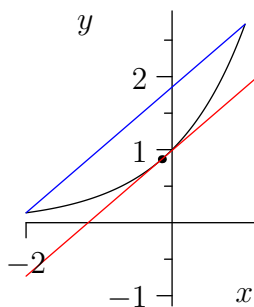
$$g(a) = f(a) - 0 - f(a) = 0 \quad \text{et} \quad g(b) = f(b) - \frac{f(b) - f(a)}{b - a}(b - a) - f(a) = 0$$

En appliquant le théorème de Rolle à la fonction g , on conclut qu'il existe un $c \in]a, b[$ tel que $g'(c) = 0$. Mais

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

ce qui démontre le théorème. □

Remarque. Pour paraphraser le résultat précédent : il existe toujours un point du graphe en lequel la tangente est parallèle à la droite liant $(a, f(a))$ et $(b, f(b))$.



Résultat XI.5 (Théorème des accroissements finis généralisé). Soient f et g des fonctions continues sur $[a, b]$ et dérivable dans $]a, b[$. Alors il existe un nombre $c \in]a, b[$ tel que

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Démonstration. Considérons h définie par $h(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}g(x)$. Posons $\lambda = \frac{f(b) - f(a)}{g(b) - g(a)}$ pour simplifier l'écriture. Nous allons appliquer le théorème de Rolle à cette fonction (exercice : terminer la preuve avec ces seules informations sans lire la suite).

Nous avons

$$h(a) = f(a) - \lambda g(a)$$

$$h(b) = f(b) - \lambda g(b)$$

et donc

$$h(b) - h(a) = f(b) - f(a) - \lambda(g(b) - g(a)) = f(b) - f(a) - (f(b) - f(a)) = 0.$$

Dès lors il existe $c \in]a, b[$ tel que

$$h'(c) = 0.$$

Or h peut se dériver explicitement grâce aux règles de calcul :

$$h'(x) = f'(x) - \lambda g'(x)$$

ce qui, pour $x = c$, fournit la relation demandée (puisque λ est justement le membre de droite de l'égalité à démontrer). □

3 Croissance, décroissance et monotonie.

Définition XI.6. Soit f une fonction réelle et $A \subset \text{dom } f$. Cette fonction est

- * *croissante* sur A si pour tout x, y dans A : $x \leq y \Rightarrow f(x) \leq f(y)$;
- * *décroissante* sur A si pour tout x, y dans A : $x \leq y \Rightarrow f(x) \geq f(y)$;
- * *strictement croissante* sur A si pour tout x, y dans A : $x < y \Rightarrow f(x) < f(y)$;
- * *strictement décroissante* sur A si pour tout x, y dans A : $x < y \Rightarrow f(x) > f(y)$.

Une fonction *monotone*, qui peut l'être strictement ou pas, est une fonction qui est soit croissante sur son domaine, soit décroissante sur son domaine.

Exemple. La fonction réelle $x \mapsto x^2$ est (monotone et) croissante sur $\mathbb{R}^+$, (monotone et) décroissante sur $\mathbb{R}^-$, mais n'est pas monotone sur $\mathbb{R}$.

Dans le théorème suivant, nous dirons « presque tout point » pour dire « en tout point sauf éventuellement un nombre fini ».

Résultat XI.7. Soit f une fonction continue sur un intervalle I . Supposons de plus que f est dérivable en presque tout point de I .

- * Si $f'(x) > 0$ pour presque tout $x \in I$, alors f est strictement croissante sur I ;
- * $f'(x) \geq 0$ pour presque tout $x \in I$ si et seulement si f est croissante sur I .

Similairement, lorsque les inégalités sont renversées la fonction f est (strictement) décroissante.

Le théorème ci-dessus reste vrai en particulier si on prend une fonction définie et dérivable sur un intervalle (quelconque, y compris non-borné).

Exemple. La fonction $x \mapsto \sqrt[3]{x}$ est continue sur $\mathbb{R}$ et a pour dérivée $x \mapsto \frac{1}{3\sqrt[3]{x^2}}$ qui est clairement strictement positive, **sauf en** 0 où elle n'existe pas. On peut cependant en déduire que la fonction de départ (racine cubique) est strictement croissante.

Démonstration. Par facilité, dans la preuve nous supposons que f est dérivable partout sur l'intérieur de l'intervalle.

Montrons le premier point : si $f'(x) > 0$ pour tout x , alors f est strictement croissante. Prenons donc $x < y$. Comme f est continue sur l'intervalle $[x, y]$ et dérivable sur l'ouvert $]x, y[$, on peut appliquer la formule des accroissements finis, c'est-à-dire qu'il existe $c \in]x, y[$ avec

$$f'(c) = \frac{f(y) - f(x)}{y - x}.$$

Or $f'(c) > 0$ et $y - x > 0$, dès lors $f(y) > f(x)$ également, ce qui prouve la croissance stricte.

Le même argument que précédemment montre que si $f'(x) \geq 0$ pour tout $x \in I$, alors f est croissante sur I .

Réciproquement, si f est croissante, prenons x dans I et montrons que $f'(x) \geq 0$. Supposons par l'absurde que $f'(x) < 0$. Alors il existe y tel que

$$\frac{f(y) - f(x)}{y - x} < 0$$

d'après le lemme suivant le résultat « Extrema et dérivée ». Ceci prouve que $f(y) - f(x)$ et $y - x$ ont signe contraire, ce qui est impossible si f est croissante. $\square$

Exemple. Considérons la fonction f définie par

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

On remarque que f s'annule en $x = 0$, nulle part ailleurs.

Sa dérivée vaut

$$f'(x) = \frac{1-x^2}{(x^2+1)^2}$$

On en déduit que la dérivée s'annule en $x = \pm 1$, est positive entre ces valeurs et négative en dehors.

Elle admet donc un minimum en -1 (valeur : $f(-1) = -1/2$), et un maximum en 1 (valeur : $1/2$).

On note de plus que f est une fonction impaire.

Enfin, on note que si $x \rightarrow +\infty$, alors $f(x) = \frac{1/x}{1+1/x^2} \rightarrow 0/1 = 0$.

Toutes ces informations donnent une idée approximative du graphe. Pouvez-vous l'esquisser ?

4 Règle de l'Hospital

Résultat XI.8. Supposons que f et g soient dérivables sur $]a-\delta, a+\delta[$ et $g'(x) \neq 0$ en tout $x \in]a-\delta, a+\delta[$ sauf peut-être pour $x = a$. Si

$$\lim_{x \rightarrow a} f(x) = 0, \quad \lim_{x \rightarrow a} g(x) = 0 \quad \text{et} \quad \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} = L \in \mathbb{R} \cup \{\pm\infty\}$$

alors

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

(en particulier, le quotient $\frac{f(x)}{g(x)}$ admet une limite pour $x \rightarrow a$).

Remarque. On peut adapter le résultat aux cas où a est remplacé par a^- ou a^+ ou $+\infty$ ou $-\infty$. Aussi aux cas où $\lim_{x \rightarrow a^+} f(x) = +\infty$ ou $-\infty$, $\lim_{x \rightarrow a^+} g(x) = \infty$ ou $-\infty$.

Ébauche de preuve. Une généralisation du théorème des accroissements finis montre l'existence, pour tout x , de c_x entre a et x vérifiant

$$\frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)}.$$

Comme on prend $x \rightarrow a$, on a forcément $c_x \rightarrow a$, d'où le résultat : si la limite du membre de droite existe, alors elle est forcément égale à la limite du membre de gauche. $\square$

Remarque. Dans le cas très particulier où $f(x)$ et $g(x)$ sont en fait dérivables également au point a et $g'(a) \neq 0$, alors le théorème est très simple à montrer puisque

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

suit simplement des règles de calcul !

Exemple. La limite suivante vaut 1 :

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x}$$

par application de la règle de l'Hospital. (En fait c'est de la triche, car pour montrer que $\sin' = \cos$, il faut préalablement calculer cette limite !)

La limite suivante se calcule également grâce à la règle :

$$\lim_{x \rightarrow 0} \frac{\cos(x) - 1}{\sin(x)} = \lim_{x \rightarrow 0} \frac{-\sin(x)}{\cos(x)} = 0$$

(Voyez-vous comment résoudre cette limite sans utiliser les dérivées ?)

Exemple. * Exemple avec une limite infinie en l'infini :

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

- * Exemple avec une limite 0/0 en $a = 1$

$$\begin{aligned}\lim_{x \rightarrow 1} \frac{\sqrt[3]{x^3 + 7} - 2}{x^3 + 3x^2 - 2x - 2} &= \frac{0}{0} \lim_{x \rightarrow 1} \frac{1/3(x^3 + 7)^{-2/3} 3x^2}{3x^2 + 6x - 2} \\ &= \lim_{x \rightarrow 1} \frac{(x^3 + 7)^{-2/3} x^2}{3x^2 + 6x - 2} \\ &= \frac{1/4}{7} = \frac{1}{28}\end{aligned}$$

- * Un exemple sans fraction (a priori)

$$\lim_{x \rightarrow 0} x \ln(x) = \lim_{x \rightarrow 0} \frac{\ln(x)}{1/x} = \lim_{x \rightarrow 0} \frac{1/x}{-1/x^2} = \lim_{x \rightarrow 0} -x = 0$$

- * Un exemple dans un exposant

$$\lim_{x \rightarrow 0} x^x = \lim_{x \rightarrow 0} \exp(\ln(x^x)) = \lim_{x \rightarrow 0} \exp(x \ln(x)) = \exp 0 = 1$$

5 Approximations polynomiales de Taylor

Bien que cela peut sembler être un paradoxe, toutes les sciences exactes sont dominées par l'idée d'approximation.
—Bertand Russell

Revenons sur l'idée d'approximer une fonction par une fonction plus simple. Nous avons déjà vu l'approximation affine, nous allons maintenant parler d'approximation par des fonctions polynomiales de degré plus élevé.

On a vu que la dérivée première d'une fonction f en un point a nous donnait une approximation de f par une fonction linéaire autour de a :

$$f(x) \approx f(a) + f'(a)(x - a)$$

et que de connaître la dérivée seconde de f en a nous permettait d'affiner cette approximation "du premier ordre" en une approximation "du second ordre" par une fonction quadratique autour de a :

$$f(x) \approx f(a) + f'(a)(x - a) + f''(a) \frac{(x - a)^2}{2}$$

Nous allons voir maintenant comment ceci se généralise.

Polynômes de Taylor

Considérons une fonction f dérivable n fois dans l'intervalle ouvert $]a - \delta, a + \delta[$ centré en a , de rayon $\delta > 0$. Le polynôme

$$T_{f,a,n}(x) = \sum_{k=0}^n f^{(k)}(a) \frac{(x-a)^k}{k!} = f(a) + f'(a)(x-a) + f''(a) \frac{(x-a)^2}{2} + \dots + f^{(n)}(a) \frac{(x-a)^n}{n!}$$

est le *polynôme de Taylor* d'ordre n de f au point a . Rappelons que $k! = k(k-1)(k-2) \dots 1$ est la factorielle de $k \in \mathbb{N}$ (par convention, $0! = 1$), et que $f^{(k)}$ désigne la dérivée k -ème de f . Le polynôme de Taylor d'ordre n de f est un polynôme de degré n dont la dérivée d'ordre k au point a est égale à la dérivée de f au point a , pour $k = 0, \dots, n$. C'est le seul tel polynôme.

L'erreur commise en approximant $f(x)$ par son polynôme de Taylor d'ordre n est notée $R_n(x)$, et s'appelle le *reste* d'ordre n de f . Donc :

$$R_n(x) = f(x) - T_n(x)$$

Notons que le reste possède un signe : si $R_n(x) > 0$ alors $T_n(x) < f(x)$ (sous-approximation), et si $R_n(x) < 0$ alors $T_n(x) > f(x)$ (sur-approximation). Étant donné que les dérivées de f et de T_n en a coïncident jusqu'à l'ordre n , on voit que la fonction reste satisfait

$$R_n(a) = R'_n(a) = \dots = R_n^{(n)}(a) = 0$$

Intuitivement, cette fonction est « très plate » autour de a . Voyons deux exemples.

polynôme de Taylor

reste

Exemple. Approximons $f(x) = \sin x$ autour de $a = 0$. Le calcul des dérivées de f est aisé :

$$f'(x) = \cos x \quad f''(x) = -\sin x \quad f'''(x) = -\cos x \quad f^{(4)}(x) = \sin x \quad \dots$$

et ainsi de suite. Evaluons maintenant f et ses dérivées en $a = 0$:

$$f(0) = 0 \quad f'(0) = 1 \quad f''(0) = 0 \quad f'''(0) = -1 \quad f^{(4)}(0) = 0 \quad \dots$$

et ainsi de suite. On voit que $f^{(k)}(a)$ est nulle si k est pair et alterne entre $+1$ et -1 pour k impair. On peut alors très facilement écrire les premiers polynômes de Taylor de $f(x) = \sin x$ en $a = 0$:

$$T_0(x) = 0$$

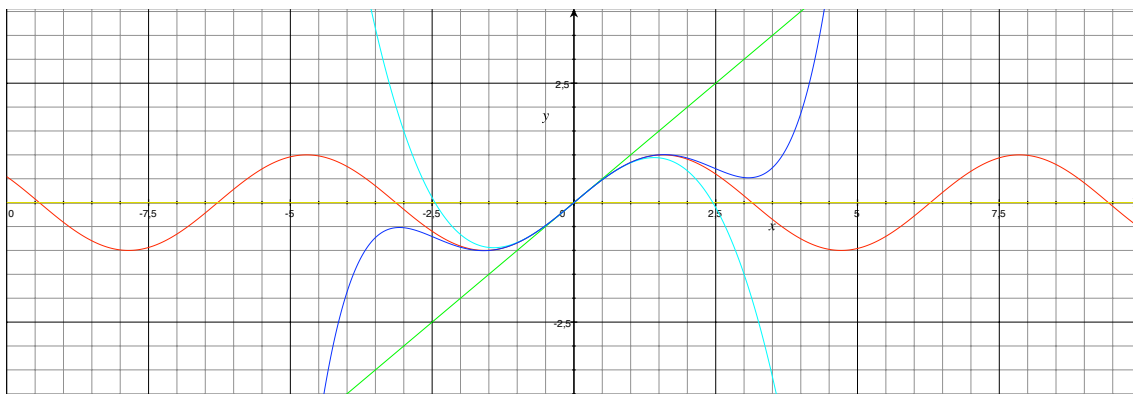
$$T_1(x) = 0 + 1 \cdot x$$

$$T_2(x) = 0 + 1 \cdot x + 0$$

$$T_3(x) = 0 + 1 \cdot x + 0 + (-1) \cdot \frac{x^3}{3!}$$

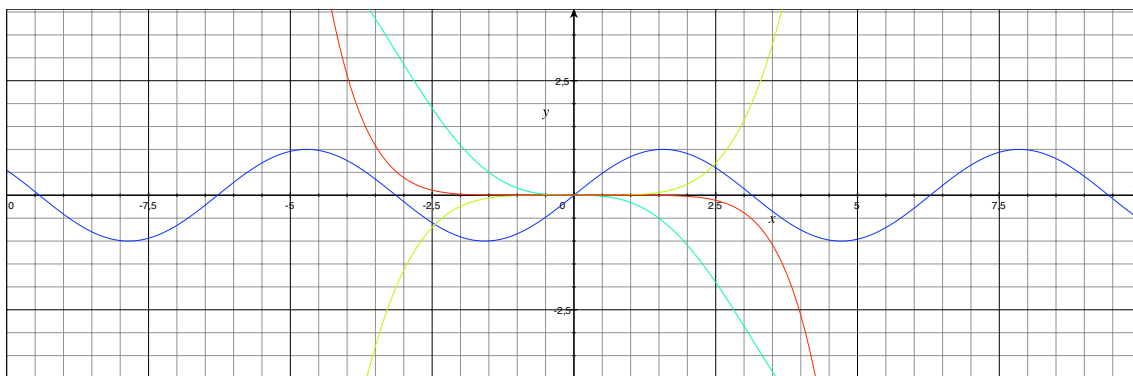
$$T_4(x) = 0 + 1 \cdot x + 0 + (-1) \cdot \frac{x^3}{3!} + 0$$

$$T_5(x) = 0 + 1 \cdot x + 0 + (-1) \cdot \frac{x^3}{3!} + 0 + 1 \cdot \frac{x^5}{5!}$$



On constate que les approximations se font meilleures au fur et à mesure que l'ordre n augmente. Remarquez qu'on a du mal à distinguer f de T_5 pour x à une distance 2 ou moins de a (c'est-à-dire $x \in [-2, 2]$) sur ces dessins.

Voyons maintenant à quoi ressemble la fonction reste $R_n = f - T_n$:



Comme promis, la fonction reste R_n se fait d'autant plus "plate" que n augmente.

Exemple. Prenons maintenant $f(x) = \ln x$ et $a = 1$. (Remarquons que cela reviendrait au même de prendre $f(x) = \ln(x+1)$ et $a = 0$.) Calculons les dérivées de f :

$$f'(x) = x^{-1} \quad f''(x) = -x^{-2} \quad f'''(x) = 2x^{-3} \quad f^{(4)}(x) = -6x^{-4} \quad \dots$$

On devine une formule générale pour $f^k(x)$, valable pour $k \geq 1$:

$$f^k(x) = (-1)^{k-1} (k-1)! x^{-k}$$

Évaluons maintenant les dérivées de f en $a = 1$:

$$f(1) = 0 \quad f'(1) = 1 \quad f''(1) = -1 \quad f'''(1) = 2 \quad f''''(1) = -6 \quad \dots$$

en général (pour $k \geq 1$) :

$$f^k(1) = (-1)^{k-1} (k-1)!$$

Les polynômes de Taylor de $f(x) = \ln x$ en $a = 1$ sont donc :

$$\begin{aligned} T_0(x) &= 0 \\ T_1(x) &= 0 + 1 \cdot (x-1) \\ T_2(x) &= 0 + 1 \cdot (x-1) - 1 \frac{(x-1)^2}{2} \\ T_3(x) &= 0 + 1 \cdot (x-1) - 1 \frac{(x-1)^2}{2} + 2 \frac{(x-1)^3}{6} \\ T_4(x) &= 0 + 1 \cdot (x-1) - 1 \frac{(x-1)^2}{2} + 2 \frac{(x-1)^3}{6} - 6 \frac{(x-1)^4}{24} \\ T_5(x) &= 0 + 1 \cdot (x-1) - 1 \frac{(x-1)^2}{2} + 2 \frac{(x-1)^3}{6} - 6 \frac{(x-1)^4}{24} + 24 \frac{(x-1)^5}{120} \end{aligned}$$

Après simplification, on voit la formule générale :

$$T_n(x) = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 + (-1)^{n-1} \dots \pm \frac{1}{n}(x-1)^n$$

Théorème de Taylor et formule du reste

La puissance du polynôme de Taylor tient en le théorème suivant, qui exprime que le polynôme de Taylor est effectivement la meilleure approximation polynomiale de f près de a en un certain sens :

Résultat XI.9. *Le polynôme de Taylor $T_{f,a,n}$ est l'unique polynôme de degré inférieur à n pour lequel la fonction de reste, R_n , est un « petit o » de $(x-a)^n$.*

D'après le théorème ci-dessus, on peut donc écrire :

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

ou encore

$$f(a+\delta x) = f(a) + f'(a)\delta x + o(\delta x)$$

ou encore

$$f(x+\delta x) = f(x) + f'(x)\delta x + o(\delta x).$$

De plus, par un raisonnement similaire à celui qui mène au théorème des accroissements finis, on peut obtenir une formule pour le reste $R_n(x)$:

Résultat XI.10 (Formule du reste de Lagrange). *Si f est dérivable $(n+1)$ -fois dans un intervalle $I =]a-\delta, a+\delta[$, alors pour chaque $x \in I$:*

$$f(x) = f(a) + f'(a)(x-a) + \dots + f^n(a) \frac{(x-a)^n}{n!} + f^{n+1}(t) \frac{(x-a)^{n+1}}{(n+1)!}$$

où t est un certain nombre entre a et x (qui dépend donc de x !), c'est-à-dire

$$R_n(x) = f^{n+1}(t) \frac{(x-a)^{n+1}}{(n+1)!}$$

pour un t entre a et x .

Voyons ce signifie ce théorème dans le cas le plus simple, quand $n = 0$:

$$f(x) = f(a) + f'(t)(x - a)$$

pour un t entre a et x . Si on prend $x > a$ cela signifie qu'il existe $t \in]a, x[$ tel que

$$f'(t) = \frac{f(x) - f(a)}{x - a}$$

ce n'est rien d'autre que le théorème des accroissements finis !

Le théorème de Taylor nous permet d'estimer l'erreur commise en approximant $f(x)$ par $T_n(x)$.

Exemple. Pour $f(x) = \ln x$ et $a = 1$. On a, en se basant sur les calculs précédents, pour $x = 2$:

$$\ln 2 = T_n(2) + R_n(2) = 1 - \frac{1}{2} + \frac{1}{3} - \dots + (-1)^{n-1} \frac{1}{n} + R_n(2)$$

avec

$$R_n(2) = (-1)^n n! t^{-n-1} \frac{1}{(n+1)!} = \frac{(-1)^n}{n+1} t^{-n-1}$$

pour un t entre $a = 1$ et $x = 2$. En particulier, t est positif et donc le *signe* de l'erreur est $(-1)^n$. C'est-à-dire, selon la parité de n on aura une sous-approximation ou une sur-approximation de $\ln 2$. Le problème, c'est qu'on ne sait pas *exactement* évaluer l'erreur parce qu'on ne sait pas ce qu'est t . On sait juste que $t \in (1, 2)$ et que l'erreur est proportionnelle à t^{-n-1} , qui est une fonction *décroissante* de t . Donc, au pire, t sera proche de 1 et on aura une erreur proche de $\frac{(-1)^n}{n+1}$. Dans tous les cas, on peut écrire

$$|R_n(2)| < \frac{1}{n+1}.$$

Comment choisir n , le nombre de termes dans notre approximation, pour être sûr de calculer *une* — ne soyons pas trop gourmands — décimale correcte de $\ln 2$? Si on prend

$$\frac{1}{n+1} \leq \frac{1}{10} \iff n+1 \geq 10 \iff n \geq 9$$

on a

$$|R_n(2)| < 10^{-1}$$

Calculons $T_9(2)$:

$$T_9(2) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} = 0,745634920\dots$$

qui est une sur-approximation, puis $T_{10}(2)$:

$$T_{10}(2) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \frac{1}{10} = 0,645634920\dots$$

qui est une sous-approximation. Ces deux nombres sont à une distance moins de 10^{-1} de $\ln 2$. Mais quel est celui dont la première décimale est correcte? Il faut plus de termes, en fait plus de 70 (!), pour conclure que $\ln 2 = 0,6\dots$

Exemple. Dans le cas du sinus, on peut calculer $\sin(1)$ avec une bonne précision bien plus rapidement. Le reste s'écrit toujours

$$R_n(x) = f^{n+1}(t) \frac{(1-0)^{n+1}}{(n+1)!}$$

mais cette fois $f^{n+1}(t)$ est toujours inférieur à 1 en valeur absolue (car les dérivées successives de $\sin$ sont $\cos$, $-\sin$, $-\cos$, $\sin$, etc.). On en déduit

$$|R_n(x)| \leq \frac{1}{(n+1)!}$$

ce qui diminue très rapidement : $\frac{1}{2}, \frac{1}{6}, \frac{1}{24}, \frac{1}{120}, \frac{1}{720}, \frac{1}{5040}$, il suffit donc de quelques termes pour obtenir une précision de plusieurs décimales.

Que retenir de ceci? Pour certaines fonctions, l'approximation de Taylor n'est « bonne » que dans un petit intervalle autour de a . Pour d'autres fonctions comme la fonction sinus, les choses se passent beaucoup mieux.

Comment choisir a ? Il faut choisir le point a pour que les calculs soient faisables. On connaît la valeur d'un sinus et de ses dérivées en 0 (ou en $\pi/2$, ou d'autres valeurs), mais pas en 1 par exemple.

6 Étude de fonctions

étude de fonction

Le but d'une *étude de fonction* est de rassembler des informations utiles sur une fonction, dans le but de mieux comprendre cette fonction. On peut par exemple vouloir dessiner un graphe approximatif, ou alors déterminer la (dé)croissance de la fonction étudiée. Voyons d'abord quelques propriétés intéressantes à mettre en exergue.

Parité

paire

Définition XI.11. Une fonction partielle $f : \mathbb{R} \rightarrow \mathbb{R}$ est dite *paire* si pour tout $x \in \text{dom } f$:

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

Dire qu'une fonction est paire signifie que son graphe est symétrique par rapport à l'axe oy .

impaire

Définition XI.12. Une fonction partielle $f : \mathbb{R} \rightarrow \mathbb{R}$ est dite *impaire* si pour tout $x \in \text{dom } f$:

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

Une fonction est impaire si et seulement si son graphe est symétrique par rapport à l'origine o .

Attention ! La plupart des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ ne sont ni paires ni impaires. C'est une erreur de penser que les fonctions impaires sont celles qui ne sont pas paires. Ceci est valable pour les nombres entiers, mais plus valable quand on parle de *fonctions*. Derrière cette constatation se cache une difficulté des mathématiques : le vocabulaire employé peut très bien changer de sens quand on change de contexte. Une autre différence avec les entiers : la somme de deux fonctions impaires est impaire !

Remarque. Les définitions de fonctions paires et impaires données ci-dessous sont valables à condition que leur domaine soit symétrique par rapport à 0. C'est le cas de $\mathbb{R}$, par exemple.

Remarque. On constate que pour toute fonction f dont le domaine est symétrique, si on définit $g(x) = \frac{f(x)+f(-x)}{2}$ et $h(x) = \frac{f(x)-f(-x)}{2}$, alors

1. ces fonctions vérifient $f = g + h$,
2. g est paire, et
3. h est impaire.

En conséquence de quoi, toute fonction (dont le domaine est symétrique, c'est-à-dire $x \in \text{dom } f \iff (-x) \in \text{dom } f$) est la somme d'une fonction paire et d'une fonction impaire.

Périodicité

période

Définition XI.13. Un nombre $p > 0$ est une *période* d'une fonction $f : A \rightarrow \mathbb{R}$ si pour tout $x \in \mathbb{R}$:

$$f(x+p) = f(x).$$

périodique

Une fonction est *périodique* si elle admet au moins une période. Dans le cas où l'ensemble des périodes de f possède un minimum, on appellera ce minimum *la* période de f . Sinon, on évitera de parler de "la" période de f (quelle serait, par exemple, "la" période d'une fonction constante?).

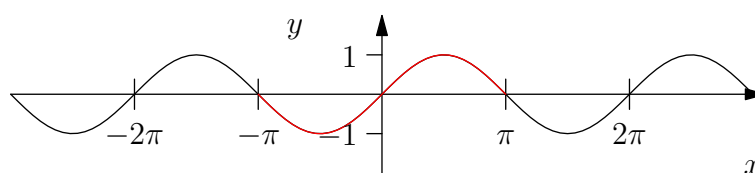
la

Exemple. La fonction sinus et la fonction cosinus

$$\sin : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \sin x$$

$$\cos : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \cos x$$

ont pour période 2π .



Exemple. (Ceci est un exemple tordu.) La fonction caractéristique des nombres rationnels admet $\mathbb{Q}_0^+$ comme ensemble de périodes. Cependant, dans le cadre de ce cours, hors les fonctions constantes et ce cas spécifique, les fonctions périodiques admettront généralement une période minimale.

Les fonctions périodiques peuvent être combinées pour en obtenir d'autres. C'est souvent possible de trouver *une* période d'une telle fonction mais trouver *la* période (si elle existe) peut se révéler plus difficile. Par exemple, la période des fonctions $\cos$ et $\sin$ est 2π . En conséquence de quoi, 2π est certainement *une* période de la fonction $\tan$, mais en réalité *la* période est π . Il n'existe pas de méthode générale pour déterminer *la* période (mais regarder l'ensemble des points où la fonction s'annule, par exemple, donne une idée).

Le graphe d'une fonction admettant p pour période se répète tous les p : traduire le graphe de f de p unités à droite (ou à gauche) ne change pas celui-ci.

Quelques règles utiles, si p est une période de f , q une période de g , h une fonction quelconque, et c une constante réelle :

1. p est une période de $h \circ f$;
2. $f \circ h$ n'est pas a priori périodique (mais peut l'être) ;
3. $x \mapsto f(x+c)$ est périodique de période p (un des rares cas où l'on est sûr de *la* période !);
4. $x \mapsto f(cx)$ est périodique de période $\frac{p}{c}$ (un autre rare cas) ;
5. $f+g$, fg , f/g admettent $\text{ppcm}(p,q)$ (si il existe, voir ci-dessous) comme période, mais il arrive qu'elle ne soit pas minimale.

Connaître *une* période p est déjà utile : on sait alors qu'on peut se restreindre à un intervalle de longueur p pour étudier la fonction, le reste étant reproduit à l'identique.

Définition XI.14. Si $a, b > 0$, leur « plus petit multiple commun », noté $\text{ppcm}(a,b)$ s'il existe, est le plus petit réel positif qui soit à la fois multiple entier de a et de b . Lorsque le ppcm n'existe pas, on dit que a et b sont incommensurables.

Exemple. Par exemple, $\text{ppcm}(1,2) = 2$ car les multiples entiers de 1 sont $1, 2, 3, 4, \dots$, et les multiples de 2 sont $2, 4, 6, 8, \dots$; le plus petit commun aux deux est bien 2.

De même $\text{ppcm}(\frac{1}{2}, \frac{1}{3}) = 1$.

Par contre $\text{ppcm}(\pi, 1)$ n'existe pas : 1 et π sont incommensurables. En effet, s'il existait m et n des entiers tels que $m1 = n\pi$, cela voudrait dire que π serait rationnel !

Quelques autres, sans calcul :

- * $\text{ppcm}(3,6) = 6$
- * $\text{ppcm}(4,6) = 12$
- * $\text{ppcm}(2\pi, 3\pi) = 6\pi$
- * $\text{ppcm}(4\pi, 3\pi) = 12\pi$
- * $\text{ppcm}(2\pi/3, \pi) = 2\pi$

Exemple. * $\sin x + \cos x$ est périodique de période 2π . Pour le prouver : on voit que cette fonction vaut 0 uniquement en $x = 3\pi/4 + k\pi$, donc la période ne peut être qu'un multiple de π . Comme π n'est pas une période (pourquoi?), et que 2π en est une, c'est bien 2π la plus petite !

* Le produit $f(x) = \sin x \cos x$ admet 2π pour période, mais *la* période de f est π . Pour le voir, on écrit $f(x) = \frac{1}{2} \sin(2x)$ (un cas rare!).

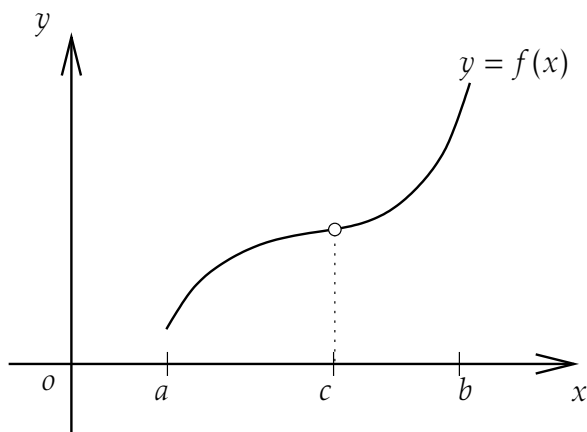
* Une période de la fonction $f : x \mapsto \frac{\sin(x)}{\sin(3x) + \tan(x)}$ est donnée comme suit : une période de $\sin(3x)$ est $\frac{2\pi}{3}$, une période de $\tan(x)$ est π , dès lors une période du dénominateur est $\text{ppcm}(\pi, \frac{2\pi}{3}) = 2\pi$. Le numérateur est de période 2π , dès lors une période du quotient est 2π . Il suffit donc d'étudier la fonction sur l'intervalle $[0, 2\pi]$ pour connaître la fonction dans sa globalité.

