

Cours de mathématiques

Partie I – Les fondements

MPSI 4

Alain TROESCH

Version du:

27 août 2020

Table des matières

1 Logique et raisonnements	7
I Rudiments de logique	8
I.1 Formule propositionnelles, prédictats	8
I.2 Quantificateurs	10
I.3 Négations	11
II Raisonnements, et principes de rédaction	12
II.1 Composition d'un texte mathématique	12
II.2 Comment construire une démonstration	12
II.3 Le Modus ponens	14
II.4 Démonstration par la contraposée	14
II.5 Disjonction des cas	15
II.6 Analyse-Synthèse	15
II.7 Raisonnement par récurrence	16
II.8 Principe de la descente infinie	19
2 Ensembles	21
I Théorie intuitive des ensembles	22
I.1 Définition intuitive	22
I.2 Ensemble des parties d'un ensemble	24
I.3 Opérations sur les parties d'un ensemble	25
I.4 Union et intersection d'une famille de sous-ensembles	29
I.5 Partitions	30
I.6 Produit cartésien	31
I.7 Fonction caractéristique (ou indicatrice)	31
II Paradoxes ensemblistes et axiomatisation	32
II.1 La crise des fondements	32
II.2 Tentatives d'axiomatisation	33
3 Applications	37
I Qu'est-ce qu'une application ?	37
II Image directe, image réciproque	41
III Injectivité, surjectivité, bijectivité	44
IV Cardinal d'un ensemble	48
IV.1 Ensembles finis et cardinal	48
IV.2 Règles de calcul sur les cardinaux	50

IV.3	Comparaison des cardinaux en cas d'injectivité et surjectivité	51
IV.4	Introduction à la dénombrabilité (HP)	52
4 Sommes		55
I	Manipulation des signes \sum et \prod	55
I.1	Définition des notations	55
I.2	Changements d'indice	58
I.3	Additivité par rapport aux bornes	59
I.4	Linéarité	60
I.5	Sommes télescopiques	61
I.6	Cas des produits	62
I.7	Sommes multiples	62
I.8	Produits de sommes	64
II	Sommes classiques à connaître	65
II.1	Somme des puissances d'entiers	65
II.2	Sommes géométriques	66
II.3	Formule du binôme	68
II.4	Retour sur les sommes de puissances d'entiers	70
5 Relations		71
I	Définitions générales	71
I.1	Relations	71
I.2	Définition de quelques propriétés sur les relations	72
II	Relations d'équivalence	73
II.1	Définitions et exemples	73
II.2	Classes d'équivalence, ensembles quotients	74
II.3	Congruences	76
III	Relations d'ordre	76
III.1	Définitions générales	77
III.2	Minimalité, maximalité	79
III.3	Le lemme de Zorn (HP)	82
6 Les nombres réels		83
I	Un mot sur \mathbb{N} et \mathbb{Z}	83
I.1	Les entiers naturels	83
I.2	Les entiers relatifs	86
II	De \mathbb{Q} à \mathbb{R}	86
II.1	Construction de \mathbb{Q}	86
II.2	Relation d'ordre dans \mathbb{Q}	87
II.3	De l'existence de nombres non rationnels	88
II.4	L'ensemble \mathbb{R}	89
III	Nombres réels	90
III.1	Rappels sur les opérations et les inégalités	90
III.2	Division euclidienne dans \mathbb{R}	94
III.3	Densité de \mathbb{Q} et $\mathbb{R} \setminus \mathbb{Q}$ dans \mathbb{R}	95
III.4	Nombres transcendants	96
III.5	Partie entière, partie décimale	97
III.6	Représentation décimale	98
IV	Intervalles	101
IV.1	Description des intervalles	101
IV.2	Intervalles et topologie	103
V	Droite achevée $\overline{\mathbb{R}}$	105

7 Le corps \mathbb{C} des complexes	107
I Les nombres complexes : définition et manipulations	107
I.1 Définition, forme algébrique	107
I.2 Module	110
II Trigonométrie	111
II.1 Cercle trigonométrique, formules de trigonométrie	111
II.2 L'exponentielle complexe et applications à la trigonométrie	116
III Racines d'un nombre complexe	120
III.1 Racines n -ièmes	120
III.2 Cas des racines carrées : expression sous forme algébrique	122
IV Nombres complexes et géométrie	124
IV.1 Affixes	124
IV.2 Alignement, orthogonalité, angles	124
IV.3 Transformations du plan	125
IV.4 Isométries et similitudes	127
IV.5 Caractérisation de certains objets géométriques	128

Préface

Ce cours est l'aboutissement de plusieurs années d'enseignement en MPSI au lycée Louis-Le-Grand. Il est basé sur les programmes actuels de la classe de MPSI, ce qui n'exclut pas certaines digressions hors programmes, dûment signalées (ou non). Ces digressions peuvent être des avances sur le programme de deuxième année, ou tout simplement des développements permettant de comprendre à quoi peuvent servir les notions introduites, ou introduisant des outils un peu plus sophistiqués que ceux du programme, replaçant ainsi les résultats du programmes dans un contexte plus vaste. Ces développements sont utiles aux meilleurs étudiants pour prendre un peu de hauteur sur les notions étudiées en MPSI, et les comprendre au-delà de ce qui est demandé. Pour d'autres étudiants, en revanche, il pourra être plus profitable de ne pas se focaliser dessus dans un premier temps, quitte à y revenir plus tard, lorsqu'ils auront acquis de l'aisance avec les notions de ces chapitres, par exemple lors des révisions précédant l'entrée en deuxième année.

En écrivant ce cours, il n'était pas dans mes objectifs de rédiger un cours complet. Notamment, les développements des exemples et surtout les démonstrations des résultats ont été volontairement omis. Cependant sous la pression de mes élèves, j'ai rajouté dans cette version des « éléments de preuve », donnant l'idée générale et le schéma des différentes démonstrations, sans entrer dans les détails techniques. Ces éléments de preuve apparaissent en grisé afin de ne pas charger visuellement le polycopié. Leur but est double :

- Permettre une certaine autonomie dans la découverte du cours : ces éléments de preuve sont conçus comme des indications permettant de développer ensuite soi-même la démonstration des résultats, comme un exercice. L'idéal est de pouvoir faire cette préparation avant que le cours soit fait en classe ; sinon pendant le cours, mais les temps de réflexion sont plus courts.
- Faciliter les révisions, sachant que les preuves complètes sont prises par les élèves sur des feuilles séparées du polycopié : lors des révisions, ces éléments doivent permettre de se remémorer rapidement les grandes lignes des preuves, sans avoir à ressortir ses notes. Cherchez alors à développer les détails des preuves par vous-même (par écrit une fois, et les fois suivantes, au moins dans votre tête) ; si vous coincez, le recours aux notes prises pendant le cours s'impose.

Ce cours est constitué de 4 parties. La première, appelée « Fondements », regroupe toutes les bases logiques et ensemblistes des mathématiques, ainsi que l'étude des nombres réels et complexes. La seconde partie est le cours d'analyse, commençant par de l'analyse concrète (étude de fonctions, pratique de l'intégration etc.) puis l'étude des suites, des approximations (Taylor, développements limités) et des séries. La troisième partie est le cours d'algèbre, regroupant l'étude des structures algébriques (groupes, anneaux, corps), avec notamment l'étude des anneaux de polynômes, puis l'étude de l'algèbre linéaire sous son aspect vectoriel et matriciel, et enfin l'algèbre bilinéaire. Enfin une quatrième et courte partie est consacrée aux probabilités.

Je remercie tous les élèves de MPSI 4 (Bestial!!!) que j'ai eus depuis que je suis au lycée Louis-Le-Grand. Vos remarques et vos nombreuses questions sont à la base de nombreuses améliorations de ces notes de cours. Par ailleurs, l'intérêt et, pour certain, la passion que vous montrez pour les mathématiques m'ont beaucoup motivé pour vous donner le meilleur de moi-même. J'ai beaucoup appris de vous, aussi bien du point de vue pédagogique que du point de vue mathématique. N'est-ce par là l'idéal de l'enseignement, quand l'enseignement devient échange ?

Je n'oublierai jamais aucun de vous, que vous ayez été l'étoile filant bien au-dessus de moi, ou l'élève avançant avec beaucoup plus de peine. Vous avez tous été formidables.

Et que mes élèves actuels et mes futurs élèves se rappellent d'une chose : que vous soyez premier ou dernier de la classe, votre place est bien là, en MPSI 4, et vous la méritez. Souvenez-vous en lorsque, à certains moments de l'année, vous aurez l'impression d'être un peu perdus.

1

Logique et raisonnements

La logique est la jeunesse des mathématiques

(Bertrand Russell)

La logique est l'hygiène des mathématiques

(André Weil)

La logique n'a ni à inspirer l'invention, ni à l'expliquer ; elle se contente de la contrôler et de la vérifier.

(Louis Couturat)

En effet, l'effet fait le même effet à la cause que l'effet que la cause lui a causé de faire

(Professeur Shadoko par Jacques Rouxier)

Ce chapitre a pour but d'introduire les concepts fondamentaux des mathématiques, à savoir les bases-même du raisonnement mathématique : Le but n'est pas l'étude de la logique formelle, ni même la présentation rigoureuse de cette logique formelle, mais de voir comment des rudiments de la théorie de la logique permettent une mise en forme rigoureuse de la structure de la pensée et du cheminement logique. Cependant, cette structuration ne peut en rien remplacer l'intuition comme le dit si bien René Thom :

Car le monde des Idées excède infiniment nos possibilités opératoires, et c'est dans l'intuition que réside l'ultima ratio de notre foi en la vérité d'un théorème - un théorème étant, selon une étymologie aujourd'hui bien oubliée, l'objet d'une vision.

(René Thom)

Et une dernière citation de René Thom également, pour vous inciter à ne pas abuser du formalisme stérile (le grand danger après un chapitre de logique) :

Ah oui ! Je pense que le langage est un très bon outil. Je me suis battu en mathématique contre les formalisateurs. Les formalisateurs sont des gens qui vous disent tout le temps : « Oh ! Le langage naturel est horrible, il tolère toute espèce d'ambiguïté, c'est impossible de faire des mathématiques avec ça ». Moi, je n'ai jamais fait que du langage naturel en mathématiques, plus quelques symboles de temps en temps

(René Thom)

I Rudiments de logique

I.1 Formules propositionnelles, prédictats

La logique propositionnelle est l'étude des formules abstraites qu'on peut écrire à partir d'un certain nombre de variables propositionnelles, représentées par des lettres. Nous nous contenterons d'une définition restant assez vague, l'objet n'étant pas l'étude de la logique formelle, mais une bonne structuration de la pensée et de la démarche scientifique.

Définition 1.1.1 (Formule propositionnelle)

Une *formule propositionnelle* est une formule liant des propositions élémentaires représentées par des lettres (ou variables propositionnelles), à l'aide d'un certain nombre de symboles représentant des opérations logiques :

- \wedge : et
- \vee : ou
- \implies : implique
- \iff : équivalent
- \neg : non

À part \neg qui se met devant une unique proposition, les autres symboles permettent de lier 2 propositions. Un parenthésage rigoureux est nécessaire afin de rendre l'expression non ambiguë quant à l'ordre des opérations à effectuer.

Exemple 1.1.2 (Formules propositionnelles)

Dans cet exemple, P , Q , R désignent des variables propositionnelles.

1. Ceci est une formule : $((P \implies Q) \vee Q) \implies ((R \wedge P) \iff \neg Q)$.
On n'affirme pas si elle est vraie ou fausse.
2. Ceci n'est pas une formule : $(P \implies) \vee R \wedge$

Chaque variable propositionnelle peut prendre une valeur de vérité : V (Vrai) ou F (Faux). Suivant les valeurs de vérité prises par les différentes variables propositionnelles intervenant dans la formule, une formule pourra alors être vraie ou fausse, ce qu'on déterminera en suivant les règles intuitives de véracité liées aux symboles de connection utilisés et rappelées ci-dessous :

Définition 1.1.3 (Définition de l'interprétation sémantique des connecteurs logiques)

Soit P , Q deux variables propositionnelles. Les tables de vérité des formules $\neg P$, $(P \vee Q)$, $(P \wedge Q)$, $(P \implies Q)$ et $(P \iff Q)$ sont définies par :

P	Q	$(P \vee Q)$	P	Q	$(P \wedge Q)$	P	Q	$(P \implies Q)$	P	Q	$(P \iff Q)$
V	V	V	V	V	V	V	V	V	V	V	V
V	F	V	V	F	F	V	F	F	V	F	F
F	V	V	F	V	F	F	V	V	F	V	F
F	F	F	F	F	F	F	F	V	F	F	V

Ces tables définissent en fait le sens logique des connecteurs.

Remarque 1.1.4

1. La table de vérité de l'implication se comprend bien en considérant sa négation : dire qu'une implication $P \implies Q$, est fausse, c'est dire que malgré le fait que l'hypothèse P soit vraie, la conclusion Q est fausse.

2. Ainsi, dire que $P \implies Q$ est vraie ne sous-entend nullement la véracité de P . En particulier, « $P \implies Q$ » n'est pas équivalent à « P donc Q », qui affirme la véracité de P . Il convient donc de faire attention à la rédaction : **le symbole « \implies » ne peut pas remplacer le mot « donc »**
3. La même remarque vaut pour l'équivalence.
4. Par ailleurs, puisque si P est faux, $P \implies Q$ est toujours vrai, pour montrer que $P \implies Q$ est vrai, il suffit de se placer dans le cas où P est vrai : on suppose que P est vrai, on montre que Q aussi. Cela correspond à l'interprétation « Si P est vrai, alors Q est vrai ». En revanche, on n'a pas de contrainte lorsque P est faux.
5. Ne pas confondre :
 - P est une condition suffisante à Q : $P \implies Q$;
 - P est une condition nécessaire à Q : $Q \implies P$;
 - P est une condition nécessaire et suffisante à Q : $P \iff Q$.
6. Pour montrer une équivalence $P \iff Q$, n'oubliez pas de montrer les *deux* implications $P \implies Q$ et $Q \implies P$. N'oubliez pas la réciproque !

Exemple 1.1.5

- « n est multiple de 6 » est une pour que n soit pair mais pas une
- $x = 1$ est une pour que $x^2 = 1$, mais pas une En revanche, si x est réel, $x = 1$ est une pour que $x^3 = 1$.
- Si f est dérivable sur \mathbb{R} , $f'(0) = 0$ est une pour que f admette un extremum local en 0, mais ce n'est pas une

Définition 1.1.6 (Formules équivalentes)

Deux formules A et B sont dites équivalentes (on notera $A \equiv B$) si elles prennent la même valeur de vérité l'une et l'autre, quelle que soit la distribution de vérités donnée sur l'ensemble des variables propositionnelles intervenant dans ces formules. Autrement dit, elles sont vraies et fausses sous les mêmes conditions sur les variables propositionnelles. On note alors $A \equiv B$.

Définition 1.1.7 (Tautologie)

Ce sont des formules toujours vraies (pour toute distribution de vérité).

On rappelle ci-dessous les équivalences et tautologies les plus importantes, formant la base du raisonnement et des manipulations ensemblistes.

Proposition 1.1.8 (Quelques équivalences ou tautologies)

A, B, C, \dots désignent des variables propositionnelles.

1. $(A \wedge B) \wedge C = A \wedge (B \wedge C)$ (associativité). On notera simplement $A \wedge B \wedge C$, et on peut généraliser à davantage de termes.
2. De même pour $A \vee B \vee C$
3. $A \wedge (B \vee C) \equiv (A \wedge B) \vee (A \wedge C)$ (distributivité)
4. $A \vee (B \wedge C) \equiv (A \vee B) \wedge (A \vee C)$ (distributivité)
5. $(A \wedge (A \implies B)) \implies B$ est une tautologie (modus ponens)
6. $(A \implies B) \equiv (B \vee \neg A)$
7. $(A \implies B) \equiv (\neg B \implies \neg A)$ (contraposée)

▫ Éléments de preuve.

On peut dresser des tables de vérité pour s'en assurer, mais il est beaucoup plus important d'avoir bien compris ces propriétés sous l'angle de la logique intuitive. On peut aussi remarquer que 6 entraîne 7. ▷

I.2 Quantificateurs

Dans un texte mathématique élaboré, Les variables propositionnelles représentent des propositions mathématiques élémentaires : des formules, des équations, des faits mathématiques etc. Ces énoncés nécessitent s'exprimer souvent à l'aide de variables mathématiques, vouées à prendre des valeurs dans un ensemble. Deux propriétés particulières liés à une formule utilisant une variable peuvent être particulièrement intéressantes :

- le fait que la formule soit vraie pour toutes les valeurs possibles de la variable dans un ensemble donné
- le fait que la formule soit vraie pour au moins une valeur de x .

Pour formuler ces propriétés, on introduit deux symboles, appelés quantificateurs :

Définition 1.1.9 (Quantificateurs)

Soit $F(x)$ une propriété dépendant d'une variable x .

- Le quantificateur universel \forall :
 $\forall x, F(x)$ est satisfait si et seulement si pour tout valeur possible prise par x , $F(x)$ est vraie.
- Le quantificateur existentiel \exists :
 $\exists x, F(x)$ est satisfait si et seulement si il existe x tel que $F(x)$ soit satisfait. Dans ce cas, même s'il n'est bien souvent pas possible d'expliquer x , on peut se donner (c'est-à-dire fixer) un x qui satisfait $F(x)$ (donc choisir un x convenable).

Remplacer une variable quantifiée par une autre (indépendante des autres variables intervenant dans la formule) ne change pas le sens de la proposition. Ainsi $\forall x, P(x)$, et $\forall y, P(y)$ sont équivalents.

Attention à bien remplacer toutes les occurrences de la variable x dans le champ d'action du quantificateur. On dit dans ce cas que x est une **variable muette**.

Avertissement 1.1.10

Attention ! En général, on ne peut pas intervertir \exists et \forall !

Remarque 1.1.11

- Quelle différence de sens faites vous entre les deux formules $\forall x, \exists y, P(x, y)$ et $\exists y, \forall x, P(x, y)$.
- L'une des deux formules implique l'autre. Laquelle ?
- Peut-on intervertir deux symboles \exists ? deux symboles \forall ?

Avertissement 1.1.12

Attention au parenthésage et à la distribution !

Exemple 1.1.13

1. Les deux formules suivantes sont-elles équivalentes : $\forall x, (P(x) \implies Q)$ et $(\forall x P(x)) \implies Q$? Trouver une formule parenthésée comme la deuxième, équivalente à la première.
2. Les deux formules $\forall x(P(x) \wedge Q(x))$ et $(\forall x P(x)) \wedge (\forall x Q(x))$ sont-elles équivalentes ? et $\forall x(P(x) \vee Q(x))$ et $(\forall x P(x)) \vee (\forall x Q(x))$? Même question avec des symboles \exists .

Avertissement 1.1.14

Ne jamais utiliser les symboles de quantification dans une phrase : il s'agit d'un symbole mathématique, pas d'une abréviation.

I.3 Négations

Dans de nombreuses occasions, par exemple pour mener des démonstrations par l'absurde ou par la contraposée, il est important de savoir nier une expression mathématique (c'est-à-dire exprimer son contraire) de façon efficace et rapide (et sans erreur!).

Cette négation peut se faire de façon purement formelle, en utilisant les règles de négation suivantes :

Propriétés 1.1.15 (Négation d'une formule)

On a les équivalences de formules suivantes (P et Q sont des formules) :

1. $\neg\neg P \equiv P$
2. $\neg(P \vee Q) \equiv \neg P \wedge \neg Q$ (*loi de De Morgan*)
3. $\neg(P \wedge Q) \equiv \neg P \vee \neg Q$ (*loi de De Morgan*)
4. $\neg(P \implies Q) \equiv P \wedge \neg Q$
5. $\neg(P \iff Q) \equiv (P \iff \neg Q) \equiv ((\neg P) \iff Q)$.

◊ Éléments de preuve.

La démonstration se fait en comparant les tables de vérité. La 4 peut aussi se déduire de l'équivalence $P \implies Q \equiv \neg P \vee Q$. Mais encore une fois, il est plus important d'avoir bien compris ces propriétés sous l'angle de la logique intuitive. ▷

Propriétés 1.1.16 (Négation des quantificateurs)

- (i) $\neg(\forall x P) \equiv \exists x (\neg P)$
- (ii) $\neg(\exists x P) \equiv \forall x (\neg P)$.

◊ Éléments de preuve.

On se contente ici de l'interprétation intuitive évidente. Pour plus de rigueur, il faudrait une définition un peu plus formelle des quantificateurs. ▷

Remarque 1.1.17

Pour contredire une propriété universelle $\forall x, P(x)$, il suffit de trouver UN contre-exemple.

Remarque 1.1.18

Si l'ensemble des valeurs x possibles est fini, disons x_1, \dots, x_n , alors

- $\forall x, P(x)$ équivaut à $P(x_1) \wedge \dots \wedge P(x_n)$
- $\exists x, P(x)$ équivaut à $P(x_1) \vee \dots \vee P(x_n)$

De nombreuses propriétés impliquant \forall et \exists sont à voir comme des généralisations de propriétés similaires sur \wedge et \vee . Vérifiez par exemple la compatibilité de la description ci-dessus avec les propriétés de négation.

Exemples 1.1.19

Niez les propositions suivantes :

1. $((A \implies B) \wedge C) \vee \neg B$
2. $((A \iff B) \vee C) \implies B \iff A$

Exemple 1.1.20

Définition de la continuité d'une fonction f en x_0 , puis négation. Exemple de la fonction H de Heaviside.

II Raisonnements, et principes de rédaction

II.1 Composition d'un texte mathématique

Un texte mathématique est constitué de :

1. **définitions** : des descriptions de certains objets constituant les briques de la théorie. C'est à voir comme un raccourci de langage.
2. **résultats** : des énoncés mettant en jeu les objets définis dans la théorie, et donnant des propriétés vérifiées par ces objets. Un résultat s'énonce sous la forme $A \implies B$. On distingue :
 - les *axiomes* : des résultats qui sont des vérités fondamentales de la théorie, et qu'on ne démontre pas (à considérer comme le cahier des charges de la théorie : on impose ces résultats, il n'y a donc pas besoin de les montrer) ;
 - les *théorèmes* : les résultats les plus significatifs, démontrés à partir des axiomes et de résultats démontrés antérieurement ;
 - les *propositions* : des résultats de moindre envergure ;
 - les *lemmes* : des résultats à voir comme des étapes vers des résultats plus consistants (résultats préliminaires, mais pouvant avoir leur intérêt en soi)
 - les *corollaires* : des conséquences assez immédiates d'autres résultats, par exemple des cas particuliers intéressants ;
3. **démonstrations** : des justifications de la véracité des résultats.
4. **conjectures** : des énoncés qu'on pense être vrais, mais qu'on n'a pas encore réussi à prouver.