Concavité

La dérivée seconde d'une fonction f en un point c (quand elle existe) indique la « concavité » du graphe de f au point $(c, f(c))$. On parle de concavité « tournée vers le haut » quand $f''(c) > 0$, et « tournée vers le bas » quand $f''(c) < 0$.

Lorsque f'' change de signe en c (et en général s'y annule), alors la concavité du graphe de f change en c et on dit que c est un *point d'inflexion* de f .

point d'inflexion



Asymptotes

On distingue trois types d'asymptotes : les asymptotes « verticales », « horizontales » et « obliques ». La différence entre asymptote horizontale et oblique tient cependant à peu de choses comme nous le verrons. Considérons une fonction réelle f .

La droite d'équation $x = a$ est une *asymptote verticale* au graphe de f si

$$\lim_{x \rightarrow a^-} f(x) = \pm\infty \quad \text{ou} \quad \lim_{x \rightarrow a^+} f(x) = \pm\infty$$

(Ces limites n'ont de sens que lorsque a est adhérent au domaine de f .)

La droite d'équation $y = b$ est une *asymptote horizontale* au graphe de f lorsque

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

(Ces limites n'ont de sens que lorsque $\mp\infty$ est adhérent au domaine de f . On parle parfois d'asymptote horizontale à gauche dans le premier cas, à droite dans le second.)

La droite d'équation $y = mx + b$ où $m \neq 0$ est une *asymptote oblique* au graphe de f si

$$\lim_{x \rightarrow -\infty} [f(x) - (mx + b)] = 0 \quad \text{ou} \quad \lim_{x \rightarrow +\infty} [f(x) - (mx + b)] = 0$$

C'est le fait que $m \neq 0$ qui justifie l'adjectif « oblique ». Lorsque $m = 0$, on retrouve le cas d'une asymptote horizontale.

La proposition suivante permet de calculer m et b lorsqu'ils existent.

Résultat XI.15. La droite $y = mx + b$, où $m \neq 0$, est *asymptote oblique* au graphe de f pour x tendant vers $+\infty$ si et seulement si

$$\lim_{x \rightarrow +\infty} \frac{f(x)}{x} = m \neq 0 \quad \text{et} \quad \lim_{x \rightarrow +\infty} [f(x) - mx] = b$$

Exemple. Le cas de la fonction $\ln$ est intrigant car

$$\lim_{x \rightarrow \infty} \frac{\ln x}{x} = 0$$

(en application de la règle de l'Hospital), mais

$$\lim_{x \rightarrow \infty} \ln x = +\infty$$

dès lors il n'y a en réalité aucune asymptote en l'infini : ni horizontale, ni oblique ! (Par contre il y a bien une asymptote verticale en 0.)

asymptote
verticale

asymptote
horizontale

asymptote
oblique

Cahier des charges d'une étude de fonction

Voici la liste des informations à rassembler dans une étude de fonction utile :

1. Le domaine de f (au cas où f est définie par une expression et que ce domaine n'est pas explicitement donné)?
2. Un tableau de signe de f .
3. f est-elle paire? impaire?
4. f est-elle périodique? Si elle existe, donner la période de f .
5. Déterminer l'ensemble des points où f est dérivable et calculer f' en ces points.
6. Faire un tableau de signe de f' .
7. Déterminer l'ensemble des points où f est deux fois dérivable et calculer f'' en ces points.
8. Faire un tableau de signe de f'' .
9. Combiner les tableaux de signe de f' et f'' pour établir un tableau de variation de f . Quels sont les intervalles de croissance et décroissance? Quels sont les extrema locaux? Quels sont les intervalles de concavité et convexité? Quels sont les points d'inflexion?
10. Déterminer les éventuelles asymptotes de f .
11. Esquisser le graphe de f .

Chapitre XII

Fonction réelles d'une variable réelle : primitives

1 Motivation

1.1 Équation aux dérivées pour une population

Question. On considère une population de bactéries au cours du temps, et on note $p(t)$ le nombre de bactéries à l'instant t . On suppose que le taux de croissance de ces bactéries est lié à deux facteurs : le nombre de bactéries, et les ressources disponibles. On peut modéliser cela par l'équation

$$p'(t) = p(t)(K - p(t))$$

où K représente les ressources disponibles.

Quelles sont les fonctions qui vérifient cette équation en tout t ?

Remarque. Une telle équation est une équation différentielle.

Réponse. Supposons que la population ne disparaît pas ($p(t) \neq 0$) et ne dépasse jamais les ressources ($p(t) \neq K$), alors :

$$\frac{p'(t)}{p(t)(K - p(t))} = 1$$

Notons $f(x) = \frac{1}{x(K-x)}$. Alors l'équation se ré-écrit :

$$f(p(t))p'(t) = 1$$

C'est-à-dire, en notant F une fonction telle que $F' = f$, on aurait

$$(F(p(t)))' = 1.$$

Si on savait trouver des fonctions dont la dérivée est une fonction donnée, on pourrait terminer ces calculs !

1.2 Équation aux dérivées pour le ressort

* L'accélération a est la dérivée de la vitesse, et la vitesse est la dérivée de la position.

$$a(t) = v'(t) = x''(t)$$

* L'équation de Newton dit : la force est proportionnelle à l'accélération.

Question. Quel est le mouvement d'un ressort ?

Réponse. Modélisons le ressort comme suit : on définit l'élongation $e(t)$ du ressort à l'instant t comme la différence entre sa position au repos et sa position à l'instant t .

* Si le ressort est étiré, $e(t) > 0$;

* si le ressort est comprimé, $e(t) < 0$.

Supposons que la force soit proportionnelle à l'élongation, mais opposée.

Réponse. On suppose donc qu'il existe $k > 0$ avec

$$F = -ke(t)$$

Par Newton :

$$ma = -ke(t)$$

Or $e(t) = x(t) - x_{\text{repos}}$, donc $e'(t) = x'(t)$ et $e''(t) = x''(t)$. On a alors :

$$me''(t) = -ke(t)$$

Si seulement nous pouvions résoudre cela ! Mais comment ?

2 Définition

Définition XII.1. Une *primitive* de f est une fonction F dérivable sur $\text{dom } f$ telle que :

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

Si F est une primitive de f sur I alors $F+C$, où C est une constante arbitraire, est aussi une primitive de f sur I . Ceci montre donc que la notion de primitive n'est pas unique. Cependant, une propriété remarquable et extrêmement importante des primitives est le résultat suivant :

Résultat XII.2. Si F et G sont deux primitives de f sur un même intervalle ouvert I , alors il existe une constante $C \in \mathbb{R}$ telle que $G(x) = F(x) + C$ pour tout x dans I .

Démonstration. La fonction $h := G - F$ a une dérivée nulle sur I . Supposons par l'absurde que h ne soit pas constante, c'est-à-dire qu'il existe $x, y \in I$ avec $h(x) \neq h(y)$. Alors d'après le théorème de Rolle, $\frac{h(x)-h(y)}{x-y} = h'(c) = 0$, ce qui prouve $h(x) = h(y)$, et est une contradiction. Donc h est une constante, disons C , c'est-à-dire $G(x) - F(x) = C$; ce que nous voulions démontrer. $\square$

Les constantes apparaissant de la sorte s'appellent parfois *constante d'intégration*. Dès lors, pour une fonction f définie sur un intervalle, il n'y a peut-être pas une seule primitive, mais quand on en connaît une, disons F , on les connaît toutes : il s'agit de l'ensemble $\{F + C \text{ t.q. } C \in \mathbb{R}\}$.

Remarque. Si f est définie sur un domaine plus compliqué qu'un intervalle alors chaque « morceau » du domaine a sa propre constante d'intégration.

Exemple. On peut montrer (pouvez-vous le faire ?) que les primitives de la fonction

$$f : \mathbb{R}_0 \rightarrow \mathbb{R} : x \mapsto \frac{1}{x}$$

sont toutes les fonctions $F : \mathbb{R}_0 \rightarrow \mathbb{R}$ telles que

$$F(x) = \begin{cases} \ln(x) + C_1 & \text{si } x > 0 \\ \ln(-x) + C_2 & \text{si } x < 0. \end{cases}$$

On écrit en général plutôt $F(x) = \ln|x| + C$, ce qui n'est pas tout à fait vrai puisqu'on manque des possibilités (la constante peut être différente sur $\mathbb{R}^+$ et sur $\mathbb{R}^-$), mais cela suffit à la plupart des applications concrètes.

Remarque. Rechercher les primitives d'une fonction donnée $f : I \rightarrow \mathbb{R}$, c'est résoudre une équation du type

$$y'(x) = f(x) \quad x \in I$$

où y est une *fonction inconnue*, et f est la fonction donnée. C'est notre premier exemple d'*équation différentielle* : une équation dont l'inconnue n'est pas un nombre mais une fonction, et qui fait intervenir la dérivée de cette fonction.

Si F est une primitive de f , on écrit

$$\int f(x) dx = F(x) \quad \left(\text{ou parfois : } \int f = F \right)$$

ou, lorsqu'on cherche *toutes* les primitives :

$$\int f(x) dx = F(x) + C \quad \left(\text{ou parfois : } \int f = F + C \right)$$

et on dit que F est une *intégrale indéfinie* de f . Trouver une primitive d'une fonction est une opération qui inverse la dérivation. On a le schéma suivant pour une fonction f et sa primitive F :

 intégrale
indéfinie

$$F + C \xrightarrow{\frac{d}{dx}} f \quad F + C \xleftarrow{\int dx} f.$$

Remarque. De la même manière que $f'(x)$ ou $\frac{d}{dx}f(x)$ désignent « la dérivée de f évaluée au point x », l'expression $\int f(x) dx$ désigne « une primitive de f évaluée a point x ». On écrira parfois

$$\int_{|x=a} f(x) dx$$

pour exprimer qu'on évalue la primitive donnée au point a .

Exemple. La fonction $f(x) = x^2$ admet $F(x) = \frac{x^3}{3}$ comme primitive, car

$$F'(x) = \left(\frac{x^3}{3} \right)' = \frac{1}{3} (x^3)' = \frac{1}{3} \cdot 3x^2 = x^2 = f(x)$$

Donc on peut écrire

$$\int x^2 dx = \frac{x^3}{3} + C$$

ce qui est simplement une façon d'écrire qu'on a là *toutes* les primitives de $f(x) = x^2$.

Remarque. Dans la plupart des cas, nous chercherons les primitives d'une *fonction élémentaire* (c'est-à-dire obtenue en combinant des fonctions exponentielle, logarithme, racine n -ème, trigonométriques et trigonométriques inverses en effectuant des sommes, produits, quotients et compositions). Bien que dans nos exemples, les primitives obtenues sont *encore* des fonctions élémentaires, c'est loin d'être toujours le cas. On peut démontrer (mais pas dans le cadre de ce cours) que les primitives d'une fonction élémentaire « prise au hasard » ne sont presque sûrement jamais élémentaires.

 fonction
élémentaire

Exemple. L'intégrale indéfinie

$$\int e^{-x^2} dx$$

n'est *pas* une fonction élémentaire. Pourtant les primitives existent, et voici d'ailleurs le graphe de l'une d'elle :

2.1 Notations

Si f est une fonction, on notera $\int f$ ses primitives. C'est-à-dire $\int f$ représente toute fonction vérifiant $(\int f)' = f$.

On notera parfois aussi :

$$\int f(x) dx$$

et la notation $\int dx f(x)$ existe également.

Remarquons que si on note $F'(x) = \frac{dF}{dx}$, alors ces notations ont des similitudes. On dit aussi F est une *intégrale indéfinie* de f .

 intégrale
indéfinie

Remarque. On écrira parfois

$$\int_{x=a} f(x) dx$$

pour exprimer qu'on évalue la primitive donnée au point a .

Exemple.

$$\int_{x=3} x^2 dx = \frac{x^3}{3} \Big|_{x=3} = \frac{3^3}{3} = 9$$

Remarque (*). Cette notation est cependant très mauvaise, car la valeur finale dépend de la primitive trouvée ! Par exemple $\frac{x^3}{3} + 1$ est aussi une primitive de x^2 , donc on aurait alors

$$\int_{x=3} x^2 dx = \left(\frac{x^3}{3} + 1 \right) \Big|_{x=3} = \frac{3^3}{3} + 1 = 10 \quad \text{Absurde !}$$

3 Intégration immédiate

Nous allons maintenant passer en revue une série de techniques d'intégration. Dans les cas les plus simples, on voit « directement » une fonction F dont la dérivée est la fonction f donnée. On obtient alors toutes les primitives de f . C'est ce qu'on appelle l'*intégration immédiate*.

En réécrivant les formules de dérivation vues précédemment, on obtient les formules d'intégration immédiate suivantes.

$$\begin{aligned} \int x^m dx &= \frac{x^{m+1}}{m+1} + C & m \in \mathbb{N}, \quad x \in \mathbb{R} \\ \int x^{-m} dx &= \frac{x^{-m+1}}{-m+1} + C & m \in \{2, 3, \dots\}, \quad x \in \mathbb{R}_0 \\ \int x^r dx &= \frac{x^{r+1}}{r+1} + C & r \neq -1, \quad x \in \mathbb{R}_0^+ \\ \int \frac{1}{x} dx &= \ln|x| + C & x \in \mathbb{R}_0 \\ \int \sin x dx &= -\cos x + C & x \in \mathbb{R} \\ \int \cos x dx &= \sin x + C & x \in \mathbb{R} \\ \int \frac{1}{\cos^2 x} dx &= \operatorname{tg} x + C & x \in \mathbb{R} \setminus \left\{ \dots, -\frac{3\pi}{2}, -\frac{\pi}{2}, \frac{\pi}{2}, \frac{3\pi}{2}, \dots \right\} \\ \int \frac{1}{\sin^2 x} dx &= -\operatorname{cotg} x + C & x \in \mathbb{R} \setminus \{ \dots, -2\pi, -\pi, 0, \pi, 2\pi, \dots \} \\ \int e^x dx &= e^x + C & x \in \mathbb{R} \\ \int b^x dx &= \frac{b^x}{\ln b} + C & b > 0, b \neq 1, \quad x \in \mathbb{R} \\ \int \frac{1}{\sqrt{1-x^2}} dx &= \arcsin x + C & x \in]-1, 1[\\ \int \frac{1}{1+x^2} dx &= \operatorname{arctg} x + C & x \in \mathbb{R} \end{aligned}$$

Remarque (*). Dans tous ces exemples, on a écrit une seule constante d'intégration C , même dans les cas où il y a plusieurs constantes d'intégration (une par « morceau » du domaine).

Sinus et cosinus hyperboliques

Il existe deux fonctions élémentaires qui se révèlent utiles dans le calcul de primitives : la fonction *cosinus hyperbolique*

cosinus
hyperbolique

$$\operatorname{ch} : \mathbb{R} \rightarrow [1, +\infty[: x \mapsto \operatorname{ch} x = \frac{e^x + e^{-x}}{2}$$

et la fonction *sinus hyperbolique*

sinus
hyperbolique

$$\operatorname{sh} : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \operatorname{sh} x = \frac{e^x - e^{-x}}{2}$$

Ces fonctions ont les propriétés suivantes. Pour tout $x \in \mathbb{R}$:

$$\begin{aligned} \operatorname{ch}^2 x - \operatorname{sh}^2 x &= 1 \\ \operatorname{ch}(-x) &= \operatorname{ch} x && (\operatorname{ch} \text{ est paire}) \\ \operatorname{sh}(-x) &= -\operatorname{sh} x && (\operatorname{sh} \text{ est impaire}) \\ (\operatorname{ch} x)' &= \operatorname{sh} x \\ (\operatorname{sh} x)' &= \operatorname{ch} x. \end{aligned}$$

Exercice. Montrer les faits suivants :

1. ch est strictement positive sur $\mathbb{R}$,
2. sh est strictement croissante sur $\mathbb{R}$, et
3. ch est strictement décroissante sur $\mathbb{R}^-$ et strictement croissante sur $\mathbb{R}^+$.

En déduire que sh est une bijection de $\mathbb{R}$ dans $\mathbb{R}$, et ch (restreinte à $\mathbb{R}^+$) une bijection de $\mathbb{R}^+$ dans $[1, +\infty[$.

Les fonctions cosinus et sinus hyperbolique ont des fonctions réciproques :

$$\operatorname{argch} : [1, +\infty[\rightarrow \mathbb{R}^+ : x \mapsto \operatorname{argch} x$$

et

$$\operatorname{argsh} : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \operatorname{argsh} x$$

dont les dérivées sont

$$(\operatorname{argch} x)' = \frac{1}{(\operatorname{ch})'(\operatorname{argch} x)} = \frac{1}{\operatorname{sh}(\operatorname{argch} x)} = \frac{1}{\sqrt{x^2 - 1}}$$

pour $x \in]1, +\infty[$, et

$$(\operatorname{argsh} x)' = \frac{1}{(\operatorname{sh})'(\operatorname{argsh} x)} = \frac{1}{\operatorname{ch}(\operatorname{argsh} x)} = \frac{1}{\sqrt{1 + x^2}}$$

pour $x \in \mathbb{R}$.

Ce qui donne les formules d'intégration immédiate ci-dessous :

$$\begin{aligned} \int \operatorname{ch} x \, dx &= \operatorname{sh} x + C && x \in \mathbb{R} \\ \int \operatorname{sh} x \, dx &= \operatorname{ch} x + C && x \in \mathbb{R} \\ \int \frac{1}{\sqrt{x^2 - 1}} \, dx &= \operatorname{argch} x + C && x \in]1, +\infty[\\ \int \frac{1}{\sqrt{1 + x^2}} \, dx &= \operatorname{argsh} x + C && x \in \mathbb{R} \end{aligned}$$

4 Linéarité

Supposons que f et g sont des fonctions admettant des primitives F et G . C'est-à-dire $F' = f$ et $G' = g$. Si $a, b \in \mathbb{R}$ sont des constantes, alors

$$(aF + bG)' = aF' + bG' = af + bg$$

et par conséquent

$$\int af + bg = a \int f + b \int g$$

Exemple.

$$\begin{aligned} \int (x^2 + x^3 + \sqrt{x}) dx &= \int x^2 dx + \int x^3 dx + \int \sqrt{x} dx \\ &= \frac{x^3}{3} + \frac{x^4}{4} + \frac{x^{3/2}}{3/2} + C \\ &= \frac{x^3}{3} + \frac{x^4}{4} + \frac{2\sqrt{x^3}}{3} + C \end{aligned}$$

- * Par linéarité
- * Par primitivisation immédiate

Exemple. Considérons la fonction polynomiale P définie par

$$P(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n,$$

où $a_0, a_1, \dots, a_n \in \mathbb{R}$ sont des constantes. Par intégration immédiate,

$$\begin{aligned} \int dx &= x + C \\ \int x dx &= \frac{x^2}{2} + C \\ &\vdots \\ \int x^n dx &= \frac{x^{n+1}}{n+1} + C \end{aligned}$$

D'où, par linéarité,

$$\begin{aligned} \int P(x) dx &= \int (a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) dx \\ &= a_0 \int dx + a_1 \int x dx + a_2 \int x^2 dx + \cdots + a_n \int x^n dx \\ &= a_0x + a_1 \frac{x^2}{2} + a_2 \frac{x^3}{3} + \cdots + a_n \frac{x^{n+1}}{n+1} + C \end{aligned}$$

Exercice. Si $a \in \mathbb{R}$ est une constante, déterminer

$$\int a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots + a_n(x-a)^n dx.$$

Exemple.

$$\begin{aligned} \int \left(e^x - \frac{1}{x} \right) dx &= \int e^x dx - \int \frac{1}{x} dx \\ &= e^x - \ln|x| + C \end{aligned} \quad (x \in \mathbb{R}_0)$$

5 Intégration par changement de variable

Rappelons que si f et g sont dérivables, alors $g \circ f$ l'est également et

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

Donc

$$\int g'(f(x))f'(x)dx = g(f(x)) + C.$$

Si on écrit $h = g'$, de sorte que $g = \int h$, on peut alors écrire deux formules utiles :

$$\int h(f(x)) \cdot f'(x) dx = \int_{u=f(x)} h(u) du$$

et

$$\int h(u) du = \int_{x=f^{-1}(u)} h(f(x))f'(x) dx.$$

où la barre verticale indique qu'après avoir intégré, on remplace u par $f(x)$ (première formule) ou x par $f^{-1}(u)$ (seconde formule).

Afin de se rappeler ces deux formules, on pourra écrire $u = f(x)$ et $du = f'(x)dx$. Cette seconde égalité en particulier n'a aucun sens ! Néanmoins c'est un moyen mnémotechnique classique et efficace pour réaliser les changements de variable.

Remarque. Rappelons :

1. Ne jamais mélanger ancienne et nouvelle variable au sein de l'intégrale !
2. Si l'ancienne variable s'appelle x et la nouvelle s'appelle u : dx apparaissait au début, il *doit* être remplacé par *une* occurrence de du . Il ne peut y avoir de du au carré (ou dans une autre fonction), de du au dénominateur, ou de « du ajouté à autre chose ».

Exemple.

$$\int 3x^2(x^3 + 5)^9 dx$$

On remarque qu'en posant $u = x^3 + 5$, on obtient $du = 3x^2 dx$. Et donc

$$\int 3x^2(x^3 + 5)^9 dx = \int \underbrace{(x^3 + 5)^9}_{=u^9} \cdot \underbrace{3x^2 dx}_{=du} = \int u^9 du$$

De $\int u^9 du = \frac{u^{10}}{10} + C$, on arrive à

$$\int 3x^2(x^3 + 5)^9 dx = \frac{(x^3 + 5)^{10}}{10} + C.$$

On vérifiera qu'en dérivant le membre de droite, on obtient bien l'intégrande dans le membre de gauche.

Exemple. Pour calculer

$$\int \sin^4 x \cos x dx$$

on pose $u = \sin x$, et on obtient

$$\int \sin^4 x \cos x dx = \int u^4 du = \frac{u^5}{5} + C = \frac{\sin^5 x}{5} + C.$$

Notons qu'il faut souvent transformer l'intégrande avant de pouvoir effectuer un changement de variable.

Remarque. Rappelons que « calculer une intégrale », par changement de variable ou toute autre méthode, est en général impossible (voir l'exemple de e^{-x^2}) et que dans les cas où c'est possible, cela requiert de la pratique et des astuces !

Exemple. Dans cet exemple, nous montrons comment la formule peut s'utiliser « à l'envers ». On considère le calcul suivant :

$$\int \frac{u^3}{\sqrt{1+u^2}} du$$

et l'on pose $u = \tan(t)$ (ce qui semble compliquer les choses plutôt que de les simplifier !) de sorte que $du = (1 + \tan^2(t))dt$ et l'intégrale devient, en utilisant l'égalité $1 + \tan^2(t) = \frac{1}{\cos^2(t)}$,

$$\int \frac{\sin^3(t)}{\cos^4(t)} dt.$$

Ce nouveau calcul de primitive semble à peine plus simple, mais une astuce classique consiste à transformer des sinus en cosinus (voir ci-dessous). Voyons cela en posant à présent $v = \cos(t)$, ce qui donne $dv = -\sin(t)dt$ et l'intégrale devient alors :

$$\int \frac{v^2 - 1}{v^4} dv = \frac{-1}{v} + \frac{1}{3v^3} = \frac{1}{v} \left(\frac{1}{3v^2} - 1 \right) = \sqrt{1+u^2} \left(\frac{u^2 - 2}{3} \right)$$

(où l'on a utilisé $\sin^3(t) = \sin(t)\sin^2(t) = -\sin(t)(\cos^2(t) - 1)$ et $v^2(u^2 + 1) = 1$).

Exemple. L'exemple précédent pouvait se résoudre également comme suit :

$$\begin{aligned} \int \frac{u^3}{\sqrt{1+u^2}} du &= \int \left(\frac{u(u^2+1)}{\sqrt{1+u^2}} - \frac{u}{\sqrt{1+u^2}} \right) du \\ &= \int \left(\sqrt{u^2+1} - \frac{1}{\sqrt{u^2+1}} \right) u du \\ &= \frac{1}{2} \int \left(\sqrt{z} - \frac{1}{\sqrt{z}} \right) dz \\ &= \frac{1}{3} \sqrt{z^3} - \sqrt{z} \\ &= \frac{1}{3} \sqrt{(1+u^2)^3} - \sqrt{1+u^2}. \end{aligned}$$

où l'on a posé $z = u^2 + 1$ de sorte que $dz = 2u du$.

Remarque (*). Il n'y a pas de « bonne manière » de résoudre un calcul de primitive : l'important est d'obtenir une expression dont la dérivée correspond à l'expression dont on part. Par exemple, $-\arccos(x)$ et $\arcsin(x)$ sont deux primitives de $\frac{1}{\sqrt{1-x^2}}$

6 Intégration par parties

Tout comme pour les techniques précédentes, le point de départ est une règle de dérivation. Si f et g sont deux fonctions dérivables, la formule de Leibniz (ou règle de dérivation d'un produit) nous dit

$$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x) \quad \text{donc} \quad f'(x)g(x) = (f(x)g(x))' - f(x)g'(x)$$

En intégrant, on trouve la *formule d'intégration par parties*

$$\int f'(x)g(x) dx = f(x)g(x) - \int f(x)g'(x) dx.$$

Ici, la difficulté est de pouvoir reconnaître quand la fonction à intégrer peut s'écrire sous la forme d'un produit $f'(x)g(x)$. Ce qu'on veut, c'est intégrer un des facteurs et dériver l'autre, de telle manière à obtenir une intégrale indéfinie plus simple.

formule
d'intégration par
parties

Exemple. Pour calculer

$$\int x \cos x \, dx$$

posons

$$g(x) = x, \quad f'(x) = \cos x$$

et donc

$$g'(x) = 1, \quad f(x) = \sin x$$

de sorte que le calcul devient

$$\begin{aligned} \int x \cos x \, dx &= (\sin x)x - \int (\sin x)1 \, dx \\ &= x \sin x - (-\cos x) + C \\ &= x \sin x + \cos x + C \end{aligned}$$

Remarquons qu'on aurait très bien pu intégrer x et dériver $\cos x$, au lieu de dériver x et intégrer $\cos x$. Mais on aurait obtenu

$$\int x \cos x \, dx = \frac{x^2}{2} \sin x - \int \frac{x^2}{2} \cos x \, dx$$

ce qui est en fait plus complexe que l'intégrale de départ et semble donc une voie sans issue.

Remarque. Cet exemple montre que le choix de $f'(x)$ et $g(x)$ est crucial dans une intégration par parties, et que si un premier essai s'est soldé par un échec, on a tout intérêt à faire un second essai en inversant les rôles de $f'(x)$ et $g(x)$.

Exemple. Calculer

$$\int \ln x \, dx$$

Remarquons que $\ln x$ est défini pour $x > 0$ donc l'intervalle où on travaille est $I =]0, \infty[$. Dans le cas qui nous occupe, l'intégrande n'est pas un produit de deux fonctions mais bien une seule fonction. Bien sûr, on peut toujours écrire

$$\int \ln x \, dx = \int 1 \ln x \, dx$$

Ce qui donne l'idée de prendre

$$\begin{aligned} g(x) &= \ln x, & f'(x) &= 1 \\ g'(x) &= \frac{1}{x}, & f(x) &= x \end{aligned}$$

Alors

$$\int \ln x \, dx = x \ln x - \int x \frac{1}{x} \, dx = x \ln x - \int 1 \, dx = x \ln x - x + C$$

pour $x \in]0, \infty[$.

7 Décomposition en fractions simples

Rappelons qu'une fonction rationnelle est une fonction du type $\frac{P(x)}{Q(x)}$ où P et Q sont des polynômes. Un résultat surprenant est que l'intégrale indéfinie d'une telle fonction,

$$\int \frac{P(x)}{Q(x)} \, dx$$

est *toujours* une fonction élémentaire (mais pas toujours une fonction rationnelle!). Ceci se démontre grâce à une technique d'intégration appelée *décomposition en fractions simples*. Cette technique ayant d'autres applications que le calcul de primitives, nous verrons en détails comment l'appliquer.

décomposition
en fractions
simples

7.1 Irréductibilité

irréductible

Définition XII.3. Un polynôme est *irréductible* s'il est non-constant et ne peut pas s'écrire sous la forme d'un produit de deux polynômes non-constants.

En d'autres termes, un polynôme est irréductible s'il n'est pas factorisable !

Résultat XII.4. Les polynômes irréductibles sont exactement la réunion des polynômes de degré 1 avec l'ensemble des polynômes de degré 2 n'ayant pas de racine (ne s'annulant pas).

Exemple. Les polynômes suivants sont irréductibles :

$$X, X-1, 3X+5, X^2+1, X^2+2X+2, \dots$$

tandis que ceux-ci ne le sont pas :

$$X^2-1, (X^2+1)^2, 3X^4+2X+1, X^4+1.$$

Remarque (*). Le résultat précédent dit en fait que tout polynôme de degré au moins égal à 3 est factorisable. Il ne dit cependant pas que ces polynômes ont forcément une racine : X^4+1 n'a bien évidemment aucune racine (dans $\mathbb{R}$). En fait, on peut vérifier que

$$X^4+1 = (X^2 - \sqrt{2}x + 1)(X^2 + \sqrt{2}x + 1).$$

7.2 Division Euclidienne

Rappelons que la division euclidienne de deux entiers m, n , $n \neq 0$, est la recherche de deux entiers d et r vérifiant $m = dn + r$ avec $0 \leq r < d$. d et r sont respectivement appelés le quotient et le reste de la division de m par n . Ces deux entiers sont uniques.

La division Euclidienne P/Q de deux polynômes P et Q a pour but de trouver deux polynômes D et R vérifiant $P = DQ + R$ et $\deg R < \deg Q$. Ces deux polynômes sont uniques !

Exemple. Par exemple la division de X^4 par $X+1$ donne :

$$X^4 = (X+1)(X^3 - X^2 + X - 1) + 1.$$

on peut obtenir cela par division Euclidienne classique (mais légèrement adaptée : ce qui est compte est le coefficient du plus haut degré)

$$\begin{array}{r} X^4 = (X+1)(X^3 - X^2 + X - 1) + 1 \\ -X^4 - X^3 \\ \hline -X^3 \\ X^3 + X^2 \\ \hline X^2 \\ -X^2 - X \\ \hline -X \\ X+1 \\ \hline 1 \end{array}$$

ou par un schéma d'Hörner (uniquement lorsque le dénominateur est de la forme $X - a$!):

$$-1 \left| \begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 \\ & -1 & 1 & -1 & 1 \\ \hline & 1 & -1 & 1 & -1 & 1 \end{array} \right.$$

7.3 Fraction simple

Définition XII.5. Une *fraction simple* est une fonction rationnelle dont le dénominateur est une puissance d'un polynôme irréductible de degré n (n vaut 1 ou 2) et dont le numérateur a un degré strictement inférieur à n .

fraction simple

Remarque (*). En d'autres termes, les fractions simples sont de la forme :

$$\frac{A}{(Bx + C)^k} \quad \text{ou} \quad \frac{Ax + B}{(Cx^2 + Dx + E)^k}$$

avec $D^2 - 4CE < 0$, pour certaines constantes réelles A, B, C, D, E et un entier $k > 0$

Remarque (*). On peut toujours ramener une fraction simple à l'une des formes suivantes :

$$\frac{u}{(x - a)^k} \quad \text{ou} \quad \frac{ux + v}{((x - b)^2 + c^2)^k}$$

pour certaines constantes réelles u, v, a, b , un réel $c > 0$ et un entier $k > 0$.

Un intérêt de ces fractions simples est qu'on peut relativement aisément en trouver des primitives, et qu'on a la proposition suivante.

Résultat XII.6. *Toute fonction rationnelle est la somme*

1. *d'un polynôme (éventuellement constant voire nul), et*
 2. *de fractions simples dont le dénominateur divise celui de la fonction rationnelle.*
- (De plus, cette décomposition est unique.)*

En pratique, pour obtenir cette décomposition en fraction simple d'une fonction rationnelle, il faudra

1. si nécessaire, appliquer l'algorithme de division Euclidienne pour obtenir la somme d'un polynôme avec une nouvelle fonction rationnelle, dont le numérateur est de degré inférieur à celui du dénominateur ;
2. factoriser le dénominateur en ses composantes irréductibles ;
3. écrire toutes les fractions simples possibles pouvant intervenir dans la décomposition, et déterminer les constantes.

7.4 Un exemple

Calculer

$$\int \frac{1}{1 - x^2} dx$$

La fonction à intégrer est définie sur $\mathbb{R} \setminus \{\pm 1\}$, qui est la réunion de trois intervalles ouverts. En factorisant le dénominateur, nous pouvons réécrire notre intégrande :

$$\frac{1}{1 - x^2} = \frac{1}{(1 - x)(1 + x)}$$

Cherchons maintenant à décomposer l'intégrande comme suit somme de fractions simples :

$$\frac{1}{(1 - x)(1 + x)} = \frac{A}{1 - x} + \frac{B}{1 + x}$$

où A et B sont des constantes à déterminer. On comprend immédiatement l'intérêt d'une telle décomposition. Si on parvient à trouver $A, B \in \mathbb{R}$ comme ci-dessus, alors

$$\begin{aligned} \int \frac{1}{1 - x^2} dx &= \int \frac{1}{(1 - x)(1 + x)} dx \\ &= \int \left(\frac{A}{1 - x} + \frac{B}{1 + x} \right) dx \\ &= A \int \frac{1}{1 - x} dx + B \int \frac{1}{1 + x} dx \\ &= -A \ln|1 - x| + B \ln|1 + x| + C \end{aligned}$$

Dans un sens, le calcul de notre intégrale indéfinie repose entièrement sur le calcul des constantes A et B de la décomposition. Voyons comment les déterminer. Repartons de l'équation

$$\frac{1}{(1-x)(1+x)} = \frac{A}{1-x} + \frac{B}{1+x}$$

que nous voulons satisfaire, et réexprimons le membre de droite :

$$\frac{A}{1-x} + \frac{B}{1+x} = \frac{A(1+x) + B(1-x)}{(1-x)(1+x)} = \frac{A+B + (A-B)x}{(1-x)(1+x)}$$

cette expression doit être égale à

$$\frac{1}{(1-x)(1+x)}$$

pour tout $x \in \mathbb{R} \setminus \{\pm 1\}$, et donc en particulier il faut

$$A + B + (A - B)x = 1$$

pour tout $x \in \mathbb{R} \setminus \{\pm 1\}$ (et par continuité : pour tout $x \in \mathbb{R}$ sans restriction.) Ceci revient à demander

$$\begin{cases} A + B = 1 \\ A - B = 0 \end{cases}$$

(on identifie les coefficients des puissances de x qui se correspondent). Ce système de deux équations à deux inconnues possède une et une seule solution, $A = B = \frac{1}{2}$. On a pu déterminer les constantes A et B qui rendent la décomposition possible. En conclusion, on trouve

$$\int \frac{1}{1-x^2} dx = -\frac{1}{2} \ln|1-x| + \frac{1}{2} \ln|1+x| + C = \ln \sqrt{\frac{|1+x|}{|1-x|}} + C$$

pour $x \in \mathbb{R} \setminus \{\pm 1\}$. Bien sûr, il est prudent de vérifier en re-dérivant !

Chapitre XIII

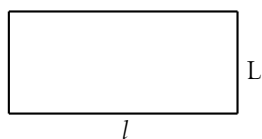
Fonction réelles d'une variable réelle : intégrales définies

Comme nous l'avons vu au chapitre précédent, l'intégrale indéfinie d'une fonction est elle-même une fonction, connue à une constante additive près.

La notion d'« intégrale *définie* d'une fonction entre deux bornes », objet de ce chapitre, est quant à elle un nombre qui s'interprète comme l'aire (dite « algébrique ») d'une certaine surface. Le « théorème fondamental du calcul différentiel et intégral » montre que les deux concepts d'intégrale sont étroitement liés.

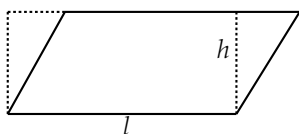
1 Introduction

L'expression de l'aire d'un rectangle plein



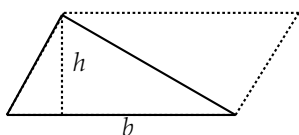
$$\text{aire} = L \times l$$

et des propriétés d'additivité de l'aire et d'invariance de l'aire par certaines transformations livrent des formules pour l'aire d'un parallélogramme plein



$$\text{aire} = l \times h$$

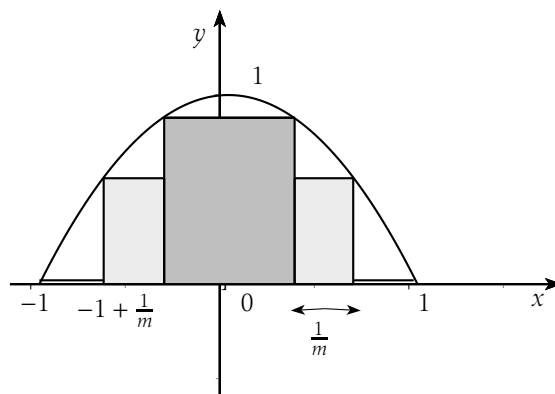
et d'un triangle plein



$$\text{aire} = \frac{b \times h}{2}$$

Pour un disque de rayon r , l'expression πr^2 donnant l'aire s'établit moins facilement. Une méthode consiste à approcher le disque par une suite de polygones réguliers inscrits et à prendre la limite des aires de ces polygones.

Comment faire alors pour l'aire sous l'arc de parabole d'équation $y = 1 - x^2$. Une situation est représentée ci-dessous, en subdivisant l'intervalle de définition $[-1, 1]$, nous obtenons une figure telle que la suivante :



Pour chaque valeur de m donnée, nous pouvons calculer l'aire de la réunion des rectangles dessinés dans cette figure et, prenant la limite lorsque $m \rightarrow \infty$, nous pouvons espérer obtenir une valeur pour l'aire. Cette valeur sera notée $\int_{-1}^1 1 - x^2 dx$. Cet espoir est effectivement fondé dans ce cas.

Plusieurs questions peuvent se poser :

- * Comment faut-il choisir la hauteur du rectangle sur chaque sous-intervalle ? Dans la figure, c'est la valeur minimale prise par la fonction sur le sous-intervalle en question qui a été choisie.
- * Comment faut-il choisir la subdivision ? Dans la figure, la subdivision choisie était régulière : la même largeur pour chaque sous-intervalle.

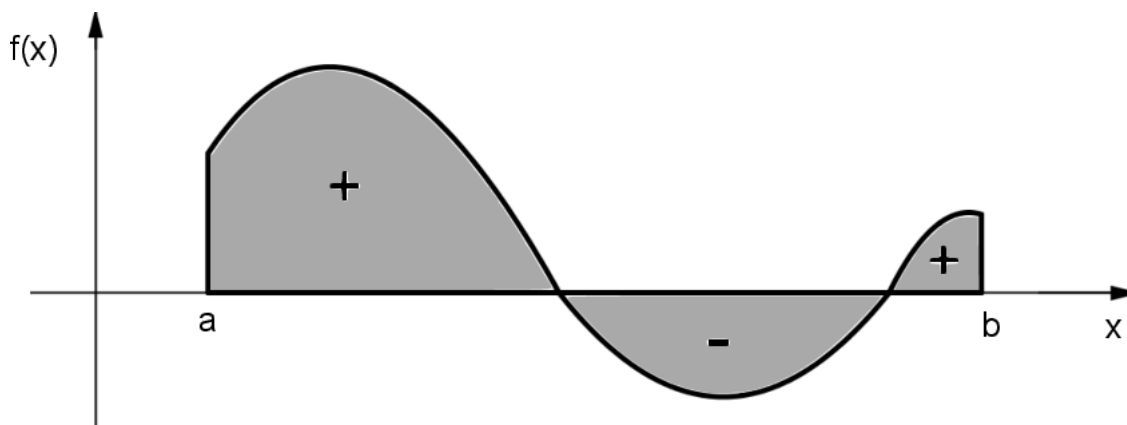
La réponse à ces questions sera « peu importe », pourvu que la fonction soit suffisamment régulière, par exemple continue.

2 Définition et propriétés

Soit a, b deux nombres réels avec $a < b$ et $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue. Nous voulons, intuitivement, définir la quantité

$$\int_a^b f(x) dx \quad (2.1)$$

pour mesurer l'*aire algébrique* de la surface comprise entre l'axe des abscisses, le graphe de f et les deux droites d'équations respectives $x = a$ et $x = b$. Le qualificatif « algébrique » signifie que la partie de cette surface située au-dessus de l'axe des abscisses contribue positivement, celle située en-dessous de l'axe des abscisses contribue négativement :

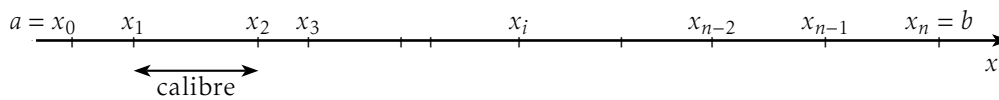


Comme nous n'avons pas jusqu'à présent de définition rigoureuse de cette aire (algébrique), nous élaborons d'abord une définition précise de l'intégrale définie.

Une *subdivision* σ d'ordre n de l'intervalle $[a, b]$ est un choix de $n + 1$ points x_i , avec n naturel positif, $i = 0, 1, \dots, n$ et

$$a = x_0 < x_1 < x_2 < \dots < x_n = b.$$

Le *calibre* de la subdivision σ est $\max_{i=1,2,\dots,n} |x_i - x_{i-1}|$. Ceci représente la taille du plus grand des sous-intervalles de la forme $[x_{i-1}, x_i]$.



Considérons les *sommes de Darboux inférieures* $s(\sigma)$ et *supérieures* $S(\sigma)$ associées à une telle subdivision :

$$s(\sigma) = \sum_{i=1}^n m_i(x_i - x_{i-1}) \quad S(\sigma) = \sum_{i=1}^n M_i(x_i - x_{i-1}) \quad (2.2)$$

sommes de
Darboux
inférieures

supérieures

où $m_i = \min\{f(x) \text{ t.q. } x \in [x_{i-1}, x_i]\}$ et $M_i = \max\{f(x) \text{ t.q. } x \in [x_{i-1}, x_i]\}$.

Pour chaque subdivision, $s(\sigma)$ et $S(\sigma)$ peuvent donc être définies. Si on considère une *suite* de subdivisions $\sigma_1, \sigma_2, \dots$, nous dirons que f est *intégrable* (au sens de Riemann, sur $[a, b]$) si $s(\sigma_i)$ et $S(\sigma_i)$ convergent toutes les deux vers le même nombre, c'est-à-dire s'il existe un réel L tel que

intégrable

$$\begin{aligned} \forall \epsilon > 0, \exists M \in \mathbb{N} \text{ t.q. } i \geq M \Rightarrow |s(\sigma_i) - L| < \epsilon \\ \text{et } \forall \epsilon > 0, \exists M \in \mathbb{N} \text{ t.q. } i \geq M \Rightarrow |S(\sigma_i) - L| < \epsilon \end{aligned}$$

et ce, quelle que soit la suite de subdivisions choisie dont le calibre tend vers 0.

Ce nombre L est l'*intégrale définie* de la fonction f de la borne initiale a à la borne finale b , il est noté

intégrale définie

$$L = \int_a^b f \quad \text{ou} \quad L = \int_a^b f(x) dx \quad (2.3)$$

(dans cette notation, le x n'a qu'une valeur décorative ou informative, dans le cas où un doute pourrait exister sur la variable de la fonction.)

Définition XIII.1. Lorsque $\int_a^b f(t) dt$ existe, on note

$$\int_b^a f(t) dt = - \int_a^b f(t) dt.$$

En d'autres termes : renverser les bornes change le signe. Cela est une convention, puisque seule la notation $\int_a^b$ avec $a < b$ avait été définie précédemment.

Résultat XIII.2. Si $f : [a, b] \rightarrow \mathbb{R}$ est continue, alors $\int_a^b f$ existe.

Résultat XIII.3. Pour $c \in \mathbb{R}$, si deux des trois intégrales suivantes existent, la troisième existe aussi et

$$\int_a^b f(t) dt = \int_a^c f(t) dt + \int_c^b f(t) dt.$$

Ceci est l'additivité par rapport au domaine d'intégration.

additivité par
rapport au
domaine
d'intégration

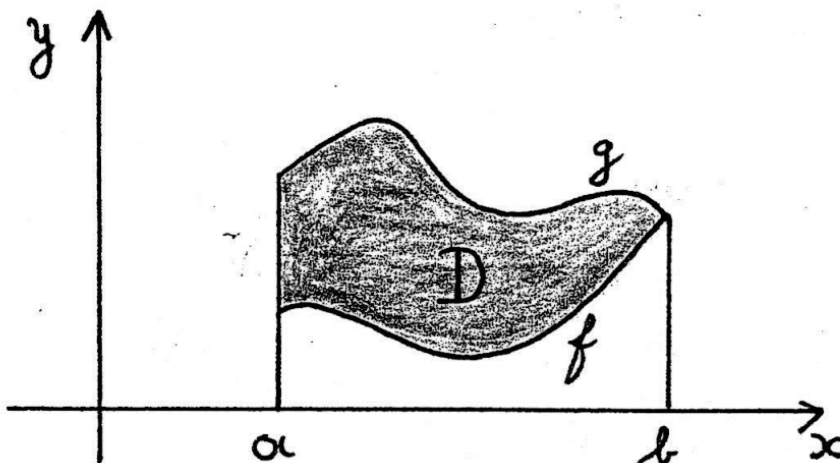
Résultat XIII.4. Considérons le sous-ensemble D du plan défini par

$$D = \{(x, y) \in \mathbb{R}^2 \text{ t.q. } x \in [a, b] \text{ et } 0 \leq f(x) \leq y \leq g(x)\}$$

où $a < b$ et f, g sont deux fonctions de $[a, b]$ vers $\mathbb{R}$ vérifiant $f(x) \leq g(x)$ pour tout $x \in [a, b]$. Si les deux intégrales définies suivantes existent, l'aire de D est

$$\int_a^b g(x) dx - \int_a^b f(x) dx.$$

Illustration du domaine D :



Une autre propriété, cette fois une propriété de linéarité :

Résultat XIII.5. Si f et g sont deux fonctions intégrables sur l'intervalle $[a, b]$ et si $\alpha, \beta \in \mathbb{R}$, alors

$$\int_a^b (\alpha f + \beta g)(x) \, dx = \alpha \int_a^b f(x) \, dx + \beta \int_a^b g(x) \, dx.$$

Résultat XIII.6. Soit f une fonction réelle sur $[a, b]$ vérifiant avec $f(x) \geq 0$ pour tout x de $[a, b]$. Alors

$$\int_a^b f(x) \, dx \geq 0.$$

De plus, si f est continue, nous avons l'équivalence :

$$\int_a^b f(x) \, dx = 0 \iff \forall x \in [a, b] : f(x) = 0.$$

Corollaire. En particulier, si $f(x) \leq g(x)$ pour tout $x \in [a, b]$, alors $\int_a^b f \leq \int_a^b g$.

3 Théorème fondamental

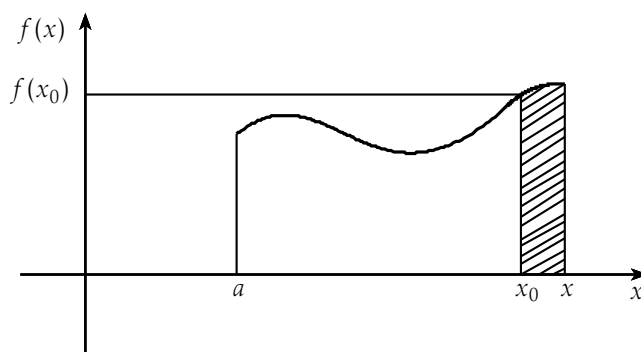
Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue ; alors pour tout $x \in [a, b]$, f est évidemment encore continue sur $[a, x]$, de sorte que nous pouvons définir

$$F(x) = \int_a^x f(t) \, dt$$

La dérivée de cette fonction F au point x_0 est donnée par :

$$F'(x_0) = \lim_{x \rightarrow x_0} \frac{\int_a^x f(t) \, dt - \int_a^{x_0} f(t) \, dt}{x - x_0} = \lim_{x \rightarrow x_0} \frac{\int_{x_0}^x f(t) \, dt}{x - x_0}$$

et s'interprète graphiquement :



le numérateur est l'aire de la surface hachurée et le dénominateur est la longueur de la « base » de cette surface. Il semble donc que $F'(x_0)$ soit le quotient, c'est-à-dire la hauteur du rectangle : approximativement $f(x_0)$. Autrement dit : F est une primitive de f . Ceci est l'objet du *théorème fondamental du calcul différentiel et intégral* :

Résultat XIII.7 (Théorème fondamental du calcul différentiel et intégral). *Soit f une fonction continue sur l'intervalle $]c, d[$, et soit a un point de $]c, d[$. Alors*

$$F :]c, d[\rightarrow \mathbb{R} : x \mapsto \int_a^x f(t) \, dt$$

est l'unique primitive de f qui s'annule en $x = a$.

Puisque deux primitives d'une même fonction sont égales à une constante près, nous obtenons que si G une primitive de f , alors $G(x) - G(a)$ est une primitive de f s'annulant en a , c'est-à-dire

$$\int_a^x f(t) \, dt = G(x) - G(a)$$

ce qui fournit un moyen de calculer le membre de gauche à condition d'avoir à disposition une primitive de f .

4 Intégration définie et indéfinie

Le calcul d'une intégrale définie s'effectue souvent par application du Théorème fondamental. En pratique, les formules de recherche de primitive (intégration par partie, par substitution) sont utilisées, et nous signalons que plusieurs de ces formules admettent des variantes pour les intégrales définies. En particulier, nous déduisons une formule pour le changement de variable en posant $t = g(x)$:

$$\int_a^b f(g(x)) \cdot g'(x) \, dx = \int_{t_1}^{t_2} f(t) \, dt$$

où $t_1 = g(a)$ et $t_2 = g(b)$.

Exemple. Soit à calculer

$$I = \int_0^2 \sin(3x+5) \, dx.$$

En posant $t = 3x + 5$, nous avons $dt = 3 \, dx$. Pour $x = 0$, il faut $t = 5$ et pour $x = 2$ il faut $t = 11$. Ainsi

$$\begin{aligned} I &= \frac{1}{3} \int_5^{11} \sin t \, dt \\ &= \frac{1}{3} [-\cos t]_5^{11} = \frac{1}{3} (-\cos 11 + \cos 5). \end{aligned}$$

5 Intégrales généralisées

Dans la définition de l'intégrale définie de Riemann

$$\int_a^b f(x) dx,$$

nous avons supposé $a, b \in \mathbb{R}$ et f fonction continue sur $[a, b]$. Les applications requièrent (au moins) deux généralisations de ce concept : lorsque f n'est pas bornée (domaine ouvert), et lorsque le domaine n'est pas borné (c'est-à-dire a ou b est $\pm\infty$).

5.1 Cas d'une fonction non bornée

Définition XIII.8. Nous dirons que l'intégrale d'une fonction f continue sur $[a, b[$ converge si la limite suivante existe (dans $\mathbb{R}$) :

$$\int_a^b f(x) dx = \lim_{u \nearrow b} \int_a^u f(x) dx.$$

Similairement pour une fonction continue sur $]a, b]$.

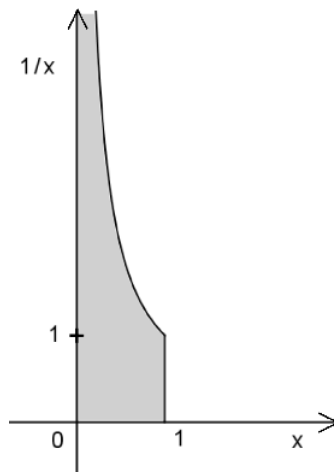
Exemple. Considérons la fonction

$$f : [0, 1] \rightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{1}{x} & \text{pour } x > 0, \\ 1 & \text{pour } x = 0. \end{cases}$$

Alors

$$\begin{aligned} \lim_{u \nearrow 0} \int_u^1 \frac{1}{x} dx \\ = \lim_{u \nearrow 0} \ln(1) - \ln(u) = +\infty \end{aligned}$$

de sorte que l'intégrale ne converge pas, plus précisément elle diverge vers $+\infty$. Géométriquement, la surface ombrée sur le graphique



n'est pas finie.

Exemple. Soit a un nombre réel, avec $a > 0$. Considérons la fonction

$$f_a :]0, 1] \rightarrow \mathbb{R} : x \mapsto \frac{1}{x^a}$$

(Le cas $a = 1$ a été traité ci-dessus.) Alors

$$\lim_{\substack{u \rightarrow 0 \\ >}} \int_u^1 \frac{1}{x^a} dx = \frac{1}{1-a} \lim_{\substack{u \rightarrow 0 \\ >}} (1 - u^{1-a})$$

de sorte que, si $a < 1$, alors l'intégrale vaut $\frac{1}{1-a}$, tandis que si $a \geq 1$, l'intégrale ne converge pas (le cas $a = 1$ ayant été fait ci-dessus.)

Définition XIII.9. Lorsque le point c près duquel f n'est pas définie (ou pas continue) est intérieur à l'intervalle d'intégration, nous posons (en notant c le point où f n'est pas définie) :

$$\int_a^b f(x) dx = \lim_{\substack{u \rightarrow c \\ <}} \int_a^u f(x) dx + \lim_{\substack{v \rightarrow c \\ >}} \int_v^b f(x) dx;$$

et l'intégrale ainsi définie converge à condition que les deux limites convergent dans $\mathbb{R}$.

Cette définition se généralise au cas d'une fonction continue sauf aux voisinages d'un nombre fini de points de l'intervalle d'intégration (il suffit de décomposer cet intervalle en sous-intervalles dont chacun ne contient qu'un seul point exceptionnel).

5.2 Cas d'un domaine non borné

Définition XIII.10. Lorsqu'on considère une fonction $f : [a, +\infty[\rightarrow \mathbb{R}$ continue, nous posons

$$\int_a^{+\infty} f(x) dx = \lim_{T \rightarrow +\infty} \int_a^T f(x) dx$$

si cette dernière limite existe.

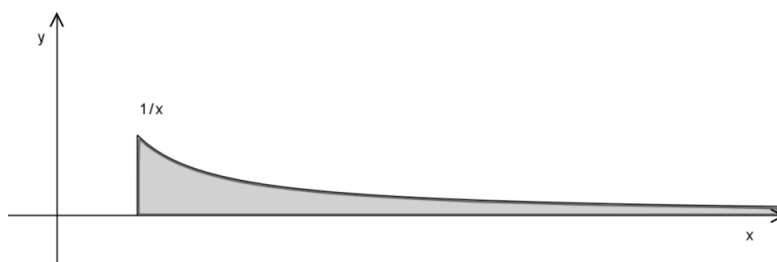
Exemple. Soit

$$f : [1, +\infty[\rightarrow \mathbb{R} : x \mapsto \frac{1}{x}.$$

Alors

$$\lim_{T \rightarrow +\infty} \int_1^T \frac{1}{x} dx = \lim_{T \rightarrow +\infty} \ln(x) - \ln(1) = +\infty.$$

de sorte que l'intégrale ne converge pas : elle diverge vers $+\infty$. Ainsi, la surface ombrée sur le graphique



est d'aire infinie.

Exercice. Soit a un nombre réel, avec $a > 0$. Considérons la fonction

$$f_a : [1, \infty[\rightarrow \mathbb{R} : x \mapsto \frac{1}{x^a}$$

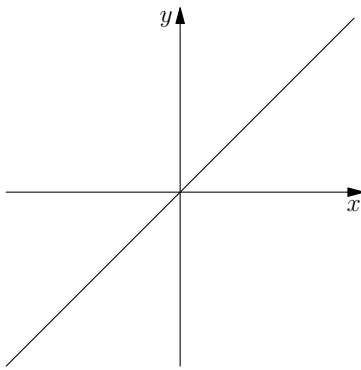
(Le cas $a = 1$ a été traité ci-dessus.) Prouvez que la fonction est intégrable sur $[1, \infty[$ si et seulement si $a > 1$.

Chapitre XIV

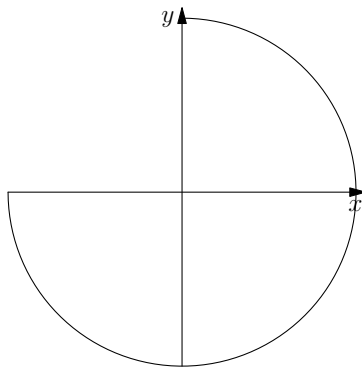
Fonctions d'une variable réelle à valeurs vectorielles

Une fonction d'une variable réelle à valeurs vectorielles est une fonction de la forme $f : A \subset \mathbb{R} \rightarrow \mathbb{R}^n$. Elle représentera, dans cette section, une courbe paramétrée dans $\mathbb{R}^n$.

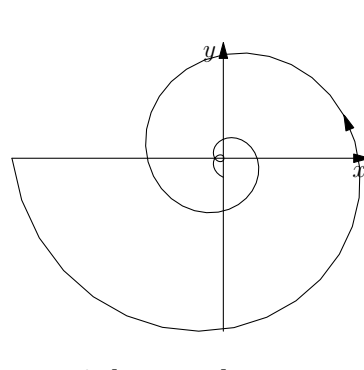
Exemple. Dans les exemples ci-dessous, c'est l'ensemble image $\text{Im } f$ qui est représenté. (Le graphe de f , quant à lui, n'a pas d'interprétation utile à ce stade.)



$$f : [-1, 1] \rightarrow \mathbb{R} : t \mapsto (t, t)$$



$$f : [-\pi, \pi/2] \rightarrow \mathbb{R} : t \mapsto (\cos t, \sin t)$$



$$f : [-5\pi, 3\pi/2] \rightarrow \mathbb{R} \\ t \mapsto (t^2 \cos t, t^2 \sin t)$$

La valeur de n indique dans quel espace la courbe vit. On peut imaginer une courbe paramétrée comme la description d'un point mobile au cours du temps. Si $n = 1$, le mobile se déplace le long d'une droite; si $n = 2$, le mobile se déplace dans un plan; si $n = 3$, le mobile se déplace dans l'espace, etc. À chaque valeur de t correspond un point. Lorsque t évolue, cela rend compte du déplacement du point.

Notons que l'ensemble $\text{Im } f$ ne donne pas une information suffisante pour pouvoir connaître le mouvement. Deux courbes peuvent provenir de deux paramétrisations différentes.

1 Régularité

Définition XIV.1. Soit $f : A \subset \mathbb{R} \rightarrow \mathbb{R}^n$ une application, $a \in A$, et $\mathbf{L} \in \mathbb{R}^n$. On dit que la limite de f en a est $\mathbf{L}$ si

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x \in A, |x - a| < \delta \Rightarrow \|f(x) - \mathbf{L}\| \leq \epsilon.$$

On écrit dans ce cas

$$\lim_{x \rightarrow a} f(x) = \mathbf{L}.$$

Définition XIV.2. La fonction $f : A \subset \mathbb{R} \rightarrow \mathbb{R}^n$ est *continue* en $a \in A$ si

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

La fonction f est dite *continue* si elle est continue en chaque point.

continue

dérivable

Définition XIV.3. Une application $f : A \subset \mathbb{R} \rightarrow \mathbb{R}$ est *dérivable* en $a \in A$ si a est intérieur à A et si la limite

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

existe. Dans ce cas, cette limite est notée $f'(a)$ est appelée la *dérivée* de f en a .

En pratique, nous avons le résultat suivant :

Résultat XIV.4. Une application $f : A \subset \mathbb{R} \rightarrow \mathbb{R}^n$ est continue (resp. dérivable) en $a \in A$ si et seulement si pour tout $i = 1 \dots n$, $f_i : A \rightarrow \mathbb{R}$ est continue (resp. dérivable) en a . De plus, si elle existe, la dérivée est donnée par

$$f'(a) = (f'_1(a), \dots, f'_n(a)).$$

2 Courbes

courbe
paramétrée

Définition XIV.5. Une *courbe paramétrée* de $\mathbb{R}^n$ est une application continue

$$\gamma : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$$

définie sur un intervalle I .

2.1 Re-paramétrisation

Intuitivement, on peut parcourir une courbe donnée de plusieurs manières : dans un sens ou dans l'autre, ou à différentes vitesses. Ceci justifie une définition :

reparamétrisa-
tion

Définition XIV.6. Si $\gamma : I \rightarrow \mathbb{R}^n$ est une courbe paramétrée, on dit qu'une courbe paramétrée $\eta : J \rightarrow \mathbb{R}^n$ est une *reparamétrisation* de γ si il existe une bijection continue $\alpha : I \rightarrow J$ telle que $\eta \circ \alpha = \gamma$.

Certaines notions, telle que la longueur d'une courbe, devraient idéalement ne pas dépendre de la paramétrisation choisie. D'autres notions en dépendent de manière évidente, comme par exemple la notion de vitesse.