Remarque 1.2.1

Un énoncé s'exprime souvent sous la forme $A \implies B$.

La proposition A regroupe les *hypothèses*

La proposition B regroupe les *conclusions*.

Ne pas oublier de bien apprendre toutes les hypothèses d'un résultat. Par exemple, considérons le théorème suivant :

Soit f une fonction dérivable sur un intervalle I . Si f' est positive sur I , alors f est croissante sur I .
 Il y a trois hypothèses dans cet énoncé : bien sûr, $f' \geq 0$, que personne n'oublie ; mais aussi f dérivable sur I (sans laquelle l'énoncé n'a pas de sens), et (plus souvent oubliée) le fait que I est un intervalle (sans quoi le résultat est faux!).

II.2 Comment construire une démonstration

Dans un exercice ou un devoir, c'est au candidat de construire soi-même la démonstration. Il est donc intéressant d'avoir une démarche permettant d'aborder ces démonstrations de façon logique et structurée. Bien entendu, l'application formelle de ces règles n'est pas suffisante, il faut à un moment de la démonstration apporter une ou plusieurs idées personnelles !

La construction rigoureuse repose sur la structure logique de l'énoncé à démontrer, en se basant sur les principes généraux suivants. Une formule mathématique étant construite en itérant des constructions élémentaires de ce type, il faut bien appliquer ces principes de démonstration à chaque étape de la construction, ce qui nécessite de dérouler la structure logique du résultat à montrer.

- **Prouver une implication $A \implies B$:**

On suppose que A est vrai, on démontre B . La rédaction commence par « *Supposons que A est vrai* ».

Dans certaines situations, il peut être plus simple de montrer l'implication contraposée (voir plus loin). Y penser si on bloque !

- **Prouver une équivalence $A \iff B$:**

On prouve en deux temps $A \implies B$ et $B \implies A$. On peut aussi raisonner par équivalences successives, mais dans ce cas, raisonner d'abord dans un sens, puis vérifier scrupuleusement qu'on peut « remonter » toutes les implications.

- **Prouver une conjonction $A \wedge B$:**

On prouve en deux temps : on prouve A , puis on prouve B .

- **Prouver une disjonction $A \vee B$:**

On prouve que $\neg A \implies B$, ce qui revient à supposer que A n'est pas vrai, et à en déduire que B est vrai. On peut bien sûr intervertir A et B : un bon choix de la propriété que l'on nie peut parfois simplifier la démonstration.

- **Prouver $\forall x A(x)$:**

La proposition A doit être vraie pour tout choix de x . On pose donc un x **supposé quelconque** (c'est-à-dire sur lequel on n'impose pas de condition), et on montre que pour ce x , $A(x)$ est vérifié. Le fait d'avoir choisi x quelconque montre qu'alors $A(x)$ est vrai pour tout x .

La démonstration débute alors systématiquement par « *Soit x ...* », puis on démontre $A(x)$.

- **Prouver $\exists x A(x)$:**

Montrer une propriété existentielle est souvent ce qu'il y a de plus délicat. Dans le meilleur des cas, on construit explicitement x qui convient. Pour s'aider à définir x convenable, on peut faire une analyse/synthèse (voir plus loin).

Exemple 1.2.2

Structure d'une démonstration adaptée à la formule :

$$\forall x, (A(x) \implies \forall y, (B(y) \vee \exists z C(z))).$$

Avertissement 1.2.3

Attention ! Utiliser de façon trop systématique et trop poussée ces différentes règles peut parfois empêcher de voir la ressemblance avec une propriété du cours, et peut nuire à l'intuition. Ne pas le faire assez nuit très souvent à la rigueur de la rédaction. Il faut donc trouver un juste milieu.

Avertissement 1.2.4

La structure logique, puis les règles de la logique formelle, ne font que structurer la démonstration. Une bonne rédaction passe par une **mise en langage de ces règles** : on rédige toujours **à l'aide de phrases**, et non par un enchaînement de formules logiques absconses !

Remarque 1.2.5 (Direction d'une preuve)

C'est le but qui dirige une preuve ! On se fixe un cap, et on le garde. À aucun moment, il ne faut le perdre du vue.

Nous voyons maintenant un certain nombre de méthodes classiques de démonstration.

II.3 Le Modus ponens.

Méthode 1.2.6 (Modus ponens)

Pour que B soit vrai, il suffit que A soit vrai et que $A \Rightarrow B$. Formellement, il s'agit d'exploiter la tautologie :

$$(A \wedge (A \Rightarrow B)) \Rightarrow B$$

Avertissement 1.2.7

Attention, $A \Rightarrow B$ n'est pas suffisant. Si on veut obtenir B , il faut aussi justifier la véracité de A ! (différence entre « \Rightarrow » et « donc »)

Le *modus ponens* est donc à voir comme une formalisation du « donc », déduction logique.

La situation typique d'utilisation du *modus ponens* est l'emploi d'un théorème : celui-ci s'écrit $A \Rightarrow B$, où A est l'hypothèse et B la conclusion. Ainsi, pour montrer B , on vérifie que l'hypothèse A est satisfaite, et on emploie le théorème $A \Rightarrow B$. Le *modus ponens* nous permet de conclure que la conclusion B est vraie aussi.

Avertissement 1.2.8

Toujours bien préciser A et $A \Rightarrow B$. En particulier, quand on utilise un théorème, toujours bien préciser le théorème utilisé d'une part (en donnant son nom), et la validité des hypothèses d'autre part.

Cela nécessite un apprentissage rigoureux du cours : la connaissance des hypothèses des théorèmes est aussi importante que la connaissance de leurs conclusions (c'est cette bonne connaissance des hypothèses qui assure aussi qu'on n'utilisera pas le théorème à tort et à travers dans des situations inadaptées).

II.4 Démonstration par la contraposée.

Il s'agit de l'utilisation de l'équivalence des deux propositions $A \Rightarrow B$ et $\neg B \Rightarrow \neg A$.

Méthode 1.2.9 (Démonstration par contraposée)

Pour montrer $A \Rightarrow B$, on peut adopter la démarche suivante : on suppose que la conclusion B est fausse, et on montre que dans ce cas, l'hypothèse A ne peut pas être vraie. Cela revient à montrer $\neg B \Rightarrow \neg A$.

Définition 1.2.10 (Contraposée)

L'expression $\neg B \Rightarrow \neg A$ s'appelle la *contraposée* de $A \Rightarrow B$.

Ce type de démonstration apparaît dans de nombreuses situations.

Exemples 1.2.11

1. Soit $n \in \mathbb{N}$. Montrer que si n^2 est pair, alors n est pair.
2. Montrer que si $x_1 + \dots + x_n = M$, alors il existe $i \in \llbracket 1, n \rrbracket$ tel que $x_i \geq \frac{M}{n}$.

Avertissement 1.2.12

Ne pas confondre la contraposée $\neg B \Rightarrow \neg A$ et l'expression $\neg A \Rightarrow \neg B$, qui n'est pas équivalente à $A \Rightarrow B$, mais à sa réciproque !

Un cas particulier important de démonstration par la contraposée est le cas de la démonstration par l'absurde. Il s'agit de la situation dans laquelle A est la propriété toujours vraie. Alors $\neg A$ est la propriété toujours fausse (il s'agit d'une contradiction).

Méthode 1.2.13 (Cas particulier : démonstration par l'absurde)

Pour démontrer B , il suffit de montrer que le fait de supposer que B est fausse conduit à une contradiction.

Là encore, les démonstrations par l'absurde interviennent dans des situations très diverses. La démonstration par l'absurde la plus connue est certainement la démonstration de Pythagore de l'irrationnalité de $\sqrt{2}$:

Exemple 1.2.14

Démonstration de l'irrationnalité de $\sqrt{2}$.

II.5 Disjonction des cas.

Ce principe de démonstration repose sur l'équivalence, déjà évoquée plus haut :

$$(A \vee B) \implies C \equiv (A \implies C) \wedge (B \implies C).$$

Méthode 1.2.15 (Disjonction des cas, ou discussion)

Pour montrer $A \vee B \implies C$, on peut séparer en deux cas : voir ce qu'il se passe sous l'hypothèse A , puis sous l'hypothèse B . Ainsi on montre que si on suppose que A est vérifié, alors C aussi, et de même, si B est vérifié, C aussi.

Un cas particulièrement important est le cas où $A \vee B$ est la proposition certaine (A et B regroupe l'ensemble de tous les cas possibles). Dans ce cas, $(A \vee B) \implies C$ équivaut à C .

Exemple 1.2.16

Montrer, sans utiliser le fait qu'il s'agit du résultat d'une somme classique, que pour tout $n \in \mathbb{N}$, $\frac{n(n+1)(2n+1)}{6}$ est entier.

II.6 Analyse-Synthèse

Ce procédé de démonstration est surtout adapté pour les problèmes existentiels (montrer l'existence d'un objet vérifiant un certain nombre de propriétés). Le principe est le suivant :

Méthode 1.2.17 (Analyse-synthèse)

- Phase d'analyse : On suppose dans un premier temps l'existence d'un objet tel que souhaité, et à l'aide des propriétés qu'il est censé vérifier, on obtient autant d'informations que possible sur la façon de construire un tel objet.
- Phase de synthèse : lorsqu'on a suffisamment d'informations sur une façon de construire l'objet recherché, on construit un objet de la sorte, de façon explicite, et on vérifie qu'il répond au problème.
- Bonus : si la phase d'analyse fournit une expression explicite de l'objet recherché, ne laissant pas le choix pour cet objet, cela fournit même l'unicité.

Remarque 1.2.18

- La phase d'analyse est la recherche de conditions nécessaires.
- La phase de synthèse est la donnée de conditions suffisantes.

Exemple 1.2.19

Soit a un réel et $I = [-a, a]$. Montrer que toute fonction $f : I \rightarrow \mathbb{R}$ s'écrit comme somme d'une fonction paire et d'une fonction impaire.

Remarque 1.2.20

Cet exemple est un cas particulier d'un exemple générique consistant à prouver qu'un espace vectoriel se décompose en somme (ici somme directe) de deux sous-espaces vectoriels. Se reporter au chapitre idoine pour ces notions.

Remarque 1.2.21

Dans le cadre d'un problème existentiel sans contrainte d'unicité, la phase d'analyse ne sert qu'à deviner une expression répondant au problème. D'un point de vue de la rigueur de rédaction, cette phase n'est pas indispensable, la phase de synthèse est suffisante pour répondre au problème existentiel. Cependant, elle permet au lecteur de mieux comprendre comment on est parvenu à l'expression voulue. Sans cette phase d'analyse, la réponse peut dans certaines situations paraître un peu parachutée.

Savoir si on rédige la phase d'analyse ou non dépend en fait de la complexité de la situation. Dans certaines situations assez simples, on comprend bien l'expression obtenue, on peut parfois même la deviner sans passer par une analyse. Dans ces cas, ne vous cassez pas la tête et faites directement la synthèse.

Avertissement 1.2.22

Soyez très précautionneux dans la rédaction d'une démonstration par analyse-synthèse. Dites bien de façon explicite qu'il s'agit d'un raisonnement de ce type. En effet, comme la phase d'analyse consiste en une recherche de conditions nécessaires, elle consiste souvent à supposer la conclusion vraie pour essayer d'obtenir le maximum d'informations sur l'expression recherchée. Un lecteur pressé (et les correcteurs au concours sont à classer dans cette catégorie) risque de prendre votre démonstration pour une pétition de principe (montrer un résultat en le supposant vrai au départ !)

II.7 Raisonnement par récurrence

Le principe de récurrence est un axiome de la construction de \mathbb{N} . Il s'énonce ainsi :

$$[\mathcal{P}(0) \wedge (\forall n \in \mathbb{N}, (\mathcal{P}(n) \implies \mathcal{P}(n+1))] \implies (\forall n \in \mathbb{N}, \mathcal{P}(n))$$

« $\mathcal{P}(0)$ » est l'initialisation, « $\forall n \in \mathbb{N}, (\mathcal{P}(n) \implies \mathcal{P}(n+1))$ » est l'hérédité (ou le caractère héréditaire ou transmissible).

Méthode 1.2.23 (Démonstration par récurrence simple)

Pour montrer une propriété $\mathcal{P}(n)$ dépendant d'un entier $n \in \mathbb{N}$, on procède suivant le schéma suivant :

- Initialisation : montrer que $\mathcal{P}(0)$ est vraie.
- Hérédité : montrer que pour tout $n \in \mathbb{N}$, $\mathcal{P}(n) \implies \mathcal{P}(n+1)$, ce qui se fait, d'après les principes développés précédemment en posant n quelconque (« Soit $n \in \mathbb{N}$ »), en supposant que pour ce n , $\mathcal{P}(n)$ est vrai, et en montrant qu'alors $\mathcal{P}(n+1)$ l'est aussi.
- Conclure, en faisant référence au principe de récurrence.

Ce principe peut s'adapter à des situations légèrement différentes :

- Le rang initial peut être un autre entier (éventuellement négatif). Cela modifie aussi alors le rang initial pour le caractère héréditaire.
- On peut faire des récurrences descendantes pour une propriété $\mathcal{P}(n)$, à démontrer sur un intervalle du type $]-\infty, n_0]$. On initialise alors avec $\mathcal{P}(n_0)$ et on montre que pour tout $n \leq n_0$, $\mathcal{P}(n) \implies \mathcal{P}(n - 1)$.
- On peut faire des récurrences bornées, pour montrer une propriété $\mathcal{P}(n)$ sur un intervalle borné, par exemple $[0, n_0]$. On initialise alors avec $\mathcal{P}(0)$ et on montre que $\mathcal{P}(n)$ implique $\mathcal{P}(n + 1)$ pour tout $n \in [0, n_0 - 1]$. D'un point de vue purement logique, il ne s'agit pas de l'utilisation du principe de récurrence, mais d'une itération (utilisation répétée, en nombre fini, du modus ponens, par transitivité de l'implication).
- Ces récurrences bornées s'adaptent aussi au cas de récurrences descendantes.
- Nous verrons un peu plus loin deux variantes du principe de récurrence : la récurrence d'ordre k , et la récurrence forte.

Exemple 1.2.24

Montrer (par récurrence) que pour tout $n \in \mathbb{N}^*$, $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$.

Avertissement 1.2.25

N'oubliez pas l'initialisation ! Prouver l'hérédité à tout rang ne suffit pas !

Exemple 1.2.26

$10^n + (-1)^n$ est-il divisible par 11 pour tout $n \in \mathbb{N}$?

Attention aussi à bien vous assurer que l'hérédité est valable à tout rang.

Exemple 1.2.27 (Le problème des crayons de couleur)

Nous montrons dans cet exemple que tout ensemble de crayons de couleur est monochrome. Nous notons pour tout $n \in \mathbb{N}^*$, $\mathcal{P}(n)$ la propriété affirmant que tout ensemble de n crayons de couleurs est constitué de crayons ayant tous la même couleur.

- La propriété $\mathcal{P}(1)$ est trivialement vraie, ce qui initialise la récurrence
- Soit $n \in \mathbb{N}^*$. Supposons que $\mathcal{P}(n)$ est vrai. Soit alors un ensemble de $n + 1$ crayons de couleur, qu'on peut supposer numérotés de 1 à $n + 1$. En appliquant l'hypothèse de récurrence aux n premiers crayons et aux n derniers crayons, le crayon 1 a la même couleur que les crayons 2 à n qui ont aussi la même couleur que le crayon $n + 1$, ce qui prouve $\mathcal{P}(n + 1)$ (figure 1.1)
- D'après le principe de récurrence, on peut donc conclure qu'il n'existe au monde que des crayons d'une même couleur.

Où est l'erreur ?

Méthode 1.2.28 (Récurrence d'ordre k)

Il s'agit d'une variante du principe de récurrence, s'exprimant ainsi :

$$((\mathcal{P}(0) \wedge \cdots \wedge \mathcal{P}(k-1)) \wedge (\forall n \in \mathbb{N}, \mathcal{P}(n) \wedge \cdots \wedge \mathcal{P}(n+k-1) \implies \mathcal{P}(n+k))) \implies \forall n \in \mathbb{N}, \mathcal{P}(n).$$

- Principe : on utilise la propriété aux k rangs précédents pour montrer l'hérédité.
- Schéma de rédaction :
 - * Initialisation : montrer $\mathcal{P}(0), \dots, \mathcal{P}(k-1)$.

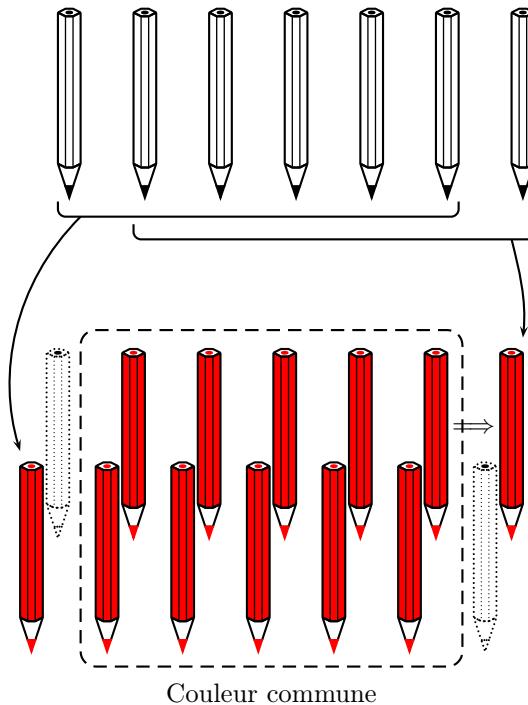


FIGURE 1.1 – Caractère héréditaire pour le problème des crayons de couleur

- * Héritage : poser $n \geq 0$, supposer $\mathcal{P}(n), \dots, \mathcal{P}(n+k-1)$, en déduire $\mathcal{P}(n+k)$.
- * Conclure en faisant appel au principe de récurrence.

Avertissement 1.2.29

Ne pas oublier d'initialiser pour les k premières valeurs ! (sinon la première implication de l'héritage n'est pas valable)

On peut bien sûr adapter le principe dans le cas où le rang initial n'est pas 0.

Exemples 1.2.30

1. Soit $(F_n)_{n \in \mathbb{N}}$ définie par $F_0 = 0$, $F_1 = 1$ et pour tout $n \geq 0$, $F_{n+2} = F_{n+1} + F_n$ (suite de Fibonacci).
Avec une mauvaise initialisation on peut « montrer » que F_n est divisible par 3, ce qui est évidemment faux !
2. Soit $n \in \mathbb{N}$. Montrer que F_n est paire si et seulement si n est multiple de 3.

Pour votre culture, ce dernier exemple est une situation particulière du résultat plus général suivant :

$$\forall (m, n) \in (\mathbb{N}^*)^2, \quad F_{m \wedge n} = F_m \wedge F_n,$$

avec $m = 3$.

Voici une dernière variante du principe de récurrence :

Méthode 1.2.31 (Récurrence forte)

La récurrence forte est basée sur la propriété formelle suivante :

$$(\mathcal{P}(0) \wedge (\forall n \geq 1, \mathcal{P}(0) \wedge \cdots \wedge \mathcal{P}(n-1) \implies \mathcal{P}(n))) \implies \forall n \in \mathbb{N}, \mathcal{P}(n).$$

- Principe : on suppose la propriété vraie à tous les rangs précédents pour la montrer à un rang donné.
- Schéma de rédaction :
 - * Initialisation pour $\mathcal{P}(0)$ (une seule valeur suffit ici)
 - * Poser $n > 0$, supposer $\mathcal{P}(k)$ vrai pour tout $k < n$, et montrer qu'alors $\mathcal{P}(n)$ est vrai.
 - * Conclure en faisant référence au principe de récurrence.

Exemple 1.2.32

Tout nombre entier $n \geq 1$ admet une décomposition en produit de nombres premiers.

Remarque 1.2.33

On retiendra du dernier exemple que le principe de récurrence forte est en particulier très utile dans de nombreuses questions liées à des propriétés de divisibilité.

Théorème 1.2.34

Les trois principes de récurrence ci-dessus sont équivalents

▫ Éléments de preuve.

La récurrence forte et la récurrence d'ordre k impliquent chacune la récurrence simple, car qui peut le plus, peut le moins. Réciproquement, poser respectivement $\mathcal{Q}(n) = \mathcal{P}(0) \wedge \cdots \wedge \mathcal{P}(n)$ ou $\mathcal{Q}(n) = \mathcal{P}(n) \wedge \cdots \wedge \mathcal{P}(n+k-1)$, et appliquer le principe de récurrence simple à \mathcal{Q} . ▷

II.8 Principe de la descente infinie

Nous terminons sur une dernière méthode classique, plus anecdotique pour une grande part des mathématiques, mais d'une telle efficacité pour certains problèmes d'arithmétique qu'on ne peut pas ne pas la mentionner.

Méthode 1.2.35 (Descente infinie)

Le principe de la descente infinie est un mélange de démonstration par l'absurde et de démonstration par récurrence. Soit $(\mathcal{P}(n))_{n \in \mathbb{N}}$ une propriété dont on veut démontrer qu'elle est fausse pour tout $n \in \mathbb{N}$: il suffit de montrer que si elle est supposée vraie à un certain rang n , il existe alors un rang $0 \leq m < n$ tel qu'elle soit encore vraie.

▫ Éléments de preuve.

En effet en itérant alors ce procédé, on pourrait construire une chaîne infinie d'entiers strictement décroissants telle que $\mathcal{P}(n)$ soit vraie, ce qui est impossible d'après la propriété fondamentale de \mathbb{N} , ce qui amène la contradiction recherchée. ▷

Remarque 1.2.36

Formellement, le principe de descente infinie est la contraposée du principe de récurrence forte appliqué à $\neg\mathcal{P}(n)$: au lieu de montrer que si pour tout $m < n$, $\neg\mathcal{P}(m)$ est vérifié alors $\neg\mathcal{P}(n)$ l'est aussi, on montre la contraposée de cette implication : si $\mathcal{P}(n)$, alors il existe $0 \leq m < n$ tel que $\mathcal{P}(m)$. L'initialisation de la récurrence découle alors de cette implication pour $n = 0$: $\mathcal{P}(0)$ ne peut pas être vraie (l'existence de $0 \leq m < n$ amenant une contradiction).

Exemple 1.2.37

- Variante de la démonstration de l'irrationnalité de $\sqrt{2}$
- Démonstration du théorème de Fermat dans le cas où $n = 4$: il n'existe pas d'entiers non nuls x , y et z tels que $x^4 + y^4 = z^4$ (exemple non développé)

Note Historique 1.2.38

- La démonstration par l'absurde est déjà connue du temps de Pythagore (preuve de l'irrationnalité de $\sqrt{2}$)
- On trouve des prémisses du raisonnement par récurrence dans les *Éléments* d'Euclide, mais cela reste très vague. La vraie naissance du raisonnement par récurrence date de 1654, lorsque Blaise Pascal écrit son *Traité du triangle arithmétique*.
- Pierre de Fermat met en place, à peu près à la même époque, le principe de la descente infinie, mélange entre le principe de récurrence et la démonstration par l'absurde. Ce type de raisonnement aussi apparaît déjà plus ou moins dans les *Éléments* d'Euclide, mais gagne vraiment sa notoriété grâce à Fermat. Il est par exemple utilisé pour montrer le grand théorème de Fermat pour l'exposant 4.