2.2 Vitesse et vecteur tangent

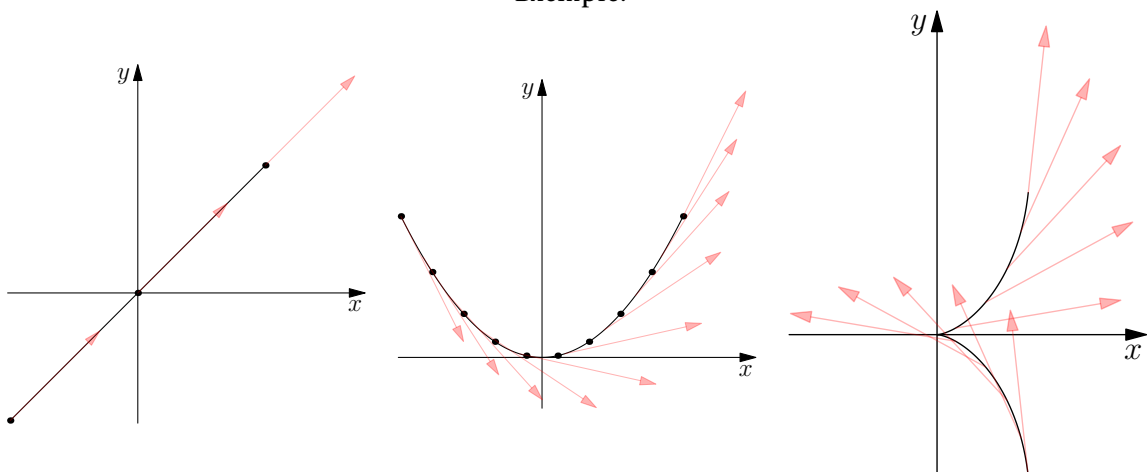
vecteur vitesse

vecteur tangent

vitesse
numériquevecteur tangent
unitaire

Définition XIV.7. Le *vecteur vitesse* ou *vecteur tangent* d'une courbe paramétrée γ à l'instant t est donné par le vecteur $\gamma'(t)$. Sa norme est parfois appelée la *vitesse numérique* à l'instant t . Le vecteur $\frac{\gamma'(t)}{\|\gamma'(t)\|}$ est appelé *vecteur tangent unitaire* à l'instant t et est noté $T(t)$.

Exemple.



2.3 Accélération et vecteur normal

Définition XIV.8. Le *vecteur accélération* de γ à l'instant t est défini par $\gamma''(t)$. Sa norme est parfois appelée l'*accélération numérique* à l'instant t .

vecteur
accélération

Attention à la définition similaire mais différente suivante, où on dérive maintenant le vecteur tangent unitaire :

accélération
numérique

Définition XIV.9. Le *vecteur normal unitaire*, s'il existe, est donné par $N(t) = \frac{T'(t)}{\|T'(t)\|}$.

vecteur normal
unitaire

Remarque (*). Le vecteur normal unitaire est effectivement normal (orthogonal) au vecteur tangent.

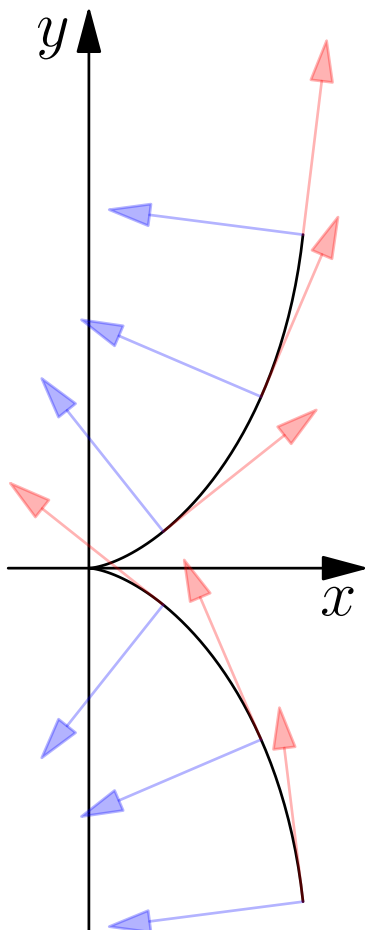
Démonstration. Puisque $T(t)$ est de norme 1, on peut dériver l'égalité

$$T(t) \cdot T(t) = 1$$

pour obtenir

$$2 T'(t) \cdot T(t) = 0$$

ce qui, en divisant par 2 et par la norme de T' , donne bien $N(t) \cdot T(t) = 0$ comme attendu. □



2.4 Vecteur binormal

Si $n = 3$, c'est-à-dire pour une courbe dans l'espace, on définit le vecteur binormal $B = T \times N$. Son utilité est la suivante :

Résultat XIV.10. La courbe γ est plane si et seulement si $B' = 0$.

Une courbe plane est une courbe dont l'image est entièrement contenue dans un plan.

2.5 Longueur d'une courbe

longueur

Si $\gamma : I \rightarrow \mathbb{R}^n$ est dérivable et si $a, b \in I$, la *longueur* de $\gamma([a, b])$ est donnée par l'intégrale :

$$\int_a^b \|\gamma'(t)\| dt$$

Idée de preuve. Subdivisons le domaine $[a, b]$ en n parties égales, et approchons la longueur par

$$\sum_{i=1}^n \|\gamma^{(i/n)} - \gamma^{((i-1)/n)}\| = \sum_{i=1}^n \frac{\|\gamma^{(i/n)} - \gamma^{((i-1)/n)}\|}{1/n} \frac{1}{n}$$

On remarque que lorsque n est grand, la fraction $\frac{\|\gamma^{(i/n)} - \gamma^{((i-1)/n)}\|}{1/n}$ s'approche de $\gamma'(i/n)$, de sorte que la somme ci-dessus correspond à l'intégrale de Riemann. $\square$

Chapitre XV

Matrices et systèmes linéaires

Les matrices sont un outil utilisé dans de nombreux domaines, comme par exemple :

- * la résolution de systèmes d'équations
- * la modélisation de certaines transformations simples (nommées « applications linéaires ») dont les rotations et les symétries.
- * de nombreux modèles de probabilités (buzzword : Google)
- * la résolution de systèmes d'équations différentielles (qui eux-même modélisent de nombreux problèmes issus de toutes les sciences)
- * Traitement d'images.

Nous nous contentons ici des définitions les plus basiques, mais leur maîtrise est indispensable.

Définition XV.1. Une *matrice* est un tableau rectangulaire de nombres.

matrice

- * Si m désigne son nombre de lignes, et n désigne son nombre de colonnes, la *taille* de la matrice est notée $m \times n$.
- * Si $n = 1$, la matrice est appelée *matrice-colonne*;
- * si par contre $m = 1$, la matrice est appelée *matrice-ligne*;
- * si enfin $m = n$, la matrice est appelée une *matrice carrée*.

taille

matrice-colonne

matrice-ligne

matrice carrée

Si M est une matrice, on note M_{ij} l'élément se trouvant à l'intersection de la i^e ligne et de la j^e colonne. On note enfin $\text{Mat}(m, n)$ l'ensemble des matrices de taille $m \times n$.

Remarque (*). Bien sûr, une matrice est déterminée exactement par ses éléments : si M et N sont des matrices de même taille vérifiant $M_{ij} = N_{ij}$ pour tout i et pour tout j , on écrit $M = N$.

Exemple. Voici un exemple « générique » de matrice.

$$\begin{pmatrix} M_{11} & \dots & M_{1n} \\ \vdots & \ddots & \vdots \\ M_{m1} & \dots & M_{mn} \end{pmatrix}$$

Exemple. Ceci sont des matrices :

$$\begin{pmatrix} 2 & 1 \\ \pi & -4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ e + \pi \end{pmatrix}, (-\pi)$$

$$\begin{pmatrix} 0 & 1 & 2 \\ 0 & -4 & 3 \\ 0 & 0 & 5 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 5 & 1 & 0 \\ 21 & 0 & 2 \end{pmatrix}, \dots$$

1 Opérations sur les matrices

1.1 Somme

Définition XV.2. Si M et N sont deux matrices de même taille, on définit leur somme $M + N$ par

$$(M + N)_{ij} = M_{ij} + N_{ij}.$$

La somme a encore la même taille que M et que N .

Exemple. Un exemple pour la somme :

$$\begin{pmatrix} 0 & 1 \\ \sqrt{2} & \pi \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 + \sqrt{2} & \pi \end{pmatrix}$$

Résultat XV.3. La matrice nulle est neutre dans l'addition matricielle.

Exemple.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

1.2 Produit

Définition XV.4. Si M est une matrice de taille $m \times n$ et N une matrice de taille $n \times p$, on définit leur produit MN par

$$(MN)_{ik} = \sum_{j=1}^n M_{ij}N_{jk} \quad \forall i = 1, \dots, m; j = 1, \dots, n.$$

Ce produit est parfois appelé le produit « ligne par colonne ».

Exemple. Un exemple pour le produit :

$$\begin{pmatrix} 0 & 1 \\ \sqrt{2} & \pi \end{pmatrix} \begin{pmatrix} 1 & 0 & \pi \\ 0 & 3 & \sqrt{2} \end{pmatrix} = \begin{pmatrix} 0 & 3 & \sqrt{2} \\ \sqrt{2} & 3\pi & 2\sqrt{2}\pi \end{pmatrix}$$

matrice nulle

Définition XV.5. Une matrice constituée uniquement de 0 est appelée *matrice nulle*. On la note parfois 0, ou encore $0_{m \times n}$ pour indiquer sa taille.

Définition XV.6. Si $\lambda \in \mathbb{R}$, et M est une matrice, le produit λM est une matrice de même taille dont les coefficients sont :

$$(\lambda M)_{ij} = \lambda M_{ij}$$

pour i, j des indices parcourant la matrice.

matrice identité

Définition XV.7. Une matrice carrée de taille $n \times n$ possédant des 0 partout, sauf sur la diagonale où il n'y a que des 1 est une *matrice identité*. On la note généralement I ou I_n .

Exemple. Les matrices identité 2×2 et 3×3 sont :

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Résultat XV.8. Si M est une matrice de taille $m \times n$, alors :

$$MI_n = M = I_m M$$

Exemple.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

1.3 Produit par un scalaire

Définition XV.9. Si $\lambda \in \mathbb{R}$, et M est une matrice, le produit λM est une matrice de même taille dont les coefficients sont :

$$(\lambda M)_{ij} = \lambda M_{ij}$$

pour i, j des indices parcourant la matrice.

Exemple.

$$5 \begin{pmatrix} 0 & 1 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 5 \\ -5 & 15 \end{pmatrix}$$

1.3.1 Déterminants

Nous allons maintenant définir une notion qui à n vecteurs de $\mathbb{R}^n$ associe l'aire (cas $n = 2$), le volume (cas $n = 3$), l'hypervolume (cas $n > 3$) orienté défini par ces vecteurs. Cette notion s'appelle le déterminant.

Remarquons que la donnée de n vecteurs de $\mathbb{R}^n$, chacun ayant n composantes, peut se résumer à la donnée d'une matrice carrée $n \times n$:

$$\begin{pmatrix} \boxed{v_{11}} & \boxed{v_{12}} & \cdots & \boxed{v_{1n}} \\ \vdots & \vdots & & \vdots \\ \boxed{v_{n1}} & \boxed{v_{n2}} & & \boxed{v_{nn}} \end{pmatrix}$$

La notion de déterminant sera vue comme étant applicable à une matrice plutôt qu'à n vecteurs. C'est la même chose.

Nous essayons de motiver le cas $n = 2$ (matrice 2×2), et nous donnerons simplement les définitions pour les autres valeurs de n .

Le cas $n = 2$ Pour des vecteurs $\mathbf{x}$ et $\mathbf{y}$ de $\mathbb{R}^2$, notons (temporairement) $A(\mathbf{x}, \mathbf{y})$ l'aire orientée du parallélogramme construit sur les vecteurs $\mathbf{x}$ et $\mathbf{y}$. Le mot « orienté » indique que l'aire pourra être positive ou négative selon la disposition relative de $\mathbf{x}$ et $\mathbf{y}$. Précisément, le signe sera positif si l'angle (orienté) est entre 0 et π , négatif si il est entre $-\pi$ et 0 (tout ceci s'entend à 2π près.) En particulier, l'ordre compte !

Exemple. L'aire engendrée par

- * $(1, 0)$ et $(0, 1)$ est 1 : positif, car l'angle entre ces vecteurs est $\frac{\pi}{2}$ (équivalent à $-\frac{3\pi}{2}$).
- * $(1, 0)$ et $(0, -1)$ est -1 : négative, car l'angle est $\frac{3\pi}{2}$ (équivalent à $-\frac{\pi}{2}$).
- * $(0, 1)$ et $(1, 0)$ est -1 : négative, pour la même raison.

Pour des vecteurs $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^2$, considérons l'aire engendrée par $\mathbf{u} + \mathbf{v}$ et $\mathbf{w}$. Dans ce cas, nous pouvons nous convaincre de la relation suivante (linéarité en le premier argument) :

$$A(\mathbf{u} + \mathbf{v}, \mathbf{w}) = A(\mathbf{u}, \mathbf{w}) + A(\mathbf{v}, \mathbf{w})$$

en admirant la figure ?? page ?? . Par ailleurs, il est clair que multiplier un des vecteurs par une constante multipliera l'aire du parallélogramme par cette constante :

$$A(k\mathbf{u}, \mathbf{v}) = kA(\mathbf{u}, \mathbf{v})$$

et enfin, puisque l'ordre compte, nous avons

$$A(\mathbf{u}, \mathbf{v}) = -A(\mathbf{v}, \mathbf{u})$$

Ceci étant acquis, nous en déduisons que pour des vecteurs (a, b) et (c, d) , nous avons :

$$\begin{aligned} A((a, b), (c, d)) &= A((a, 0) + (0, b), (c, d)) \\ &= A((a, 0), (c, d)) + A((0, b), (c, d)) \\ &= A((a, 0), (c, 0) + (0, d)) + A((0, b), (c, 0) + (0, d)) \\ &= A((a, 0), (c, 0)) + A((a, 0), (0, d)) + A((0, b), (c, 0)) + A((0, b), (0, d)) \\ &= 0 + ad - bc + 0 \end{aligned}$$

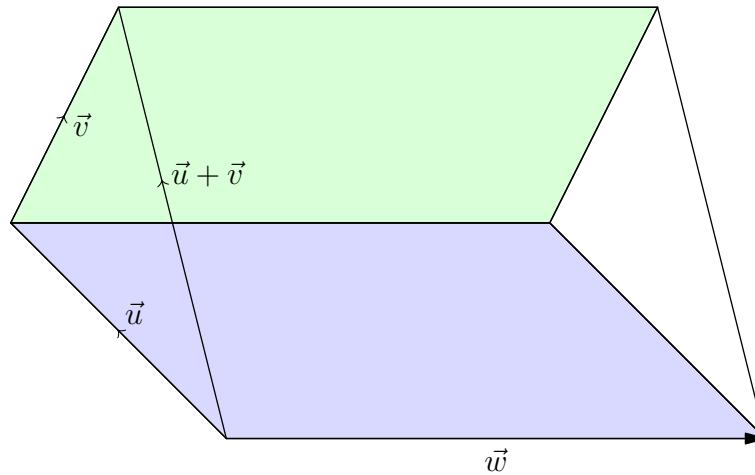


FIGURE XV1 – L'aire vérifie une propriété de linéarité : l'aire du parallélogramme construit sur $\mathbf{u} + \mathbf{v}$ et $\mathbf{w}$ est la somme des aires colorées puisque les deux triangles ont la même aire.

Résultat XV.10. L'aire orientée engendrée par les vecteurs (a, b) et (c, d) est égale à $ad - bc$.

On dira que $ad - bc$ est le *déterminant* de la matrice $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$. De manière intéressante, c'est aussi le déterminant de la matrice :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Le cas général Étant donné une matrice réelle M , on définit $\det M$, un nombre réel, son *déterminant* :

Définition XV.11. * Le déterminant d'une matrice 1×1 est égal au seul nombre de la matrice ; $\det M = M_{11}$.

* Le déterminant d'une matrice quelconque est obtenu de la manière suivante :

1. On choisit une rangée : ligne ou colonne, peu importe laquelle.
2. Pour chaque élément M_{ij} de la rangée, on calcule le déterminant de la matrice obtenue en enlevant la i^e ligne et la j^e colonne de A . (Cette nouvelle matrice est plus petite !) On note ce nombre temporairement D_{ij} .
3. On somme tous les produits $(-1)^{i+j} M_{ij} D_{ij}$ en parcourant la rangée choisie :

$$\det M = \sum_{i=1}^n (-1)^{i+j} M_{ij} D_{ij} = \sum_{j=1}^m (-1)^{i+j} M_{ij} D_{ij}$$

Et cela ne dépend pas de la rangée choisie !

Définition XV.12. Le déterminant d'une matrice $\begin{pmatrix} M_{11} & \dots & M_{1n} \\ \vdots & \ddots & \vdots \\ M_{k1} & \dots & M_{kn} \end{pmatrix}$ se note $\begin{vmatrix} M_{11} & \dots & M_{1n} \\ \vdots & \ddots & \vdots \\ M_{k1} & \dots & M_{kn} \end{vmatrix}$.

Le signe attribué alterne d'un élément au suivant, il est donc facile à se rappeler par le schéma suivant :

$$\begin{pmatrix} + & - & + & \dots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{pmatrix}$$

Exemple. Considérons la matrice, et développons selon la première colonne

$$\begin{vmatrix} 1 & -3 \\ 2 & 4 \end{vmatrix} = (-1)^{1+1} 1 \begin{vmatrix} 1 & -3 \\ 2 & 4 \end{vmatrix} + (-1)^{2+1} 2 \begin{vmatrix} 1 & -3 \\ 1 & 4 \end{vmatrix} = 1 \cdot 4 - 2 \cdot (-3) = 10$$

Remarque. Dans le cas des matrices 2×2 , on retient la formule une bonne fois pour toute :

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

« La diagonale descendante moins la diagonale montante ! »

Remarque. En général on va chercher à développer selon les rangées où il y a autant de 0 que possible afin d'éviter des calculs !

Exemple. Développons le déterminant 3×3 suivant selon la seconde colonne.

$$\begin{aligned} \begin{vmatrix} 1 & 3 & 5 \\ 2 & 0 & 1 \\ 0 & 1 & -1 \end{vmatrix} &= (-1)^{1+2} 3 \begin{vmatrix} 1 & 5 \\ 2 & 1 \end{vmatrix} + (-1)^{2+2} 0 \begin{vmatrix} 1 & 5 \\ 2 & 1 \end{vmatrix} + (-1)^{3+2} 1 \begin{vmatrix} 1 & 3 \\ 2 & 0 \end{vmatrix} \\ &= -3 \begin{vmatrix} 1 & 5 \\ 2 & 1 \end{vmatrix} + 0 - 1 \begin{vmatrix} 1 & 3 \\ 2 & 0 \end{vmatrix} \\ &= -3(-2 - 0) - (1 - 10) = 6 + 9 = 15 \end{aligned}$$

1.3.2 Transposition

Définition XV.13. La *transposée* d'une matrice M de taille $m \times n$ est la matrice tM (parfois notée M') de taille $n \times m$ vérifiant :

$$({}^tM)_{ij} = M_{ji}$$

Exemple. La transposée de ... est ...

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \Rightarrow \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \Rightarrow \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix}.$$

1.3.3 Inverse

Définition XV.14. Une matrice carrée A est *inversible* si il existe une matrice B de même taille telle que

$$AB = BA = I$$

La matrice B est appelée l'inverse de A , et on note $B = A^{-1}$.

Résultat XV.15. L'inverse de A est unique.

Démonstration. Si B et C sont deux inverses de A , alors

$$B = BI = BAC = IC = C$$

□

Exemple. La matrice identité I_3 est sa propre inverse

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

Remarque (*). Si D et E sont des matrices diagonales, leur produit est encore une matrice diagonale.

$$\begin{pmatrix} D_{11} & 0 & \dots & 0 \\ 0 & D_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & D_{nn} \end{pmatrix} \begin{pmatrix} E_{11} & 0 & \dots & 0 \\ 0 & E_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & E_{nn} \end{pmatrix} = \begin{pmatrix} D_{11}E_{11} & 0 & \dots & 0 \\ 0 & D_{22}E_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & D_{nn}E_{nn} \end{pmatrix}$$

Résultat XV.16. Si $D_{11}, \dots, D_{nn}$ sont non-nuls, l'inverse de

$$D = \begin{pmatrix} D_{11} & 0 & \dots & 0 \\ 0 & D_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & D_{nn} \end{pmatrix} \text{ est } \begin{pmatrix} D_{11}^{-1} & 0 & \dots & 0 \\ 0 & D_{22}^{-1} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & D_{nn}^{-1} \end{pmatrix}$$

Si l'un des coefficients diagonaux de D est nul, alors D n'est pas inversible.

Exemple. Soit A et B définies par

$$A = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad B = \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix}$$

Alors

$$AB = \begin{pmatrix} \cos(\theta + \beta) & -\sin(\theta + \beta) \\ \sin(\theta + \beta) & \cos(\theta + \beta) \end{pmatrix}.$$

En particulier l'inverse de A est

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

1.3.4 Trace

Définition XV.17. La *trace* d'une matrice carrée M est la somme de ses éléments diagonaux :

$$\text{Tr}(M) = M_{11} + \dots + M_{nn}$$

où $n \times n$ est la taille de la matrice.

Exemple.

$$\text{Tr} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = 5$$

1.3.5 Propriétés des opérations sur les matrices

Résultat XV.18. Si $A, D \in \text{Mat}(m, n)$, $B, C \in \text{Mat}(n, p)$ et $\lambda \in \mathbb{R}$, alors :

$$A(B + C) = AB + AC \quad (A + D)B = AB + DB \quad \lambda(AB) = (\lambda A)B = A(\lambda B)$$

Preuve de la première égalité. En position ij , le membre de gauche vaut

$$\begin{aligned} \sum_{k=1}^n A_{ik}(B + C)_{kj} &= \sum_{k=1}^n A_{ik}(B_{kj} + C_{kj}) \\ &= \sum_{k=1}^n (A_{ik}B_{kj} + A_{ik}C_{kj}) = (AB)_{ij} + (AC)_{ij} \end{aligned}$$

ce qui est égal au membre de droite. □

Résultat XV.19. Si $A \in \text{Mat}(m, n)$ et $B \in \text{Mat}(n, p)$, alors :

$${}^t(AB) = {}^tB {}^tA$$

Démonstration. On écrit les définitions :

$$({}^t(AB))_{ij} = (AB)_{ji} = \sum_{k=1}^n A_{jk} B_{ki} =$$

$$\sum_{k=1}^n ({}^tA)_{kj} ({}^tB)_{ik} = \sum_{k=1}^n ({}^tB)_{ik} ({}^tA)_{kj} = ({}^tB {}^tA)_{ij}$$

□

Résultat XV.20. Si A, B sont des matrices carrées de même taille, si $\lambda \in \mathbb{R}$, alors

$$\begin{aligned} \text{Tr}(A + B) &= \text{Tr } A + \text{Tr } B & \text{Tr}(\lambda A) &= \lambda \text{Tr}(A) \\ \text{Tr}({}^t(A)) &= \text{Tr}(A) & \text{Tr}(AB) &= \text{Tr}(BA) \end{aligned}$$

Exemple. La trace du produit n'est pas égale au produit des traces :

$$\text{Tr}\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right) = \text{Tr}\left(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right) = 0$$

Résultat XV.21. Si A, B sont des matrices inversibles de même taille, alors

$$(AB)^{-1} = B^{-1} A^{-1} \quad (A^{-1})^{-1} = A \quad ({}^tA)^{-1} = ({}^tA^{-1})$$

Démonstration. Pour la première égalité, calculons :

$$(AB)(B^{-1} A^{-1}) = ABB^{-1} A^{-1} = AIA^{-1} = AA^{-1} = I$$

De même pour $(B^{-1} A^{-1})(AB) = I$. Les autres inégalités sont laissées en exercice. □

Le déterminant d'une matrice et de sa transposée sont égaux :

Résultat XV.22.

$$\det({}^tM) = \det(M)$$

Résultat XV.23. Le déterminant d'un produit de deux matrices carrées de même taille est le produit des déterminants :

$$\det(AB) = \det(A) \det(B).$$

Remarque (*). Le déterminant d'une somme n'est pas la somme des déterminants !

1.3.6 Propriétés sur les déterminants

Résultat XV.24. Le déterminant d'une matrice triangulaire est le produit de ses coefficients diagonaux

Exemple.

$$\begin{vmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{vmatrix} = 18$$

Si M est une matrice, son déterminant ne varie pas lorsque

- * nous ajoutons à l'une des lignes, une combinaison linéaire des autres
- * nous ajoutons à l'une des colonnes, une combinaison linéaire des autres

son déterminant *change de signe* lorsque :

- * nous échangeons des lignes entre elles
- * nous échangeons des colonnes entre elles

combinaison
linéaire

Définition XV.25. Une *combinaison linéaire* de $v_1, \dots, v_k$ est

$$\lambda_1 v_1 + \dots + \lambda_k v_k$$

(Ici les $v_1, \dots, v_k$ peuvent représenter des matrices colonnes, des matrices lignes, ou même des vecteurs.)

Exemple. Ici nous remplaçons la première colonne C_1 par $C_1 - 2C_2$.

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 6 & 3 & 3 \end{vmatrix} = \begin{vmatrix} -3 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 3 & 3 \end{vmatrix} = -3 \begin{vmatrix} 1 & 0 \\ 3 & 3 \end{vmatrix} = -9$$

cofacteur

comatrice

Définition XV.26. Si A est une matrice carrée. Notons à nouveau D_{ij} le déterminant de la matrice obtenue en enlevant la i^e ligne et la j^e colonne de A . Notons $\tilde{A}_{ij} := (-1)^{i+j} D_{ij}$. Ce nombre est appelé le *cofacteur* d'indices i et j . La matrice $\tilde{A}$ est la matrice des cofacteurs, également appelée *comatrice* de A .

Résultat XV.27. Pour toute matrice carrée, nous avons :

$$A^t(\tilde{A}) = \det A I.$$

En particulier, ceci montre que si $\det A \neq 0$, l'inverse de A est donnée par $\frac{A^t(\tilde{A})}{\det A}$.

Nous citons encore le résultat suivant sans démonstration :

Résultat XV.28. Si A et B sont des matrices $n \times n$, et $k \in \mathbb{R}$, alors

- * $\det(AB) = \det A \det B$;
- * $\det(kA) = k^n \det A$;
- * $\det(A') = k \det A$, où A' est une matrice obtenue à partir de A en multipliant une ligne ou une colonne par le réel k .

Remarque (*). Avec peu d'efforts, il est possible d'utiliser le premier résultat pour démontrer les suivants.

2 Systèmes linéaires

système
d'équations

solution

linéaire

Définition XV.29. * Une équation est une égalité faisant intervenir des quantités inconnues.

- * Un *système d'équations* est une liste d'équations faisant intervenir des quantités inconnues (par exemple nommées $x_1, \dots, x_n$).
- * Une *solution* d'un système d'équations est un élément $(x_1, x_2, \dots, x_n)$ de $\mathbb{R}^n$ tel que chaque égalité est vérifiée.
- * Le système est *linéaire* si chaque équation est polynomiale de degré 1 en chacune des inconnues.

Exemple. Le système suivant est un système linéaire

$$\begin{cases} x + y = 3 \\ x - y = 1 \end{cases}$$

Il a pour unique solution $(x, y) = (2, 1)$.

Vérification. Si (x, y) est une solution, alors $2x = (x + y) + (x - y) = 3 + 1 = 4$, donc il faut $x = 2$.

Comme $x + y = 3$ il faut donc $y = 1$. D'où il **faudrait** $(x, y) = (2, 1)$ (= condition nécessaire).

Inversement, en remplaçant $x = 2$ et $y = 1$ dans les équations, les égalités sont vérifiées. La condition nécessaire est donc suffisante. $\square$

Les systèmes suivants ne sont pas linéaires :

$$\begin{cases} x^2 + y = 1 \\ x - y^2 = 2 \end{cases} \quad \text{et} \quad \begin{cases} \cos(x) + y = 1 \\ \sin(x) + y^2 = \pi \end{cases}$$

Ces systèmes sont beaucoup plus difficiles à résoudre que le système linéaire vu précédemment.

Un système linéaire de k équations avec inconnues $(x_1, \dots, x_n)$ peut s'écrire de manière générale sous la forme :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{k1}x_1 + a_{k2}x_2 + \dots + a_{kn}x_n = b_k \end{cases}$$

où les a_{ij} et b_i sont des coefficients constants (c'est-à-dire ne dépendent pas des inconnues).

En d'autres termes, on peut écrire le système linéaire ci-dessus sous la forme $A\mathbf{x} = \mathbf{b}$ où $A \in \mathbb{R}^{k \times n}$ est une matrice, $\mathbf{b} \in \mathbb{R}^{k \times 1}$ est un vecteur-colonne donné, et $\mathbf{x} \in \mathbb{R}^{1 \times n}$ le vecteur-colonne des inconnues :

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kn} \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_k \end{pmatrix}$$

2.1 Méthode du pivot de Gauß

Écrivons un système linéaire sous forme d'une *matrice augmentée*, où chaque ligne correspond à une équation du système :

matrice
augmentée

$$\left[\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{k1} & \dots & a_{kn} & b_k \end{array} \right]$$

Les opérations suivantes ne changent pas les solutions du système :

- * Multiplier (ou diviser) une ligne par un réel non-nul ;
- * Échanger plusieurs lignes entre elles ;
- * Ajouter (ou retrancher) une combinaison linéaire des autres lignes à une ligne donnée.

Et grâce à ces règles simples, nous pouvons résoudre n'importe quel système linéaire !

2.2 Un exemple de la méthode

Trouvons les solutions du système :

$$\begin{cases} x + y + z = 1 \\ x + 2y + 3z = 0 \\ 2x - y = \pi \end{cases}$$

Ré-écrivons le d'abord sous forme semi-matricielle :

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 0 \\ 2 & -1 & 0 & \pi \end{array} \right]$$

On sélectionne un élément (en général sur la diagonale) qui nous servira à annuler d'autres entrées de la matrice :

$$\left[\begin{array}{ccc|c} \boxed{1} & 1 & 1 & 1 \\ 1 & 2 & 3 & 0 \\ 2 & -1 & 0 & \pi \end{array} \right]$$

- * On remplace L_2 par $L_2 - L_1$
- * On remplace L_3 par $L_3 - 2L_1$

ce qui donne

$$\left[\begin{array}{ccc|c} \boxed{1} & 1 & 1 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & -3 & -2 & \pi-2 \end{array} \right].$$

Nous choisissons le pivot suivant, toujours sur la diagonale, et recommençons :

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & \boxed{1} & 2 & -1 \\ 0 & -3 & -2 & \pi-2 \end{array} \right].$$

* On remplace L_1 par $L_1 - L_2$, et

* On remplace L_3 par $L_3 + 3L_2$.

Le résultat est donc :

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 2 \\ 0 & \boxed{1} & 2 & -1 \\ 0 & 0 & 4 & \pi-5 \end{array} \right].$$

Le pivot suivant vaut 4 et pas 1, alors nous divisons d'abord par 4 :

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & \frac{\pi-5}{4} \end{array} \right].$$

Puis nous appliquons encore la même technique :

* $L_1 \leftarrow L_1 + L_3$

* $L_2 \leftarrow L_2 - 2L_3$

pour obtenir

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{\pi+3}{4} \\ 0 & 1 & 0 & -1 - \frac{\pi-5}{2} \\ 0 & 0 & \boxed{1} & \frac{\pi-5}{4} \end{array} \right].$$

Le système de départ possède donc les mêmes solutions que le système représenté par

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{\pi+3}{4} \\ 0 & 1 & 0 & \frac{3-\pi}{2} \\ 0 & 0 & \boxed{1} & \frac{\pi-5}{4} \end{array} \right].$$

c'est-à-dire

$$\begin{cases} x = \frac{\pi+3}{4} \\ y = \frac{3-\pi}{2} \\ z = \frac{\pi-5}{4} \end{cases}$$

Considérons le système représenté par :

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 1 & 1 & 2 \end{array} \right]$$

On échange $L_2 \leftrightarrow L_3$:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & \boxed{1} & 1 & 2 \\ 0 & 0 & 1 & 4 \end{array} \right]$$

et nous pouvons continuer comme précédemment.