2

Ensembles

Aus dem Paradies, das Cantor uns geschaffen, soll uns niemand vertreiben können.

(David Hilbert)

The finest product of mathematical genius and one of the supreme achievements of purely intellectual human activity.

(David Hilbert, à propos de la théorie de Cantor)

Pour choisir une chaussette plutôt que l'autre pour chaque paire d'une collection infinie, on a besoin de l'axiome du choix. Mais pour les chaussures, ce n'est pas la peine.

(Bertrand Russell)

La notion d'ensemble semble à première vue une notion intuitive évidente, ne nécessitant pas de précautions particulières. Cette notion intuitive est à la base de toutes les mathématiques, depuis leur origine, que ce soit l'arithmétique élémentaire (ensemble d'entiers, puis de divers autres types de nombres, utilisés depuis qu'on sait compter), la géométrie d'Euclide à nos jours (une figure est un sous-ensemble du plan), l'analyse (on étudie des fonctions définies sur des ensembles), ou l'algèbre moderne (étude des structures algébriques, définies comme des ensembles munis d'un certain nombre de lois supplémentaires).

Longtemps, les mathématiciens se sont contentés de ce point de vue intuitif, sans chercher à formaliser cette notion. Ce n'est qu'à l'aube du XX^e siècle qu'on s'est penché sur cette formalisation, qui a bien failli faire vaciller l'édifice mathématique sur ses fondations. En effet, Cantor, puis Russell au travers de son célèbre paradoxe, ont montré qu'on ne pouvait pas se contenter de cette approche intuitive, et que celle-ci amenait des contradictions si on admettait que toute collection pouvait être un ensemble : ainsi, le paradoxe de Russell montre qu'il ne peut pas exister d'ensemble des ensembles. Les mathématiciens pensèrent même un moment qu'il n'était pas possible de donner une formalisation correcte de la notion d'ensemble ; cela aurait signifié ni plus ni moins que la faillite des mathématiques. Heureusement, au prix d'une axiomatique assez lourde, les mathématiciens logiciens de l'époque ont réussi à mettre en place cette formalisation. On peut dire que cette « crise des fondements » a marqué la naissance des mathématiques et de la logique moderne, par une formalisation systématique de toutes les notions utilisées. Depuis, l'édifice mathématique a des fondements solides et ne s'assoit plus sur des sables mouvants. Même la notion d'indécidabilité, dans un premier temps assez choquante, a fini par trouver sa place dans cet édifice solide, par une latitude qu'autorisent ces résultats indécidables dans le choix de l'axiomatique initiale. Ainsi, par exemple, l'axiome du choix dont on parlera dans la suite de ce chapitre étant indécidable (pour l'axiomatique de Zermelo-Frankel), on pourra construire deux théories mathématiques, l'une incluant l'axiome du choix, l'autre, beaucoup plus pauvre, ne l'incluant pas. Ainsi, dans certaines théories, l'axiome

du choix permet d'aller un peu, ou beaucoup plus loin, en permettant notamment de construire certains objets infinis.

En ce qui nous concerne, nous nous contenterons du point de vue intuitif. Nous souleverons tout de même les problèmes que peut engendrer ce point de vue, et nous évoquerons de façon très superficielle le problème de l'axiomatisation de la théorie des ensembles.

I Théorie intuitive des ensembles

I.1 Définition intuitive

Pour définir rigoureusement la notion d'*ensemble*, il faut une axiomatique très complexe, c'est pourquoi nous admettons cette notion. Nous nous contenterons de :

Définition 2.1.1 (Ensemble, point de vue intuitif)

- Un *ensemble* E est une collection d'objets.
- Les objets dont est constituée la collection définissant E sont appelés *éléments de E* .
- On dit que x appartient à E si x est élément de E , et on note $x \in E$.

Note Historique 2.1.2

La notation \in est introduite par l'italien Peano en 1890. Il s'agit d'un epsilon, pour désigner la lettre E de « esti » (« il est » en italien)

Il existe plusieurs façons de décrire un ensemble

Définition 2.1.3 (Définitions d'ensembles)

- Une définition *par énumération* d'un ensemble E est la donnée explicite de tous les éléments de l'ensemble : $E = \{x_1, \dots, x_n\}$
- Une définition *par compréhension* d'un ensemble E est la donnée d'une propriété P caractérisant les éléments de E (parmi les éléments d'un ensemble plus gros F) : $F = \{x \in E \mid P(x)\}$
- Une définition *par induction structurelle* de E est la donnée d'un certain nombre d'éléments de E , et d'une façon de construire, étape par étape les autres éléments de E à partir de ceux donnés.
- Une définition *par constructions* (unions, intersections) est une façon de construire un ensemble à partir d'autres ensembles (ce sera étudié dans le paragraphe suivant)

Remarques 2.1.4

1. Il n'y a pas de notion d'ordre des éléments d'un ensemble E : on peut énumérer les éléments de E dans l'ordre qu'on veut.
2. Il n'y a pas de notion de multiplicité d'un élément : un élément appartient ou n'appartient pas à un ensemble, mais il ne peut pas « appartenir plusieurs fois ». S'il apparaît plusieurs fois dans une énumération des éléments de E , attention au fait qu'il s'agit bien du même élément !
3. L'énumération des éléments d'un ensemble est généralement donnée entre accolades $\{\dots\}$ pour désigner l'ensemble.

Exemples 2.1.5

1. Énumération : $\{1, 3, 7, 9\} = \{3, 9, 7, 1\} = \{1, 1, 3, 3, 3, 7, 9, 9\}$.

2. Compréhension : $\{x \in \mathbb{R} \mid \exists y \in \mathbb{N}, x^2 = y\}$, ou $\{x \in \mathbb{R} \mid x^2 - 4x + 1 \geq 0\}$
3. Induction structurelle :
 - (i) \mathbb{N} (ensemble des entiers naturels) est défini par induction structurelle à partir d'un élément initial 0 et de la construction, pour chaque élément n , d'un successeur $n + 1$.
 - (ii) L'ensemble E tel que $2, 3 \in E$ et si p et q sont dans E , pq est dans E .
 - (iii) L'ensemble des formules propositionnelles, défini comme sous-ensemble de toutes les chaînes de caractère, qui contient les variables propositionnelles, et stable par certaines contructions du type $A \vee B \dots$

Définition 2.1.6 (Sous-ensemble)

Soit E un ensemble. Un sous-ensemble de E est un ensemble F tel que tout élément de F est aussi élément de E :

$$\forall x, x \in F \implies x \in E.$$

On note $F \subset E$.

On a clairement (et cela relève en fait de l'axiomatique), pour F, G et H des sous-ensembles d'un ensemble E :

Proposition 2.1.7 (Principe de double-inclusion)

$F = G$ si et seulement si $F \subset G$ et $G \subset F$.

Proposition 2.1.8 (Transitivité)

Si $F \subset G$ et $G \subset H$ alors $F \subset H$.

Les deux propriétés précédentes nous assurent que \subset définit une relation d'équivalence sur l'ensemble des parties de E , au sens défini dans le chapitre 5.

Notation 2.1.9 (Ensemble vide)

Un ensemble vide, noté \emptyset , est un ensemble ne contenant aucun élément. Il est donc défini par la propriété :

$$\forall x, x \notin \emptyset.$$

Proposition 2.1.10

Un ensemble vide est sous-ensemble de tout ensemble E .

▫ Éléments de preuve.

En effet :

$$\forall x, x \in \emptyset \implies x \in E$$

(la source de l'implication étant toujours fausse, l'implication est vraie)

▷

Corollaire 2.1.11 (Unicité de l'ensemble vide)

Il existe un unique ensemble vide.

▫ **Éléments de preuve.**

L'existence, intuitivement évidente, provient de l'axiomatique. Pour l'unicité, si on a deux ensembles vides, ils vérifient la double-inclusion d'après la propriété précédente. ▷

Le seul sous-ensemble de \emptyset est \emptyset lui-même.

Terminologie 2.1.12 (singleton)

On appelle *singleton* un ensemble constitué d'un unique élément, donc de la forme $\{a\}$.

Définition 2.1.13 (Cardinal, notion intuitive)

Intuitivement, le cardinal d'un ensemble correspond à sa taille. Pour un ensemble fini, il s'agit du nombre de ses éléments. On note dans ce cas $\text{Card}(E)$ ou $|E|$ le cardinal de E .

Exemple 2.1.14

- $\text{Card}(\emptyset) = 0$.
- $\text{Card}(\{\emptyset\}) = 1$.

On peut définir, comme on le verra plus tard, une notion de cardinal pour des ensembles infinis, mais l'intuition en est moins évidente. Par exemple, \mathbb{N} et \mathbb{Q} ont même cardinal !

I.2 Ensemble des parties d'un ensemble

Une partie d'un ensemble est un autre nom donné à un sous-ensemble.

Notation 2.1.15 (Ensemble des parties d'un ensemble)

On note $\mathcal{P}(E)$ l'ensemble des parties de E , c'est-à-dire l'ensemble dont les éléments sont les sous-ensembles de E : $F \in \mathcal{P}(E) \iff F \subset E$.

Remarque 2.1.16

Ainsi que nous l'avons évoqué en début de section, toute collection d'éléments ne peut pas nécessairement être considérée comme un ensemble. Par exemple, on ne peut pas parler de l'ensemble des ensembles. Il n'est donc *a priori* pas évident que $\mathcal{P}(E)$ soit toujours un ensemble. En fait, cela fait partie des axiomes que l'on pose pour définir la théorie des ensembles.

Proposition 2.1.17

Pour tout ensemble E , $\emptyset \in \mathcal{P}(E)$ et $E \in \mathcal{P}(E)$. Ainsi, $\mathcal{P}(E)$ n'est jamais vide. Il contient toujours \emptyset et E .

Exemples 2.1.18

Déterminer $\mathcal{P}(E)$ dans les cas suivants :

1. $E = \emptyset$
2. $E = \{1\}$
3. $E = \{1, 2, 3\}$
4. $E = \{\emptyset, \{\emptyset\}\}$

Par commodité, nous introduisons les notations suivantes :

Notation 2.1.19 (Ensemble des parties de cardinal fixé)

Soit E un ensemble fini (à prendre au sens intuitif, ainsi que la notion de cardinal), et $k \in \mathbb{N}$. On note $\mathcal{P}_k(E)$ l'ensemble des parties de E de cardinal k .

Pour simplifier les écritures, nous utiliserons également les deux notations raccourcies suivantes, non universelles (donc à redéfinir) :

Notation 2.1.20 (Ensemble des parties de $\llbracket 1, n \rrbracket$)

Soit n et k deux entiers naturels. On note :

1. $\mathcal{P}(n) = \mathcal{P}(\llbracket 1, n \rrbracket)$
2. $\mathcal{P}_k(n) = \mathcal{P}_k(\llbracket 1, n \rrbracket)$.

I.3 Opérations sur les parties d'un ensemble

Nous étudions dans ce paragraphe les constructions classiques permettant de définir des ensembles à partir d'autres. On considère dans ce paragraphe les constructions internes à $\mathcal{P}(E)$ pour un ensemble E donné : autrement dit, à partir de deux sous-ensembles A et B de E , on construit d'autres sous-ensembles de E . Dans tout ce paragraphe, E désigne un ensemble, et A, B, C, \dots désignent des sous-ensembles de E

Définition 2.1.21 (Union, intersection, différence ensembliste, complémentation)

1. **Intersection** : l'intersection de A et B , notée $A \cap B$, est l'ensemble des éléments de E contenus à la fois dans A et dans B :

$$x \in A \cap B \iff (x \in A) \wedge (x \in B).$$

Ainsi, $A \cap B = \{x \in E \mid (x \in A) \wedge (x \in B)\}$.

2. **Union** : l'union de A et B , notée $A \cup B$, est l'ensemble des éléments de E contenus soit dans A soit dans B :

$$x \in A \cup B \iff (x \in A) \vee (x \in B).$$

3. **Déférence ensembliste** : La différence ensembliste $A \setminus B$ (A privé de B , ou A moins B) est l'ensemble des éléments de A qui ne sont pas dans B :

$$x \in A \setminus B \iff (x \in A) \wedge (x \notin B).$$

4. **Complémentation** : Si $B \subset A$, on appelle complémentaire de B dans A la différence $A \setminus B$. Plusieurs notations sont utilisées :

$$A \setminus B = A - B = \complement_A B = B^c = {}^c B = \overline{B}.$$

Les trois dernières notations ne faisant pas référence à A , elles sont utilisées lorsqu'il n'y a pas d'ambiguïté sur l'univers total A dans lequel on prend le complémentaire, généralement égal à l'ensemble E tout entier, qui est fixé dans le contexte (par exemple, la dernière notation est utilisée pour désigner l'événement contraire en probabilité).

5. **Déférence symétrique** : La différence symétrique $A \Delta B$ est l'ensemble des éléments x appartenant à l'un des deux ensembles A ou B , mais pas à l'autre :

$$A \Delta B = \{x \in E \mid x \in A \setminus B \text{ ou } x \in B \setminus A\} = A \setminus B \cup B \setminus A = (A \cup B) \setminus (A \cap B)$$

Ces constructions sont illustrées dans les figures 2.1, 2.2, 2.3 et 2.4.

Exemple 2.1.22

1. Soit $A, B \subset E$. Décrire de la façon la plus simple possible tous les sous-ensembles de E définies par les différentes zones du diagramme en patates.
2. Soit A_1, \dots, A_n des sous-ensembles de E . Combien de « zones élémentaires » définissent ces sous-ensembles ? Les décrire sous forme d'intersections.

Un certain nombre de propriétés de l'union et de l'intersection découlent de façon immédiate des règles logiques correspondantes :

Proposition 2.1.23 (Associativité, commutativité)

1. $A \cap B = B \cap A$ (*commutativité de \cap*)
2. $A \cup B = B \cup A$ (*commutativité de \cup*)
3. $(A \cap B) \cap C = A \cap (B \cap C)$ (*associativité de \cap*)
4. $(A \cup B) \cap C = A \cup (B \cup C)$ (*associativité de \cup*)

Proposition 2.1.24 (Distributivités)

L'union est distributive sur l'intersection, et réciproquement :

1. $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$ (*distributivité de \cap sur \cup*)
2. $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$ (*distributivité de \cup sur \cap*)

Les propriétés suivantes traduisent certaines propriétés de l'inclusion, vues comme relation d'ordre sur un ensemble de parties :

Proposition 2.1.25 (\cap et \cup vus comme bornes supérieures et inférieures)

1. *L'intersection de A et B est la borne inférieure de A et B dans $\mathcal{P}(E)$, c'est-à-dire le plus grand des ensembles (au sens de l'inclusion) inclus à la fois dans A et B . En d'autres termes :*
 - $A \cap B \subset A$ et $A \cap B \subset B$
 - Pour tout C , $(C \subset A) \wedge (C \subset B) \implies C \subset A \cap B$.
2. *L'union de A et B est la borne supérieure de A et B dans $\mathcal{P}(E)$, c'est-à-dire le plus petit des ensembles contenant à la fois A et B . En d'autres termes :*
 - $A \subset A \cup B$ et $B \subset A \cup B$
 - Pour tout C , $(A \subset C) \wedge (B \subset C) \implies A \cup B \subset C$.

Corollaire 2.1.26

1. $A \cap \emptyset = \emptyset$, $A \cup \emptyset = A$
2. Si $A \subset B$, $A \cap B = A$ et $A \cup B = B$.

Définition 2.1.27 (Ensembles disjoints)

Deux ensembles A et B sont disjoints si $A \cap B = \emptyset$.

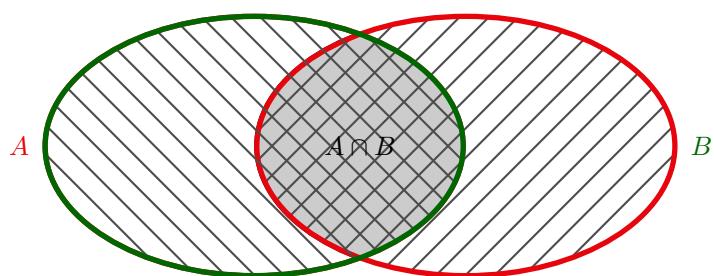


FIGURE 2.1 – Intersection de deux ensembles

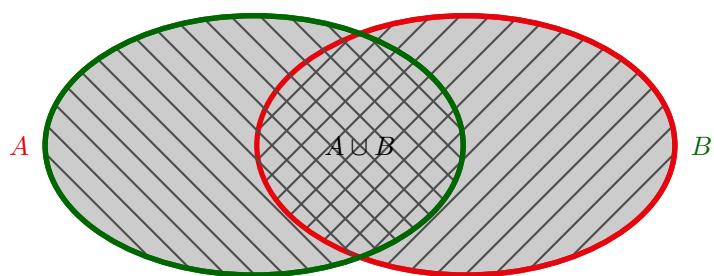


FIGURE 2.2 – Union de deux ensembles

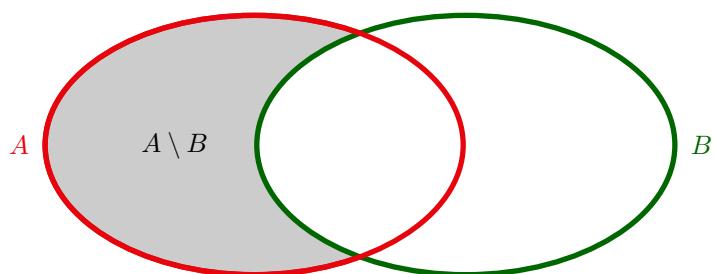


FIGURE 2.3 – Différence ensembliste

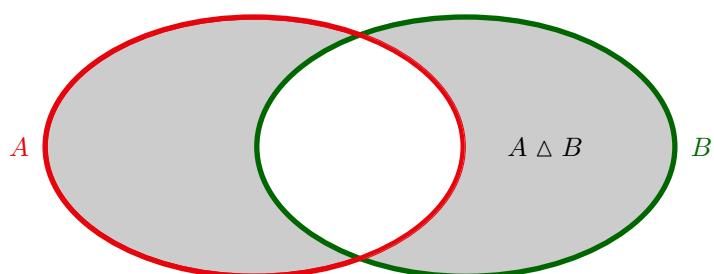


FIGURE 2.4 – Différence symétrique

Avertissement 2.1.28

Attention à ne pas confondre *disjoint* et *distinct* !

Notation 2.1.29 (union disjointe)

Si A et B sont disjoints, l'union $A \cup B$ peut être notée $A \sqcup B$. Cette notation *affirme* que A et B sont disjoints.

La complémentation correspond bien entendu à la négation ensembliste, et est donc régie par des lois similaires à celles de la négation. En particulier, les lois de De Morgan ont leur analogue ensembliste, portant d'ailleurs le même nom.

Proposition 2.1.30 (Lois de De Morgan, figure 2.5)

1. $\complement_E(A \cup B) = \complement_E A \cap \complement_E B$
2. $\complement_E(A \cap B) = \complement_E A \cup \complement_E B$

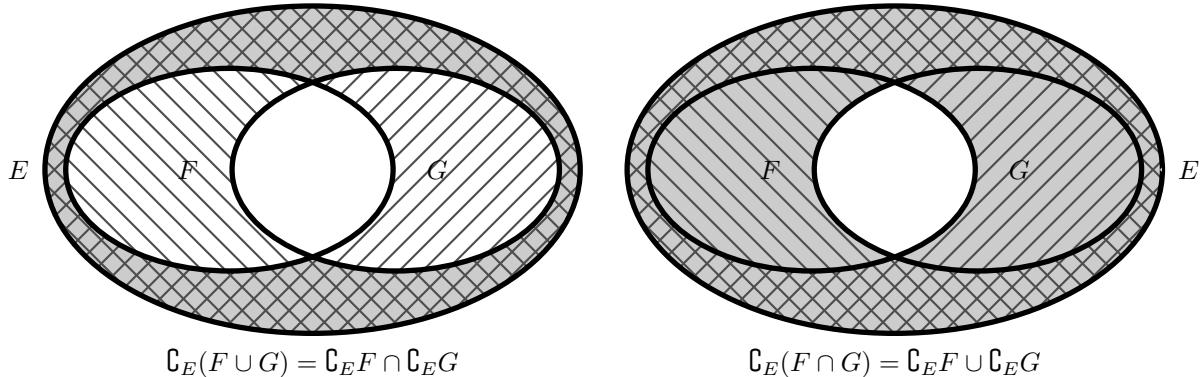


FIGURE 2.5 – Lois de De Morgan

Corollaire 2.1.31 (Décroissance du complémentaire)

L'application $A \mapsto \complement_E A$ définie sur $\mathcal{P}(E)$ ordonné par inclusion est décroissante :

$$\forall A, B \in \mathcal{P}(E), \quad A \subset B \implies \complement_E B \subset \complement_E A.$$

▫ Éléments de preuve.

Appliquer la première loi de De Morgan à $A \cup B = B$.

▷

Le complémentaire est caractérisé par la propriété suivante :

Proposition 2.1.32 (caractérisation du complémentaire)

$\complement_E A$ est l'unique sous-ensemble B de E tel que $E = A \sqcup B$.

▫ Éléments de preuve.

C'est en gros le principe du tiers exclus.

▷

On en déduit facilement :

Corollaire 2.1.33

1. $\complement_E E = \emptyset$
2. $\complement_E \emptyset = E$
3. $\complement_E \complement_E A = A$ (*la complémentation est involutive*)

Nous terminons enfin ce paragraphe par une propriété importante de la différence symétrique : si, pour construction, la commutativité de Δ est évidente, son associativité l'est beaucoup moins.

Théorème 2.1.34 (Associativité de Δ)

Soit A , B et C des sous-ensembles de E . Alors

$$(A \Delta B) \Delta C = A \Delta (B \Delta C).$$

▫ Éléments de preuve.

On peut le démontrer en manipulant les expressions logiques. On le laisse en exercice. On verra un peu plus loin comment l'utilisation des fonctions caractéristiques trivialise la question. ▷

Il faut bien garder en tête la correspondance entre opérations ensemblistes et connecteurs logiques :

$\cup \equiv \vee$	$x \in A \cup B \iff (x \in A) \vee (x \in B)$
$\cap \equiv \wedge$	$x \in A \cap B \iff (x \in A) \wedge (x \in B)$
$\complement \equiv \neg$	si $x \in A$, $(x \in \complement_A B \iff \neg(x \in B))$
$\subset \equiv \implies$	$A \subset B \iff (\forall x, x \in A \implies x \in B),$
$\Delta \equiv$ ou exclusif	$x \in A \Delta B \iff x \in A \text{ ou (exclusif) } x \in B$
$= \equiv \iff$	$A = B \iff (\forall x, x \in A \iff x \in B).$

I.4 Union et intersection d'une famille de sous-ensembles

Nous généralisons certaines constructions vues dans le paragraphe précédent au cas d'un plus grand nombre, fini ou infini d'ensembles.

Définition 2.1.35 (unions et intersections sur une famille)

Soit $(A_i)_{i \in I}$ une famille (finie ou infinie) d'ensembles. On définit alors l'union et l'intersection de cette famille par :

- $\bigcup_{i \in I} A_i = \{x \mid \exists i \in I, x \in A_i\}$
- $\bigcap_{i \in I} A_i = \{x \mid \forall i \in I, x \in A_i\}$

L'existence de tels ensembles est un axiome de la théorie des ensembles.

Proposition 2.1.36

Si $I = \llbracket 1, n \rrbracket$ (et plus généralement si I est fini, après numérotation de ses éléments), $\bigcup_{i \in I} A_i$ correspond à l'itération de l'union de 2 ensembles.

▫ Éléments de preuve.

C'est une récurrence assez immédiate. ▷

Notation 2.1.37 (Union disjointe)

Si les A_i sont deux à deux disjoints, on peut noter l'union $\bigsqcup_{i \in I} A_i$.

Proposition 2.1.38 (propriétés liées aux unions et intersections quelconques)

Un certain nombre de propriétés vues dans le paragraphe précédent se généralisent aux unions et intersections quelconques. Étant donné $(A_i)_{i \in I}$ une famille d'ensembles, et B un ensemble, tous inclus dans un ensemble E :

1. $B \cap \left(\bigcup_{i \in I} A_i \right) = \bigcup_{i \in I} (B \cap A_i)$ (*distributivité*)
2. $B \cup \left(\bigcap_{i \in I} A_i \right) = \bigcap_{i \in I} (B \cup A_i)$ (*l'autre distributivité*)
3. $\complement_E \left(\bigcup_{i \in I} A_i \right) = \bigcap_{i \in I} (\complement_E A_i)$ (*loi de De Morgan*)
4. $\complement_E \left(\bigcap_{i \in I} A_i \right) = \bigcup_{i \in I} (\complement_E A_i)$ (*loi de De Morgan*)

▫ Éléments de preuve.

Les deux premières égalités découlent de propriétés logiques similaires (ou par double-inclusion), les deux dernières de la négation des quantificateurs et des connecteurs logiques. ▷

I.5 Partitions

Définition 2.1.39 (Partition d'un ensemble)

Une partition de E est un sous-ensemble de $\mathcal{P}(E)$ dont les éléments sont des sous-ensembles non vides de E , deux à deux disjoints et d'union égale à E .

Par exemple, une partition finie, est un ensemble $\{A_1, \dots, A_n\}$ de sous-ensembles de E tels que :

1. $\forall i \in \llbracket 1, n \rrbracket, A_i \neq \emptyset,$
2. $\forall (i, j) \in \llbracket 1, n \rrbracket^2, i \neq j \implies A_i \cap A_j = \emptyset,$
3. $A_1 \cup \dots \cup A_n = E.$

Les sous-ensembles A_i sont appelés *parts* de la partition, et n est appelé *longueur* de la partition.

Puisque les parts d'une partition ne peuvent pas être vides, il n'existe qu'un nombre fini de partitions d'un ensemble donné. Ce nombre est égal, par définition, au nombre de Bell B_n .

Définition 2.1.40 (Variantes autour des partitions)

- Nous désignerons par *partition ordonnée* de E un n -uplet (A_1, \dots, A_n) tel que l'ensemble $\{A_1, \dots, A_n\}$ obtenu en oubliant l'ordre des parts soit une partition de E .
- Nous désignerons par *partage* (terminologie non standard) ou *partition à parts éventuellement vide* un ensemble $\{A_1, \dots, A_n\}$ vérifiant les points 2 et 3 de la définition d'une partition, mais pas nécessairement le point 1.
- Nous aurons alors aussi une notion de *partage ordonné*, ou de *partition ordonnée à parts éventuellement vides*.

I.6 Produit cartésien

La dernière construction que nous voyons est une construction « externe » dans le sens où, contrairement aux précédentes, tout ne se passe pas parmi les sous-ensembles d'un ensemble donné E .

Définition 2.1.41 (Produit cartésien, figure 2.6)

Soit E et F deux ensembles. Le produit cartésien de E et F , noté $E \times F$, est l'ensemble des couples (a, b) , tels que $a \in E$ et $b \in F$.

Ainsi, pour chaque élément a de E , on dispose d'une copie complète de B , identifiée au produit $\{a\} \times B$, c'est-à-dire à l'ensemble des couples (a, b) , pour b parcourant B . On peut bien entendu aussi le voir en inversant le rôle de A et B . Il faut relier cette notion à celle de **rectangle plein**, qui n'est autre que le produit cartésien de deux intervalles de \mathbb{R}

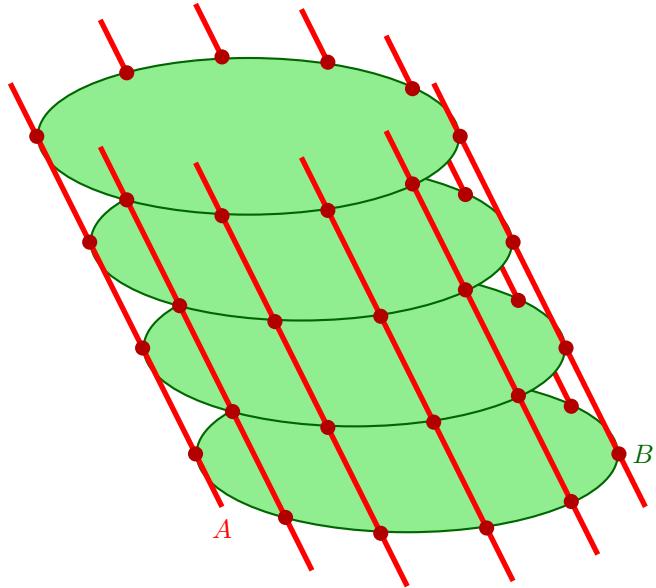


FIGURE 2.6 – Produit cartésien de deux ensembles

Notation 2.1.42

On note $E^2 = E \times E$, et plus généralement E^n pour un produit cartésien à n termes.

Proposition 2.1.43 (propriétés du produit cartésien)

Soit E , E' et F des ensembles. Alors :

1. $E \times F = \emptyset \iff (E = \emptyset) \vee (F = \emptyset)$;
2. $(E \cup E') \times F = (E \times F) \cup (E' \times F)$;
3. $(E \cap E') \times F = (E \times F) \cap (E' \times F)$;

I.7 Fonction caractéristique (ou indicatrice)

Nous introduisons dans ce paragraphe un outil souvent assez efficace pour étudier les constructions internes à $\mathcal{P}(E)$, E étant un ensemble fixé.

En admettant momentanément la notion intuitive de fonction, on peut associer, à tout sous-ensemble de E , une fonction de E dans $\{0, 1\}$.

Définition 2.1.44 (Fonction indicatrices, ou caractéristiques, d'un ensemble)

Soit E un ensemble et A un sous-ensemble de E . La fonction indicatrice (ou caractéristique) de A , notée $\mathbb{1}_A$, est la fonction de E dans $\{0, 1\}$ définie par :

$$\begin{aligned}\mathbb{1}_A : E &\longrightarrow \{0, 1\} \\ x &\mapsto \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A. \end{cases}\end{aligned}$$

On trouve parfois la notation χ_A au lieu de $\mathbb{1}_A$.

Un certain nombre de propriétés élémentaires se traduisent sur les fonctions indicatrices. Les constructions élémentaires se transcrivent également très bien au niveau des fonctions indicatrices.

Proposition 2.1.45 (propriétés des fonctions indicatrices)

Soit E un ensemble, et A, B des sous-ensembles de E . On a alors :

1. $A = \{x \in E \mid \mathbb{1}_A(x) = 1\}$.
2. $\mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B$
3. $\mathbb{1}_{\complement_E A} = 1 - \mathbb{1}_A$
4. Si A et B sont disjoints, $\mathbb{1}_{A \sqcup B} = \mathbb{1}_A + \mathbb{1}_B$
5. Dans le cas général, $\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A \cap B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \mathbb{1}_B$

Certaines propriétés se traduisent sur les fonctions indicatrices. Par exemple, A et B sont disjoints si et seulement si $\mathbb{1}_A \mathbb{1}_B = 0$, ou bien $\{A_1, \dots, A_n\}$ est une partition de E si et seulement si les A_i sont non vides et $\mathbb{1}_{A_1} + \dots + \mathbb{1}_{A_n} = 1$.

Les fonctions indicatrices peuvent donner des démonstrations rapides et formelles de certaines propriétés ensemblistes.

Exemple 2.1.46

Redémontrer la distributivité de l'intersection sur l'union à l'aide des fonctions caractéristiques.

Toute la puissance des fonctions caractéristiques apparaît dans la preuve du résultat suivant, assez pénible si vous cherchez à la faire par des moyens élémentaires, très élégante et rapide si vous utilisez convenablement les fonctions indicatrices.

Théorème 2.1.47 (Associativité de la différence symétrique)

Soit A, B et C des sous-ensembles de E . On a alors

$$(A \Delta B) \Delta C = A \Delta (B \Delta C).$$

▫ Éléments de preuve.

Il suffit de remarquer que $\mathbb{1}_{A \Delta B} \equiv \mathbb{1}_A + \mathbb{1}_B$ [2].

▷

II Paradoxes ensemblistes et axiomatisation

II.1 La crise des fondements

Note Historique 2.2.1

- En 1896, alors qu'il essaye de donner une construction de \mathbb{N} et des ordinaux plus généraux (nombres transfinis), Georg Cantor se rend compte d'une contradiction dans sa théorie des ensembles (il définit un ensemble comme une collection quelconque d'objets). En effet, il prouve que pour tout ensemble E , $\text{Card}(\mathcal{P}(E)) > \text{Card}(E)$ (dans le sens qu'il donne à la notion de cardinal, défini à l'aide des ordinaux). Cette contradiction empêche de parler de l'ensemble des ensembles : en effet, par définition, l'ensemble de ses parties vérifierait $\mathcal{P}(E) \subset E$, ce qui est incompatible avec l'inégalité ci-dessus.
Il n'en parle à personne car sa théorie est menacée, à part à David Hilbert, qui fait autorité à l'époque, et à qui il envoie une lettre.
- En 1897, Cesare Burali-Forti parvient à la même conclusion et publie son résultat, sans vraiment être persuadé que cela remet tout en cause.
- En 1901, le logicien et philosophe anglais Bertrand Russell exprime le premier paradoxe simple prouvant que toute collection ne peut pas être un ensemble. En effet, en définissant

$$E = \{\text{ensembles } F \text{ qui ne se contiennent pas eux-mêmes}\},$$

si E était un ensemble :

- * si $E \notin E$, alors par définition de E , E est élément de E , d'où une contradiction ;
- * si $E \in E$, alors, par définition de E , E n'est pas élément de E , d'où une contradiction.

Cet argument très simple montre que E ne peut pas être un ensemble.

Ce paradoxe est connu sous le nom de « paradoxe de Russell » ou parfois « paradoxe du barbier ». En effet, la situation s'apparente à celle d'un barbier qui rase tout homme qui ne se rase pas lui-même. Ce barbier peut-il être un homme ?

- Ce paradoxe a en fait été envoyé par Russell au logicien et philosophe allemand Gottlob Frege, suite à la parution du premier volume de son ouvrage *Les fondements de l'arithmétique*, pour lui prouver que son ouvrage reposait sur une contradiction. Frege publie tout de même le second volume, en lui adjoignant un appendice dans lequel il fait l'aveu et le constat sans doute les plus désarmants de toute l'histoire des mathématiques :

« Pour un écrivain scientifique, il est peu d'infortunes pires que de voir l'une des fondations de son travail s'effondrer alors que celui-ci s'achève. C'est dans cette situation inconfortable que m'a mis une lettre de M. Bertrand Russell, alors que le présent volume allait paraître »

Par la suite, Frege cessa presque entièrement ses travaux mathématiques.

- D'autres paradoxes, recherchant des formes amusantes, virent alors le jour, comme le paradoxe de Jules Richard (1905), dont un énoncé simplifié est le paradoxe de Berry (formulé par Russell en 1906, il l'attribue à un bibliothécaire londonien du nom de Berry) :

Soit E l'« ensemble » des entiers qui ne peuvent pas se définir en moins de 16 mots. Cet ensemble est non vide car son complémentaire est fini (car il y a un nombre fini de mots dans la langue française).

Par les propriétés des sous-ensembles de \mathbb{N} , cet « ensemble » admet un plus petit élément, qui peut être défini par : « le plus petit entier qui ne peut pas se définir en moins de seize mots », définition qui n'utilise que 15 mots !

II.2 Tentatives d'axiomatisation

Note Historique 2.2.2

- De nombreuses tentatives d'axiomatisation de la théorie des ensembles à la suite de cette crise des fondements, toutes n'ont pas été fructueuses
- Le choix qui s'est imposé est finalement l'axiomatique de Zermelo-Fraenkel à laquelle on ajoute ou non l'axiome du choix.

La théorie des ensembles de Zermelo-Fraenkel définit les ensembles comme étant des objets satisfaisant à un certain nombre d'axiomes. Dans cette théorie, les éléments eux-mêmes sont tous des ensembles. Les nombres (les entiers relatifs dans un premier temps) sont alors définis comme étant des ensembles en particulier, par exemple à la façon des ordinaux de Cantor.

Axiome 2.2.3 (les axiomes de la théorie de Zermelo-Fraenkel, HP)

Voici les noms des différents axiomes, et leur interprétation intuitive :

- *Axiome d'extensionnalité : c'est lui qui dit que deux ensembles sont égaux si et seulement si ils ont mêmes éléments. Cet axiome est notamment à la base du principe de double-inclusion.*
- *Axiome de la paire : il affirme l'existence des paires $\{a, b\}$, lorsque a et b sont deux ensembles. En particulier, l'axiome de la paire affirme aussi l'existence des singletons $\{a\}$, pour un ensemble a (prendre $a = b$!)*
- *Axiome de la réunion : il donne la possibilité de construire des unions des éléments d'un ensemble (ces éléments étant eux-même des ensembles)*
- *Axiome des parties : il affirme que si a est un ensemble, alors $\mathcal{P}(a)$ aussi.*
- *Schéma d'axiome (i.e. série d'axiomes) de compréhension : il permet en particulier de définir un ensemble par compréhension (comme sous-ensemble d'un ensemble donné, constitué des éléments vérifiant une certaine propriété)*
- *Axiome de l'infini : il donne l'existence de l'infini, et notamment des ensembles infinis.*
- *Axiome de fondation : il dit en particulier qu'un ensemble ne peut pas s'appartenir (on ne peut pas avoir $x \in x$).*

À ces différents axiomes, on ajoute ou non (suivant la théorie) l'axiome du choix. Cet axiome du choix est indécidable à partir des axiomes de Zermelo-Fraenkel, et il est actuellement couramment admis qu'il doit être considéré comme vrai.

Axiome 2.2.4 (Axiome du choix, HP)

Soit I un ensemble, et pour tout $i \in I$, E_i un ensemble. Alors il existe une fonction $f : I \longrightarrow \bigcup_{i \in I} E_i$ telle que pour tout $i \in I$, $f(i) \in E_i$, appelée fonction de choix.

Autrement dit, étant donné une famille d'ensembles, on peut choisir un élément dans chaque ensemble, d'où le nom de cet axiome. Évidemment, si I est fini, ce n'est pas très étonnant, et cela résulte de l'axiome de récurrence (car par définition-même de l'interprétation du symbole \exists , on peut toujours choisir **un** élément d'un ensemble non vide). N'invoquez donc l'axiome du choix que pour un choix infini, c'est dans cette situation qu'il est pertinent.

Remarque 2.2.5

1. L'axiome de la paire, couplée à l'axiome de l'union, permet de considérer $A \cup B$ pour tous ensembles A et B
2. En itérant cet argument, on peut considérer l'union de n ensembles.
3. L'intersection d'une famille quelconque (non réduite à l'ensemble vide) d'ensembles s'obtient par compréhension (axiome de compréhension, en se plaçant globalement dans un des ensembles donnés).
4. L'union d'une famille quelconque pose davantage de problèmes ; il faut déjà préciser ce qu'on entend par famille $(a_i)_{i \in I}$: il s'agit d'une application d'un ensemble I dans un autre ensemble E qui à $i \in I$ associe a_i . Ainsi, les a_i doivent eux-même être des éléments d'un ensemble. C'est le cas en particulier si les a_i sont tous des sous-ensembles d'un même ensemble B (dans ce cas, les a_i sont tous éléments de $\mathcal{P}(B)$, qui est un ensemble, d'après l'axiome des parties). Dans ce cas, l'image $\{a_i, i \in I\}$ de cette famille est un ensemble (on peut la définir par compréhension), donc on peut considérer l'union de ses éléments (axiome de l'union).
5. L'existence (et l'unicité) de l'ensemble vide provient de l'axiome de compréhension, à partir d'un ensemble A quelconque (on sait qu'il existe au moins un ensemble, d'après l'axiome de l'infini ;

de toute façon, si ce n'était pas le cas, la théorie serait bien pauvre). L'ensemble vide peut alors se définir par compréhension de la façon suivante :

$$\emptyset = \{x \in A \mid x \neq x\}.$$

6. Le produit cartésien se définit à l'aide de couples, un couple (a, b) de $A \times B$ étant défini comme la paire $\{a, \{a, b\}\}$.

Remarque 2.2.6

Interprétez la citation de Bertrand Russell en début de chapitre.

Applications

« Les mathématiciens n'étudient pas des objets mais les relations entre ces objets. »

(Henri Poincaré)

Une civilisation constituée de groupes de personnes n'interagissant pas entre eux serait assez pauvre. Ce sont les relations entre groupes d'individus qui permettent de comparer, d'apprendre, de progresser, de transmettre, que ce soient des relations internes à un groupe donné, ou des relations d'un groupe à un autre. Une civilisation sans relation est vouée à la stérilité et à l'immobilisme, et donc à l'extinction.

Il en est de même pour les ensembles mathématiques. La théorie axiomatique des ensembles est certes très belle en soi, mais si on n'y rajoute pas une couche, elle est d'un intérêt assez limité pour le mathématicien recherchant le débouché concret (on a tendance à oublier que ce débouché concret a longtemps été la motivation-même des scientifiques, y compris mathématiciens). Comme dans le cas d'une civilisation, il faut faire interagir les ensembles, il faut créer des relations permettant de comparer les éléments d'un ensemble d'une façon ou d'une autre. Ce n'est qu'ainsi qu'on peut donner vie aux ensembles.

Nous étudions dans ce chapitre un type particulier de relations entre deux ensembles. Il s'agit de la notion d'application, ou de fonction, qu'on peut considérer comme synonyme ou non, suivant qu'on est puriste ou non (dans le sens inverse).

I Qu'est-ce qu'une application ?

Définition 3.1.1 (Application, définition intuitive)

Soit E et F deux ensembles. Une application f est un objet qui à tout élément x de E associe un élément $f(x)$ de F .

C'est bien sûr cette interprétation intuitive qu'il faut garder à l'esprit pour bien comprendre ce qu'est une application. Mais cela ne décrit pas bien la nature précise d'une application. Pour savoir à quoi précisément correspond une application, il faut une définition plus rigoureuse, basée sur des objets déjà définis. Il faut donc traduire de façon ensembliste cette action d'associer un élément à un autre.

Définition 3.1.2 (Application, définition formelle ; graphe ; image)

Une application f est un triplet d'ensembles (E, F, G) tel que :

- G est un sous-ensemble de $E \times F$, appelé graphe de f
- pour tout x dans E , il existe un et un seul élément y de F tel que $(x, y) \in G$. Cet élément y est noté $f(x)$.

L'ensemble E est appelé ensemble de départ, ou ensemble source, de l'application f . L'ensemble F est appelé ensemble d'arrivée de l'application f .

Définir une application nécessite donc la donnée de E , de F et du graphe G , ce qui revient à définir, d'une façon ou d'une autre, un élément $f(x)$ pour tout $x \in E$. L'ensemble de ces données est souvent synthétisé par la notation suivante :

$$\begin{array}{rccc} f : & E & \longrightarrow & F \\ & x & \longmapsto & f(x), \end{array}$$

en remplaçant $f(x)$ par son expression, par exemple $\cos(x)$.

Avertissement 3.1.3

Une expression $f(x)$ ne désigne pas une application, mais seulement la valeur d'une application en un point. Parler de l'application, ou la fonction $x \cos(x)$ n'a pas de sens. Tout au plus est-il acceptable de parler de l'application (ou la fonction) $x \mapsto x \cos(x)$, même si les ensembles E et F sont omis dans cette notation.

Note Historique 3.1.4

- La notion est ancienne, mais reste longtemps vague et mal définie, sans réelle notation.
- En 1694, Leibniz est le premier à parler de « fonction d'un point M », puis Newton parle de « fluente du temps t ».
- Leibniz est le premier à proposer une notation, notamment dans le cas où on utilise plusieurs fonctions : il propose $\bar{x} | \underline{1}$ et $\bar{x} | \underline{2}$ pour ce qu'on noterait actuellement $f_1(x)$ et $f_2(x)$.
- Euler introduit en 1734 la notation fx , encore en vigueur actuellement (au parenthésage près).

Remarque 3.1.5 (Application ou fonction ?)

- Actuellement, une fonction f de E dans F désigne une application de E dans F , où E est un sous-ensemble de F , appelé domaine de définition de f . Ainsi, une fonction est une « application par nécessairement définie sur E tout entier ».
- L'énoncé pris dans l'autre sens est plus rigoureux : « Une application est une fonction définie sur tout son domaine ».
- Historiquement, une fonction désignait plutôt une application à valeurs numériques.
- Les deux mots sont souvent utilisés indifféremment et on ne va pas en faire toute une histoire ! En effet, cet abus de langage n'est pas très gênant :
 - * Comme vous l'avez compris, une application est une fonction, donc utiliser la terminologie « fonction » n'est pas incorrect.
 - * Inversement, une fonction de E dans F définit une unique application $\tilde{f} : E' \rightarrow F$ où E' est le domaine de définition de f . L'abus de langage consistant à parler de l'application f se comprend alors très bien en considérant la restriction de f à son domaine.