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & \boxed{1} & 1 & 2 \\ 0 & 0 & 1 & 4 \end{array} \right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & \boxed{1} & 4 \end{array} \right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 4 \end{array} \right]$$

Dès lors $x = -2, y = -2, z = 4$

Supposons maintenant que le système soit donné par

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 0 & 4 & 4 \\ 0 & 0 & 2 & 2 \end{array} \right]$$

Pas de pivot en ligne 2? Tant pis! On divise donc L_3 par 2 et on l'utilise comme pivot.

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 1 \end{array} \right]$$

Dès lors le système devient :

$$\begin{cases} x + y = -1 \\ 0 = 0 \\ z = 1 \end{cases}$$

Ce système est particulier : la seconde équation est toujours vérifiée.

On n'a donc en réalité que **deux** équations pour trois inconnues. Il y a ici une *infinité* de solutions.

infinité

Les solutions sont donc de la forme $(x, -1 - x, 1)$ avec $x \in \mathbb{R}$ (ou $(-1 - y, y, 1)$ avec $y \in \mathbb{R}$, c'est équivalent!).

3 Inversion de matrices

Considérons le calcul matriciel suivant :

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b & \lambda c \\ d & e & f \\ g & h & i \end{pmatrix}$$

Un autre :

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} d & e & f \\ a & b & c \\ g & h & i \end{pmatrix}$$

Un dernier :

$$\begin{pmatrix} 1 & 0 & 0 \\ \lambda & 1 & \mu \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{pmatrix} a & b & c \\ d + \lambda a + \mu g & e + \lambda b + \mu h & f + \lambda c + \mu i \\ g & h & i \end{pmatrix}$$

Remarque. Comme on l'a vu : les opérations permises pour résoudre un système sont des produits matriciels.

- * Partant d'une matrice M , supposons avoir une suite d'opérations transformant M en la matrice identité.
- * Notons les matrices associées à ces opérations : $P_1, P_2, P_3, \dots, P_k$.
- * Dès lors : $P_k \cdots P_2 P_1 M = I$.
- * Et donc $P_k \cdots P_2 P_1 = M^{-1}$!
- * Comme $P_k \cdots P_2 P_1 = P_k \cdots P_2 P_1 I$, il suffit d'appliquer la séquence d'opérations à la matrice identité elle-même pour avoir l'inverse de M .

Cherchons à inverser la matrice suivante :

$$M = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -2 \\ 3 & 0 & 1 \end{pmatrix}$$

Nous écrivons la matrice augmentée de l'identité :

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 3 & 0 & 1 & 0 & 0 & 1 \end{array} \right]$$

Et appliquons la méthode du pivot :

$$\begin{aligned} \left[\begin{array}{ccc|ccc} \boxed{1} & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & -6 & 4 & -3 & 0 & 1 \end{array} \right] &\Rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & -2 & 0 \\ 0 & \boxed{1} & -2 & 0 & 1 & 0 \\ 0 & 0 & -8 & -3 & 6 & 1 \end{array} \right] \\ \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & -2 & 0 \\ 0 & \boxed{1} & -2 & 0 & 1 & 0 \\ 0 & 0 & -8 & -3 & 6 & 1 \end{array} \right] &\Rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & -2 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & \boxed{1} & \frac{3}{8} & \frac{-3}{4} & \frac{-1}{8} \end{array} \right] \\ &\Rightarrow \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{-1}{8} & \frac{1}{4} & \frac{3}{8} \\ 0 & 1 & 0 & \frac{3}{8} & \frac{-1}{4} & \frac{-2}{8} \\ 0 & 0 & 1 & \frac{3}{8} & \frac{-3}{4} & \frac{-1}{8} \end{array} \right] \end{aligned}$$

Dès lors l'inverse de la matrice M est

$$M^{-1} = \begin{pmatrix} \frac{-1}{8} & \frac{1}{4} & \frac{3}{8} \\ \frac{3}{8} & \frac{-1}{4} & \frac{-2}{8} \\ \frac{3}{8} & \frac{-3}{4} & \frac{-1}{8} \end{pmatrix}$$

La méthode du Pivot de Gauß est donc une excellente méthode pour inverser des matrices si nécessaire.

Chapitre XVI

Fonctions de plusieurs variables réelles à valeurs vectorielles

Nous généralisons ici l'étude des fonctions $\mathbb{R} \rightarrow \mathbb{R}$ au cadre des fonctions $A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, c'est-à-dire définies sur un sous-ensemble $A \subset \mathbb{R}^n$ et à valeurs dans $\mathbb{R}^m$. Nous ne supposons plus, dorénavant, que les applications sont linéaires. Cependant nous étudierons comment les applications linéaires peuvent approximer les applications quelconques.

1 Une fonction comme une transformation

Souvent, et plus particulièrement lorsque $m = n$, il est utile d'imaginer une fonction comme une transformation de l'espace $\mathbb{R}^n$ considéré. Dans ce cas, la notion de graphe n'est pas la plus pertinente.

Exemple. Une rotation, une homothétie, etc. sont des transformations du plan.

Exemple. L'application $\mathbb{R} \rightarrow \mathbb{R} : x \mapsto -x$ est une transformation de la droite : c'est la symétrie par rapport à l'origine. Dans cette vision, son graphe nous distrait plutôt qu'il ne nous aide.

Les exemples mentionnés ci-avant sont en fait des exemples linéaires. Que se passe-t-il dans le cas non-linéaire ?

Exemple. Regardons la fonction $\mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^3$. Elle peut s'imaginer comme une transformation qui laisse fixe 0 ainsi que 1 et -1 , mais compresse l'intervalle $[-1, 1]$ et dilate les intervalles $]-\infty, -1]$ et $[1, \infty[$. Dans cette vision, on se rend compte que le « coefficient de dilatation » ou de « compression » n'est pas identique selon le point considéré : proche de 0, les points ont tendance à se rapprocher très fort, mais plus on s'en éloigne, plus les points ont tendances à s'éloigner les uns des autres.

1.1 Motivation et transformation de la droite

1.1.1 Dilatation et dérivée

Si $\delta > 0$, alors $I := [a, a + \delta]$ est un intervalle de longueur δ . Pour savoir comment il se compresse ou se dilate, on va estimer la taille de l'intervalle $f(I)$ par $[f(a), f(a + \delta)]$ (ce n'est pas la réalité, mais cela suffit pour illustrer) et rapporter cela à la taille de l'intervalle de départ. Cela donne le nombre $\frac{f(a + \delta) - f(a)}{\delta}$ qui indique un taux de dilatation : plus ce nombre est grand, plus l'intervalle a grandi par l'action de f .

Bien entendu, on ne s'intéresse pas à la dilatation d'un intervalle fixé, mais à la dilatation « près de a », c'est-à-dire lorsque $\delta \rightarrow 0$. Ce coefficient donne donc le nombre dérivé de f en a : $f'(a)$.

1.1.2 Dilatation et applications linéaires

Sur la droite, les applications linéaires sont les applications de la forme $L : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto cx$ pour un certain réel c (une matrice de taille 1×1 dans le langage du chapitre précédent). En d'autres termes, il s'agit exactement des dilatations ($|c| > 1$) et des compressions ($|c| < 1$) composées éventuellement avec une symétrie (lorsque $c < 0$).

Il est dès lors naturel de « classer » ou de « modéliser » la dilatation d'une fonction quelconque $f : \mathbb{R} \rightarrow \mathbb{R}$ (près d'un point a fixé) en terme d'une application linéaire. Ce classement a déjà été réalisé puisque nous savons que

$$f(x) = f(a) + f'(a)(x - a) + o(|x - a|)$$

(par définition de la dérivée, ou par le théorème de Taylor XI.9).

En d'autres termes, si nous notons $df_a : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto f'(a)x$, nous avons :

$$f(x) = f(a) + df_a(x - a) + o(|x - a|)$$

ou encore de manière équivalente

$$\lim_{x \rightarrow a} \frac{f(x) - f(a) - df_a(x - a)}{x - a} = 0.$$

1.2 Transformations générales

Considérant le cas d'une application $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ quelconque, nous allons nous intéresser pour tout $a \in A$ à l'existence d'une application linéaire df_a vérifiant

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - df_a(\mathbf{x} - \mathbf{a})}{\|\mathbf{x} - \mathbf{a}\|} = 0.$$

Cette application sera appelée la *différentielle* de f au point a .

La propriété fondamentale étant que les applications $\mathbf{x} \mapsto f(\mathbf{x})$ et $\mathbf{x} \mapsto f(\mathbf{a}) + df_a(\mathbf{x} - \mathbf{a})$ sont proches l'une de l'autre lorsque $\mathbf{x}$ est proche de $\mathbf{a}$. Nous formerons les définitions formelle dans la section 5.

2 Graphe

Rappelons que le graphe d'une application $f : A \rightarrow B$ est l'ensemble $\Gamma_f = \{(x, f(x)) \in A \times B\}$. Lorsque $A, B \subset \mathbb{R}$, il était naturel de dessiner cet ensemble dans le plan muni d'un repère cartésien. Il permet souvent d'aider l'intuition lorsqu'il s'agit de considérer une fonction comme une association d'un élément de A avec un élément de B .

Lorsque $A \subset \mathbb{R}^n$ et $B = \mathbb{R}^m$, comme c'est le cas dans la situation que nous voulons étudier, il faudrait pouvoir « dessiner » dans $\mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$. Notre esprit humain étant assez limité, il est rare d'être capable de représenter des espaces de dimension plus grande que 3 ; ceci laisse les possibilités suivantes pour $(m, n) : (1, 1), (1, 2), (2, 1)$.

Dans le premier cas, si $m = n = 1$, c'est la notion de graphe déjà vue : en général, il s'agira d'une courbe dans le plan. Pour comprendre un tel graphe, nous avons déjà vu des outils, tels que la dérivée, qui permettent d'en esquisser l'allure. Dans les deux autres cas, le graphe ne sera plus dans le plan mais dans l'espace. Pour le représenter, il faudra donc en général user d'un effet de perspective.

2.1 Application du plan vers la droite

Il s'agit ici du deuxième cas, $m = 1, n = 2$, c'est-à-dire une application $f : \mathbb{R}^2 \rightarrow \mathbb{R}$; le graphe est alors donné par

$$\Gamma_f = \{(x, y, f(x, y)) \text{ t.q. } x, y \in \text{dom } f\},$$

c'est-à-dire est l'ensemble des solutions de l'équation $z = f(x, y)$. En général, cet ensemble décrit une surface de l'espace.

Exemple. Voici quelques expressions de fonctions $\mathbb{R}^2 \rightarrow \mathbb{R}$, et l'interprétation géométrique du graphe : $f(x, y) = k$, où $k \in \mathbb{R}$ est une constante : le graphe est le plan d'équation $z = k$, c'est-à-dire un plan parallèle au plan contenant les axes de coordonnées, à « hauteur » (positive ou négative) k ;

$f(x, y) = ax + by$, où $a, b \in \mathbb{R}$ sont des constantes (il s'agit d'une forme linéaire) : le graphe est le plan d'équation $z = ax + by$;

$f(x, y) = x^2 + y^2$: la hauteur est donnée par le carré de la distance entre (x, y) et l'origine, on appelle cela un paraboloides de révolution ;

$f(x, y) = \sqrt{r^2 - x^2 - y^2}$, où $r > 0$ est une constante : il s'agit d'une surface dont les points vérifient l'équation $z^2 + x^2 + y^2 = r^2$, c'est donc un morceau de la sphère centrée en l'origine de rayon r – il s'agit de l'hémisphère supérieure.

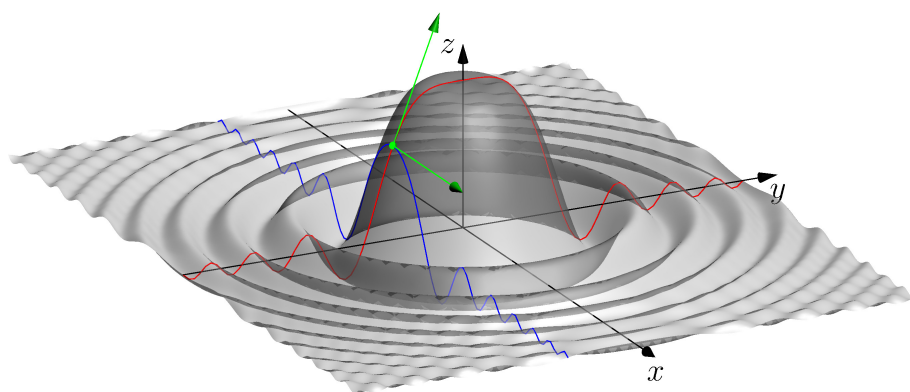
Pour aider à comprendre la structure générale d'une surface (par exemple un graphe) compliquée, il peut être utile de réaliser des « coupes transverses », que nous appellerons « traces » :

Définition XVI.1. La *trace* d'une surface dans un plan est l'intersection de la surface avec le plan.

trace

Les traces, en général des courbes, sont des objets qui « vivent » dans un plan : on peut dès lors les représenter sur la feuille simplement, sans besoin d'utiliser d'effets de perspective.

Exemple. Dans l'exemple ci-dessous, deux traces ont été dessinées. En bleu, une trace obtenue avec un plan dont la coordonnée y est constante ; en rouge une trace obtenue sur un plan dont la coordonnée x est constante égale à 0. En vert sont représentés les vecteurs tangent à ces courbes (dont la pente est la dérivée partielle correspondante – voir définition XVI.8).



(Ce dessin a été repris de Bruno Colombel sous licence CC-by-sa.)

Exemple. Considérons l'exemple du parabolôïde de révolution, c'est-à-dire $f(x, y) = x^2 + y^2$. La trace avec le plan horizontal d'équation $z = k$ du parabolôïde de révolution est l'intersection du graphe dont l'équation est $z = x^2 + y^2$, avec le plan $z = k$: il s'agit donc de la courbe d'équation $k = x^2 + y^2$ (et $z = k$). On comprend alors que cette courbe n'existe pas si $k < 0$ (car l'équation est impossible). Lorsque $k > 0$, c'est un cercle de rayon $\sqrt{k}$ et de centre $(0, 0, k)$. En d'autres termes, le graphe est constitué de cercles superposés, tous centrés autour de l'axe des z , mais de rayon variable. Remarquons que pour $k = 0$, le seul point vérifiant les conditions est l'origine $(0, 0, 0)$. Dans ce cas, donc, la trace est réduite à un point.

Regardons à présent la trace dans le plan d'équation $x = k$. Les équations de cette courbe sont données par

$$\begin{cases} z = k^2 + y^2 \\ x = k, \end{cases}$$

ce qui, en représentant z (imaginé être la hauteur) en fonction de y , donne une parabole dont le sommet (le point le plus bas) est situé en $(k, 0, k^2)$. La trace dans le plan d'équation $y = k$ est identique.

La trace dans un plan plus générique, d'équation $z = d - ax - by$ (pour certaines constantes a, b, d réelles) donne, en remplaçant z dans l'équation du parabolôïde :

$$\begin{cases} d - ax - by = x^2 + y^2 \\ z = d - ax - by \end{cases}$$

dont la première équation est en général l'équation d'un cylindre : $(x + \frac{a}{2})^2 + (y + \frac{b}{2})^2 = d + \frac{a^2+b^2}{4}$ sauf si $d \leq -\frac{a^2+b^2}{4}$, auquel cas c'est une droite verticale (cas égalité) ou même l'ensemble vide (inégalité stricte). Dès lors la trace est vue comme l'intersection d'un cylindre avec le plan : une ellipse (ou un cercle).

Remarque (*). Parfois, on appelle « courbe de niveau » une trace horizontale. Cependant, plus souvent, ces termes correspondent à une autre notion, comme ci-dessous.

Définition XVI.2. La courbe de niveau k de la fonction $f : \text{dom } f \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ est constituée de l'ensemble

$$\{\mathbf{x} \in \text{dom } f \text{ t.q. } f(\mathbf{x}) = k\}.$$

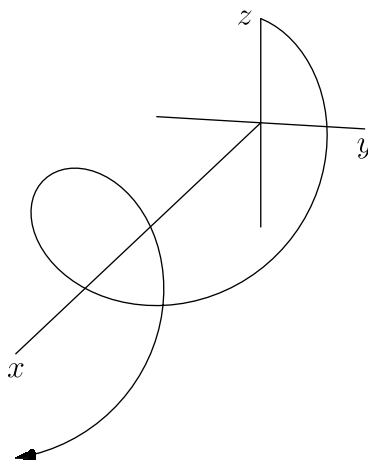
En d'autres termes, la courbe de niveau est un sous-ensemble du domaine dont l'image par f correspond à la trace avec le plan horizontal d'équation $z = k$.

2.2 Applications de la droite vers le plan

Il s'agit ici du « troisième cas » : on considère une application $A \subset \mathbb{R} \rightarrow \mathbb{R}^2$. Nous avons déjà vu qu'une manière naturelle de considérer une telle application est de la voir comme la paramétrisation d'une courbe du plan. Cependant, il est aussi possible d'en dessiner le graphe, ce qui reviendra à dessiner l'ensemble

$$\{(x, f_1(x), f_2(x)) \text{ t.q. } x \in A\}$$

Par exemple pour $f(x) = (\cos t, \sin t)$ (paramétrisation du cercle), on obtiendrait une hélice



La courbe paramétrée, dans ce cas le cercle, est obtenue en projetant le graphe sur le plan des deux dernières coordonnées.

3 Limites et continuité

Comme dans le cas des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, nous allons définir divers niveaux de régularité ; le premier est le cas des fonctions continues, ce qui impose d'introduire la notion de limite et d'adhérence. Contrairement au cas des fonctions d'une variable réelle à valeurs vectorielles, il n'est pas aisé de se ramener simplement à la notion déjà vue pour les fonctions réelles.

Définition XVI.3. Un point $\mathbf{a} \in \mathbb{R}^n$ est adhérent à un sous-ensemble $A \subset \mathbb{R}^n$ si pour tout δ strictement positif, il existe $\mathbf{x} \in A$ tel que $\|\mathbf{a} - \mathbf{x}\| < \delta$. On note $\text{adh } A$ l'ensemble des points adhérents à A .

Définition XVI.4. Soit $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, $\mathbf{a} \in \text{adh } A$ et $\mathbf{L} \in \mathbb{R}^m$; on dit que la limite de f en $\mathbf{a}$ est $\mathbf{L}$ si

$$\forall \epsilon > 0, \exists \delta > 0 : \forall \mathbf{x} \in A, \|\mathbf{x} - \mathbf{a}\| < \delta \Rightarrow \|f(\mathbf{x}) - \mathbf{L}\| < \epsilon.$$

Cette situation est notée $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = \mathbf{L}$.

Exemple. Si $f(x, y) = \frac{x^3}{x^2 + y^2}$, alors $\lim_{(x, y) \rightarrow 0} f(x, y) = 0$. Prouvons cela de deux manières :

Via la définition. Nous voulons montrer

$$\forall \epsilon > 0, \exists \delta > 0 : \forall (x, y) \in \mathbb{R}^2, \|(x, y) - (0, 0)\| < \delta \Rightarrow \|f(x, y) - 0\| < \epsilon$$

c'est-à-dire

$$\forall \epsilon > 0, \exists \delta > 0 : \forall (x, y) \in \mathbb{R}^2, \sqrt{x^2 + y^2} < \delta \Rightarrow \left| \frac{x^3}{x^2 + y^2} \right| < \epsilon$$

Or $|x^3| = |x|x^2 \leq |x|x^2 + |x|y^2 = |x|(x^2 + y^2)$ puisque $|x|y^2 \geq 0$. Par ailleurs, similairement, $|x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2}$. Dès lors,

$$|f(x, y) - 0| = \left| \frac{x^3}{x^2 + y^2} \right| \leq |x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2} = \|(x, y)\|.$$

Nous concluons que si $\epsilon > 0$ est fixé, alors il existe $\delta > 0$, par exemple défini par $\delta := \epsilon$, tel que si $\|(x, y)\| < \delta$, alors $|f(x, y)| \leq \|(x, y)\| < \epsilon$, ce que nous voulions voir. $\square$

Définition XVI.5. Une application $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ est dite *continue au point a* si

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a}).$$

continue au point a

Globalement, f est dite continue si elle est continue en tout point de son domaine.

Remarque (*). Ces définitions généralisent les définitions déjà données, en d'autres termes si $n = 1$ on retrouve les définitions déjà faites dans le cas d'une variable réelle à valeur réelle (si $m = 1$) et à valeur vectorielle (si $m > 1$), ni plus ni moins.

Il serait aisé, à partir de ces définitions, de prouver que les projections sur les axes de coordonnées sont continues :

Résultat XVI.6. Les applications $(x, y) \mapsto x$ et $(x, y) \mapsto y$ sont continues sur $\mathbb{R}^2$.

Résultat XVI.7 (Règles de calcul). Nous énonçons ces règles dans le cas de la continuité, mais elles sont également valables dans le calcul de limites :

- * La somme, la différence, le quotient et le produit de deux applications continues en $\mathbf{a}$ est encore continu en $\mathbf{a}$ (si le dénominateur ne s'annule pas en $\mathbf{a}$, dans le cas du quotient.)
- * La somme, la différence, le quotient et le produit de deux applications continues sur leur domaine est encore continu sur son domaine.

De même pour la composition :

- * La composée d'une application f continue en $\mathbf{a}$ avec une application continue en $f(\mathbf{a})$ est encore continue en $\mathbf{a}$
- * La composée de deux applications continues est continue sur son domaine.

En particulier, nous avons le résultat suivant :

Corollaire. Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est continue en $\mathbf{a}$ (ou $\mathbf{a}$ admet une limite $\mathbf{L}$), alors pour toute courbe $\gamma : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ telle que $\gamma(0) = \mathbf{a}$, il faut $\lim_{t \rightarrow 0} f(\gamma(t)) = f(\mathbf{a})$ (ou $\mathbf{L}$).

Remarque (*). Ce résultat dit que si la limite existe, alors elle vaut la même valeur quelle que soit la direction de laquelle on approche le point $\mathbf{a}$. En particulier, une « fonction en colimaçon » autour de l'axe des z (pensons à une rampe d'accès de parkings superposés) n'est pas continue en $(0, 0)$:

Exemple. Considérons $f(x, y) = \frac{xy}{x^2+y^2}$. Prenons d'abord $\gamma(t) = (t, 0)$ de sorte que $f(\gamma(t)) = \frac{t \cdot 0}{t^2+0^2} = 0$ pour tout $t \neq 0$. Bien sûr $\lim_{t \rightarrow 0} f(\gamma(t)) = \lim_{t \rightarrow 0} 0 = 0$. Prenons maintenant $\gamma(t) = (t, t)$, de sorte que $f(\gamma(t)) = \frac{1}{2}$ pour tout $t \neq 0$. La limite est désormais différente, donc f n'admet pas de limite en $(0, 0)$.

Par contre, f est continue partout ailleurs par application des règles de calcul.

Pour mieux comprendre f , considérons $\gamma(t) = (r \cos(t), r \sin(t))$. Alors

$$f(\gamma(t)) = \cos(t) \sin(t) = \frac{1}{2} \sin(2t),$$

de sorte que la hauteur sur le graphe ne dépend que de l'angle t (et pas de la distance r), et varie comme le sinus de l'angle double.

4 Dérivabilité

Si $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, la notion de dérivée définie via $f(\mathbf{x} + \delta)$ n'a pas de sens, car ici $\mathbf{x} \in \mathbb{R}^n$ et $\delta \in \mathbb{R}$. Il y a deux manières naturelles de s'en sortir : prendre δ dans $\mathbb{R}^n$ ou faire en sorte de restreindre le nombre de variables de f . Le premier cas ne conduit pas immédiatement à une notion intéressante, regardons alors le second. « Restreindre le nombre de variables » veut dire qu'on ne laisse varier $x \in \mathbb{R}^n$ que dans une direction, afin de laisser une fonction d'une seule variable réelle.

Définition XVI.8. Soit $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ une application, $\mathbf{a} \in A$ et $\mathbf{v} \in \mathbb{R}^n$. La *dérivée directionnelle* de f au point $\mathbf{a}$ dans la direction $\mathbf{v}$ est le nombre réel donné par

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})}{t}.$$

Si $\mathbf{v} = \mathbf{e}_i$ pour i entre 1 et n , on parle de la *i^e dérivée partielle* en $\mathbf{a}$. On la note parfois

- * $\partial_i f(\mathbf{a})$ ou,
- * lorsque $\mathbf{x} = (x_1, \dots, x_n)$, on la note parfois $\frac{\partial f}{\partial x_i}(\mathbf{a})$ ou,
- * lorsque $\mathbf{x} = (x, y)$, on la note parfois $\frac{\partial f}{\partial x}(\mathbf{a})$ (pour la première dérivée partielle), $\frac{\partial f}{\partial y}(\mathbf{a})$ (pour la seconde dérivée partielle), etc.

La matrice *jacobienne* de f au point $\mathbf{a}$ est la matrice composée de toutes les dérivées partielles :

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

Lorsque $m = 1$, la jacobienne ne comporte qu'une seule ligne, et on peut donc l'assimiler à un vecteur qui est appelé le *gradient* de f au point $\mathbf{a}$:

$$\nabla f(\mathbf{a}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{a}) \right) \in \mathbb{R}^n.$$

Exemple. Si $f(x, y) = x^2 y^3 + \cos(y)$, les dérivées partielles sont

$$\frac{\partial f}{\partial x}(a, b) = 2ab^3 \qquad \frac{\partial f}{\partial y}(a, b) = 3a^2 b^2.$$

Le gradient est donc, au point (a, b) , donné par $(2ab^3, 3a^2 b^2)$.

Plus généralement, la dérivée directionnelle dans la direction (u, v) au point (a, b) est

$$\frac{\partial f}{\partial (u, v)}(a, b) = \lim_{t \rightarrow 0} \frac{f((a, b) + t(u, v)) - f(a, b)}{t} = \lim_{t \rightarrow 0} \frac{(a + tu)^2 (b + tv)^3}{t} = 2ab^3 u + 3a^2 b^2 v$$

5 Différentiabilité

Une autre notion de « dérivabilité » s'obtient comme annoncé en section 1.2.

Définition XVI.9. Soit $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ une application et $\mathbf{a}$ un point intérieur au domaine A de f . On dit que f est *différentiable* en $\mathbf{a}$ s'il existe une application linéaire $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ vérifiant

différentiable

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - T(\mathbf{x} - \mathbf{a})}{\|\mathbf{x} - \mathbf{a}\|} = 0.$$

Cette application sera appelée la *différentielle* de f au point $\mathbf{a}$ et notée $df_{\mathbf{a}}$.

différentielle

Résultat XVI.10. La somme, différence, le produit, le quotient et la composée de fonctions différentiables sont différentiables lorsque cela a du sens (par exemple, pour le quotient, il faut que le dénominateur ne s'annule pas).

Résultat XVI.11. Si f est différentiable en $\mathbf{a}$, alors elle admet des dérivées directionnelles dans toutes les directions et de plus

$$df_{\mathbf{a}}(\mathbf{v}) = \frac{\partial f}{\partial \mathbf{v}}(\mathbf{a})$$

pour tout vecteur $\mathbf{v} \in \mathbb{R}^n$. De plus, la matrice associée à l'application linéaire $df_{\mathbf{a}}$ est donnée par la matrice jacobienne de f en $\mathbf{a}$.

Remarque (*). La notion de différentiabilité est plus forte que simplement demander que toutes les dérivées directionnelles (et a fortiori les dérivées partielles) existent. Précisément, on peut créer des exemples de fonctions dont toutes les dérivées directionnelles en un point donné existent mais qui n'est pas différentiable en ce point.

Résultat XVI.12. Si une application $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ est différentiable au point $\mathbf{a}$, alors $\text{grad } f(\mathbf{a})$ est perpendiculaire à la courbe de niveau $f(\mathbf{a})$ au point $\mathbf{a}$. Par ailleurs, $\text{grad } f(\mathbf{a})$ donne la direction de plus grande pente, c'est-à-dire pour laquelle $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a})$ est maximale.

Exemple. Ce dernier point est de nature particulièrement géométrique : si on imagine que $f(x, y)$ donne la hauteur d'un point situé au dessus du point (x, y) du plan, alors $\nabla f(x, y)$ donnera effectivement la direction dans laquelle la pente sera la plus grande, tandis que la direction orthogonale au gradient donnera la direction dans laquelle la pente sera nulle : la courbe de niveau.

6 Règles de dérivation

Nous avons déjà mentionné le fait que la somme, différence, le produit, le quotient et la composée de fonctions différentiables sont différentiables lorsque cela a du sens. Il se fait qu'en grattant un peu, ce résultat permet d'obtenir une formule particulièrement intéressante.

Soit $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ et $g : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ deux applications avec $g(I) \subset A$, de sorte à pouvoir les composer. Supposons encore que g est différentiable (c'est-à-dire dérivable) en un point $t \in I$, et que f est différentiable en le point $g(t) \in A$. Dans ce cas, $f \circ g : I \rightarrow \mathbb{R}$ est dérivable en t , mais nous avons également une expression :

$$(f(g(t)))' = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(g(t)) g'_i(t)$$

ce qu'on peut ré-écrire

$$(f(g(t)))' = \nabla f(g(t)) \cdot g'(t).$$

Remarque (*). Cette égalité s'obtient à partir de la règle de calcul suivante sur les différentielles $d(f \circ g)(t) = df(g(t)) \circ dg(t)$. Paraphrasons : « La différentielle de la composée est la composée des différentielles » (en faisant attention au point auquel on calcule la différentielle).

7 Dérivabilité et différentiabilité aux ordres supérieurs

Pour généraliser les notions de dérivées secondes, il est très simple de parler des dérivées partielles d'ordre 2, 3, etc. puisqu'il suffit de dériver successivement par rapport à chacune des variables.

Exemple. Si $f(x, y) = x^2 \cos(y)$, alors

$$\frac{\partial f}{\partial x}(x, y) = 2x \cos(y) \quad \frac{\partial f}{\partial y}(x, y) = -x^2 \sin(y)$$

et les dérivées secondes sont données en dérivant les expressions ci-dessus :

$$\begin{aligned} \frac{\partial^2 f}{\partial x \partial x}(x, y) &= 2 \cos y & \frac{\partial^2 f}{\partial x \partial y}(x, y) &= -2x \sin y \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) &= -2x \sin y & \frac{\partial^2 f}{\partial y \partial y}(x, y) &= -x^2 \cos y. \end{aligned}$$

Remarque (*). Afin de raccourcir l'écriture, on notera $\frac{\partial^2 f}{\partial x^2}(x, y)$ au lieu de $\frac{\partial^2 f}{\partial x \partial x}(x, y)$, et similairement pour y .

Définition XVI.13. Si $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ admet des dérivées partielles secondes en un point, sa *matrice Hessienne* est

$$\begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}.$$

Une manière similaire de faire pour la différentiabilité conduit malheureusement à une notion de « différentielle seconde » plus complexe à imaginer, aussi nous n'en parlons pas. Néanmoins la différentiabilité est une notion utile, et il se fait qu'on peut parler de fonctions « deux fois différentiables » sans parler de « différentielle seconde ». Voici un résultat faisant le lien :

Résultat XVI.14. Si une application f admet des dérivées partielles en tout point (x, y) d'un domaine A , et que les applications

$$(x, y) \mapsto \frac{\partial f}{\partial x}(x, y) \quad \text{et} \quad (x, y) \mapsto \frac{\partial f}{\partial y}(x, y)$$

sont continues, alors f est différentiable en tout point (intérieur) de A .