Lorsque f est une application d'un sous-ensemble de \mathbb{R} dans \mathbb{R} , le graphe de f est un sous-ensemble de \mathbb{R}^2 . On visualise une fonction f en représentant son graphe dans le plan muni d'un repère (figure 3.1)

Lorsque E et F sont des ensembles finis, on peut représenter l'application f sous forme d'un *diagramme sagittal* : on représente les éléments de E d'un côté, ceux de F d'un autre côté, l'application f est alors représentée par une série de flèches reliant les éléments x de E à leur image $f(x)$ dans F .

Par exemple, l'application de $[1, 6]$ dans $[1, 4]$ définie par $f(1) = 3, f(2) = 1, f(3) = 1, f(4) = 4, f(5) = 3$ et $f(6) = 2$ sera représentée par le diagramme sagittal de la figure 3.2

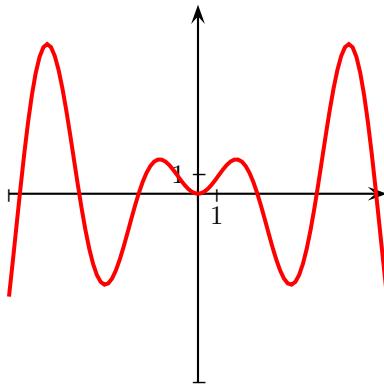
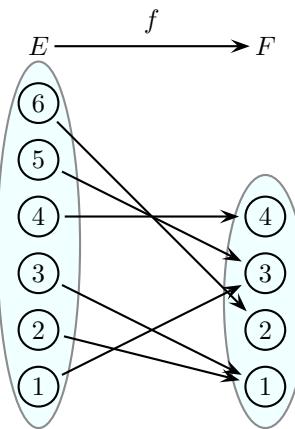
FIGURE 3.1 – Graphe de $x \mapsto x \sin x$ 

FIGURE 3.2 – Diagramme sagittal

Exemples 3.1.6

1. Application identique (identité)
2. Soit $E \subset F$. L'injection canonique $i : E \rightarrow F$.
3. Soit $E \subset F$. La fonction indicatrice de E dans F , $\mathbb{1}_E : F \rightarrow \{0, 1\}$ (la terminologie « fonction » renvoie ici à l'acceptation historique)
4. La projection canonique $p_E : E \times F \rightarrow E$.

Notation 3.1.7

On note F^E l'ensemble des applications de E dans F .

Nous avons déjà rencontré un type d'applications pour lesquelles la dépendance par rapport à la variable est notée indiciellement :

Définition 3.1.8 (Familles et suites)

- Une *famille* $(x_i)_{i \in I}$ d'éléments d'un ensemble E , indexée par un ensemble I , est une application $x : I \longrightarrow E$. Dans cette situation, on utilise plutôt une notation indicielle x_i à la place de $x(i)$. L'ensemble I est appelé l'ensemble des indices de la famille $(x_i)_{i \in I}$
- Une *suite* d'éléments d'un ensemble E est une famille indexée par \mathbb{N} , ou éventuellement par un

ensemble $\{n \in \mathbb{N} \mid n \geq n_0\}$. Nous étudierons notamment les suites réelles ($E = \mathbb{R}$) et les suites complexes ($E = \mathbb{C}$)

Remarque 3.1.9

Une suite est donc un cas particulier de famille. Dans ce cas particulier, on dispose d'une notion d'ordre des éléments, induit par l'ordre usuel des entiers relatifs : on range alors naturellement les éléments de la suite dans l'ordre de leur indice. En général, on ne peut faire cela pour une famille quelconque. C'est ce qui fait qu'on obtient souvent pour les suites des résultats et des théories qu'on ne peut pas ou pas facilement généraliser aux familles quelconques. On trouvera plus tard cette situation dans le problème de la sommation (séries convergentes *versus* familles sommables), où l'ordre de sommation a son importance.

Voici quelques moyens de définir une application ou une fonction (la liste n'est pas exhaustive) :

- **Définition par une formule explicite :**

$$\forall x \in \mathbb{R}, \quad f(x) = \frac{\sin x}{1 + x^2}$$

- **Définition par disjonction de cas (application définie par morceaux) :**

$$\forall x \in \mathbb{R}, \quad f(x) = \begin{cases} e^x & \text{si } x < 0 \\ 0 & \text{si } x = 0 \\ \frac{1}{x} & \text{si } x > 0. \end{cases}$$

Attention dans ce cas :

- * si les différents cas ne couvrent pas la totalité du domaine, cela définit une fonction et non une application ;
- * si les différents cas sont redondants (par exemple définition pour $x \leq 0$ et $x \geq 0$), il faut vérifier la cohérence aux points redondants (dans l'exemple, vérifier que la valeur définie pour $f(0)$ est la même dans le cas $x \leq 0$ et dans le cas $x \geq 0$).

- **Définition par récurrence :**

$$u_0 = 2 \quad \text{et} \quad \forall n \in \mathbb{N}, \quad u_{n+1} = 3u_n^2 - 2.$$

Évidemment, ce n'est possible que pour définir une suite ! On peut généraliser cela à un domaine E défini par induction structurelle, en calquant la définition de la fonction sur la description des constructions définissant E et les éléments initiaux de E .

- **Définition implicite :**

La fonction f est définie comme étant, pour une valeur de x donnée, l'unique solution d'une certaine équation. Par exemple :

$$\forall x \in \mathbb{R}_+, \quad \int_0^{f(x)} e^{xt^2} dt = x$$

Cela définit le plus souvent une fonction et non une application, l'équation pouvant ne pas avoir de solution pour certaines valeurs de x

- **Définition par une équation fonctionnelle ou différentielle :**

Exemple : il existe une unique application dérivable f de \mathbb{R} dans \mathbb{R} telle que $f(0) = f'(0) = 1$ et pour tout $(x, y) \in \mathbb{R}^2$, $f(x+y) = f(x)f(y)$ (laquelle ?)

Exemple : il existe une unique application f dérivable sur \mathbb{R} , telle que $f(0) = 1$ et $f' = f$ (laquelle ?)

Définition 3.1.10 (composition d'applications)

Soit $f : E \rightarrow F$ et $g : F \rightarrow G$. Alors, l'application composée de f et de g est l'application, notée $g \circ f$, définie par :

$$\begin{aligned} g \circ f : E &\longrightarrow G \\ x &\longmapsto g(f(x)). \end{aligned}$$

Remarque 3.1.11

1. Est-ce dans le contexte décrit ci-dessus, la notation $g(f)$ est correcte ? Sur quel ensemble g devrait-elle être définie pour donner un sens à cette notation ?
2. On peut composer des fonctions également. Le domaine de définition est alors constitué des éléments x de D_f tels que $f(x) \in D_g$.

Définition 3.1.12 (Restriction, prolongement, corestriction)

Soit E et F deux ensembles, et soit $E' \subset E$ et $F' \subset F$.

1. Soit $f : E \rightarrow F$ une application. La restriction de f à E' , notée $f|_{E'}$, est l'unique application de E' dans F telle que pour tout $x \in E'$, $f|_{E'}(x) = f(x)$.

On peut aussi définir la restriction d'une fonction : $f|_{E'}$ ne sera alors définie qu'aux points de E' en lesquels f l'est.

2. Soit $g : E' \rightarrow F$. Un prolongement de g à E est une application $f : E \rightarrow F$ telle que $g = f|_{E'}$. Un tel prolongement n'est bien entendu pas unique en général.

3. Soit $f : E \rightarrow F$ une application. La corestriction de f à F' est la fonction $f|^{F'}$ non définie en x tel que $f(x) \notin F'$, et telle que $f|^{F'}(x) = f(x)$ si $f(x) \in F'$.

On peut définir de la même manière la corestriction d'une fonction.

Avertissement 3.1.13

En général, la corestriction d'une application n'est pas une application, mais seulement une fonction au sens évoqué ci-dessus. On donne ci-après une condition pour que ce soit une application.

Exemples 3.1.14

1. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f(x) = x^2$. Décrire sa restriction à \mathbb{R}^+ , sa corestriction à $[1, +\infty[$
2. Comparer la restriction à E' de la corestriction à F' de $f : E \rightarrow F$ à la corestriction à F' de la restriction à E' de f .

II Image directe, image réciproque

Soit E et F deux ensembles, et $f : E \rightarrow F$ une application de E vers F .

Définition 3.2.1 (Image d'une application)

L'image de f est l'ensemble $\text{Im}(f) = \{y \in F \mid \exists x \in E, f(x) = y\}$. C'est l'ensemble des points de F qui sont images d'un certain point x .

La notion d'image permet notamment de donner une condition d'existence d'une corestriction :

Proposition 3.2.2 (Corestriction d'une application)

Soit $f : E \rightarrow F$ une application et $F' \subset F$. La corestriction de f est une application si et seulement si $\text{Im}(f) \subset F'$.

Définition 3.2.3 (Image directe)

1. Soit $E' \subset E$ un sous-ensemble de E . L'image directe de E' par f est l'ensemble :

$$\tilde{f}(E') = \{y \in F \mid \exists x \in E', f(x) = y\}.$$

C'est l'ensemble des valeurs qui sont images d'un x de E' .

2. Cette construction définit une application « image directe » :

$$\begin{aligned}\tilde{f} : \mathcal{P}(E) &\longrightarrow \mathcal{P}(F) \\ E' &\longmapsto f(E')\end{aligned}$$

Définition 3.2.4 (antécédent d'un élément)

- Soit $y \in F$. Un antécédent par f de y est un élément x de E tel que $f(x) = y$.
- On note $\widehat{f^{-1}}(\{y\})$ (ou lorsqu'il n'y a pas d'ambiguïté, $f^{-1}(\{y\})$), ou encore, de façon abusive, $f^{-1}(y)$ l'ensemble des antécédents de y par f . C'est un sous-ensemble de E .

Remarque 3.2.5

Un élément de F peut ne pas avoir d'antécédent par f , ou peut en avoir un seul, ou plusieurs.

Proposition 3.2.6 (Partage associé à une application)

Soit f une application de E dans F . Alors les ensembles $f^{-1}(\{y\})$, $y \in F$ sont deux à deux disjoints, et $\bigcup_{y \in F} f^{-1}(\{y\}) = E$.

Ainsi, $\{f^{-1}(\{y\}), y \in F\}$ est une partition à parts éventuellement vides de E , ou, selon une terminologie non universelle introduite dans le chapitre précédent, un partage de E .

▫ Éléments de preuve.

Le fait que les parts sont disjointes provient de l'unicité de l'image, le fait que l'union est E provient de l'existence de l'image. Que dire si f est une fonction ? ▷

L'ensemble $f^{-1}(\{y\})$ des antécédents de y correspond à un cas particulier d'une notion plus générale :

Définition 3.2.7 (Image réciproque)

1. Soit $F' \subset F$ un sous-ensemble de F . L'image réciproque de F' est l'ensemble : $\widehat{f^{-1}}(F') = \{x \in E \mid f(x) \in F'\}$ C'est l'ensemble des éléments dont l'image est dans F' , ou, autrement dit, l'ensemble des antécédents d'éléments de F' .
2. Cette construction définit une application :

$$\begin{aligned}\widehat{f^{-1}} : \mathcal{P}(F) &\longrightarrow \mathcal{P}(E) \\ F' &\longmapsto f^{-1}(F')\end{aligned}$$

Ainsi, par définition, $\boxed{x \in f^{-1}(F') \iff f(x) \in F'}$. En d'autres termes $\widehat{f^{-1}}(F')$ est l'ensemble des éléments de E qui sont antécédents d'au moins (et dans ce cas, d'exactement – pourquoi?) un élément de F' .

Remarque 3.2.8

Les tildes et chapeaux sur \tilde{f} et $\widehat{f^{-1}}$ sont présents uniquement pour distinguer ces applications (définies sur l'ensemble des parties de E et F respectivement) des applications f et f^{-1} (cette dernière désignant la fonction réciproque de f , n'existant que si f est bijective, notion définie ci-dessous), définies sur E et F respectivement.

On verra plus tard qu'il ne peut en fait pas y avoir d'ambiguïté sur la signification de $f^{-1}(F')$ (image reciproque par f ou image directe par f^{-1}). Nous nous autoriserons donc, et ce dès maintenant, à alléger les notations, en désignant simplement par $f(E')$ et $f^{-1}(F')$ les images directes et réciproques par f .

Proposition 3.2.9 (Croissance des images directes et réciproques)

Soit f une application de E dans F . Alors \tilde{f} et $\widehat{f^{-1}}$ sont croissantes pour l'inclusion. Autrement dit :

- Pour tous E', E'' dans $\mathcal{P}(E)$ tels que $E' \subset E''$, on a $f(E') \subset f(E'')$;
- Pour tous F', F'' dans $\mathcal{P}(F)$ tels que $F' \subset F''$, on a $f^{-1}(F') \subset f^{-1}(F'')$;

▫ Éléments de preuve.

Intuitivement évident. Revenir aux éléments pour le justifier rigoureusement. ▷

Proposition 3.2.10 (Images directes et réciproques d'unions ou intersections)

Soit E et F deux ensembles, $f : E \rightarrow F$ une application, et $(E_i)_{i \in I}$ une famille de sous-ensembles de E , $(F_i)_{i \in I}$ une famille de sous-ensembles de F . Alors :

1. $f\left(\bigcup_{i \in I} E_i\right) = \bigcup_{i \in I} f(E_i)$
2. $f\left(\bigcap_{i \in I} E_i\right) \subset \bigcap_{i \in I} f(E_i)$ (attention ! ce n'est qu'une inclusion !)
3. $f^{-1}\left(\bigcup_{i \in I} F_i\right) = \bigcup_{i \in I} f^{-1}(F_i)$
4. $f^{-1}\left(\bigcap_{i \in I} F_i\right) = \bigcap_{i \in I} f^{-1}(F_i)$.

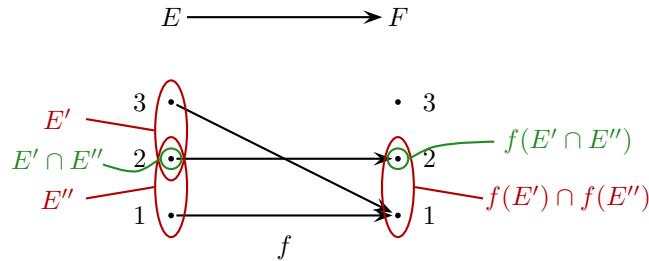
Exemple 3.2.11

On peut avoir une inclusion stricte $f(E' \cap E'') \subset f(E') \cap f(E'')$. Pour un exemple, voir figure 3.3

En retenir que tout se passe bien avec les images réciproques, mais qu'il faut prendre un peu plus de précautions avec les images directes.

Remarque 3.2.12 (Transition vers la section suivante)

À quelle condition suffisante l'image d'une intersection est-elle l'intersection des images ? Est-ce une condition nécessaire ?

FIGURE 3.3 – Exemple dans lequel $f(E' \cap E'') \neq f(E') \cap f(E'')$

III Injectivité, surjectivité, bijectivité

Définition 3.3.1 (Injectivité, surjectivité, bijectivité)

Soit $f : E \rightarrow F$ une application. On dit que :

1. f est injective ssi tout élément de F admet au plus un antécédent par f
2. f est surjective ssi tout élément de F admet au moins un antécédent par f
3. f est bijective ssi f est injective et surjective.

Remarque 3.3.2

Ainsi, f est bijective si et seulement si tout élément de F admet exactement un antécédent. Pour cette raison, une bijection est parfois aussi appelée correspondance 1 à 1 (issu de la terminologie anglaise « one-to-one correspondance »).

Terminologie 3.3.3 (permutation, groupe symétrique)

Une bijection de E dans lui-même est appelée *permutation* de E . On note $\mathfrak{S}(E)$ l'ensemble des permutations de E . Si $E = \{1, \dots, n\}$, on note \mathfrak{S}_n au lieu de $\mathfrak{S}(E)$. L'ensemble \mathfrak{S}_n est appelé *n-ième groupe symétrique*.

La notation \mathfrak{S} est l'écriture gothique de la lettre S . On trouve parfois aussi plus simplement la notation S_n au lieu de \mathfrak{S}_n pour désigner le groupe symétrique (et $S(E)$ au lieu de $\mathfrak{S}(E)$, notation du programme officiel). On trouve également parfois Σ_n et $\Sigma(E)$

Exemples 3.3.4

1. Soit $E \subset F$. L'injection $i : E \rightarrow F$ est
2. Soit E et F deux ensembles, $F \neq \emptyset$. La projection $p_E : E \times F \rightarrow E$ est
3. La fonction identité $E \rightarrow E$ est
4. À quelle condition la fonction caractéristique $\mathbb{1}_F$ d'un sous-ensemble F de E est-elle surjective ? injective ?
5. La fonction $x \mapsto x^2$ est :
 - si elle est vue comme fonction de \mathbb{R} dans \mathbb{R}_+ ;
 - si elle est vue comme fonction de \mathbb{R}_+ dans \mathbb{R}
 - si elle est vue comme fonction de \mathbb{R}_+ dans \mathbb{R}_+
 - si elle est vue comme fonction de \mathbb{R} dans \mathbb{R} .

Il est donc important de porter une attention particulière aux domaines de départ et d'arrivée dans l'étude des propriétés d'injectivité et de surjectivité. D'autre part, constatez sur cet exemple qu'une fonction peut n'être ni injective ni surjective.

En adoptant la notion intuitive de cardinal d'un ensemble G comme étant le nombre d'éléments de G (éventuellement infini), on obtient :

Proposition 3.3.5 (CNS d'injectivité)

Soit $f : E \rightarrow F$ une application. Les propositions suivantes sont équivalentes :

- (i) f est injective ;
- (ii) $\forall y \in F, \text{Card}(f^{-1}(\{y\})) \leq 1$,
- (iii) $\forall (x, y) \in E^2, x \neq y \implies f(x) \neq f(y)$,
- (iv) $\forall (x, y) \in E^2, f(x) = f(y) \implies x = y$.

▫ Éléments de preuve.

(i) équivaut à (ii) par définition, (iii) équivaut à (iv) par contraposition, (i) implique (iv) en considérant l'image réciproque de $f(x)$, et (iv) implique (ii) en contraposant : si (ii) est faux, on trouve facilement x et y contredisant (iv). ▷

Proposition 3.3.6 (CNS de surjectivité)

Soit $f : E \rightarrow F$ une application. Les propositions suivantes sont équivalentes :

- (i) f est surjective ;
- (ii) $\forall y \in F, \text{Card}(f^{-1}(\{y\})) \geq 1$;
- (iii) $\forall y \in F, \exists x \in E, f(x) = y$;
- (iv) $\text{Im}(f) = F$.

Corollaire 3.3.7 (Partition associée à une surjection)

Lorsque l'application $f : E \rightarrow F$ est surjective, le partage de E associé à f est une partition.

▫ Éléments de preuve.

En effet, d'après le point (ii), les parts sont non vides. ▷

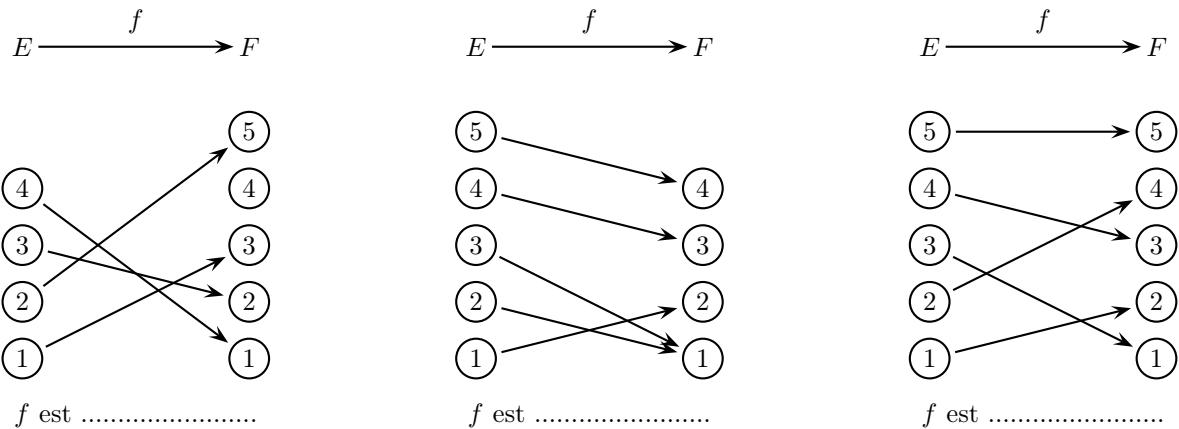
Dans le cas d'ensembles finis, on peut illustrer ces notions sur des diagrammes sagittaux (figure 3.4).

Remarque 3.3.8

Intuitivement, si $f : E \rightarrow F$ est une injection, il y a plus d'éléments dans F que dans E . C'est l'inverse si f est une surjection, et c'est E et F ont même cardinal si f est une bijection. C'est vrai si E et F sont finis. Attention, aux cas où E et F sont infinis : la situation peut parfois être contraire à l'intuition.

Exemples 3.3.9

1. Principe de l'hôtel de Hilbert : soit $x \notin \mathbb{N}$. Alors \mathbb{N} et $\mathbb{N} \cup \{x\}$ peuvent être mis en bijection (on dit qu'ils ont même cardinal).
2. La numérotation en diagonales de $\mathbb{N} \times \mathbb{N}$ est une bijection de \mathbb{N} sur \mathbb{N}^2 .
3. $\tan :]-\frac{\pi}{2}, \frac{\pi}{2}[\rightarrow \mathbb{R}$ est une bijection (rappel : pour tout x pour lequel $\cos(x) \neq 0$, $\tan(x) = \frac{\sin(x)}{\cos(x)}$). Pourtant, il y a « plus » d'éléments dans \mathbb{R} que dans $] -\frac{\pi}{2}, \frac{\pi}{2}[$.

FIGURE 3.4 – Diagramme sagittal d'une fonction f injective, surjective ou bijective

4. Soit $f : [0, 1]^2 \rightarrow [0, 1]$ définie de la façon suivante : étant donné x et y dans $[0, 1]$, on note x_i le i -ième chiffre de x après la virgule dans son écriture décimale, et de même pour y_i . On rappelle que si x est décimal, x admet deux écritures décimales, l'une qui termine par une infinité de 0, l'autre qui termine par une infinité de 9 (car $0.999999\dots = 1$). On choisit dans ce cas, afin que les x_i et les y_i soient définis de façon unique, la représentation terminant par des 0. On définit alors f sur le couple (x, y) par :

$$f(x, y) = 0.x_1y_1x_2y_2x_3y_3\dots$$

Il s'agit donc du réel de $[0, 1]$ obtenu en alternant les chiffres de x et ceux de y .

La fonction f est injective. Est-elle surjective ?

5. On peut de même construire une fonction $g : [0, 1] \rightarrow [0, 1]^2 \setminus \{(1, 1)\}$ par :

$$g(x) = (0.x_1x_3x_4\dots; 0.x_2x_4x_6\dots),$$

les x_i représentant encore les chiffres de x dans l'écriture décimale, avec les mêmes conventions pour les réels décimaux.

La fonction f est surjective. Est-elle injective ?

Les exemples précédents permettent d'établir (avec la notion de cardinal définie plus loin) que \mathbb{N} est de même cardinal que \mathbb{N}^2 , que $]-\frac{\pi}{2}, \frac{\pi}{2}[$ est de même cardinal de \mathbb{R} (en fait, tout intervalle non vide et non restreint à un singleton est de même cardinal que \mathbb{R}), et, plus surprenant encore, que $[0, 1]^2$ est de même cardinal que $[0, 1]$ (de quoi il ressort que \mathbb{R}^2 est de même cardinal que \mathbb{R} !)

Proposition 3.3.10 (Composée d'injections, surjections, bijections)

Soit $f : E \rightarrow F$ et $g : F \rightarrow G$ deux applications. Alors :

1. *si f et g sont injectives, alors $g \circ f$ aussi ;*
2. *si f et g sont surjectives, alors $g \circ f$ aussi ;*
3. *si f et g sont bijectives, alors $g \circ f$ aussi ;*

▫ Éléments de preuve.

Pour le point (i), considérer x et y tels que $g \circ f(x) = g \circ f(y)$, et obtenir d'abord $f(x) = f(y)$ puis $x = y$. Le point (ii) est évident, le point (iii) est la combinaison de (i) et (ii). ▷

Cette proposition admet une réciproque partielle :

Proposition 3.3.11 (« Dé-composée » d'injections, surjections, bijections)

Soit $f : E \rightarrow F$ et $g : F \rightarrow G$ deux applications. Alors :

1. si $g \circ f$ est injective, alors f est
2. si $g \circ f$ est surjective, alors g est

▫ Éléments de preuve.

1. Considérer x et y tels que $f(x) = f(y)$, et se ramener à l'utilisation de l'injectivité de $g \circ f$.
2. Considérer un antécédent par $g \circ f$ de $z \in G$. Comment en déduire un antécédent par g de z ?

▷

Dans une structure algébrique, l'inversibilité d'un élément a se traduit par l'existence de b tel que $ab = ba = 1$ (ou 1 est le neutre multiplicatif de la structure). La propriété suivante est donc à mettre en rapport avec une propriété d'inversibilité.

Théorème 3.3.12 (Caractérisation de la bijectivité)

Soit $f : E \rightarrow F$ une application. L'application f est bijective si et seulement s'il existe $g : F \rightarrow E$ telle que $g \circ f = \text{id}_E$ et $f \circ g = \text{id}_F$, donc si f est inversible.

De plus, dans ce cas, la fonction g est unique.

▫ Éléments de preuve.

Sens direct par analyse/synthèse : l'analyse montre que $g(y)$ doit être l'unique antécédent par f de y .

Sens réciproque par la proposition 3.3.11.

▷

Avertissement 3.3.13

Attention, il ne suffit pas que $g \circ f = \text{id}_E$ ou $f \circ g = \text{id}_F$ pour obtenir la bijectivité : il faut avoir les deux égalités.

Définition 3.3.14 (Application réciproque)

Dans la situation du théorème précédent, l'application g telle que $g \circ f = \text{id}_E$ et $f \circ g = \text{id}_F$ est appelée *application réciproque* de f , et est notée f^{-1} .

Dans le cas où f est bijective, la notation f^{-1} peut désigner à la fois la fonction image réciproque et la fonction réciproque. Appliquée à un sous-ensemble de F , il peut s'agir de l'image réciproque de ce sous-ensemble ou de son image directe par la fonction réciproque f^{-1} . La proposition 3.3.16 montre que l'ambiguïté des notations n'est pas gênante. Sa démonstration est basée sur le lemme suivant, dont la démonstration est élémentaire :

Lemme 3.3.15 (Image réciproque d'un élément par une bijection)

Soit f une bijection. Alors

$$\widehat{f^{-1}}(\{y\}) = \{f^{-1}(y)\} = \widetilde{f^{-1}}(\{y\})$$

Proposition 3.3.16 (Image réciproque de f et image directe de f^{-1})

Soit f une bijection. Alors $\widehat{f^{-1}} = \widehat{f}^{-1}$.

Le théorème 3.3.12 peut se décomposer en caractérisation de l'injectivité et de la surjectivité séparément, mais cela nécessite l'axiome du choix, pour l'une des deux caractérisations.

Théorème 3.3.17 (Caractérisation de l'injectivité et de la surjectivité, HP)

Soit $f : E \rightarrow F$ une fonction.

1. (avec AC) f est surjective si et seulement s'il existe une application $g : F \rightarrow E$ telle que $f \circ g = \text{id}_F$, donc si et seulement si f est inversible à droite.
2. Si $E \neq \emptyset$, f est injective si et seulement s'il existe une application $g : F \rightarrow E$ telle que $g \circ f = \text{id}_E$, donc si et seulement si f est inversible à gauche.

▫ Éléments de preuve.

Un peu le même principe que pour la bijectivité, mais dans un cas, on peut avoir plusieurs choix possibles (d'où la nécessité de l'AC), dans l'autre, on peut ne pas avoir d'antécédent (attribuer à ces éléments une image quelconque choisie à l'avance). ▷

IV Cardinal d'un ensemble

On montre dans ce paragraphe comment la notion de bijection permet de formaliser la définition de cardinal d'un ensemble.

IV.1 Ensembles finis et cardinal

On ne peut pas définir directement le cardinal d'un ensemble quelconque, sans introduire de nouveaux objets. En revanche, on peut définir sans difficulté des classes de cardinaux (donc définir une condition pour que deux ensembles aient même cardinal, ce qui permet ensuite de grouper les ensembles en classes d'ensembles de même cardinal).

Définition 3.4.1 (Définition de la cardinalité selon Frege)

On dit que deux ensembles E et F ont *même cardinal* s'il existe une bijection de E à F . On note $\text{Card}(E) = \text{Card}(F)$.

Définition 3.4.2 (Ensemble fini, cardinal d'un ensemble fini)

Un ensemble est fini s'il existe un entier n et une bijection $f : \llbracket 1, n \rrbracket \rightarrow E$, ou de façon équivalente une bijection $g : E \rightarrow \llbracket 1, n \rrbracket$. On dit dans ce cas que E est de cardinal n , et on note $\text{Card}(E) = n$, ou $|E| = n$.

Lemme 3.4.3

Soit A et B deux ensembles disjoints, C et D deux ensembles disjoints tels qu'il existe une bijection $f_1 : A \rightarrow C$ et une bijection $f_2 : B \rightarrow D$. Alors il existe une bijection $f : A \sqcup B \rightarrow C \sqcup D$.

▫ Éléments de preuve.

Construire f coïncidant avec f_1 sur A et avec f_2 sur B . Cette application f est bien définie car A et B sont disjoints. On peut construire la réciproque de la même manière, ou prouver facilement l'injectivité et la surjectivité. ▷

Lemme 3.4.4

Si A est fini et $a \notin A$, alors $A \cup \{a\}$ est fini.

▫ Éléments de preuve.

Avec le lemme précédent, en considérant $f_2 : \{n+1\} \rightarrow \{a\}$. ▷

Proposition 3.4.5

Un sous-ensemble de $\llbracket 1, n \rrbracket$ est fini.

▫ Éléments de preuve.

Récurrence sur n . Pour montrer $\mathcal{P}(n+1)$ à partir de $\mathcal{P}(n)$, pour $E \subset \llbracket 1, n+1 \rrbracket$, discuter suivant que $n+1 \in E$ ou non. ▷

Proposition 3.4.6 (Caractérisation des ensembles finis)

Soit E un ensemble. Les propriétés suivantes sont équivalentes :

- (i) E est fini
- (ii) il existe $n \in \mathbb{N}$ et une surjection $\llbracket 1, n \rrbracket \rightarrow E$
- (iii) il existe $n \in \mathbb{N}$ et une injection $E \rightarrow \llbracket 1, n \rrbracket$

▫ Éléments de preuve.

Pour chacune des propriétés (ii) et (iii), l'une des implications de l'équivalence à (i) est évidente. L'autre s'obtient par restriction ou corestriction, en se ramenant à la proposition 3.4.5. ▷

Proposition 3.4.7 (Sous-ensemble d'un ensemble fini)

Un sous-ensemble F d'un ensemble fini E est aussi fini.

▫ Éléments de preuve.

Utiliser la caractérisation. On peut aussi se ramener directement à 3.4.5. ▷

Théorème 3.4.8 (Unicité du cardinal)

Soit $n \neq m$. Il n'existe pas de bijection de $\llbracket 1, n \rrbracket$ dans $\llbracket 1, m \rrbracket$. Ainsi, un ensemble E ne peut pas être à la fois de cardinal n et de cardinal m .

▫ Éléments de preuve.

Par symétrie, il suffit de montrer que si $n > m$, il n'existe pas d'injection de $\llbracket 1, n \rrbracket$ dans $\llbracket 1, m \rrbracket$. Raisonner par récurrence sur n en remarquant que la restriction à $\llbracket 1, n \rrbracket$ d'une injection de $\llbracket 1, n+1 \rrbracket$ dans $\llbracket 1, m \rrbracket$ est encore une injection. En déduire une injection de $\llbracket 1, n \rrbracket$ dans $\llbracket 1, m-1 \rrbracket$. ▷

Exemples 3.4.9

1. $|E| = 0$ si et seulement si $E = \emptyset$,
2. $|\llbracket 1, n \rrbracket| = n$.

IV.2 Règles de calcul sur les cardinaux

Proposition 3.4.10 (Cardinal d'une union disjointe)

Soit A, B, A_1, \dots, A_n des ensembles finis.

1. Si $A \cap B = \emptyset$, alors $|A \sqcup B| = |A| + |B|$.
2. Plus généralement, si pour tout $(i, j) \in \llbracket 1, n \rrbracket^2$ tel que $i \neq j$, $A_i \cap A_j = \emptyset$, alors

$$|A_1 \sqcup \cdots \sqcup A_n| = |A_1| + \cdots + |A_n|.$$

▫ Éléments de preuve.

Le point 2 s'obtient du point 1 par récurrence. Le point 1 nécessite la construction d'une bijection de $\llbracket 1, n+m \rrbracket \rightarrow A \sqcup B$ à partir de 2 bijections de $\llbracket 1, n \rrbracket \rightarrow A$ et $\llbracket 1, m \rrbracket \rightarrow B$. ▷

Proposition 3.4.11 (Cardinal d'un complémentaire)

Si $A \subset B$, alors $|\complement_B A| = |B| - |A|$.

▫ Éléments de preuve.

Écrire l'un des trois ensembles en jeu comme union disjointe des deux autres et appliquer la proposition précédente. ▷

Corollaire 3.4.12 (Cardinal d'un sous-ensemble)

Si $A \subset B$, alors $|A| \leq |B|$, avec égalité si et seulement si $A = B$.

▫ Éléments de preuve.

L'inégalité découle immédiatement de la proposition précédente. Le cas d'égalité résulte du fait qu'alors $\complement_B A$ est de cardinal 0. ▷

Proposition 3.4.13 (Cardinal d'une union quelconque)

Soit A et B des ensembles finis. On a :

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

▫ Éléments de preuve.

Là encore, cela découle de la propriété des sommes indexées sur une union quelconque. ▷

Plus généralement, on a la formule suivante, due à Moivre :

Théorème 3.4.14 (Formule du crible de Poincaré, ou formule d'inclusion-exclusion, HP)

Soit A_1, \dots, A_n des ensembles finis. Alors :

$$|A_1 \cup \cdots \cup A_n| = \sum_{k=1}^n \left((-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} |A_{i_1} \cap \cdots \cap A_{i_k}| \right) = \sum_{\substack{I \subset \llbracket 1, n \rrbracket \\ I \neq \emptyset}} (-1)^{|I|-1} \left| \bigcap_{i \in I} A_i \right|.$$

▫ Éléments de preuve.

Il s'agit d'une récurrence sur n . On applique le cas $n = 2$ à $A_1 \cup \dots \cup A_n$ et A_{n+1} , on utilise l'hypothèse de récurrence pour le calcul de $|A_1 \cup \dots \cup A_n|$ ainsi que pour le calcul de $|(A_1 \cup \dots \cup A_n) \cap A_{n+1}|$ après avoir distribué l'intersection. Remarquer qu'on obtient tous les termes voulus, les indices (ensemblistes $I \subset \llbracket 1, n+1 \rrbracket$) étant triés en 3 catégories :

- les ensembles I non vides ne contenant pas $n+1$ (la somme obtenue à partir de $A_1 \cup \dots \cup A_n$)
- les ensembles I non réduits à un élément et contenant $n+1$ (la somme obtenue à partir de $(A_1 \cup \dots \cup A_n) \cap A_{n+1}$)
- Le sous-ensemble $I = \{n+1\}$ (le terme $|A_{n+1}|$ apparaissant lors de la première étape.)

▷

Proposition 3.4.15 (Cardinal d'un produit cartésien)

1. Soit A et B deux ensembles finis. Alors $|A \times B| = |A| \times |B|$.
2. Plus généralement, soit A_1, \dots, A_n des ensembles finis. Alors

$$|A_1 \times \dots \times A_n| = \prod_{i=1}^n |A_i|.$$

▫ Éléments de preuve.

Le point 2 s'obtient du 1 par récurrence.

Pour le point 1, découper en tranche, ou utiliser une bijection de $\llbracket 0, nm-1 \rrbracket$ dans $\llbracket 0, n-1 \rrbracket \times \llbracket 0, m-1 \rrbracket$ définie par une opération arithmétique classique.

▷

IV.3 Comparaison des cardinaux en cas d'injectivité et surjectivité

Des résultats liés aux cardinaux des sous-ensembles d'un ensemble fini, on déduit assez facilement la comparaison des cardinaux de E et de F en cas d'existence d'une injection ou d'une surjection entre les deux.

Proposition 3.4.16 (Injectivité, surjectivité, bijectivité et cardinal)

Soit E et F deux ensembles finis, et soit $f : E \rightarrow F$ une application. Alors :

1. Si f est injective, $\text{Card}(E) \leq \text{Card}(F)$
2. Si f est surjective, $\text{Card}(E) \geq \text{Card}(F)$
3. Si f est bijective, $\text{Card}(E) = \text{Card}(F)$.

▫ Éléments de preuve.

Restreindre ou corestreindre de façon adéquate, de sorte à se ramener à une bijection.

▷

On en déduit notamment une caractérisation des applications bijectives entre ensembles de même cardinal :

Corollaire 3.4.17 (Caractérisation des bijections)

Soit A et B deux ensembles finis de même cardinal, et $f : A \rightarrow B$. Alors les 3 propriétés suivantes sont équivalentes :

- (i) f est bijective
- (ii) f est injective
- (iii) f est surjective

▫ Éléments de preuve.

On se rend facilement compte qu'il suffit de montrer que $(ii) \iff (iii)$.

Si f est injective, remarquer que $|\text{Im}(f)| = |A|$ et utiliser le cas d'égalité des cardinaux des sous-ensembles.

Si f n'est pas injective, montrer que $|\text{Im}(f)| < |A|$, en construisant une surjection d'un sous-ensemble strict de A sur $\text{Im}(f)$. ▷

IV.4 Introduction à la dénombrabilité (HP)

Tout ce paragraphe est du ressort du programme de Spé, à part le dernier point. On utilisera dans ce paragraphe l'existence du minimum d'une partie non vide de \mathbb{N} (ce qui est un axiome de la construction de \mathbb{N} ; nous en reparlerons dans un chapitre ultérieur)

Définition 3.4.18 (Dénombrabilité)

Un ensemble E est *dénombrable* s'il peut être mis en bijection avec \mathbb{N} . On dit qu'il est au plus dénombrable s'il est fini ou dénombrable.

Lemme 3.4.19

Un sous-ensemble E non fini de \mathbb{N} est dénombrable.

▫ Éléments de preuve.

Construire par récurrence une application injective de $\llbracket 0, n \rrbracket$ dans E , en rajoutant à chaque étape le minimum des éléments non encore utilisés. Cela définit une bijection $\mathbb{N} \rightarrow E$. ▷

Ainsi, tout sous-ensemble de \mathbb{N} est au plus dénombrable.

Lemme 3.4.20 (Caractérisation des ensembles au plus dénombrables)

Soit E un ensemble non vide. Les trois propriétés suivantes sont équivalentes :

- (i) *E est au plus dénombrable*
- (ii) *il existe une fonction injective $f : E \rightarrow \mathbb{N}$.*
- (iii) *il existe une fonction surjective $f : \mathbb{N} \rightarrow E$.*

▫ Éléments de preuve.

- (i) \implies (ii) : considérer une bijection $E \rightarrow F \subset \mathbb{N}$, et composer par l'injection canonique.
- (ii) \implies (iii) : c'est l'inversibilité à gauche d'une injection
- (iii) \implies (ii) : c'est l'inversibilité à droite d'une surjection. Pourquoi l'axiome du choix n'est-il pas nécessaire ici ?
- (ii) \implies (i) : corestriction à l'image.

▷

Nous utilisons le fait, déjà démontré en exercice, que \mathbb{N}^2 est dénombrable (*via* la numérotation en diagonale).

Proposition 3.4.21 (Construction d'ensembles dénombrables)

1. *Si E et F sont au plus dénombrables, alors $E \times F$ est au plus dénombrable.*

2. Plus généralement, un produit $E_1 \times \cdots \times E_n$ d'un nombre fini d'ensembles au plus dénombrables est au plus dénombrable. Si l'un d'eux est dénombrable et les autres non vides, alors le produit est dénombrable.
3. Une union d'un nombre au plus dénombrable d'ensembles au plus dénombrables est au plus dénombrable. Si l'un des ensembles est dénombrable, l'union est dénombrable.

△ Éléments de preuve.

1. Construire une surjection par composition : $\mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N} \rightarrow E \times F$
2. Récurrence
3. Supposons les ensembles indexés sur I dénombrable. Construire une surjection

$$\mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N} \rightarrow I \times \mathbb{N} \rightarrow \bigcup_{i \in I} A_i.$$

▷

Corollaire 3.4.22

1. L'ensemble \mathbb{Z} est dénombrable.
2. L'ensemble \mathbb{N}^p est dénombrable.
3. L'ensemble \mathbb{Q} des rationnels est dénombrable.
4. L'ensemble $\mathbb{Z}[x]$ des fonctions polynomiales à coefficients entiers est dénombrable.

Il existe des ensembles infinis non dénombrables. Par exemple :

Théorème 3.4.23

L'ensemble des réels \mathbb{R} est non dénombrable.

△ Éléments de preuve.

Simon, considérer une énumération $(x_n)_{n \in \mathbb{N}^*}$ de tous les réels, et considérer un y tel que le n -ième chiffre après la virgule de y soit distinct de 0, 9 et du n -ième chiffre de x_n . Le réel y peut-il être dans l'énumération ? Pourquoi prendre la précaution de prendre les chiffres distincts de 0 et 9 ? Cette construction est-elle dépendante de l'axiome du choix ?

▷

Pour la culture, nous terminons cette étude par un théorème de Cantor, affirmant notamment qu'étant donné un ensemble, il existe toujours un ensemble de cardinal plus gros. En particulier, il existe une infinité de cardinaux infinis différents.

Définition 3.4.24

On dit que $\text{Card}(E) \leq \text{Card}(F)$ si et seulement s'il existe une injection de E dans F , et $\text{Card}(E) < \text{Card}(F)$ si et seulement si $\text{Card}(E) \leq \text{Card}(F)$ et $\text{Card}(E) \neq \text{Card}(F)$.

Théorème 3.4.25 (Cantor, 1891, HP)

Pour tout ensemble X , on a $\text{Card}(X) < \text{Card}(\mathcal{P}(X))$.

△ Éléments de preuve.

S'il existe une surjection $f : X \rightarrow \mathcal{P}(X)$, considérer $Y \subset X$ constitué des éléments x tels que $x \notin f(x)$, et obtenir une contradiction en considérant l'apparence de c à Y , c étant un élément de X tel que $f(c) = Y$.

▷

Comme on l'a signalé plus haut, ce résultat entre en contradiction avec l'existence de l'ensemble des ensembles.

4

Sommes

« La totalité est plus que la somme des parties »

(Aristote)

« $1 + 2 + 3 + 4 + \dots = -\frac{1}{12}$ »

(Leonhard Euler)

Introduction

Le but de ce chapitre est d'introduire et systématiser l'usage du signe \sum pour désigner une somme d'éléments. Dans la mesure du possible, l'utilisation de cette notation est préférable à celle utilisant des petits points, bien moins rigoureuse.

I Manipulation des signes \sum et \prod

Nous nous intéressons dans ce paragraphe à la définition de la somme des éléments d'une famille finie $(x_i)_{i \in I}$ de réels ou complexes (ou plus généralement des objets qu'on sait sommer de façon commutative et associative). Le cas de familles infinies relève de techniques plus fines, car comme est amené à faire une infinité d'opérations, il faut s'assurer que le procédé « converge ». Nous reparlerons de cela plus tard.

I.1 Définition des notations

On rappelle qu'une loi $+$ définie sur un ensemble E est dite commutative si pour tous x, y de E , $x + y = y + x$, et associative si pour tous x, y, z de E , $x + (y + z) = (x + y) + z$. De même pour une loi multiplicativa \times .

Définition 4.1.1 (signes \sum et \prod : définition générale)

Soit I un ensemble fini et $(a_i)_{i \in I}$ une famille de nombres réels ou complexes.

- L'expression $\sum_{i \in I} a_i$ désigne la somme de tous les éléments a_i , pour $i \in I$. Ainsi, si $I = \{i_1, \dots, i_n\}$,

$$\sum_{i \in I} a_i = a_{i_1} + \dots + a_{i_n}.$$

- L'expression $\prod_{i \in I} a_i$ désigne de même le produit de tous les éléments a_i , pour $i \in I$.

- Lorsque $I = \llbracket n, m \rrbracket$, où n et m sont deux entiers tels que $n \leq m$, on note $\sum_{i \in \llbracket n, m \rrbracket} a_i = \sum_{i=n}^m a_i$.

Cette définition ne dépend pas du choix d'une numérotation des éléments de I , du fait de la commutativité et de l'associativité de la somme et du produit dans \mathbb{C} . Il s'agit en fait de l'utilisation de la propriété de commutativité généralisée, que nous démontrerons plus rigoureusement lors de l'étude des structures algébriques.

Remarques 4.1.2

- La lettre i utilisée pour énumérer les éléments de I résulte évidemment d'un choix arbitraire : on peut remplacer cette lettre par toute autre lettre n'ayant pas de signification externe à la somme. On dit que i est une *variable muette*. Ainsi :

$$\sum_{i \in I} a_i = \sum_{j \in I} a_j = \sum_{\beta \in I} a_\beta$$

En revanche, $\sum_{n \in [1, n]} a_n$ n'a pas de sens.