Dès lors, si une fonction admet des dérivées partielles *secondes* continues en tout point, nous obtenons une notion de différentiabilité du second ordre. Plus précisément, nous dirons qu'une application est de *classe* C^1 si elle admet des dérivées partielles continues (c'est-à-dire lorsque le résultat précédent s'applique), et qu'elle est de classe C^2 si elle admet des dérivées partielles secondes continues. Ainsi de suite, on peut définir les fonctions de classe C^k pour toutes les valeurs de k . Si une fonction est de classe C^k pour tout k , on dit aussi qu'elle est de classe C^∞ .

Remarque (*). La plupart des fonctions avec lesquelles nous voulons travailler (polynômes, exponentielles, etc.) sont de classe C^∞ sur leur domaine. Les fonctions « racines » font exception.

Résultat XVI.15. Si une fonction est de classe C^2 , alors sa matrice Hessienne est symétrique.

Dit autrement : dans ce cas, les dérivées secondes mixtes commutent. C'est bien évidemment le cas de toutes les fonctions qui vont nous intéresser.

8 Optimisation à plusieurs variables

Pour $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$, on peut définir la notion de maximum ou de minimum (local, global, strict) comme dans le cas des fonctions réelles (voir définition XI.1). On a alors un résultat similaire à celui des fonctions réelles :

matrice
Hessienne

classe

Résultat XVI.16. Lorsque $m = 1$, si $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ est différentiable en $\mathbf{a}$ et admet un extremum en ce point, alors $\nabla f(\mathbf{a})$ est le vecteur nul.

Regardons plus attentivement le cas $n = 2$.

Résultat XVI.17. Si $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ est différentiable en (a, b) et vérifie $\nabla f(a, b) = \mathbf{0}$, alors notons δ le déterminant de sa matrice Hessienne en (a, b) .

- * Si $\delta > 0$ et si $\frac{\partial^2 f}{\partial x^2}(a, b) > 0$, alors (a, b) est un minimum local strict de f ;
- * Si $\delta > 0$ et si $\frac{\partial^2 f}{\partial x^2}(a, b) < 0$, alors (a, b) est un maximum local strict de f ;
- * Si $\delta < 0$, alors (a, b) n'est ni un minimum, ni un maximum de f .

Ce critère permet de se passer de la notion de « tableau de signe » qui serait d'ailleurs fort délicate à définir à plusieurs variables.

Chapitre XVII

Intégrales multiples

1 Définition et motivation

Soit $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction. Dans le cas des fonctions d'une variable réelle à valeurs réelles (chapitre XIII), nous nous intéressons à mesurer l'aire sous la courbe. Ici, nous nous intéressons à mesurer le volume sous le graphe de f (en général c'est une surface).

À cette fin, nous pouvons procéder à une définition similaire à celle qui a été faite dans le cas des fonctions réelles : nous divisons le domaine A en petits rectangles R_{ij} de côtés Δx_i et Δy_j , et sur chacun des rectangles nous calculons le volume « au dessus du rectangle » V_{ij} par la formule :

$$V_{ij} = \text{hauteur}_{ij} \Delta x_i \Delta y_j$$

où la hauteur $_{ij}$ (qui est négative si f est négative) peut être approximée, soit par excès en regardant la plus grande valeur atteinte par f sur le rectangle :

$$M_{ij} = \max\{f(x, y) \text{ t.q. } (x, y) \in R_{ij}\}$$

soit par défaut, en regardant la plus petite valeur atteinte par f sur le rectangle :

$$m_{ij} = \min\{f(x, y) \text{ t.q. } (x, y) \in R_{ij}\}.$$

Il faut ensuite sommer sur l'ensemble des rectangles pour obtenir une valeur par excès $\bar{S}(f)$ ou par défaut $\underline{S}(f)$ du volume recherché. Si enfin, lorsque la taille des rectangles tend vers 0, les valeurs de $\bar{S}$ et $\underline{S}$ tendent l'une vers l'autre, on appelle cette valeur commune l'*intégrale de la fonction* et elle se note :

intégrale de la fonction

$$\iint_A f$$

ou, si l'on désire explicitement écrire les variables :

$$\iint_A f(x, y) dx dy.$$

Certains écrivent cette dernière égalité en écrivant d'abord l'*élément d'aire* $dx dy$, c'est-à-dire sous la forme :

élément d'aire

$$\iint_A dx dy f(x, y).$$

Une fonction pour laquelle le processus décrit ci-dessus donne effectivement un nombre réel est dite *intégrable* (au sens de Riemann). Cette notion d'intégrabilité à deux variables peut facilement s'étendre au cas général des fonctions à plusieurs variables. Lorsqu'il y a trois variables, la notation devient $\iiint_A f$, etc.

intégrable

Précisons ce que nous calculons avec une intégrale : si $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ est une fonction, alors l'intégrale de cette fonction représente le volume « algébrique » sous la surface d'équation $z = f(x, y)$, c'est-à-dire

$$\iint_A f(x, y) dx dy = \text{vol}(S^+) - \text{vol}(S^-)$$

où $\text{vol } S^\pm$ représente le volume de l'ensemble $S^\pm$, qui sont des ensembles définis par

$$S^+ = \left\{ (x, y, z) \text{ t.q. } \begin{array}{l} (x, y) \in A, \text{ et} \\ 0 \leq z \leq f(x, y) \end{array} \right\} \quad S^- = \left\{ (x, y, z) \text{ t.q. } \begin{array}{l} (x, y) \in A, \text{ et} \\ 0 \geq z \geq f(x, y) \end{array} \right\}$$

Le mot « algébrique » se réfère au fait qu'il y a contribution négative lorsque la fonction est négative.

2 Propriétés

Bien que toutes les fonctions ne sont pas intégrables, nous avons un théorème :

Résultat XVII.1. Si $A \subset \mathbb{R}^2$ est un domaine borné « suffisamment régulier », et $f : A \rightarrow \mathbb{R}$ est continue, alors f est intégrable.

Bien sûr, comme dans le cas des fonctions d'une variable, nous avons des propriétés d'additivité :

Résultat XVII.2. Si $A, B \subset \mathbb{R}^2$ sont des ensembles bornés, et $A \cap B = \emptyset$, alors

$$\iint_{A \cup B} f = \iint_A f + \iint_B f.$$

De même, si $\alpha, \beta \in \mathbb{R}$ et $f, g : A \rightarrow \mathbb{R}$ alors

$$\iint_A (\alpha f + \beta g) = \alpha \iint_A f + \beta \iint_A g$$

En pratique, on peut souvent calculer ces intégrales en se ramenant au cas des intégrales à une seule variable (et donc au théorème fondamental du calcul différentiel et intégral). Nous devons introduire quelques définitions :

Définition XVII.3. Un domaine $A \subset \mathbb{R}^2$ est dit *verticalement simple* s'il existe $a, b \in \mathbb{R}$ et des fonctions $g_1, g_2 : [a, b] \rightarrow \mathbb{R}$ telles que

$$A = \left\{ (x, y) \in \mathbb{R}^2 \text{ t.q. } x \in [a, b], g_1(x) \leq y \leq g_2(x) \right\}.$$

Le domaine est *horizontalement simple* s'il existe $c, d \in \mathbb{R}$ et des fonctions $h_1, h_2 : [c, d] \rightarrow \mathbb{R}$ telles que

$$A = \left\{ (x, y) \in \mathbb{R}^2 \text{ t.q. } y \in [c, d], h_1(y) \leq x \leq h_2(y) \right\}.$$

Résultat XVII.4 (Théorème de Fubini, version faible). Soit $A \subset \mathbb{R}^2$ et $f : A \rightarrow \mathbb{R}$.

* Si A est *verticalement simple*, alors

$$\iint_A f = \iint_A f(x, y) dx dy = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x, y) dy \right) dx.$$

* Si A est *horizontalement simple* et $f : A \rightarrow \mathbb{R}$, alors

$$\iint_A f = \iint_A f(x, y) dx dy = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x, y) dx \right) dy.$$

Exemple. Considérons l'ellipse $E = \left\{ \frac{x^2}{u^2} + \frac{y^2}{v^2} \leq 1 \right\}$ ($u, v > 0$ sont fixés). Calculons

$$\iint_E 1$$

verticalement
simple

horizontalement
simple

de la sorte :

$$\begin{aligned}
 \int_{-u}^u \int_{-\sqrt{1-\frac{x^2}{u^2}}}^{\sqrt{1-\frac{x^2}{u^2}}} 1 \, dy \, dx &= \int_{-u}^u 2\sqrt{1-\frac{x^2}{u^2}} \, dx \\
 &= \frac{2u}{u} \int_{-u}^u \sqrt{u^2-x^2} \, dx \\
 &= \frac{4u}{u} \int_0^u \sqrt{u^2-x^2} \, dx \\
 &= \frac{4u}{u} \frac{\pi u^2}{4} \quad \text{cette étape est justifiée ci-dessous} \\
 &= \pi u^2
 \end{aligned}$$

où $\int_0^u \sqrt{u^2-x^2} \, dx = \frac{\pi u^2}{4}$ puisque c'est l'aire d'un quart de cercle de rayon u .

3 Coordonnées cartésiennes

Nous avons introduit l'intégration à plusieurs variables pour calculer des volumes sous des graphes, plus précisément si $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ est une fonction, alors l'intégrale de cette fonction représente le volume « algébrique » sous la surface d'équation $z = f(x, y)$. Géométriquement, les coordonnées (x, y, z) représentent ici des coordonnées cartésiennes.

Regardons un cas particulier : pour $A \subset \mathbb{R}^2$, considérons

$$\{(x, y, z) \text{ t.q. } (x, y) \in A, z \in [0, 1]\}.$$

C'est un solide dont la base est A et dont la hauteur est 1. Son volume est donc, numériquement, égal à l'aire de la base (multiplié par 1). En d'autres termes, nous avons :

Résultat XVII.5. Si $A \subset \mathbb{R}^2$, alors $\iint_A 1$ représente l'aire de A .

Exemple. Soit $f : A := [0, 1] \times [0, \pi/2] \rightarrow \mathbb{R} : (x, y) \mapsto 1$. Alors

$$\iint f = \iint 1 \, dx \, dy = \int_0^1 \left(\int_0^{\pi/2} dy \right) dx = \pi/2$$

est l'aire du domaine de f (décrit en coordonnées cartésiennes) : c'est l'aire d'un rectangle.

4 Coordonnées polaires

Soit $g : B \subset \mathbb{R}^2 \rightarrow \mathbb{R} : (\rho, \theta) \mapsto g(\rho, \theta)$ une fonction. Alors l'intégrale

$$\iint_B g(\rho, \theta) \rho \, d\rho \, d\theta$$

représente le volume « algébrique » sous le graphe d'équation $z = g(\rho, \theta)$, où ρ et θ sont ici des coordonnées polaires du plan. (On dira aussi que (ρ, θ, z) sont des *coordonnées cylindriques* de l'espace – voir section 6.)

coordonnées
cylindriques

Exemple. Soit $g : A := [0, 1] \times [0, \pi/2] \rightarrow \mathbb{R} : (\rho, \theta) \mapsto 1$, c'est-à-dire *exactement* la même fonction que dans l'exemple de la section précédente (seul le nom des variables a été changé, mais il importe peu).

Alors $\iint g(\rho, \theta) \rho \, d\rho \, d\theta = \pi/4$ est l'aire du domaine de g , *vu comme représentant des coordonnées polaires sur le plan*, c'est-à-dire un quart de disque de rayon 1.

Remarque (*). Attention, les notations ont été changées mais là n'est pas l'important. Ce qui importe est que ρ ait été ajouté dans l'intégrale. On dira parfois que $\rho \, d\rho \, d\theta$ est l'élément d'aire des coordonnées polaires.

Insistons :

- * $\iint_A 1 \, d\rho \, d\theta$ est l'aire du morceau de plan dont l'ensemble des coordonnées cartésiennes se trouve dans A , c'est un rectangle; tandis que
- * $\iint_A \rho \, d\rho \, d\theta$ est l'aire du morceau de plan dont l'ensemble des coordonnées polaires est dans A , c'est un disque.

4.1 Justification heuristique

On aimerait justifier le fait que si le domaine représente un ensemble de couples de coordonnées polaires, alors l'intégrale ci-dessus est bien le volume annoncé.

Pour cela, il faut se rappeler qu'en coordonnées cartésiennes, nous avons partitionné le domaine en petits rectangles de côtés Δx_i et Δy_j , sur chacun desquels nous construisons un parallélépipède rectangle dont la hauteur (notée m_{ij} ou M_{ij}) était une approximation de $f(x, y)$, de sorte que leur volume valait hauteur $_{ij} \Delta x_i \Delta y_j$.

En coordonnées polaires, une approche similaire conduit à considérer, non pas des rectangles, mais des secteurs de couronnes. Précisément si (ρ_0, θ_0) est un point, on construit un petit ensemble de points dont les coordonnées polaires (ρ, θ) vérifient

$$\rho_0 - \frac{\Delta \rho}{2} \leq \rho \leq \rho_0 + \frac{\Delta \rho}{2} \quad \text{et} \quad \theta_0 - \frac{\Delta \theta}{2} \leq \theta \leq \theta_0 + \frac{\Delta \theta}{2}.$$

Cet ensemble a pour aire

$$\frac{\Delta \theta}{2\pi} \left(\pi \left(\rho_0 + \frac{\Delta \rho}{2} \right)^2 - \pi \left(\rho_0 - \frac{\Delta \rho}{2} \right)^2 \right) = \rho_0 \Delta \theta \Delta \rho.$$

C'est le ρ_0 apparaissant dans cette formule qui justifie la présence de ρ dans la formule annoncée.

Remarque (*). Le point (ρ_0, θ_0) a été pris au milieu du secteur; un autre choix conduit à une formule un peu moins simple d'apparence, mais qui conduit bien sûr à la même formule finale pour le volume.

4.2 Exemple d'intégration en coordonnées polaires

Pour trouver l'aire de l'intersection des disques de rayon $1/2$ centrés en $(0, 0)$ et $(1/2, 0)$ respectivement, nous allons déterminer leurs équations en coordonnées polaires.

Le disque centré en l'origine a pour équation $\rho \leq 1/2$. Le disque centré en $(1/2, 0)$ n'a pas d'équation polaire évidente mais son équation cartésienne est

$$\left(x - \frac{1}{2}\right)^2 + y^2 \leq \frac{1}{4},$$

ce qui se simplifie en $x^2 + y^2 \leq x$. En remplaçant alors (x, y) par $(\rho \cos \theta, \rho \sin \theta)$, on trouve l'équation

$$\rho^2 \leq \rho \cos \theta,$$

c'est-à-dire $\rho \leq \cos \theta$. L'ensemble de points recherché (l'intersection des disques) est donc décrit en coordonnées polaires par l'ensemble

$$B := \left\{ (\rho, \theta) \text{ t.q. } 0 \leq \rho \leq \frac{1}{2}, \rho \leq \cos \theta, \theta \in]-\pi, \pi] \right\}.$$

Nous savons que pour trouver l'aire, il faut calculer $\iint_B \rho \, d\rho \, d\theta$.

Remarque (*). Nous avons introduit précédemment la notion de domaine verticalement simple et horizontalement simple. Bien sûr, nous pouvons utiliser cette notion même lorsque les variables s'appellent ρ et θ , cependant les mots « vertical » et « horizontal » se réfèrent à l'interprétation géométrique des coordonnées cartésiennes. Nous dirons donc plutôt « simple en la coordonnée ρ » ou « simple en la coordonnée θ ».

L'ensemble B est simple en la coordonnée θ : la contrainte $\rho \leq \cos \theta$ est équivalente à $-\arccos \rho \leq \theta \leq \arccos \rho$ dès lors nous pouvons intégrer par rapport à θ et puis par rapport à ρ :

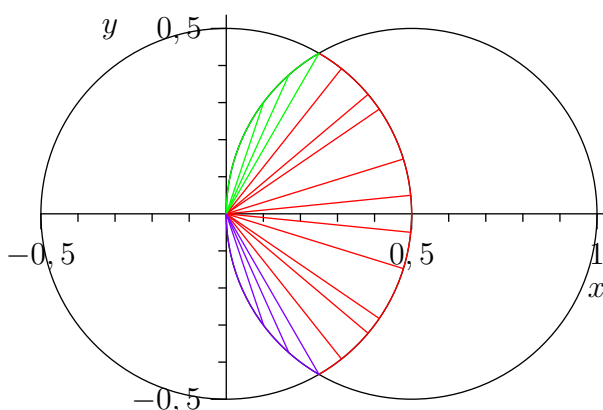
$$\int_0^{1/2} \int_{-\arccos \rho}^{\arccos \rho} \rho d\theta d\rho.$$

Le résultat est $\frac{\pi}{6} - \frac{\sqrt{3}}{8}$.

Une autre approche eut été d'intégrer d'abord par rapport à ρ . Ceci est faisable, mais le bord du domaine est moins agréable à écrire : il faut scinder en trois morceaux. Ceci conduit à écrire

$$\int_{-\pi/2}^{-\pi/3} \int_0^{\cos \theta} \rho d\rho d\theta + \int_{-\pi/3}^{\pi/3} \int_0^{1/2} \rho d\rho d\theta + \int_{\pi/3}^{\pi/2} \int_0^{\cos \theta} \rho d\rho d\theta$$

dont la valeur numérique est évidemment la même que l'intégrale précédente.



5 Coordonnées quelconques

Rappelons qu'un système de coordonnées sur un sous-ensemble du plan est une bijection entre ce sous-ensemble, et une partie de $\mathbb{R}^2$. Les systèmes de coordonnées cartésiennes (définies sur tout le plan) et polaires (définies sur tout le plan excepté l'origine) sont nos favoris. En la présence de *deux* systèmes de coordonnées, chaque point du plan possède donc *deux paires* de coordonnées. Le passage d'une paire à l'autre fournit une nouvelle bijection, appelée l'application de changement de coordonnées. Utilisons ces notions pour calculer des volumes dans un système de coordonnées fixé.

Résultat XVII.6. De manière générale si un point donné a des coordonnées (par exemple polaires) (u, v) , et si nous notons

$$F(u, v) = (F_1(u, v), F_2(u, v))$$

l'application donnant les coordonnées cartésiennes correspondant à ce même point (« application de changement de coordonnées »), alors on peut écrire l'égalité

$$\iint_A f(x, y) dx dy = \iint_{F^{-1}(A)} f(F(u, v)) |\det \text{Jac } F(u, v)| du dv.$$

où $F^{-1}(A)$ est l'ensemble des couples (u, v) représentant des points dont les coordonnées cartésiennes $F(u, v) = (x, y)$ sont dans A :

$$F^{-1}(A) := \{(u, v) \text{ t.q. } F(u, v) \in A\}.$$

Ici, $F : \text{dom } F \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$ est une application de deux variables à valeurs dans $\mathbb{R}^2$, car nous parlons de coordonnées sur le plan. Une égalité tout à fait similaire peut être écrite avec une application $F : \text{dom } F \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, il suffit d'ajuster le nombre de variables.

Exemple. Pour les coordonnées polaires, on a $F : \mathbb{R}^+ \times]-\pi, \pi] \rightarrow \mathbb{R}^2 : (\rho, \theta) \mapsto (\rho \cos \theta, \rho \sin \theta)$, de sorte que $\text{Jac } F_{(\rho, \theta)}$ est donnée par

$$\begin{pmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{pmatrix}$$

dont le déterminant est ρ , ce qui ré-affirme la formule déjà donnée.

6 Coordonnées cylindriques

Les *coordonnées cylindriques* sont définies par l'application de changement de coordonnées

$$F : \mathbb{R}^+ \times]-\pi, \pi] \times \mathbb{R} \rightarrow \mathbb{R}^3 : (\rho, \theta, z) \mapsto (\rho \cos \theta, \rho \sin \theta, z)$$

Dans ce cas, nous avons $\det \text{Jac } F = \rho$ (tout comme dans le cas des coordonnées polaires du plan), et nous avons l'interprétation suivante :

$$\iiint_A f(x, y, z) dx dy dz = \iiint_{F^{-1}(A)} f(F(\rho, \theta, z)) |\det \text{Jac } F_{(\rho, \theta, z)}| d\rho d\theta dz.$$

est « l'hypervolume algébrique » sous le « graphe » de la fonction f . Cet ensemble n'est pas représentable (c'est un ensemble de $\mathbb{R}^4$). Cependant, dans le cas particulier où $f(x, y, z) = 1$, l'intégrale ci-dessus représente le volume de l'ensemble des points de l'espace dont les coordonnées cartésiennes sont dans A .

7 Coordonnées sphériques

Les coordonnées sphériques sont définies par l'application de changement de coordonnées

$$F : \mathbb{R}^+ \times]-\pi, \pi] \times [0, \pi] \rightarrow \mathbb{R}^3 : (\rho, \theta, \varphi) \mapsto (\rho \cos(\theta) \sin(\varphi), \rho \sin(\theta) \sin(\varphi), \rho \cos \varphi)$$

Ici, θ est parfois nommée *longitude* et φ est alors la *colatitude* ou *angle azimuthal*.

Dans ce cas, nous avons $|\det \text{Jac } F| = \rho^2 \sin(\varphi)$, et nous avons l'interprétation suivante :

$$\iiint_A f(x, y, z) dx dy dz = \iiint_{F^{-1}(A)} f(F(\rho, \theta, \varphi)) \rho^2 \sin(\varphi) d\rho d\theta d\varphi.$$

est « l'hypervolume algébrique » sous le « graphe » de la fonction f . Cet ensemble n'est pas représentable. Cependant, dans le cas particulier où $f(x, y, z) = 1$, l'intégrale ci-dessus représente le volume de l'ensemble des points de l'espace dont les coordonnées cartésiennes sont dans A .

Remarque (*). Attention, le nom des variables peut différer selon les conventions. La distance au centre est généralement notée ρ ou r , ce qui ne porte pas à confusion, mais les variables angulaires sont souvent échangées. Il est important de vérifier la signification de chaque variable.

Exemple. Notons $A = F(B)$, où $B = [0, 1] \times [0, 2\pi[\times [0, \pi]$. Alors on peut vérifier que B représente la boule de rayon 1 en coordonnées sphériques. Dès lors $A = \{(x, y, z) \text{ t.q. } x^2 + y^2 + z^2 \leq 1\}$ représente la boule en coordonnées cartésiennes. On en déduit que le volume de la boule vaut

$$\begin{aligned} \iiint_A 1 \, dx \, dy \, dz &= \iiint_B \rho^2 \sin \varphi \, d\rho \, d\theta \, d\varphi \\ &= \int_0^1 \int_0^{2\pi} \int_0^\pi \rho^2 \sin \varphi \, d\varphi \, d\theta \, d\rho \\ &= \int_0^1 \int_0^{2\pi} [-\rho^2 \cos \varphi]_0^\pi \, d\theta \, d\rho \\ &= 2 \int_0^1 \int_0^{2\pi} \rho^2 \, d\theta \, d\rho \\ &= 4\pi \int_0^1 \rho^2 \, d\rho = \frac{4}{3}\pi \end{aligned}$$

comme attendu.

Chapitre XVIII

Équations différentielles ordinaires

Une *équation différentielle ordinaire* (EDO) est une équation faisant intervenir une *fonction* inconnue, généralement notée y dans ce cours. L'équation prend alors de manière générale la forme

équation
différentielle
ordinaire

$$F(x, y, y', \dots, y^{(n)}) = 0$$

où F est une fonction de $n+1$ variables, y est la fonction inconnue et x est la variable de cette fonction. Cette notation est un abus de notation courant – plus précisément une *solution* de l'équation ci-dessus est une fonction $y : I \rightarrow \mathbb{R}$ définie sur un intervalle I , continue sur cet intervalle, et vérifiant

$$F(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0$$

pour toute valeur de $x \in I$. *Résoudre* une équation différentielle est trouver l'ensemble de ses solutions.

Résoudre

Si F dépend effectivement de la dernière variable (la dérivée d'ordre n), on dit que l'équation est d'ordre n . Trouver l'ensemble des solutions d'une EDO donnée est en général un problème compliqué.

Exemple. Quelques exemples d'équations différentielles :

- * $y'' + y = 0$ est une équation du deuxième ordre ($n = 2$);
- * $y' = ky(1 - \frac{y}{K})$ est une équation du premier ordre ($n = 1$);
- * $y' = f(x)$, où f est une fonction donnée, est une équation du premier ordre. Cet exemple montre que la recherche de primitive est un exemple basique d'équation différentielle. En pratique, la recherche de primitive est un outil permettant d'écrire des solutions explicites d'équations différentielles.

1 Problème de Cauchy

En général, une équation différentielle admet plusieurs (une infinité de) solutions. Résoudre une équation différentielle, cela revient à trouver *toutes* les solutions. On s'intéresse parfois à une ou plusieurs *solutions particulières*, comme par exemple les solutions vérifiant certaines conditions supplémentaires.

Exemple. Trouver l'ensemble des solutions de l'équation $y' = x \exp(-y)$ vérifiant $y'(0) = 0$.

Définition XVIII.1. Un *problème de Cauchy* est la donnée d'une équation différentielle d'ordre n et d'une *condition initiale* de la forme

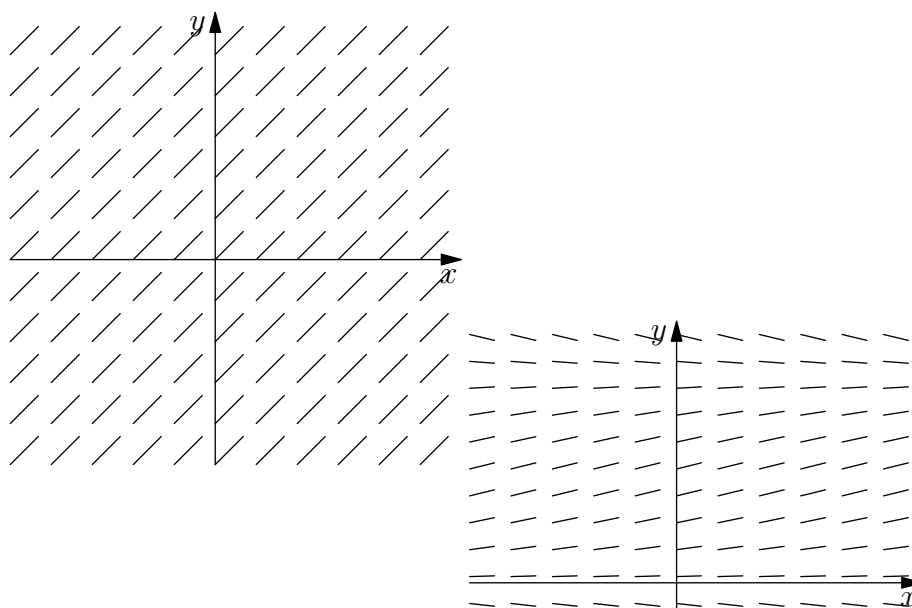
problème de
Cauchy

$$\begin{cases} y(x_0) = y_0, \\ y'(x_0) = y'_0, \\ \vdots \\ y^{(n-1)'}(x_0) = y_0^{(n-1)'}. \end{cases}$$

condition initiale

où $x_0, y_0, y'_0, \dots, y_0^{(n-1)'}$ sont des constantes.

Remarque (*). Dans les cas que nous rencontrerons, il y a une unique solution pour un problème de Cauchy donné.

FIGURE XVIII.1 – Champ des pentes pour l'équation $y' = 1$ et pour l'équation $y' = y(1 - y)$

2 Équations du premier ordre

forme normale

Une équation du premier ordre est sous *forme normale* si elle s'écrit

$$y' = f(x, y)$$

pour une certaine fonction f .

Exemple. * L'équation $y' = xy$ est sous forme normale ;

* l'équation $yy' = x$ ne l'est pas.

2.1 Interprétation géométrique

Supposons avoir une solution y de l'équation $y' = f(x, y)$, c'est-à-dire $y'(x) = f(x, y(x))$ en tout point x de l'intervalle de définition de y . Considérons le point $(x, y(x))$ du graphe de y . En ce point, le graphe admet une tangente dont la pente est $y'(x)$, c'est-à-dire $f(x, y(x))$. En d'autres termes, la pente d'une solution de l'équation différentielle ne dépend que du point du plan en lequel cette solution passe. Nous pouvons alors, pour chaque point du plan, dessiner la tangente que devrait avoir une solution passant par ce point. Ceci donne lieu à un *champ des pentes* dont la figure 2.1 donne deux exemples. De la sorte, une solution est une courbe qui reste tangente aux pentes en tous ses points (voir figure 2.1).

champ des pentes

2.2 Équations à variables séparables

variables séparables

Si une EDO peut s'écrire sous la forme $q(y)y' + p(x) = 0$ pour certaines fonctions p et q données, on dit alors que le problème est à *variables séparables*. La forme $q(y)y' + p(x) = 0$ est appelée une équation à variables séparées.

On peut résoudre une telle équation par la méthode suivante : soit P et Q des primitives de p et q respectivement ; alors la dérivée de $Q(y(x)) + P(x)$ vaut $q(y(x))y'(x) + p(x)$ c'est-à-dire 0 d'après l'équation. Dès lors il existe une constante réelle C telle que

$$Q(y(x)) + P(x) = C.$$

Il suffit ensuite de résoudre l'équation pour trouver $y(x)$ en fonction de x , ce qui doit se faire au cas par cas.

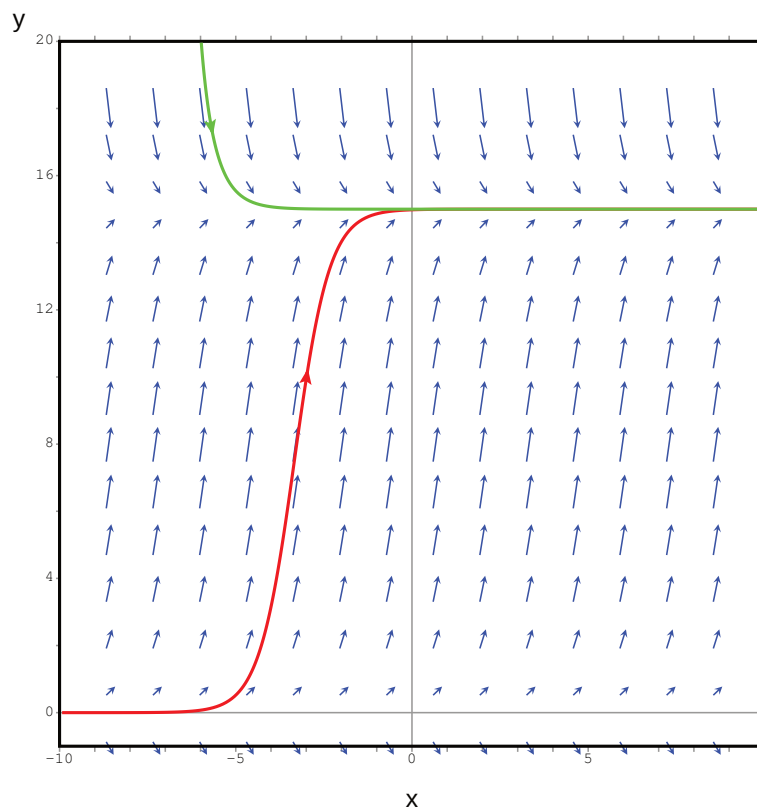


FIGURE XVIII2 – Graphe réalisé avec le logiciel Maxima

```
plotdf(k*y*(1-y/K), [parameters, "k=2, K=15"], [sliders, "k=1:4, K=10:100"], [y, -1, 20])$
```

Remarque (*). En pratique, pour résoudre $q(y)y' + p = 0$, nous écrirons souvent

$$\int q(y(x))y'(x)dx = - \int p(x)dx$$

c'est-à-dire, en posant $z = y(x)$

$$\int q(z)dz = - \int p(x)dx.$$

Pour éviter d'introduire une variable z qui joue le rôle de y , il sera fréquent d'écrire simplement $y = y(x)$ à la place, c'est-à-dire :

$$\int q(y)dy = - \int p(x)dx.$$

tout en se rappelant, après avoir trouvé les primitives, que y dépend en réalité de x .