- Par usage, il est assez rare de définir une loi additive non commutative et non associative. Ainsi, en général, à partir du moment où on dispose d'une loi $+$, on pourra construire des sommes de familles finies de la sorte.
- En revanche, beaucoup de lois multiplicatives ne sont pas commutatives. Par exemple le produit matriciel. Dans ce cas, le produit tel que défini plus haut est mal défini. Il faut se donner un ordre dans lequel faire les produits (donc un ordre sur les éléments de I). C'est le cas par exemple si on considère un sous-ensemble $[0, n]$ de \mathbb{N} .
- Une numérotation des éléments de I consiste en une bijection $\varphi : [1, n] \rightarrow I$. Une telle bijection définit donc une numérotation $I = \{i_1, \dots, i_n\}$ où $i_k = \varphi(k)$. C'est l'existence d'une telle bijection qui d'ailleurs permet aussi de définir rigoureusement la notion de cardinal.
- Via la bijection décrite ci-dessus, on voit que toute somme pourrait s'indexer sur un intervalle d'entiers :

$$\sum_{i \in I} a_i = \sum_{k=1}^n a_{\varphi(k)}.$$

L'indexation ensembliste n'est qu'une commodité permettant parfois d'avoir à expliciter une bijection.

Une somme se définit donc en ajoutant les termes les uns aux autres au fur et à mesure. On peut ainsi la définir par récurrence sur le nombre d'éléments de I , en isolant un élément pour se ramener à une somme plus petite : pour $i_0 \in I$:

$$\sum_{i \in I} a_i = \left(\sum_{i \in I \setminus \{i_0\}} a_i \right) + a_{i_0}.$$

La difficulté consiste alors à vérifier que cette définition est indépendante du choix de i_0 (démonstration laissée en exercice, basée évidemment sur la commutativité, puisque les choix des i_0 pour se ramener à la somme vide qui initialise cette définition détermine l'ordre de sommation)

Dans le cas d'une somme indexée sur des entiers consécutifs, il y a un choix naturel :

$$\sum_{i=1}^{n+1} a_i = \left(\sum_{i=1}^n a_i \right) + a_{n+1}.$$

Cette définition itérative permet de bien comprendre la nécessité de la convention suivante, correspondant à l'initialisation de la récurrence :

Convention 4.1.3 (somme vide, produit vide)

Lorsque $I = \emptyset$, on pose par convention :

$$\sum_{i \in \emptyset} a_i = 0 \quad \text{et} \quad \prod_{i \in \emptyset} a_i = 1.$$

Ainsi, si $p < n$, $\llbracket n, p \rrbracket$ est vide, donc $\sum_{i=n}^p a_i = 0$. Par exemple, $\sum_{i=2}^1 i^2 = 0$.

Remarque 4.1.4

On notera qu'en général, une somme n'est pas forcément prise sur un ensemble d'entiers successifs, ni même sur un ensemble d'entiers. La seule condition est que l'**ensemble des indices soit fini**. On étudiera le cas où l'ensemble des indices est \mathbb{N} dans le chapitre sur les séries, et plus généralement le cas d'une somme sur une famille dénombrable (cas des familles sommables).

Note Historique 4.1.5

Le signe \sum a été introduit par le mathématicien suisse Leonhard Euler en 1755, le symbole \prod date de Gauss, mais on en trouve trace chez Descartes. Mais leur usage ne s'est pas répandu immédiatement, et de nombreux mathématiciens ont continué à utiliser des points de suspension (par exemple Abel au début du 19^e siècle)

Exemples 4.1.6

$$1. \sum_{k=1}^4 k(k-1) = 1(1-1) + 2(2-1) + 3(3-1) + 4(4-1) = 2 + 6 + 12 = 20.$$

$$2. \sum_{i \in \{2,3,5\}} i^2 = 2^2 + 3^2 + 5^2 = 4 + 9 + 25 = 38.$$

3. si $E = \{(i, j) \in \mathbb{N}^2 \mid i + j = 5\} = \{(0, 5), (1, 4), \dots, (5, 0)\}$, alors

$$\sum_{(i,j) \in E} \frac{i}{j+1} = \frac{0}{6} + \frac{1}{5} + \frac{2}{4} + \frac{3}{3} + \frac{4}{2} + \frac{5}{1} = \frac{87}{10}.$$

$$4. \sum_{i \in E} 1 = 1 + \dots + 1 = |E| \text{ (autant de termes 1 que d'éléments dans } E\text{).}$$

Remarque 4.1.7

Ce dernier exemple pourrait être pris comme définition du cardinal. Revenant à une numérotation définie par une bijection $\varphi : \llbracket 1, n \rrbracket \rightarrow E$, il vient donc :

$$|E| = \sum_{k=1}^n 1 = n.$$

Ainsi, on retrouve la définition usuelle du cardinal, comme étant l'unique n tel qu'il existe une bijection de $\llbracket 1, n \rrbracket$ sur E (définition de Frege)

Mettons à part un dernier exemple, définissant une suite de nombres que nous utiliserons souvent :

Définition 4.1.8 (Factorielle)

Soit $n \in \mathbb{N}$. On définit la factorielle de n , notée $n!$, par :

$$n! = \prod_{k=1}^n k.$$

Ainsi, $n!$ peut être définie par récurrence par $0! = 1$ (convention du produit vide) et la relation :

$$n! = n \times (n - 1)!.$$

Dans ce qui suit, on exprime quelques règles de manipulation du symbole \sum . La plupart de ces règles ont leurs analogues pour les produits.

I.2 Changements d'indice

Du fait même de la définition à l'aide d'une numérotation (donc d'une bijection), la somme est invariante par bijection. C'est ce qu'exprime le théorème suivant :

Théorème 4.1.9 (Changements d'indice)

Soit I et J deux ensembles, et $f : I \rightarrow J$ une bijection de I sur J . Alors, pour toute famille $(b_j)_{j \in J}$,

$$\sum_{j \in J} b_j = \sum_{i \in I} b_{f(i)}$$

Même règle pour les produits.

▫ Éléments de preuve.

Il s'agit simplement du fait que si $\psi : \llbracket 1, n \rrbracket \rightarrow I$ est une bijection, servant de numérotation pour la deuxième somme, on retrouve la première somme en considérant la numérotation définie par la bijection $f \circ \psi$. ▷

Cette formule ne fait que traduire le fait que lorsque i parcourt I , les $f(i)$ prennent une et une seule fois chaque valeur j de J , du fait de la bijectivité de f .

En appliquant ce résultat à la bijection réciproque f^{-1} (obtenue en remontant les flèches), on a alors aussi, pour toute famille $(a_i)_{i \in I}$:

$$\sum_{i \in I} a_i = \sum_{j \in J} a_{f^{-1}(j)}$$

Le cas le plus fréquent est celui où la bijection est donnée par une translation sur des ensembles d'entiers consécutifs :

Corollaire 4.1.10 (changements d'indices par translation)

Soit n, p et ℓ trois entiers tels que $n \leq p$. Soit $(a_i)_{i \in \llbracket n, p \rrbracket}$ une famille. Alors :

$$\sum_{i=n}^p a_i = \sum_{i=n-\ell}^{p-\ell} a_{i+\ell}.$$

Exemple 4.1.11

$$\sum_{k=0}^{n-1} (k+1)^2 = \sum_{k=1}^n k^2.$$

I.3 Additivité par rapport aux bornes

La propriété suivante est à rapprocher de la relation de Chasles pour les intégrales.

Lemme 4.1.12

Soit n et m deux entiers tels que $n \leq m$. Alors

$$\sum_{k=1}^m a_k = \sum_{k=1}^n a_k + \sum_{k=n+1}^m a_k.$$

▫ Éléments de preuve.

C'est une récurrence immédiate sur $m \geq n$, à n fixé. L'initialisation est la convention des sommes vides. ▷

Proposition 4.1.13 (Somme indexée dans une union disjointe)

On suppose que $I = I_1 \sqcup I_2$ est une union disjointe, et I fini. Alors :

$$\sum_{i \in I} a_i = \sum_{i \in I_1} a_i + \sum_{i \in I_2} a_i.$$

▫ Éléments de preuve.

À partir de bijections $[1, n] \rightarrow I_1$ et $[1, m] \rightarrow I_2$, construire une bijection $[1, n+m] \rightarrow I_1 \sqcup I_2$ (remarquez qu'on a déjà construit une telle bijection ; dans quel contexte ?) ▷

Exemples 4.1.14

1. Soit $E = \{0, \dots, 2n-1\}$. En écrivant E sous la forme de l'union de ses éléments pairs et de ses éléments impairs, calculer $\sum_{i=0}^{2n-1} \left\lfloor \frac{i}{2} \right\rfloor$, où $\lfloor x \rfloor$ désigne la partie entière de x .
2. En considérant $\sum_{i \in I_1 \sqcup I_2} 1$, on retrouve la propriété des cardinaux des unions disjointes.

Avertissement 4.1.15

Attention à prendre une union *disjointe*, sinon on somme deux fois chaque élément indexé par un indice de l'intersection. Sinon, considérer la propriété suivante.

Proposition 4.1.16

Si l'union $I_1 \cup I_2$ n'est pas nécessairement disjointe,

$$\sum_{i \in I_1 \cup I_2} a_i = \sum_{i \in I_1} a_i + \sum_{i \in I_2} a_i - \sum_{i \in I_1 \cap I_2} a_i.$$

▫ Éléments de preuve.

Se ramener à la propriété précédente en considérant l'union disjointe $(I_1 \setminus I_2) \sqcup (I_1 \cap I_2) \sqcup (I_2 \setminus I_1)$. ▷

Remarque 4.1.17

Là encore, en considérant $\sum_{i \in I_1 \cup I_2} 1$, on retrouve la propriété du cardinal d'une union quelconque.

Plus généralement, on obtient le résultat suivant :

Proposition 4.1.18 (Sommation par groupement de termes)

Soit I un ensemble fini et I_1, \dots, I_n deux à deux disjointx tels que $I_1 \sqcup \dots \sqcup I_n = I$ (si les I_k sont de plus non vides, la donnée de tels ensembles est ce qu'on appelle une partition de I). Soit $(a_i)_{i \in I}$ une famille. Alors

$$\sum_{i \in I} a_i = \sum_{i \in I_1} a_i + \sum_{i \in I_2} a_i + \dots + \sum_{i \in I_n} a_i = \sum_{j=1}^n \sum_{i \in I_j} a_i.$$

⊣ Éléments de preuve.

On a déjà utilisé le cas $n = 3$ dans la propriété précédente. Le cas général s'obtient par récurrence immédiate. ▷

I.4 Linéarité**Proposition 4.1.19 (Linéarité du symbole \sum)**

Soit I un ensemble fini et $(a_i)_{i \in I}$ et $(b_i)_{i \in I}$ deux familles (réelles ou complexes), et λ, μ deux nombres réels ou complexes. Alors :

- $\sum_{i \in I} a_i + \sum_{i \in I} b_i = \sum_{i \in I} (a_i + b_i).$
- $\lambda \sum_{i \in I} a_i = \sum_{i \in I} \lambda a_i.$
- En combinant les deux égalités : $\lambda \sum_{i \in I} a_i + \mu \sum_{i \in I} b_i = \sum_{i \in I} (\lambda a_i + \mu b_i).$

⊣ Éléments de preuve.

Se ramener par énumération au cas $I = \llbracket 1, n \rrbracket$, puis récurrence sur n . ▷

Cette proposition énonce le fait que \sum est une « forme linéaire sur l'espace vectoriel des familles indexées par un ensemble fini donné I . » (voir chapitre *Espaces vectoriel et Applications linéaires*)

Corollaire 4.1.20 (somme de termes constants)

Soit E un ensemble fini et a un nombre réel ou complexe. Alors :

$$\sum_{i \in E} a = a \cdot \sum_{i \in E} 1 = a \cdot |E|.$$

Exemple 4.1.21

$$\sum_{k=0}^n k(k+1) = \sum_{k=0}^n k^2 + \sum_{k=0}^n k = \frac{n(n+1)(2n+1)}{6} + \frac{n(n+1)}{2}.$$

Remarque 4.1.22

Attention à prendre des sommes indexées sur le *même* ensemble ! Si ce n'est pas le cas, on ne peut regrouper les sommes que sur l'intersection des indices, en laissant chacun dans leur somme les éléments indexés hors de cette intersection :

$$\sum_{i \in I_1} a_i + \sum_{i \in I_2} b_i = \sum_{i \in I_1 \cap I_2} (a_i + b_i) + \sum_{i \in I_1 \setminus (I_1 \cap I_2)} a_i + \sum_{i \in I_2 \setminus (I_1 \cap I_2)} b_i.$$

Exemples 4.1.23

$$1. \sum_{i=0}^n a_i + \sum_{i=1}^{n+1} b_i = a_0 + \sum_{i=1}^n (a_i + b_i) + b_{n+1}.$$

$$2. \text{ Transformation d'Abel : Soit } (a_n)_{n \in \mathbb{N}} \text{ et } (b_n)_{n \in \mathbb{N}} \text{ deux suites, et pour tout } n \in \mathbb{N}, B_n = \sum_{k=0}^n b_k.$$

Montrer que

$$\sum_{k=0}^n a_k b_k = \sum_{k=0}^{n-1} B_k (a_k - a_{k+1}) + a_n B_n.$$

I.5 Sommes télescopiques

Définition 4.1.24 (somme télescopique)

On dit qu'une somme $\sum_{k=0}^n a_k$ est télescopique si pour tout $k \in \{0, \dots, n\}$, on peut écrire de façon simple a_k sous la forme $a_k = b_{k+1} - b_k$.

Les sommes télescopiques se calculent facilement. La technique utilisée (séparer la somme en deux et faire un changement d'indice sur une des deux sommes) est à retenir : elle s'adapte à des situations plus générales.

Proposition 4.1.25 (calcul des sommes télescopiques)

Soit $\sum (b_{k+1} - b_k)$ une somme télescopique. Alors :

$$\sum_{k=0}^n (b_{k+1} - b_k) = b_{n+1} - b_0.$$

▫ Éléments de preuve.

Soit par récurrence, soit en séparant par linéarité, en faisant un changement d'indice sur l'une des deux sommes, en regroupant sur les indices communes, et en simplifiant. ▷

Exemples 4.1.26

$$1. \text{ Calculer } \sum_{k=0}^n k \cdot k!.$$

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

$$3. \text{ Trouver un polynôme } P \text{ de degré 2 tel que pour tout } x \in \mathbb{R}, P(x+1) - P(x) = x. \text{ En déduire } \sum_{k=n}^m k.$$

4. Trouver des réels (a, b, c) tels que pour tout $k \in \mathbb{N}^*$,

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

En déduire l'expression de $\sum_{k=1}^n \frac{1}{k(k+1)(k+2)}$. Le retrouver par un télescopage simple.

I.6 Cas des produits

On peut adapter ces résultats au cas des produits :

Les règles similaires pour le produit sont :

Proposition 4.1.27 (Règles pour les produits)

Avec des notations cohérentes, on obtient les règles suivantes :

- Si $I_1 \cap I_2 = \emptyset$, $\prod_{i \in I_1} a_i \prod_{i \in I_2} a_i = \prod_{i \in I_1 \cup I_2} a_i$.
- $\left(\prod_{i \in I} a_i\right)^\lambda \left(\prod_{i \in I} b_i\right)^\mu = \prod_{i \in I} (a_i^\lambda b_i^\mu)$.
- $\prod_{i \in I} a = a^{|I|}$.

Définition 4.1.28 (produit télescopique)

On dit qu'un produit $\prod_{k=0}^n a_k$ est télescopique si pour tout $k \in \{0, \dots, n\}$, on peut écrire de façon simple a_k sous la forme $a_k = \frac{b_{k+1}}{b_k}$.

Proposition 4.1.29 (calcul des produits télescopiques)

Soit $\prod \frac{b_{k+1}}{b_k}$ un produit télescopique. Alors :

$$\prod_{k=0}^n \frac{b_{k+1}}{b_k} = \frac{b_{n+1}}{b_0}.$$

I.7 Sommes multiples

Nous étudions maintenant les sommes multiples. Certaines familles peuvent être indexées sur un produit cartésien (ou au moins un sous-ensemble). Soit K un sous-ensemble de $I \times J$, et $(a_{i,j})_{(i,j) \in K}$ une famille doublement indexée (*i.e.* indexée sur un produit cartésien). Le but est d'étudier la somme $\sum_{(i,j) \in K} a_{i,j}$ en se ramenant à des sommes portant sur un seul des deux indices. Pour cela, on introduit la notion de « coupe » de l'ensemble K .

Définition 4.1.30 (coupes d'un sous-ensemble de $I \times J$, voir figure 4.1)

Soit $K \subset I \times J$.

- Soit $i \in I$; on définit la coupe de K suivant i :

$$K_{i,\bullet} = \{j \in J \mid (i, j) \in K\}.$$

- Soit $j \in J$; on définit la coupe de K suivant j par :

$$K_{\bullet,j} = \{i \in I \mid (i,j) \in K\}$$

Définissons également :

- $K'_{i,\bullet} = \{(i,j) \mid j \in K_{i,\bullet}\} = K \cap (\{i\} \times J) = \{(i,j) \mid (i,j) \in K, j \in J\}$
- $K'_{\bullet,j} = \{(i,j) \mid i \in K_{\bullet,j}\} = K \cap (I \times \{j\}) = \{(i,j) \mid (i,j) \in K, i \in I\}$.

Ainsi, $K_{i,\bullet}$ est le projeté sur J de $K'_{i,\bullet}$ et $K_{\bullet,j}$ est le projeté sur I de $K'_{\bullet,j}$.

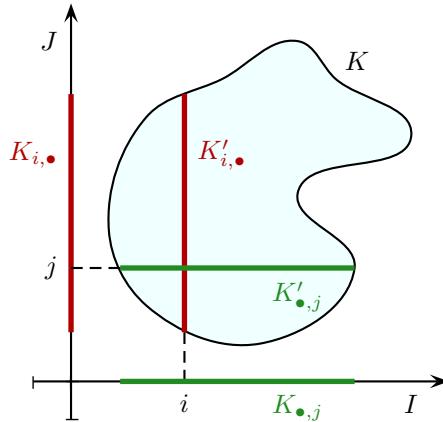


FIGURE 4.1 – Coupes d'un ensemble

Théorème 4.1.31 (Interversion de signes somme)

Soit $K \subset I \times J$, et $(a_{i,j})_{(i,j) \in K}$ une famille indexée sur K . Alors :

$$\begin{aligned} \sum_{(i,j) \in K} a_{i,j} &= \sum_{i \in I} \sum_{j \in K_{i,\bullet}} a_{i,j} = \sum_{i \in I} \sum_{(i,j) \in K'_{i,\bullet}} a_{i,j} \\ &= \sum_{j \in J} \sum_{i \in K_{\bullet,j}} a_{i,j} = \sum_{j \in J} \sum_{(i,j) \in K'_{\bullet,j}} a_{i,j}. \end{aligned}$$

◊ Éléments de preuve.

En effet, $(K'_{i,\bullet})_{i \in I}$ forme un partage de K . Est-ce une partition ? ▷

Dans la pratique ce résultat est très fréquemment utilisé pour intervertir des signes somme (passer du deuxième terme au troisième terme de cette égalité) lorsque les bornes de l'indice de la somme interne dépendent de l'indice de la somme externe (dans ce cas, on essaie de voir la somme interne comme la somme sur une certaine coupe).

En particulier, lorsque $K = I \times J$, on obtient $K_{i,\bullet} = J$ pour tout i , et $K_{\bullet,j} = I$ pour tout J . Ainsi :

Corollaire 4.1.32 (somme sur un pavé)

Soit une famille $(a_{i,j})_{(i,j) \in I \times J}$. Alors

$$\sum_{i \in I} \sum_{j \in J} a_{i,j} = \sum_{j \in J} \sum_{i \in I} a_{i,j} = \sum_{(i,j) \in I \times J} a_{i,j}$$

Une récurrence immédiate permet de généraliser à une somme sur $I_1 \times I_2 \times \cdots \times I_n$: on peut alors enchaîner les n sommes dans l'ordre qu'on veut.

Avertissement 4.1.33

Attention, en général, si on n'est pas sur un pavé, il ne suffit pas d'intervertir purement et simplement les signes sommes. Dans la plupart des cas, une telle interversion amène d'ailleurs à une expression qui n'a pas de sens.

Ceci est un peu plus qu'un exemple. C'est une technique incontournable :

Exemple 4.1.34 (somme sur un triangle, à savoir faire !)

Montrer que $\sum_{i=0}^n \sum_{j=0}^i a_{i,j} = \sum_{j=0}^n \sum_{i=j}^n a_{i,j}$.

I.8 Produits de sommes

Dans la formule de la double somme, si le terme général s'exprime comme produit de deux termes, chacun d'entre eux ne dépendant que d'un des deux indices, on obtient l'expression du produit de deux sommes :

Proposition 4.1.35 (Produit de deux sommes)

$$\left(\sum_{i \in I} a_i \right) \left(\sum_{j \in J} b_j \right) = \sum_{i \in I} \sum_{j \in J} a_i b_j = \sum_{(i,j) \in I \times J} a_i b_j.$$

▫ Éléments de preuve.

Partant du milieu (égal au terme de droite), factoriser d'abord la somme interne par a_i (à i fixé). On peut ensuite factoriser la somme externe par $\sum b_j$. ▷

Remarque 4.1.36

Avec les a_i et b_j tous égaux à 1, on retrouve l'expression du cardinal d'un produit cartésien.

Remarque 4.1.37

Attention, si on effectue le produit de deux sommes indexées sur le même ensemble, et pour lesquels le même indice est utilisé, pensez à d'abord rendre les indices indépendants (les indices étant muets, changez l'un des deux afin d'avoir deux indices différents) :

$$\left(\sum_{i=1}^n a_i \right) \left(\sum_{i=1}^n b_i \right) = \left(\sum_{i=1}^n a_i \right) \left(\sum_{j=1}^n b_j \right) = \sum_{(i,j) \in \llbracket 1, n \rrbracket^2} a_i b_j.$$

Avertissement 4.1.38

Attention, en revanche,

$$\left(\sum_{i=1}^n a_i \right) \left(\sum_{i=1}^n b_i \right) \neq \sum_{i=1}^n a_i b_i,$$

Le terme de gauche est la somme sur un carré alors que la somme de droite n'est la somme que sur la diagonale de ce carré !

Le produit de deux sommes se généralise de la façon suivante :

Théorème 4.1.39 (Distributivité généralisée)

$$\prod_{k=1}^n \left(\sum_{i=1}^{m_k} a_{k,i} \right) = \sum_{(i_1, \dots, i_n) \in \llbracket 1, m_1 \rrbracket \times \dots \times \llbracket 1, m_n \rrbracket} a_{1,i_1} \cdots a_{n,i_n}.$$

▫ Éléments de preuve.