Remarque (*). Rappelons nous de la notation $\frac{dy}{dx}$ pour désigner y' , et remarquons que l'équation

$$q(y)y' = -p(x)$$

se résout en écrivant

$$\int q(y)dy = - \int p(x)dx.$$

Ceci donne un moyen mnémotechnique agréable : il suffit d'écrire $y' = \frac{dy}{dx}$, de multiplier par dx et d'intégrer.

Exemple. Attention cependant que la remarque ci-dessus ne fonctionne que dans le cas où l'équation est sous forme séparée, par exemple on ne peut *pas* résoudre $y' + yx = 0$ en écrivant

$$\text{Ceci est faux} \Rightarrow \int dy + \int yx dx = 0. \quad (2.1)$$

Cela n'a pas beaucoup de sens d'intégrer yx par rapport à x , car cela voudrait dire intégrer $y(x)x$ sans connaître y . Ceci n'est pas possible.

Pour s'en sortir ici, il faut écrire $\frac{y'}{y} + x = 0$ et donc calculer :

$$\int \frac{dy}{y} + \int x dx = 0$$

pour obtenir une bonne solution. On obtient alors :

$$\ln|y| + \frac{x^2}{2} = C$$

d'où

$$y(x) = \tilde{C} \exp\left(-\frac{x^2}{2}\right).$$

où C est une constante, et $\tilde{C}$ est une autre constante.

Remarque (*). Dans l'exemple précédent, le lecteur attentif remarquera que les valeurs absolues ont disparu. Notons simplement que si on a $|y(x)| = \tilde{C} \exp\left(-\frac{x^2}{2}\right)$, c'est que $y(x) = \pm \tilde{C} \exp\left(-\frac{x^2}{2}\right)$. Il suffit donc d'ajuster la constante $\tilde{C}$ dans notre expression pour tenir compte de ce signe $\pm$.

Exemple. Considérons l'équation $y' = ky\left(1 - \frac{y}{K}\right)$. Elle est à variables séparables, et nous pouvons écrire :

$$\frac{y'}{y\left(1 - \frac{y}{K}\right)} = k$$

ce que nous résolvons :

$$\int \frac{dy}{y\left(1 - \frac{y}{K}\right)} = kx + C.$$

Le membre de gauche se résout en décomposant en fractions simples :

$$\int \left(\frac{1}{y} + \frac{1}{K-y} \right) dy = kx + C$$

dont on déduit

$$\ln|y| + \ln|K-y| = kx + C$$

d'où $y(K-y) = \tilde{C} \exp(kx)$. (On pourrait maintenant résoudre cette équation du second degré en y pour trouver y explicitement.)

2.3 Équations du type homogène

Une *équation différentielle du type homogène* est une équation qui s'écrit sous forme normale $y' = f(x, y)$ avec comme condition supplémentaire que $f(ax, ay) = f(x, y)$ pour tout $a \in \mathbb{R}$.

Une telle équation se résout en changeant de fonction inconnue : $y(x) = xu(x)$. u est la nouvelle fonction inconnue.

Exemple. L'équation suivante est du type homogène :

$$y' = \frac{x+2y}{2x-y}$$

Posons $y = xu$, d'où $y' = u + xu'$, et l'équation devient :

$$u + xu' = \frac{x+2xu}{2x-xu} = \frac{1+2u}{2-u}$$

ce qu'on transforme en

$$xu' = \frac{1+2u}{2-u} - u = \frac{u^2+1}{2-u}$$

équation
différentielle du
type homogène

et par conséquent (ceci est une équation à variables séparables)

$$\frac{2-u}{u^2+1}u' = \frac{1}{x}$$

d'où,

$$\int \frac{2-u}{u^2+1} du = \int \frac{dx}{x}.$$

À ce point, on dit que l'équation différentielle est résolue à *quadrature* près. C'est-à-dire à une intégrale près. Dans le cas présent, on peut trouver des primitives explicites :

quadrature

$$2 \arctan u - \frac{1}{2} \ln(u^2 + 1) = \ln|x| + C.$$

À ce point-ci, on dit que nous avons une solution sous *forme implicite* : nous n'avons pas une formule pour $u(x)$ (ou $y(x)$) mais nous savons qu'il « suffit » de résoudre une équation classique (non-différentielle) et isoler u pour obtenir une telle formule. Dans le cas présent, nous ne tenterons pas d'obtenir une solution explicite.

forme implicite

2.4 Équations linéaires

Une équation différentielle est dite *linéaire* si elle est écrite sous la forme :

linéaire

$$q(x)y' + p(x)y = f(x) \quad (\text{EL})$$

où p, q, f sont des fonctions données.

- * Les fonctions p et q sont appelés les « coefficients ». Ces coefficients sont parfois constants (ne dépendent pas de x) mais en général ce sont des fonctions de x .
- * La fonction f est appelée le « second membre » de l'équation. Si $f(x) = 0$ pour tout x , on dit de l'équation qu'elle est *homogène*.

homogène

Pour résoudre une équation linéaire, nous procédons en quatre étapes faciles :

Étape L1 – Résolution de l'équation linéaire homogène associée (ÉLHA) On appelle « équation linéaire homogène associée » (ÉLHA) l'équation suivante :

$$q(x)y' + p(x)y = 0 \quad (\text{ÉLHA})$$

C'est une équation différentielle, mais dont les solutions nous seront utiles pour exprimer les solutions de l'équation linéaire (EL) de départ.

Cette équation est à variables séparables, dès lors nous pouvons en écrire explicitement les solutions en vertu de la méthode décrite précédemment :

$$y_{\text{SGEH}}(x) = C \exp(-\Phi(x)) \quad \text{où } \Phi(x) = \int \frac{p(x)}{q(x)} dx$$

où C est une constante d'intégration. L'acronyme SGEH signifie : solution générale de l'équation homogène. « Générale » car nous avons de la sorte *toute* les solutions de l'équation (ÉLHA).

Étape L2 – Solution particulière de l'équation linéaire (SPEL)

Motivation Rappelons que notre but est de trouver *toutes* les solutions de l'équation linéaire (EL). Cependant, supposons avoir *une* solution de cette équation, notons-la y_{SPEL} . Soit y une autre solution de (EL), et soit $z = y - y_{\text{SPEL}}$. Alors nous pouvons calculer $q(x)z' + p(x)z$:

$$\begin{aligned} q(x)z' + p(x)z &= q(x)y' - q(x)y'_{\text{SPEL}} + p(x)y - p(x)y_{\text{SPEL}} \\ &= (q(x)y' + p(x)y) - (q(x)y'_{\text{SPEL}} + p(x)y_{\text{SPEL}}) \\ &= f(x) - f(x) \\ &= 0 \end{aligned}$$

En d'autres termes, z est une solution de l'(ÉLHA). Or nous connaissons ces solutions. Nous en déduisons qu'il existe C tel que

$$z(x) = C \exp(-\Phi(x))$$

ou en d'autres termes :

$$y(x) = y_{\text{SPEL}}(x) + \underbrace{C \exp(-\Phi(x))}_{y_{\text{SGEH}}(x)}$$

Ceci justifie de chercher *une* solution particulière, y_{SPEL} . Pour ce faire, nous voyons pour le moment deux méthodes.

Méthode de la génération spontanée Il arrive dans certains cas qu'une solution particulière apparaisse de manière évidente. Nul besoin dans ce cas d'aller chercher plus loin.

Exemple. L'équation $\exp x^2 y' + y = 1$ admet une solution évidente : $y(x) = 1$. En effet, $y'(x) = 0$ et donc l'équation est satisfaite.

Méthode de variation de la constante (à une variable) Dans tous les cas où la méthode précédente ne fonctionne pas, nous pouvons appliquer une méthode plus calculatoire.

Cherchons une solution particulière de la forme

$$y_{\text{SPEL}} = u(x) \exp(-\Phi(x)),$$

où u est une nouvelle fonction inconnue. Nous reconnaissons ici l'expression de y_{SGEH} où nous avons remplacé la constante par une fonction inconnue. Pour cette raison, la méthode s'appelle *méthode de la variation de la constante*.

Pour que la fonction y_{SPEL} ci-dessus soit solution de l'équation (EL), il faut que l'équation soit vérifiée :

$$q(x)y'_{\text{SPEL}} + p(x)y_{\text{SPEL}} = f(x)$$

c'est-à-dire

$$q(x)u' \exp(-\Phi(x)) - q(x)u \exp(-\Phi(x)) \frac{p}{q} + p(x)u \exp(-\Phi(x)) = f(x)$$

ou encore, puisque deux termes se simplifient :

$$q(x)u' \exp(-\Phi(x)) = f(x)$$

dont on tire

$$u = \int \frac{f(x) \exp(\Phi(x))}{q(x)} dx.$$

On en déduit une expression pour y_{SPEL} en y remplaçant la valeur de u que nous venons de trouver.

Remarque (*). Généralement, pour résoudre les exercices, il est suggéré de procéder à la substitution $y_{\text{SPEL}} = u(x) \exp(-\Phi(x))$ indiquée au début du paragraphe et de faire le calcul ci-dessus « à la main » plutôt que d'utiliser la formule pour u que nous venons d'obtenir. La raison est que cette formule est complexe à retenir et à manipuler, alors que dans les cas pratiques soumis aux étudiants il y a souvent des simplifications dès le départ.

Étape L3 – Écrire la solution générale de l'équation linéaire D'après les raisonnements précédents, la solution générale de l'équation linéaire (EL) de départ est donnée par

$$y(x) = y_{\text{SGEH}}(x) + y_{\text{SPEL}}(x).$$

Étape L4 – Déterminer les constantes d'intégration grâce aux conditions initiales. Si des conditions initiales ont été fournies dans l'exercice, voici venu le temps de les prendre en compte. Rappelons que la solution générale est donnée par

$$y(x) = y_{\text{SGEH}}(x) + y_{\text{SPEL}}(x)$$

et contient une constante d'intégration (venant de y_{SGEH}). Ceci veut dire que pour trouver la solution particulière vérifiant les conditions initiales, il faut déterminer la constante d'intégration.

Remarque (*). En général la solution particulière trouvée à l'étape L2 n'est *pas* la solution particulière qui vérifie les conditions initiales.

méthode de la variation de la constante

3 Équations linéaires du second ordre à coefficients constants

Une *équation linéaire du second ordre à coefficients constants* est une équation de la forme :

$$ay'' + by' + cy = f(x)$$

équation linéaire
du second ordre
à coefficients
constants

pour certaines constantes a, b, c et une fonction donnée f . La résolution d'une telle équation se fait en quatre étapes notées L1, L2, L3 et L4, comme dans le cas des équations linéaires du premier ordre.

3.1 Méthode de résolution

Étape L1 Il faut résoudre l'équation linéaire homogène associée (ÉLHA) : $ay'' + by' + cy = 0$.

Remarque (*). Dans le cas des équations du premier ordre, cela revenait à résoudre une équation à variables séparables. Dans le cas présent, la méthode ne s'applique pas.

À une telle équation on associe son *polynôme caractéristique* : $a\lambda^2 + b\lambda + c$. Selon le nombre de racines de ce polynôme, la solution prendra une forme différente.

polynôme
caractéristique

Résultat XVIII.2. Les solutions de l'ÉLHA sont :

Cas 1 Si $b^2 - 4ac > 0$, on note $\lambda_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$ et $\lambda_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$ les deux racines du polynôme caractéristique. La solution générale de l'équation homogène (SGEH) est alors :

$$y_{\text{SGEH}}(x) = C_1 \exp(\lambda_1 x) + C_2 \exp(\lambda_2 x)$$

où C_1 et C_2 sont des constantes d'intégration.

Cas 2 Si $b^2 - 4ac = 0$, on note $\lambda = \frac{-b}{2a}$. La solution générale de l'équation homogène (SGEH) est alors :

$$y_{\text{SGEH}}(x) = (C_1 + C_2 x) \exp(\lambda x)$$

où C_1 et C_2 sont des constantes d'intégration.

Cas 3 Si $b^2 - 4ac < 0$, on note $\rho = \frac{-b}{2a}$ et $\omega = \frac{1}{2a} \sqrt{4ac - b^2}$. La solution générale de l'équation homogène (SGEH) est alors :

$$y_{\text{SGEH}}(x) = (C_1 \cos(\omega x) + C_2 \sin(\omega x)) \exp(\rho x)$$

où C_1 et C_2 sont des constantes d'intégration.

Nous ne démontrons pas ce théorème, mais il est facile (quoiqu'un peu long) de vérifier que les solutions suggérées sont bien des solutions. Vérifier qu'il n'y en a pas d'autres est un peu plus astucieux.

Remarque (*). Dans le théorème ci-dessus, la solution y_{SGEH} est toujours de la forme :

$$y_{\text{SGEH}}(x) = C_1 g_1(x) + C_2 g_2(x)$$

où g_1 et g_2 sont données par le théorème et dépendent du cas dans lequel on se trouve. C_1 et C_2 sont les constantes d'intégration.

Étape L2 Comme dans le cas du premier ordre, il faut trouver une solution particulière y_{SPEL} de l'équation linéaire (EL) de départ. Si la méthode de la génération spontanée s'applique, inutile de chercher plus loin ! Dans le cas contraire, deux solutions s'offrent à nous : généraliser la méthode de la variation de la constante, ou faire de la divination avancée.

Variation des constantes Cette méthode a l'avantage de s'appliquer dans tous les cas, mais le désavantage d'être souvent très calculatoire. Nous cherchons une solution particulière de la forme :

$$y_{\text{SPEL}}(x) = u(x)g_1(x) + v(x)g_2(x)$$

où g_1 et g_2 sont les fonctions données par le théorème et la remarque ci-dessus, et u, v sont de nouvelles fonctions inconnues. Notons cela de manière compacte $y_{\text{SPEL}} = ug_1 + vg_2$. Supposons en outre que $u'g_1 + v'g_2 = 0$.

Alors

$$y'_{\text{SPEL}} = u'g_1 + ug'_1 + v'g_2 + vg'_2 = ug'_1 + vg'_2$$

et donc

$$y''_{\text{SPEL}} = (ug'_1 + vg'_2)' = u'g'_1 + v'g'_2 + ug_1'' + vg_2''$$

ce qui permet de calculer :

$$ay''_{\text{SPEL}} + by'_{\text{SPEL}} + cy_{\text{SPEL}} = a(u'g'_1 + v'g'_2 + ug_1'' + vg_2'') + b(ug'_1 + vg'_2) + c(ug_1 + vg_2)$$

Puisque g_1 et g_2 sont des solutions de l'ÉLHA, elles vérifient $ag'' + bg' + cg = 0$, dès lors nous avons la simplification suivante :

$$ay''_{\text{SPEL}} + by'_{\text{SPEL}} + cy_{\text{SPEL}} = a(u'g'_1 + v'g'_2)$$

et ceci doit être égal à f . En d'autres termes, si nous trouvons u, v telles que :

$$\begin{cases} u'g_1(x) + v'g_2(x) = 0 \\ u'g'_1(x) + v'g'_2(x) = \frac{f(x)}{a} \end{cases}$$

alors y_{SPEL} définie plus haut sera bien une solution particulière de l'équation linéaire de départ.

Ce système de deux équations se résout pour donner :

$$\begin{cases} u' = \frac{-f(x)g_2(x)}{a(g_1(x)g'_2(x) - g_2(x)g'_1(x))} \\ v' = \frac{f(x)g_1(x)}{a(g_1(x)g'_2(x) - g_2(x)g'_1(x))} \end{cases}$$

Dès lors, nous avons une solution y_{SPEL} à quadrature près (il faut encore intégrer pour obtenir u et v).

Divination avancée Lorsque le second membre $f(x)$ est suffisamment simple, il est possible de deviner une solution en résolvant de petits systèmes algébriques.

Exemple. Si l'équation linéaire de départ est $y'' + y' + y = \cos x$ par exemple, on comprend rapidement qu'en remplaçant y par une combinaison de cosinus et de sinus, le membre de gauche s'évaluera à une combinaison de sinus et cosinus. L'espoir est de pouvoir choisir la combinaison qui donnera précisément $\cos x$ à l'arrivée.

Les diverses méthodes ci-dessous dépendent de la forme de f .

Polynômes Si f est un polynôme de degré n , on cherche y_{SPEL} sous la forme d'un polynôme de degré :

- * n si $c \neq 0$
- * $n + 1$ si $c = 0, b \neq 0$
- * $n + 2$ si $c = 0, b = 0, a \neq 0$.

Exponentielles Si f a la forme $f(x) = A \exp(kx)$, pour des constantes A et k , alors on essaye y_{SPEL} de la forme :

- * $\beta \exp(kx)$, si $ak^2 + bk + c \neq 0$ (c'est-à-dire k n'est pas racine du polynôme caractéristique),
- * $\beta x \exp(kx)$, si k est racine simple du polynôme $a\lambda^2 + b\lambda + c$,
- * $\beta x^2 \exp(kx)$, si k est racine double du polynôme $a\lambda^2 + b\lambda + c$.

Trigono-exponentielles Si $f(x) = (A \cos(kx) + B \sin(kx)) \exp(lx)$ pour certaines constantes A, B, k, l , alors on essaye y_{SPEL} de la forme :

- * $(\alpha \cos(kx) + \beta \sin(kx)) \exp(lx)$ si $k \neq \rho$ ou $l \neq \omega$ (ρ et ω sont des quantités données dans le théorème pour l'équation linéaire homogène ÉLHA.)
- * $x(\alpha \cos(kx) + \beta \sin(kx)) \exp(lx)$ sinon (c'est-à-dire $k = \rho$ et $l = \omega$).

Étape L3 La solution générale de l'équation linéaire (EL) de départ est donnée par

$$y(x) = y_{\text{SGEH}}(x) + y_{\text{SPEL}}(x).$$

Étape L4 Si des conditions initiales ont été fournies dans l'exercice, voici venu le temps de les prendre en compte. Rappelons que la solution générale est donnée par

$$y(x) = y_{\text{SGEH}}(x) + y_{\text{SPEL}}(x)$$

et contient deux constantes d'intégration (venant de y_{SGEH}). Ceci veut dire que pour trouver la solution particulière vérifiant les conditions initiales, il faut déterminer les constantes d'intégration.

Remarque (*). En général la solution particulière trouvée à l'étape L2 n'est *pas* la solution particulière qui vérifie les conditions initiales.

3.2 Exemples

Résolution d'une équation homogène Résolvons l'équation $y'' - 3y' + 2y = 0$, avec comme conditions initiales $x_0 = 0, y_0 = 1, y'_0 = 0$. Remarquons que cette équation linéaire est déjà homogène, donc il n'y a que les étapes L1 et L4 à considérer.

Le polynôme caractéristique est $\lambda^2 - 3\lambda + 2$ dont les racines sont 2 et 1. Ceci nous place dans le cas 1 du théorème XVIII.2. La solution générale de l'ÉLHA (qui est, dans notre cas, également notre équation de départ) est donc donnée par :

$$y(x) = C_1 \exp(2x) + C_2 \exp x.$$

Les étapes L2 et L3 sont inutiles : nous avons déjà résolu l'équation linéaire puisqu'elle était homogène. Pour faire bonne mesure, nous pouvons néanmoins signaler que la solution nulle est une solution particulière de l'équation.

L'étape L4 nous dicte de prendre en compte les conditions initiales. Dès lors il faut avoir :

$$\begin{cases} y(0) = 1 \\ y'(0) = 0 \end{cases}$$

c'est-à-dire

$$\begin{cases} C_1 + C_2 = 1 \\ 2C_1 + C_2 = 0 \end{cases}$$

dont on tire $C_1 = -1$ et $C_2 = 2$. La solution (unique) du problème de Cauchy initialement posé est donc

$$y(x) = -\exp(2x) + 2\exp x.$$

Résolution d'une équation linéaire non-homogène Nous considérons cette fois l'équation linéaire suivante :

$$y'' + y = \cos(2x).$$

Pour l'étape L1 : le polynôme caractéristique est $\lambda^2 + 1$, ce qui nous place dans le cas 3 du théorème XVIII.2, avec $\rho = 0$ et $\omega = 1$. Dès lors :

$$y_{\text{SGEH}}(x) = C_1 \cos(x) + C_2 \sin(x)$$

Pour l'étape L2, nous remarquons que nous sommes dans la recette trigono-exponentielle (3.1) avec $k = 2$ et $l = 0$, premier sous-cas car $\omega \neq k$. Dès lors nous essayons une solution particulière de la forme :

$$y_{\text{SPEL}} = \alpha \cos(2x) + \beta \sin(2x)$$

et nous imposons que cette fonction soit solution de l'équation linéaire de départ. Pour cela nous devons calculer :

$$\begin{aligned} y'_{\text{SPEL}} &= -2\alpha \sin(2x) + 2\beta \cos(2x) \\ y''_{\text{SPEL}} &= -4\alpha \cos(2x) - 4\beta \sin(2x) \end{aligned}$$

et remplacer dans l'équation :

$$y_{\text{SPEL}}''(x) + y_{\text{SPEL}}(x) = \cos(2x) \iff -3\alpha \cos(2x) - 3\beta \sin(2x) = \cos(2x).$$

Ceci est vérifié pour $\alpha = \frac{-1}{3}$ et $\beta = 0$, dès lors

$$y_{\text{SPEL}}(x) = \frac{-\cos(2x)}{3}$$

est notre solution particulière.

L'étape L3 est une étape de récapitulation : la solution générale de l'équation linéaire de départ est

$$y(x) = C_1 \cos(x) + C_2 \sin(x) - \frac{1}{3} \cos(2x)$$

L'étape L4 prend les conditions initiales en compte, mais nous n'avons donné aucune conditions initiales.

Résolution d'une autre équation linéaire Nous considérons cette fois l'équation linéaire suivante :

$$y'' - 9y = x + \exp(2x) - \sin(2x) + \exp(3x)$$

Le polynôme caractéristique est $\lambda^2 - 9$, ce qui nous place dans le cas 1 du théorème XVIII.2. Dès lors :

$$y_{\text{SGEH}}(x) = C_1 \exp(3x) + C_2 \exp(-3x)$$

Pour l'étape L2, nous remarquons que le second membre comporte plusieurs termes. Dans cette situation, nous allons analyser chaque terme, lui appliquer la recette correspondante, et faire la somme des candidats solutions particulières :

- * pour le terme x , c'est un polynôme de degré 1. La recette correspondante indique de chercher une solution sous forme d'un polynôme de degré 1, c'est-à-dire de la forme $\alpha x + \beta$;
- * le terme $\exp(2x)$ est une exponentielle, il faut chercher une solution de la forme $\gamma \exp(2x)$;
- * le terme $\sin(2x)$ est trigono-exponentiel, il faut chercher une solution de la forme $\delta \cos(2x) + \epsilon \sin(2x)$;
- * le terme $\exp(3x)$ est une exponentielle et 3 est racine simple du polynôme caractéristique, donc il faut chercher une solution de la forme $\phi x \exp(3x)$.

Pour résumer, nous cherchons une solution particulière de la forme :

$$y_{\text{SPEL}}(x) = (\alpha x + \beta) + \gamma \exp(2x) + (\delta \cos(2x) + \epsilon \sin(2x)) + \phi x \exp(3x)$$

et nous imposons que cette fonction soit solution de l'équation linéaire de départ. Pour cela nous calculons

$$y_{\text{SPEL}}''(x) = 4\gamma \exp(2x) - 4(\delta \cos(2x) + \epsilon \sin(2x)) + \phi(6 + 9x) \exp(3x)$$

et ré-écrivons l'équation :

$$\begin{aligned} 4\gamma \exp(2x) - 4(\delta \cos(2x) + \epsilon \sin(2x)) + \phi(6 + 9x) \exp(3x) \\ - 9((\alpha x + \beta) + \gamma \exp(2x) + (\delta \cos(2x) + \epsilon \sin(2x)) + \phi x \exp(3x)) \\ = x + \exp(2x) - \sin(2x) + \exp(3x) \end{aligned}$$

ce qui se simplifie en :

$$-9\alpha x - 9\beta - 5\gamma \exp(2x) - 13\delta \cos(2x) - 13\epsilon \sin(2x) + 6\phi \exp(3x) = x + \exp(2x) - \sin(2x) + \exp(3x)$$

et est donc vérifié dès que

$$\begin{array}{lll} \alpha = \frac{-1}{9} & \beta = 0 & \gamma = \frac{-1}{5} \\ \delta = 0 & \epsilon = \frac{1}{13} & \phi = \frac{1}{6}. \end{array}$$

Dès lors

$$y_{\text{SPEL}}(x) = -\frac{1}{9}x + \frac{-1}{5}\exp(2x) + \frac{1}{13}\sin(2x) + \frac{1}{6}x\exp(3x)$$

est notre solution particulière.

L'étape L3 est une étape de récapitulation : la solution générale de l'équation linéaire de départ est

$$y(x) = C_1 \exp(3x) + C_2 \exp(-3x) - \frac{1}{9}x + \frac{-1}{5}\exp(2x) + \frac{1}{13}\sin(2x) + \frac{1}{6}x\exp(3x)$$

L'étape L4 prend les conditions initiales en compte, mais nous n'avons donné aucune conditions initiales. Elles serviraient, sinon, à déterminer les constantes C_1 et C_2 .

Chapitre XIX

Nombres complexes

1 Motivations

Rappelons que le plan muni de coordonnées cartésiennes est en bijection avec $\mathbb{R}^2$. C'est-à-dire qu'à chaque point du plan on associe exactement une paire de coordonnées cartésiennes, et réciproquement à chaque paire de coordonnées (un élément de $\mathbb{R}^2$) on associe un point du plan.

Rappelons encore qu'à un point de coordonnées (x, y) on peut également associer les coordonnées polaires (ρ, θ) où $\rho = \sqrt{x^2 + y^2}$ représente la distance par rapport à l'origine, et θ représente l'angle (orienté) entre le demi-axe des abscisses positives et la demi-droite issue de l'origine passant par (x, y) . Le lien entre les deux types de coordonnées est donné par

$$\begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \end{cases}$$

Considérons l'homothétie de centre $(0, 0)$ et de rapport k . Elle se décrit simplement en coordonnées cartésiennes par $(x, y) \mapsto (kx, ky)$. En coordonnées polaires, elle s'écrit $(\rho, \theta) \mapsto (k\rho, \theta)$ lorsque $k > 0$, et elle s'écrit $(\rho, \theta) \mapsto (-k\rho, \theta + \pi)$ lorsque $k < 0$.

Si on considère maintenant la rotation de centre $(0, 0)$ et d'angle φ , elle s'écrit naturellement en coordonnées polaires : $(\rho, \theta) \mapsto (\rho, \theta + \varphi)$. Pour obtenir sa version cartésienne $(x, y) \mapsto (x', y')$, il suffit d'écrire les coordonnées cartésiennes du point d'arrivée et de développer grâce aux formules trigonométriques :

$$\begin{aligned} (x', y') &= (\rho \cos(\theta + \varphi), \rho \sin(\theta + \varphi)) \\ &= (\rho \cos(\theta) \cos(\varphi) - \rho \sin(\theta) \sin(\varphi), \rho \cos(\theta) \sin(\varphi) + \rho \sin(\theta) \cos(\varphi)) \\ &= (x \cos(\varphi) - y \sin(\varphi), x \sin(\varphi) + y \cos(\varphi)) \end{aligned}$$

L'application $(x, y) \mapsto (x \cos(\varphi) - y \sin(\varphi), x \sin(\varphi) + y \cos(\varphi))$ représente donc la rotation d'angle φ en coordonnées cartésiennes.

Définissons un « produit complexe » sur les points de $\mathbb{R}^2$: $(a, b)(x, y) = (ax - by, ay + bx)$. Cette formule est particulièrement intéressante car lorsque $(a, b) = (\cos \varphi, \sin \varphi)$, la multiplication par (a, b) revient à faire la rotation d'angle φ ; et lorsque $(a, b) = (k, 0)$, cela revient à l'homothétie de rapport k .

Afin d'alléger la notation, nous allons noter $a + bi$ au lieu de (a, b) . En particulier le couple $(a, 0)$ est simplement noté $a + 0i$ ou a , tandis que le couple $(0, 1)$ est simplement noté i . Nous ré-écrivons alors la formule du produit complexe :

$$(a + ib)(x + iy) = ax - by + i(ay + bx)$$

Dorénavant, ces éléments seront appelés « nombres complexes ».

Remarque (*). Ceci laisse craindre une certaine ambiguïté : si a est un nombre réel, faut-il voir a comme un nombre réel ou comme le nombre complexe $a + 0i$?

Heureusement, que a soit vu comme un réel ou un complexe, le résultat est toujours le même. Par exemple en écrivant le produit complexe $(a + 0i)(c + 0i)$, la réponse est $ac + 0i$. En d'autres termes le produit complexe étend le produit réel, ce qui justifie la notation.

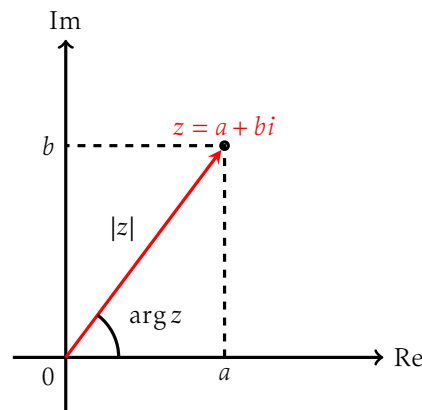


FIGURE XIX1 – Illustration de nombres complexes dans le plan de Gauß (voir plus loin pour les définitions de $|z|$ et $\arg z$)

2 Définitions

Résumons ce que nous avons introduit jusqu'à présents :

Définition XIX.1. L'ensemble des *nombres complexes* est l'ensemble noté

$$\mathbb{C} = \{a + bi \text{ t.q. } a, b \in \mathbb{R}\}.$$

Si $z = a + bi$ est un nombre complexe avec $a, b \in \mathbb{R}$, alors a est sa *partie réelle* et b sa *partie imaginaire*.

En d'autres termes, lorsque nous dirons « Soit z un nombre complexe, ... » nous supposons donc nous donner $z = a + bi$ pour certains réels a, b .

Résultat XIX.2. Deux nombres complexes $a + bi$ et $c + di$ sont égaux si et seulement si $a = c$ et $b = d$.

Démonstration. Ceci découle du fait que $a + ib$ n'est autre que (a, b) . Dès lors l'égalité $a + bi = c + di$ se ré-écrit $(a, b) = (c, d)$ d'où le résultat. $\square$

Définition XIX.3. Un nombre complexe dont la partie réelle est nulle est appelé un nombre *imaginaire pur*.

Remarque (*). Un nombre complexe dont la partie imaginaire est nulle est constitué d'un seul nombre réel. Le produit de deux tels réels est encore un nombre réel. Ceci permet d'écrire $\mathbb{R} \subset \mathbb{C}$: l'ensemble des nombres réels est vu comme le sous-ensemble de $\mathbb{C}$ correspondant aux nombres complexes dont la partie imaginaire est nulle.

2.1 Plan de Gauß

Nous avons introduit les nombres complexes comme des points du plan, avec une multiplication complexe héritée des formules pour les homothéties et rotations. Dans cette vision, on appelle ce plan le *plan de Gauß*. La plupart des opérations sur les nombres complexes auront leur interprétation géométrique dans ce plan.

3 Opérations

3.1 Somme

Définition XIX.4. La *somme* de deux nombres complexes $z = a + bi$ et $z' = c + di$ est le nombre complexe noté $z + z' = (a + c) + (b + d)i$. En d'autres termes : la somme de deux nombres complexes s'obtient en additionnant les parties réelles et les parties imaginaires.

Notons que la définition ci-dessus est cohérente avec la définition des couples de réels : $(a, b) + (c, d) = (a + c, b + d)$.

Résultat XIX.5. L'addition des nombres complexes a les propriétés suivantes :

- * $(z + z') + z'' = z + (z' + z'')$ (associativité);
- * $z + z' = z' + z$ (commutativité);
- * $z + 0 = z$ (existence d'un neutre : 0);
- * $z + (-z) = 0$ (existence d'un opposé);

où z, z', z'' sont des nombres complexes quelconques, et $-z$ est le complexe obtenu à partir de $z = a + ib$ comme suit : $-z = -a - ib$.

Notons que la propriété d'associativité permet de ne pas s'inquiéter de l'ordre dans lequel on réalise les opérations, tout comme pour les nombres réels : $(2 + 3) + 4 = 5 + 4 = 9$ ou $2 + (3 + 4) = 2 + 7 = 9$.

Démonstration. Prouvons la commutativité : soit $z = a + ib$ et $z' = c + id$. Alors

$$z + z' = (a + c) + i(b + d) \quad \text{et} \quad z' + z = (c + a) + i(d + b)$$

qui sont bien identiques, puisque l'addition des réels est elle-même commutative.

Les autres propriétés peuvent se prouver similairement. □

Exemple. * $(1 + i) + (2 + i) = 3 + 2i$

$$* (-1 + i) + (3 - i) = 2$$

$$* -1 + 2i + 1 + 2i = 4i$$

3.2 Produit

Définition XIX.6. Le produit sur $\mathbb{C}$ est défini par :

$$(a + ib)(x + iy) = ax - by + i(ay + bx).$$

En particulier, $i^2 = (0 + i)(0 + i) = (0 - 1) + 0i = -1$. Ceci est évidemment une relation fondamentale : il s'agit d'un nombre dont le carré est strictement négatif, ce qui n'a aucun sens dans les nombres réels.

Exemple. * $12(2 + i) = 24 + 12i$

$$* (2 - i)(4 + i) = 9 - 2i$$

$$* (1 + i)(1 + i) = 2i$$

$$* (1 + i)(1 - i) = 2$$

Résultat XIX.7. Le produit des nombres complexes a les propriétés suivantes :

- * $(zz')z'' = z(z'z'')$ (associativité);
- * $zz' = z'z$ (commutativité);
- * $1z = z$ (existence d'un neutre : 0);
- * $zz^{-1} = 1$ (existence d'un inverse);
- * $z(z' + z'') = zz' + zz''$ (distributivité par rapport à l'addition).

où z, z', z'' sont des nombres complexes quelconques, et z^{-1} est le complexe obtenu à partir de $z = a + ib$ comme suit : $z^{-1} = \frac{a - ib}{a^2 + b^2}$.

Démonstration. Faisons, pour l'exemple, la commutativité : notons $z = a + ib$ et $z' = x + iy$ et calculons

$$(a + ib)(x + iy) = ax - by + i(ay + bx)$$

tandis que

$$(x + iy)(a + ib) = xa - yb + i(ya + xb).$$

Les membres de droite sont égaux (car la multiplication réelle est elle-même commutative) dès lors nous avons le résultat.

Les autres propriétés se prouvent similairement. Pour l'existence de l'inverse, il suffit de vérifier que la formule donnée pour z^{-1} donne le résultat escompté. Le lecteur est également invité à consulter la section 3.6 à ce sujet. □

Remarque (*). En particulier, vue la propriété de distributivité, on peut « oublier » la formule du produit complexe et se contenter de distribuer et d'utiliser la règle $i^2 = -1$:

$$(a + ib)(x + iy) = ax + ai y + ibx + ibiy = ax + i(ay + bx) + byi^2 = ax - by + i(ay + bx).$$

3.3 Parties réelles et imaginaires

Définition XIX.8. Si $z = a + ib$, on définit sa *partie réelle* et sa *partie imaginaire* respectivement par :

$$\operatorname{Re} z := a$$

$$\operatorname{Im} z := b.$$

3.4 Conjugaison

Les notions de somme et produit de nombres complexes étendent des opérations similaires sur les nombres réels. L'opération de conjugaison est quant à elle complètement nouvelle :

Définition XIX.9. Si $z = a + ib$ est un nombre complexe, on définit son *conjugué* par

$$\bar{z} = a - ib.$$

Exemple. * $\overline{2+i} = 2-i$

$$* \overline{2-i} = 2+i$$

$$* \overline{3+2i(1-5i)} = \overline{3+2i-10i^2} = \overline{13+2i} = 13-2i$$

Résultat XIX.10. La conjugaison complexe vérifie :

$$* \overline{z + z'} = \bar{z} + \bar{z'}$$

$$* z = \overline{\bar{z}} \text{ (involutivité)}$$

$$* \overline{zz'} = \bar{z}\bar{z'}$$

$$* \overline{z^2} = \bar{z}^2; \overline{z^3} = \bar{z}^3; \text{ etc.}$$

Démonstration. Notons $z = a + ib$ et $z' = x + iy$ et prouvons la première propriété :

$$\overline{z + z'} = \overline{(a + ib) + (x + iy)} = \overline{a + x + i(b + y)} = a + x - i(b + y) = a - ib + x - iy = \bar{z} + \bar{z'}.$$

Pour la deuxième de ces propriétés : si $z = a + ib$, alors $\bar{z} = a - ib$, et donc $\overline{\bar{z}} = a - i(-b) = a + ib$.

Pour la troisième, nous avons d'une part :

$$\overline{(a + ib)(x + iy)} = \overline{ax - by + i(ay + bx)} = ax - by - i(ay + bx)$$

et d'autre part :

$$\overline{a + ib} \overline{x + iy} = (a - ib)(x - iy) = ax - (-b)(-y) + i(a(-y) + (-b)x) = ax - by - i(ay + bx)$$

ce qui montre l'égalité.

La dernière de ces propriétés est une application de la précédente en prenant $z' = z$, puis $z' = z^2$, etc. $\square$

3.5 Module

La conjugaison complexe a une propriété que nous n'avons pas encore mentionnée : si $z = a + ib$, alors

$$z\bar{z} = (a + ib)(a - ib) = a^2 + b^2.$$

Le produit d'un nombre complexe avec son conjugué est donc un nombre réel, et ce réel est nul si et seulement si z est le nombre complexe nul (c'est-à-dire $a = b = 0$).

Définition XIX.11. Le *module* de $z = a + bi$ est le nombre réel noté $|z| := \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$.

En conséquence de quoi, le module d'un nombre complexe est toujours positif. Géométriquement, il s'agit précisément de la norme du vecteur (a, b) , c'est-à-dire la distance entre l'origine et le nombre complexe z vu comme point du plan (voir figure XIX1).

Résultat XIX.12. *Les propriétés suivantes sont vérifiées pour tous $z, z' \in \mathbb{C}$:*

- * $|zz'| = |z||z'|$
- * $|z + z'| \leq |z| + |z'|$
- * $|z| \geq 0$
- * $|z| = 0 \iff z = 0$

Démonstration. La première propriété se vérifie par un calcul (à ne pas confondre avec l'inégalité de Cauchy-Schwarz!) : notons $z = a + ib$ et $z' = c + id$

$$zz' = ac - bd + i(ad + bc)$$

dès lors

$$|zz'|^2 = (ac - bd)^2 + (ad + bc)^2 = a^2c^2 - 2abcd + b^2d^2 + a^2d^2 + b^2c^2 + 2abcd$$

et par ailleurs

$$|z|^2 = a^2 + b^2 \qquad |z'|^2 = c^2 + d^2 \qquad |z||z'|^2 = a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2$$

ce qui est bien la même chose.

La deuxième propriété est l'inégalité triangulaire (puisque le module, géométriquement, est la norme du vecteur correspondant!).

Les deux dernières propriétés ont déjà été justifiées précédemment. □

Remarque (*). En particulier, si z est réel (c-à-d. sa partie imaginaire est nulle), le module de z est simplement la valeur absolue. Ceci explique pourquoi la notation $|\cdot|$ est la même pour la valeur absolue et le module.

3.6 Inverse

Nous avons déjà vu que si $z = a + ib$, alors $\frac{a-ib}{a^2+b^2}$ est son inverse. Cet inverse est généralement noté z^{-1} ou $\frac{1}{z}$. La formule est vraie, comme nous pouvons le vérifier en calculant le produit, mais elle semble compliquée.

Cette formule découle en fait de celle définissant le module : puisque

$$z\bar{z} = a^2 + b^2$$

nous en déduisons

$$z \frac{\bar{z}}{a^2 + b^2} = 1$$

ce qui justifie la formule pour l'inverse :

$$z^{-1} = \frac{1}{z} = \frac{\bar{z}}{a^2 + b^2}$$

qui n'est autre que ce que nous avons donné. Nous pourrions écrire également :

$$z^{-1} = \frac{1}{z} = \frac{\bar{z}}{z\bar{z}}$$

ce qui est très simple à mémoriser.

3.7 Exponentiation

Nous avons les propriétés suivantes :

$$\begin{aligned}(a + ib)^2 &= a^2 - b^2 + 2iab \\ (a + bi)^3 &= a^3 - 3ab^2 + i(3ba^2 - b^3) \\ &\vdots\end{aligned}$$

Les formules, même si elles deviennent rapidement complexes à retenir, découlent seulement de la distributivité.

3.8 Racines carrées, cubiques, ...

Si $r \in \mathbb{R}^+$, $\sqrt{r}$ est l'unique nombre réel positif dont le carré vaut r . Dans $\mathbb{C}$, il n'existe pas de notion de « positif » ou « négatif », et la notation $\sqrt{z}$ n'est en général pas définie pour $z \in \mathbb{C}$.

Cependant, pour un $z \in \mathbb{C}$ donné, nous pouvons parler des racines carrées (ou cubiques, etc.) de z . Un complexe $a + ib$ sera une racine carrée de $z = x + iy$ si $(a + ib)^2 = x + iy$. En d'autres termes, il faut et suffit que $a^2 - b^2 = x$ et $2ab = y$.

De même pour les racines cubiques, et de manière générale pour les racines n^{e} .

Cette méthode est cependant longue, et nous verrons une méthode plus géométrique dans le plan de Gauß en voyant la forme polaire en section 4.

3.9 Argument

Nous avons déjà vu que le module d'un nombre complexe z est la distance entre le point du plan représenté par z et l'origine (le nombre complexe nul). Ceci appelle une autre définition :

Définition XIX.13. L'argument d'un nombre complexe z est l'angle polaire de z vu comme un point du plan.

Les coordonnées polaires sont mentionnées en §2.3, page 54 (Chapitre ??).

On note généralement $\arg z$ cet angle, mais nous savons que la mesure de cet angle n'est défini qu'à un multiple entier de 2π près ! L'argument du nombre complexe nul n'est pas défini. La mesure de cet angle, de la même manière que pour les coordonnées polaires, est choisie généralement entre 0 et 2π : c'est sa *détermination principale*, cependant il arrivera souvent d'utiliser des angles entre $-\pi$ et π .

Exemple. La détermination principale des arguments de certains nombres complexes :

- * $\arg i = \frac{\pi}{2}$;
- * $\arg r = 0$ si r est un réel positif;
- * $\arg r = \pi$ si r est un réel strictement négatif.
- * $\arg(1 + i) = \frac{\pi}{4}$;

Exemple. Le nombre complexe $z = \cos \theta + i \sin \theta$ a pour argument θ .

Résultat XIX.14. Pour des complexes z et z' nous avons (à un multiple entier de 2π près) :

$$\arg(zz') = \arg z + \arg z' \quad (\text{cas particulier : } \arg(-z) = \pi + \arg z)$$

De façon générale, l'argument de z se détermine via les formules

$$\frac{a}{|z|} = \cos(\arg(z)) \quad \frac{b}{|z|} = \sin(\arg(z))$$

ou encore par la formule

$$\frac{b}{a} = \tan(\arg(z)) \quad \text{en vérifiant le quadrant.}$$

La vérification du quadrant vient de ce que la tangente ne détermine l'angle qu'à π près, mais que le signe de a et b permettent de lever cette indétermination.

4 Forme polaire ou trigonométrie

Dans le plan de Gauß, un complexe z représente un point. Ce point admet des coordonnées cartésiennes : la partie réelle et la partie imaginaire de z . Il admet également des coordonnées polaires : le module et l'argument de z . Le module est généralement noté ρ et l'argument est souvent noté θ ; ils permettent de déterminer univoquement ce complexe grâce à la formule

$$z = a + bi = \rho(\cos(\theta) + i \sin(\theta))$$

où $\rho = |z|$ et $\theta = \arg z$, et $a, b \in \mathbb{R}$. La notation $a + ib$ est généralement appelée *forme normale* ou *forme cartésienne*, et la seconde est la *forme polaire* ou *forme trigonométrique*, ou encore (voir ci-dessous) *forme exponentielle*.

Une notation courante et très commode est d'écrire z sous la forme d'une exponentielle complexe, $z = \rho \exp(i\theta)$, où par définition

$$\exp(i\theta) := \cos(\theta) + i \sin(\theta).$$

Cette notation est appelée *formule d'Euler*.

Remarque. Avec cette définition, la notation $z = \rho \exp(i\theta)$ garde un sens si ρ est négatif, mais alors $|z| = -\rho$ et $\arg z = \theta + \pi$

D'après ce que nous avons vu dans l'introduction (le produit complexe correspond à des rotations et dilatations), la forme polaire d'un complexe est pratique pour calculer des produits et des puissances de nombres complexes :

Résultat XIX.15. Si $z = \rho \exp(i\theta)$ et $z' = \rho' \exp(i\theta')$ sont deux nombres complexes, et n un entier, on a :

1. $\bar{z} = \rho \exp(-i\theta)$;
2. $\frac{1}{\rho \exp(i\theta)} = \frac{1}{\rho} \exp(-i\theta)$;
3. $(\rho \exp(i\theta))(\rho' \exp(i\theta')) = \rho\rho' \exp(i(\theta + \theta'))$;
4. $(\rho \exp(i\theta))^n = \rho^n \exp(in\theta)$.

La dernière formule prend parfois le nom de « De Moivre ». On comprend l'intérêt de la notation exponentielle, puisque les opérations décrites deviennent de simples applications des règles sur les exposants.

Démonstration. Vérifions la formule du produit (troisième propriété) :

$$\begin{aligned} (\rho \exp(i\theta))(\rho' \exp(i\theta')) &= \rho(\cos(\theta) + i \sin(\theta))\rho'(\cos(\theta') + i \sin(\theta')) \\ &= \rho\rho'(\cos(\theta)\cos(\theta') - \sin(\theta)\sin(\theta') + i(\cos(\theta)\sin(\theta') + \sin(\theta)\cos(\theta'))) \\ &= \rho\rho'(\cos(\theta + \theta') + i \sin(\theta + \theta')) \\ &= \rho\rho' \exp(i(\theta + \theta')). \end{aligned}$$

Montrons également la formule de De Moivre. À cette fin, nous utilisons un raisonnement appelé *réurrence* ou *induction* : avant tout, il est évident que cette formule est vraie pour $n = 1$. Supposons maintenant la formule vraie pour $n = k$, où k est un entier fixé. Montrons que sous cette hypothèse (appelée « hypothèse de réurrence »), la formule reste vraie pour $n = k + 1$:

$$\begin{aligned} (\rho \exp(i\theta))^{k+1} &= (\rho \exp(i\theta))(\rho \exp(i\theta))^k \\ &= (\rho \exp(i\theta))(\rho^k \exp(ik\theta)) \\ &= (\rho\rho^k \exp(i(\theta + k\theta))) \\ &= \rho^{k+1} \exp(i(k+1)\theta) \end{aligned}$$

ce que nous voulions démontrer. Le fait que « si la formule est vraie pour $n = k$ alors elle est vraie pour $n = k + 1$ » est appelé « étape d'induction ». Dès lors, puisque la formule est vraie pour $n = 1$ (nous l'avons remarqué plus tôt), elle est forcément vraie pour $n = 1 + 1 = 2$ d'après l'étape d'induction. En

forme normale

forme
cartésienne

forme polaire

forme
trigonométriqueforme
exponentielle

formule d'Euler

réurrence

induction

conséquence de quoi elle doit être vraie pour $n = 2 + 1 = 3$, toujours pas l'étape d'induction. Nous en concluons *par récurrence* que cette formule est vraie pour tout entier $n \geq 1$.

Pour montrer que la formule est vraie pour *tous* les entiers, il suffirait de vérifier une nouvelle étape d'induction : « si la formule est vraie pour $n = k$, alors elle est vraie pour $n = k - 1$ ». Ceci permettrait –mais nous ne le faisons pas– de procéder à une induction « dans l'autre sens », permettant de montrer les cas $n = 0, -1, -2, \dots$ successivement. $\square$

Exemple. Voyons quelques formes polaires :

$$\begin{aligned} 1 &= \exp(i0) \\ i &= \exp\left(i\frac{\pi}{2}\right) \\ -1 &= \exp(i\pi) \\ 1 + \sqrt{3}i &= 2\exp\left(i\frac{\pi}{3}\right) \end{aligned}$$

4.1 Fonctions trigonométriques

Par définition des parties réelles et imaginaires, nous avons :

$$z = \Re z + i \Im z$$

et par la définition du conjugué, nous avons :

$$\bar{z} = \Re z - i \Im z$$

Nous en déduisons, par addition et soustraction :

$$z + \bar{z} = 2\Re z \qquad z - \bar{z} = 2i \Im z$$

Les fonctions trigonométriques cosinus et sinus peuvent donc être ré-écrites de la manière suivante :

$$\cos(x) = \frac{\exp(ix) + \exp(-ix)}{2} \qquad \sin(x) = \frac{\exp(ix) - \exp(-ix)}{2i}$$

car $\exp(ix) = \cos x + i \sin x$ dans notre notation.

4.2 Exponentielle complexe

Nous avons introduit la notation $\exp(ix) = \cos x + i \sin x$ pour tout $x \in \mathbb{R}$. De manière plus générale, si $z = a + ib$ est un nombre complexe, nous pouvons définir :

$$\exp z = \exp(a + ib) = \exp(a) \exp(ib)$$

Cette nouvelle application $\exp \cdot : \mathbb{C} \rightarrow \mathbb{C} : z \mapsto \exp z$ vérifie l'identité $\exp z \exp z' = \exp(z + z')$ pour tous complexes z, z' .

5 Racine n^e d'un complexe

Rappelons que lorsque r est un nombre réel, on note $\sqrt{r}$ l'unique nombre positif dont le carré vaut r . Évidemment, $-\sqrt{r}$ est également un nombre dont le carré est r , mais la définition du symbole $\sqrt{}$ est sans ambiguïté.

Dans le cadre des nombres complexes, par contre, il n'existe pas de notion « agréable » de positivité. Dès lors, on ne parlera jamais de *la* racine d'un nombre complexe, mais *des* racines.

Soit un entier $n \geq 1$ et z un complexe. On appelle *racine n^e* de z tout nombre complexe w tel que $w^n = z$. Pour les déterminer, nous avons la propriété suivante :

racine n^e

Résultat XIX.16. Si z est donné sous forme polaire $\rho \exp(i\theta)$ ($\rho \geq 0$), alors les racines n^e de z sont les nombres $w_0, \dots, w_{n-1}$ définis par

$$w_k = \sqrt[n]{\rho} \exp\left(i \frac{\theta + 2k\pi}{n}\right) \quad \text{où } k = 0, \dots, n-1$$

En particulier, si $z \neq 0$, il y a exactement n racines n^e de z .

Démonstration. Il y a deux choses à démontrer : d'une part que w_k est bien une racine n^e de z , et d'autre part qu'il n'y en a pas d'autres.

Pour la première partie, il est simple de vérifier que $w_k^n = \rho \exp(i(\theta + 2k\pi)) = z$ puisque

$$\exp(i(\theta + 2k\pi)) = \cos(\theta + 2k\pi) + i \sin(\theta + 2k\pi) = \cos(\theta) + i \sin(\theta) = \exp(i\theta).$$

Pour la seconde partie, considérons $w = r \exp(i\varphi)$ une racine n^e de z . Alors $w^n = z$, et en particulier $r^n = |w^n| = |z| = \rho$. Ceci montre déjà que $r = \sqrt[n]{\rho}$. Par ailleurs, puisque $w^n = z$ et $\rho = r^n$, il faut également $\exp(in\varphi) = \exp(i\theta)$. C'est-à-dire :

$$\cos(n\varphi) = \cos(\theta) \quad \sin(n\varphi) = \sin(\theta)$$

Or deux angles ne peuvent avoir même sinus et même cosinus que si ils sont égaux à 2π près : dès lors,

$$n\varphi = \theta + 2k\pi \text{ pour un certain } k \in \mathbb{Z}.$$

d'où $\varphi = \frac{\theta + 2k\pi}{n}$ comme annoncé. Reste à vérifier qu'il suffit de prendre $k = 0, \dots, n-1$. Pour cela, remarquons simplement l'égalité :

$$\begin{aligned} w_{k+n} &= \sqrt[n]{\rho} \exp\left(i \frac{\theta + 2(k+n)\pi}{n}\right) \\ &= \sqrt[n]{\rho} \exp\left(i \left(\frac{\theta + 2k\pi}{n} + 2\pi\right)\right) \\ &= \sqrt[n]{\rho} \exp\left(i \frac{\theta + 2k\pi}{n}\right) = w_k \end{aligned}$$

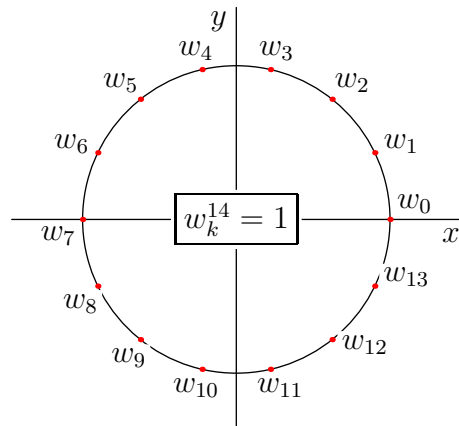
c'est-à-dire que pour toute valeur de k , w_k se trouve déjà parmi $w_0, w_1, \dots, w_{n-1}$. □

5.1 Racines de l'unité

Un cas particulier est $z = 1$. Dans ce cas, ses racines n^e sont :

$$w_k = \exp\left(i \frac{2k\pi}{n}\right) \quad \text{où } k = 0, \dots, n-1$$

et nous pouvons les représenter géométriquement :



6 Interprétations géométriques

Chaque opération s'interprète géométriquement, soit sous forme cartésienne, soit sous forme polaire.

6.1 Somme

Soit $z_0 \in \mathbb{C}$. L'application $z \mapsto z_0 + z$ correspond à une translation.

6.2 Produit

Soit $z_0 \in \mathbb{C}$ s'écrivant $z_0 = \rho \exp(i\theta)$. L'application $z \mapsto z_0 z$ correspond à une transformation résultant de :

- * une homothétie de rapport ρ , et
- * une rotation d'angle θ .

6.3 Parties réelles et imaginaires

Les parties réelles et imaginaires d'un nombre complexes correspondent aux projections sur l'axe réel (horizontal) et l'axe imaginaire (vertical) respectivement.

6.4 Conjugaison

La conjugaison $z \mapsto \bar{z}$ correspond à la symétrie orthogonale par rapport à l'axe réel (horizontal).

6.5 Module

Le module de z correspond à la distance entre l'origine et le point z considéré. En particulier, un cercle centré en z_0 de rayon r est décrit par

$$\{z \in \mathbb{C} \text{ t.q. } |z - z_0| = r\}$$

le disque ouvert correspondant est

$$\{z \in \mathbb{C} \text{ t.q. } |z - z_0| < r\}.$$

6.6 Inverse

L'application $z \mapsto z^{-1}$ est la composée d'une symétrie et d'une homothétie d'après la formule $z^{-1} = \frac{\bar{z}}{|z|^2}$.

La formule sous forme polaire est probablement plus parlante : $z^{-1} = \frac{1}{\rho} \exp(-i\theta)$ si $z = \rho \exp(i\theta)$.

6.7 Exponentiation

Les applications du type $z \mapsto z^n$ ne se décrivent pas de manière simple comme transformation du plan (il ne s'agit pas d'une rotation), mais l'image de z peut néanmoins se décrire d'après la forme trigonométrique : $(\rho \exp(i\theta))^2 = \rho^2 \exp(2i\theta)$, c'est-à-dire que l'angle est doublé et le module pris au carré. De même pour les autres valeurs de n .

Le cas particulier où $\rho = |z| = 1$ est le plus simple à imaginer : il s'agit simplement de doubler (ou tripler, etc. selon la valeur de n) l'angle du nombre complexe z considéré.

6.8 Racines carrées, cubiques, ...

Cette opération est d'une certaine manière réciproque de la précédente, sauf qu'il y a plusieurs réponses possibles. On ne peut dès pas parler de « l'application $z \mapsto \sqrt[n]{z}$ ».

6.9 Argument

L'argument de z a déjà été décrit géométriquement : c'est l'angle entre le demi-axe réel positif et la demi-droite issue de l'origine passant par z .

7 Équations

Dès lors que nous avons de « nouveaux nombres » nous pouvons résoudre des équations qui les impliquent.

7.1 Premier degré

Les équations polynomiales du premier degré offrent peu d'intérêt nouveau : considérons l'équation d'inconnue z suivante :

$$az + b = 0$$

où $a, b \in \mathbb{C}$ sont des constantes données. Alors sa solution est $z = -\frac{b}{a}$, car les règles d'addition de multiplication par les complexes sont identiques aux règles sur les réels.

7.2 Second degré

Considérons l'équation

$$az^2 + bz + c = 0$$

alors, exactement comme dans le cas réel, nous transformons cette équation en :

$$a \left(\left(z + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a^2} \right) = 0$$

avec pour différence qu'il n'est pas nécessaire de discuter le signe de $b^2 - 4ac$, puisque les solutions sont toujours données par :

$$z = \frac{-b \pm \omega}{2a}$$

où ω est une racinée carrée de $b^2 - 4ac$ (c'est-à-dire $\omega^2 = b^2 - 4ac$).

Remarque (*). Si $b^2 - 4ac = 0$, alors ceci ne produit qu'une seule racine, dite « racine double » ou « de multiplicité 2 ». Si nous admettons qu'une telle racine compte effectivement pour deux (on dira qu'on les compte « avec leur multiplicité »), alors tout polynôme de degré 2 admet toujours exactement 2 racines.

7.3 Degrés supérieurs

Si $P(z) = a_n z^n + \dots + a_1 z + a_0$ est un polynôme de degré n à coefficients complexes, alors nous avons :

Définition XIX.17. Une racine r de P est de *multiplicité* m si on peut écrire $P(z) = (z - r)^m Q(z)$, où Q est un polynôme (forcément de degré $n - m$) dont r n'est pas une racine. multiplicité

Résultat XIX.18. Tout polynôme de degré n admet exactement n racines complexes (comptées avec leur multiplicité).

8 Complexes et équations différentielles linéaires à coefficients constants

Considérons l'équation différentielle ordinaire suivante :

$$a_n y^{n'} + a_{n-1} y^{(n-1)'} + \dots + a_1 y' + a_0 y = f(x)$$

où $a_0, \dots, a_n$ sont des constantes (réelles ou complexes) et f est une fonction donnée.

Nous avons déjà étudié les cas $n = 1$ et $n = 2$ précédemment (voir section 3) dans le cas réel, mais nous allons maintenant discuter brièvement le cas général à l'aide des nombres complexes.

8.1 Fonctions réelles à valeurs complexes

Une fonction à valeurs complexes est une fonction de la forme $y : \mathbb{R} \rightarrow \mathbb{C} : x \mapsto y(x)$. Comme $\mathbb{C}$ est identifiable à $\mathbb{R}^2$, un certain nombre de définitions se transposent aisément du cas « fonction réelle à valeurs vectorielles » au cas « fonction réelle à valeurs complexes ». En particulier, la notion de dérivée : pour dériver une fonction à valeurs complexes, il suffit de dériver en considérant que i est une constante. Ceci revient effectivement à dériver « composante par composante ».

Exemple. Dérivons la fonction y définie par $y(x) = \exp(\lambda x)$, où $x \in \mathbb{R}$ et $\lambda \in \mathbb{C}$. Écrivons $\lambda = a + ib$:

$$y(x) = \exp(\lambda x) = \exp(ax)(\cos bx + i \sin bx)$$

et donc sa dérivée vaut :

$$y'(x) = a \exp(ax)(\cos bx + i \sin bx) + \exp(ax)(-b \sin bx + bi \cos bx)$$

ce que nous pouvons ré-écrire

$$y'(x) = a \exp(ax)(\cos bx + i \sin bx) + bi \exp(ax)(i \sin bx + \cos bx)$$

et donc

$$y'(x) = (a + bi) \exp(ax)(\cos bx + i \sin bx) = \lambda \exp(\lambda x) = \lambda y(x).$$

Cet exemple montre que la fonction exponentielle suit la même propriété de dérivation que l'exponentielle réelle.

D'autres fonctions peuvent être considérées, mais c'est de cet exemple précis dont nous avons besoin dans la suite.

8.2 Résolution d'une équation homogène avec les nombres complexes

La première étape d'une résolution d'équation linéaire est la résolution de l'équation homogène associée. Faisons cela :

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0$$

Cherchons une solution de la forme $y(x) = \exp(\lambda x)$. Puisque $y'(x) = \lambda y(x)$, nous avons $y''(x) = \lambda^2 y(x)$, et ainsi de suite. Dès lors ce y sera solution si et seulement si :

$$a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0 = 0$$

ce qui est une équation polynomiale de degré n . Elle admet donc n solutions (comptées avec multiplicités). Cette équation est appelée « équation caractéristique ».

Mentionnons à présent le résultat que nous venons presque de démontrer :

Résultat XIX.19. *Les solutions de l'équation*

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0$$

sont les combinaisons linéaires de n fonctions déterminées comme suit : si $\lambda_1, \dots, \lambda_k$ sont les solutions distinctes de l'équation caractéristique, de multiplicité respective $m_1, \dots, m_k$, alors il faut considérer

$$\begin{aligned} & \exp(\lambda_1 x), x \exp(\lambda_1 x), x^2 \exp(\lambda_1 x), \dots, x^{m_1-1} \exp(\lambda_1 x) \\ & \vdots \\ & \exp(\lambda_k x), x \exp(\lambda_k x), x^2 \exp(\lambda_k x), \dots, x^{m_k-1} \exp(\lambda_k x). \end{aligned}$$