Récurrence sur n .

▷

II Sommes classiques à connaître

De nombreuses sommes se calculent en se ramenant à des sommes connues. Pour cette raison, il est important d'avoir un catalogue de sommes (finies) qu'on sait calculer. Ces sommes sont celles de ce paragraphe, auxquelles s'ajoutent les sommes télescopiques, et certaines sommes obtenues par des techniques d'analyse (dérivation des sommes géométriques par exemple).

II.1 Somme des puissances d'entiers

Proposition 4.2.1 (somme des puissances d'entiers, petits exposants)

Pour tout $n \in \mathbb{N}$,

- $\sum_{k=1}^n k^0 = \sum_{k=1}^n 1 = n$
- $\sum_{k=1}^n k = \frac{n(n+1)}{2}$.
- $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$.
- $\sum_{k=1}^n k^3 = \left(\frac{n(n+1)}{2} \right)^2$.

▫ Éléments de preuve.

D'innombrables (indénombrables ?) possibilités. Géométriquement pour certaines (voir ci-dessous), ou classiquement par récurrence sur n , ou encore de proche en proche, en augmentant petit à petit la valeur de k (voir méthode ci-dessous), ou encore par manipulations de sommes doubles à intervertir etc.

▷

On donne une interprétation géométrique des cas des exposants 1 et 3 dans la figure 4.2

Remarque 4.2.2

En utilisant les formules précédentes, on obtient sans peine la somme des nombres impairs consécutifs :

$$\sum_{k=1}^n (2k - 1) = n^2.$$

Cette formule a une interprétation géométrique toute simple, déjà connue des Grecs antiques (Euclide) : elle est donnée en figure 4.3.

$$\text{côté : } 1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

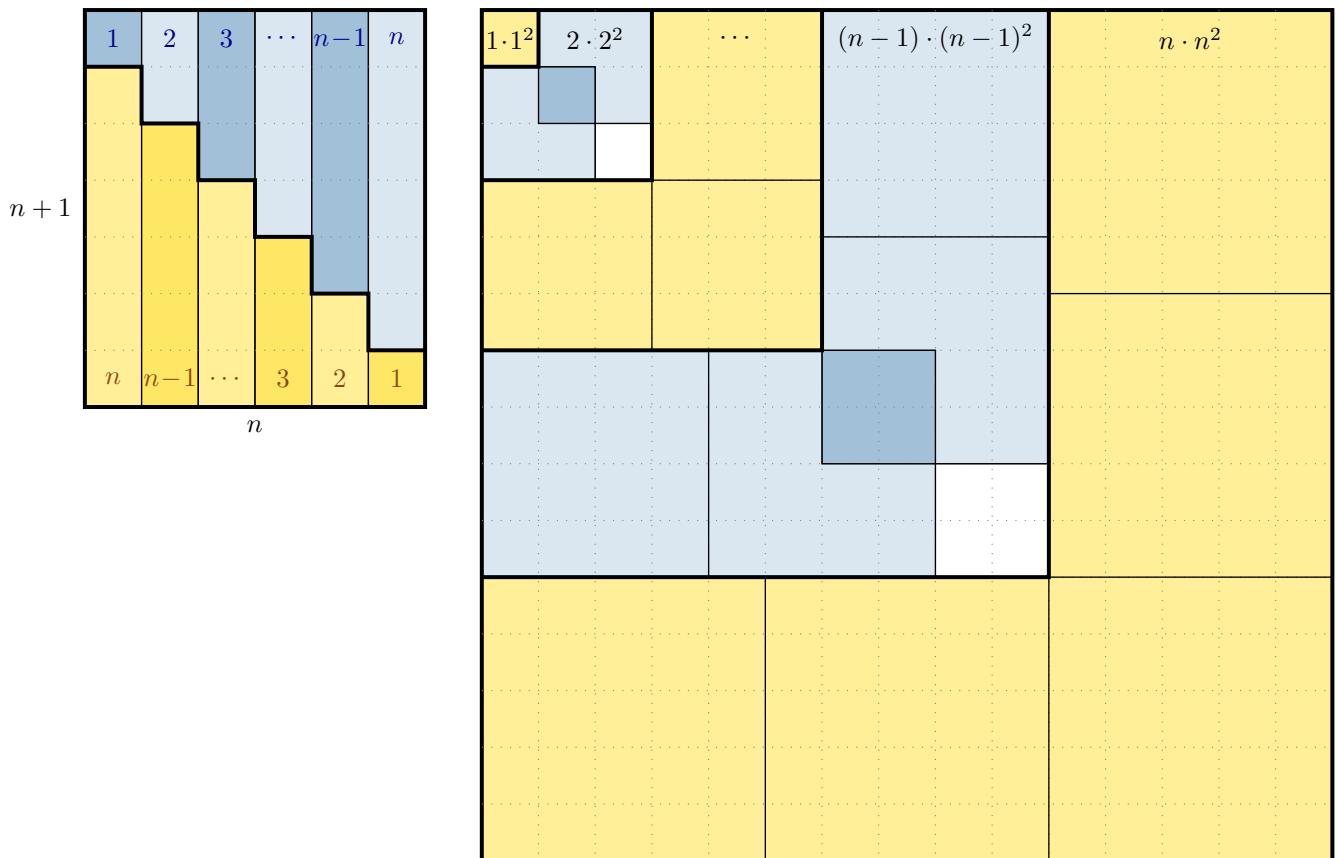


FIGURE 4.2 – Illustration géométrique de $\sum_{k=1}^n k$ et $\sum_{k=1}^n k^3$.

Exemples 4.2.3

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

II.2 Sommes géométriques

La deuxième grande famille de sommes à bien connaître est la famille des sommes géométriques.

Proposition 4.2.4 (Sommes géométriques)

Soit $z \in \mathbb{C}$ et $n \in \mathbb{N}$. Alors :

- si $z = 1$, $\sum_{k=0}^n z^k = n + 1$;
- si $z \neq 1$, $\sum_{k=0}^n z^k = \frac{1 - z^{n+1}}{1 - z}$.

▫ Éléments de preuve.

Dans le cas $z \neq 1$, calculer $(1 - z) \sum_{k=0}^n z^k$ par télescopage. ▷

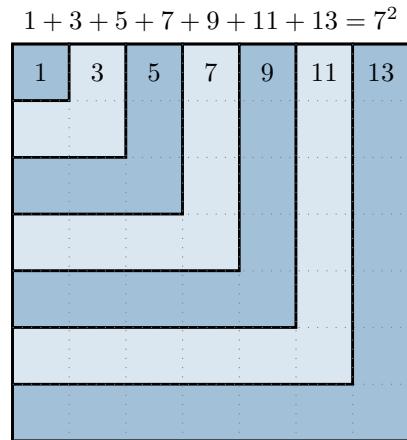


FIGURE 4.3 – Illustration géométrique de la somme des entiers impairs

Exemples 4.2.5

1. Calculer $\sum_{k=0}^{4n} (2i)^k$.
2. Calculer $\sum_{k=0}^n p^{2k}$ en fonction de $p \in \mathbb{R}$ et $n \in \mathbb{N}$.
3. Calculer, pour $n \geq 4$, $\sum_{k=6}^{n+2} e^{-3k}$.

En passant à la limite lorsque n tend vers $+\infty$, on obtient :

Théorème 4.2.6 (Séries géométriques)

Soit $z \in \mathbb{C}$. La série $\sum z^n$ converge si et seulement si $|z| < 1$, et dans ce cas, $\sum_{n=0}^{+\infty} z^n = \frac{1}{1-z}$

Proposition 4.2.7 (Factorisations de $a^n - b^n$ et $a^n + b^n$, Bernoulli)

Soit a et b des nombres complexes et n un entier. Alors :

- $a^n - b^n = (a - b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \cdots + ab^{n-2} + b^{n-1}) = (a - b) \sum_{k=0}^{n-1} a^{n-1-k}b^k$;
- En particulier, pour $b = 1$, on obtient :

$$a^n - 1 = (a - 1)(a^{n-1} + a^{n-2} + \cdots + a + 1) = (a - 1) \sum_{k=0}^{n-1} a^k;$$

- si n est impair, alors :

$$a^n + b^n = (a + b)(a^{n-1} - a^{n-2}b + a^{n-3}b^2 - \cdots - ab^{n-2} + b^{n-1}) = (a + b) \sum_{k=0}^{n-1} (-1)^k a^{n-1-k}b^k;$$

▫ Éléments de preuve.

Encore du télescopage.

▷

Exemple 4.2.8

1. Que retrouve-t-on pour $n = 2$?
2. Quelles sont les racines dans \mathbb{C} du polynôme $1 + X + X^2$? du polynôme $1 + X + X^2 + X^3$? Et plus généralement ?

II.3 Formule du binôme**Définition 4.2.9 (Coefficient binomial, définition algébrique)**

On définit, pour tout $n \in \mathbb{N}$, et $k \in \llbracket 0, n \rrbracket$,

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

Pour toute autre valeur de $(n, k) \in \mathbb{Z}^2$, on pose par convention $\binom{n}{k} = 0$.

Proposition 4.2.10 (interprétation combinatoire des coefficients binomiaux, admis)

$\binom{n}{k}$ est le nombre de sous-ensembles à k éléments d'un ensemble à n élément. Autrement dit, si $|E| = n$,

$$|\mathcal{P}_k(E)| = \binom{n}{k}.$$

Proposition 4.2.11 (symétrie des coefficients binomiaux)

De façon évidente, $\binom{n}{k} = \binom{n}{n-k}$.

Exemple 4.2.12

Les valeurs particulières suivantes sont à bien graver dans votre mémoire :

- $\binom{n}{0} = \binom{n}{n} = 1$
- $\binom{n}{1} = \binom{n}{n-1} = n$
- $\binom{n}{2} = \binom{n}{n-2} = \frac{n(n-1)}{2}$.

Les deux formules suivantes sont essentielles, et peuvent être démontrées algébriquement ou par un argument combinatoire :

Théorème 4.2.13 (Formule de Pascal)

Pour tout $(n, k) \neq (0, 0)$,

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

Théorème 4.2.14 (Formule comité-président)

Pour tout $(n, k) \in \mathbb{Z}$,

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

Vous comprendrez le nom de cette formule lorsque nous ferons de la combinatoire.

Grâce à la formule de Pascal, on peut construire les coefficients binomiaux par ligne (une ligne représentant une valeur de n donnée), de proche en proche. Cette construction est particulièrement utile lorsqu'on recherche des coefficients binomiaux pour une valeur raisonnablement petite de n (disons $n \leq 10$), par exemple en vue d'utiliser la formule du binôme de Newton (présentée un peu plus loin). La construction de ce triangle, appelé triangle de Pascal, est expliquée dans la figure 4.4.

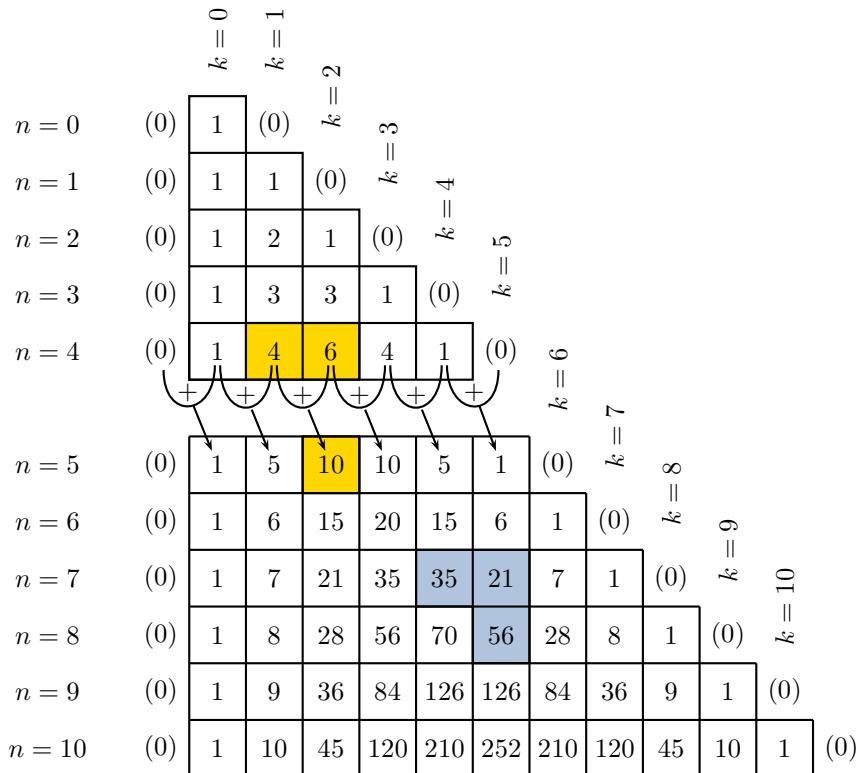


FIGURE 4.4 – Triangle de Pascal pour le calcul de $\binom{n}{p}$

Les coefficients des 6 premières lignes ($n \in \llbracket 0, 5 \rrbracket$) sont à connaître par cœur.

Note Historique 4.2.15

Le triangle des coefficients binomiaux, que nous appelons communément « triangle de Pascal » était en fait connu depuis bien longtemps déjà lorsque Blaise Pascal s'y intéressa : on y trouve mention déjà chez Halayudha, mathématicien indien du 10^e siècle, ainsi qu'en Chine au 13^e siècle. La contribution de Pascal a essentiellement été de démontrer en 1654 un grand nombre de propriétés de ce triangle, jusque-là admises. C'est d'ailleurs à cette occasion qu'il a mis au point le principe de la démonstration par récurrence !

Théorème 4.2.16 (Formule du binôme de Newton)

Soit a et b deux nombres complexes, et $n \in \mathbb{N}$. Alors

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

▫ Éléments de preuve.

Récurrence sur n , en utilisant la formule de Pascal.

▷

Note Historique 4.2.17

La formule du binôme était elle aussi déjà connue depuis longtemps, notamment en Inde et en Chine, en relation avec le triangle de Pascal. L'apport de Newton est bien réel et important, mais ne concerne pas les exposants entiers naturels. Isaac Newton a en fait généralisé cette formule pour tout exposant réel, sous forme d'une somme infinie. Pour les entiers négatifs, c'est ce qu'on appelle la « formule du binôme négatif ».

Plus tard dans l'année, nous reprendrons les différentes identités de cette section pour les voir d'un point de vue combinatoire.

II.4 Retour sur les sommes de puissances d'entiers

Notons, pour tout $(n, p) \in (\mathbb{N}^*)^2$,

$$S_n(p) = \sum_{k=1}^n k^p.$$

On peut calculer $S_n(p)$ de proche en proche à l'aide de la formule du binôme (rappelée un peu plus loin).

Méthode 4.2.18 (Calcul de proche en proche des $S_n(p)$)

Pour calculer $S_n(p)$ en fonction des sommes précédentes, considérer la somme :

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

La somme S peut être calculée d'une part comme somme télescopique, d'autre part à l'aide du développement du binôme (ce qui fait partir les termes d'exposant $p+1$). Isoler ensuite les termes d'exposant p .

Relations

« Ce n'est pas un lemme, et il n'est pas de moi »

(Max Zorn, à propos du « lemme de Zorn »)

Comme on l'a vu, les applications, ou plus généralement les fonctions, sont une façon de mettre en relation deux ensembles, donc de faire interagir nos objets formels que sont les ensembles. Mais le rôle des deux ensembles ainsi mis en relation n'est pas symétrique, du fait de la contrainte d'unicité de l'image, alors qu'on n'a pas la contrainte d'unicité de l'antécédent. Nous voyons dans ce chapitre comment définir une notion plus générale, permettant de retrouver cette symétrie perdue.

Ayant déjà étudié les relations définissant les applications ou les fonctions (relations applicationnelles ou fonctionnelles), nous étudierons plus précisément les deux autres types importants de relations à bien connaître : les relations d'équivalence, et les relations d'ordre.

I Définitions générales

I.1 Relations

Définition 5.1.1 (relation binaire)

- Une *relation binaire* entre deux ensembles E et F est un sous-ensemble G de $E \times F$.
- On note souvent $x\mathcal{R}y$ pour dire que $(x, y) \in G$, et on dit que x est en relation avec y . On parle alors de la relation \mathcal{R} .
- Certains types de relation sont aussi notés $x \equiv y$, ou $x \sim y$, ou $x \leq y \dots$

Exemples 5.1.2

1. appartenance à une droite
2. L'inclusion entre ensemble. L'appartenance entre ensembles.
3. Les congruences d'entiers.

Comme pour les fonctions, on peut représenter une relation par un diagramme sagittal. Par exemple, la relation représentée par la figure 5.1 est la relation entre $\llbracket 1, 5 \rrbracket$ et $\llbracket 1, 4 \rrbracket$ définie par le sous-ensemble $G = \{(1, 2), (1, 3), (2, 3), (2, 4), (4, 1), (5, 1), (5, 2), (5, 4)\}$, donc la relation définie par $1\mathcal{R}2, 1\mathcal{R}3, 2\mathcal{R}3, 2\mathcal{R}4, 4\mathcal{R}1, 5\mathcal{R}1, 5\mathcal{R}2, 5\mathcal{R}4$, les autres paires n'étant pas en relation.

On peut également représenter cette relation sous forme d'un tableau à double-entrée, en décidant de représenter par deux signes distinctifs le fait que $x\mathcal{R}y$ soient en relation ou non (par exemple par une croix ou par rien) Ainsi, la relation précédente est représentée par le tableau de la figure 5.2

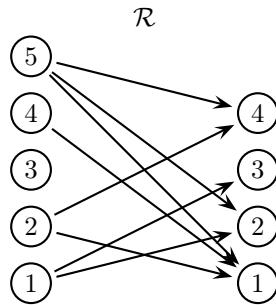


FIGURE 5.1 – Diagramme sagittal d'une relation

<i>F</i>	1	2	3	4
<i>E</i>				
1		×	×	
2	×			×
3				
4	×			
5	×	×		×

FIGURE 5.2 – Représentation tabulaire (ou matricielle) d'une relation

Remarque 5.1.3

Les fonctions et les applications sont des cas particuliers de relation binaire. Une relation binaire définissant une fonction est appelée relation fonctionnelle.

Plus précisément :

Définition 5.1.4 (relation fonctionnelle (resp. applicationnelle))

Une relation \mathcal{R} entre E et F est *fonctionnelle* (resp. *applicationnelle*) si pour tout x de E il existe au plus un (resp. un et un seul) y de F tel que $x\mathcal{R}y$.

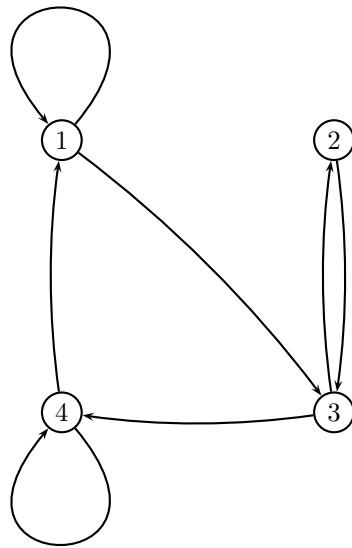
Lorsque \mathcal{R} est une relation entre E et lui-même, on dit que \mathcal{R} est une relation sur E . Dans ce cas, on dispose d'une troisième représentation possible, correspondant à la représentation sagittale dans laquelle on a identifié les éléments des deux ensembles de départ et d'arrivée. On obtient de la sorte un graphe orienté dont les sommets sont les éléments de E . Par exemple, la relation définie sur $\llbracket 1, 4 \rrbracket$ par $1\mathcal{R}1, 1\mathcal{R}3, 2\mathcal{R}3, 3\mathcal{R}2, 3\mathcal{R}4, 4\mathcal{R}1$ et $4\mathcal{R}4$ est représentée par le graphe de la figure 5.3

I.2 Définition de quelques propriétés sur les relations

On définit maintenant un certain nombre de propriétés susceptibles d'être satisfaites par une relation d'un ensemble E dans lui-même. Les 4 premières sont à bien connaître, les 2 dernières sont plus anecdotiques.

Définition 5.1.5 (reflexivité, symétrie, antisymétrie, transitivité et al.)

Soit \mathcal{R} une relation sur E . On dit que :

FIGURE 5.3 – Graphe d'une relation sur E

- \mathcal{R} est réflexive si pour tout $x \in E$, $x\mathcal{R}x$;
- \mathcal{R} est symétrique si pour tout $(x, y) \in E^2$, $x\mathcal{R}y \implies y\mathcal{R}x$;
- \mathcal{R} est antisymétrique si pour tout $(x, y) \in E^2$, $(x\mathcal{R}y) \wedge (y\mathcal{R}x) \implies (x = y)$;
- \mathcal{R} est transitive si pour tout $(x, y, z) \in E^3$, $(x\mathcal{R}y) \wedge (y\mathcal{R}z) \implies (x\mathcal{R}z)$.
- \mathcal{R} est irréflexive (ou antiréflexive) si pour tout x , on a $\neg(x\mathcal{R}x)$
- \mathcal{R} est asymétrique si $x\mathcal{R}y \implies \neg(y\mathcal{R}x)$.

Remarque 5.1.6

Comment qualifiez-vous une relation antisymétrique et irréflexive ?

Nous allons maintenant définir deux types de relations que l'on rencontre fréquemment.

II Relations d'équivalence

II.1 Définitions et exemples

Définition 5.2.1 (relation d'équivalence)

Une relation d'équivalence sur E est une relation réflexive, symétrique et transitive. On note souvent une relation d'équivalence $x \equiv y$ ou $x \sim y$.

Exemples 5.2.2

1. Égalité.
2. Congruences modulo n dans \mathbb{N} (notation $\equiv_{[n]}$ ou $\equiv ...[n]$).
3. Congruences dans \mathbb{R} .
4. Appartenance à la même partie d'une partition.
5. La relation définissant \mathbb{Q} sur $\mathbb{Z} \times \mathbb{N}^*$: $(p, q) \equiv_{\mathbb{Q}} (p', q')$ si et seulement $pq' = p'q$ (ces deux couples vont définir le même rationnel).
6. Conjugaison dans \mathfrak{S}_n .
7. Multiplication par un scalaire non nul dans \mathbb{R}^n ou \mathbb{C}^n .

Une relation d'équivalence sur un ensemble fini peut être représenté par son graphe orienté (figure 5.4). Ce graphe possède un certain nombre de caractéristiques :

- Le graphe se décompose en un certain nombre de blocs non reliés les uns les autres, les points au sein d'un même bloc étant reliés (on parle de composantes connexes du graphe)
- Chaque point appartient à un bloc (il y est éventuellement seul). Ainsi, les blocs forment une partition de E .
- À l'intérieur de chaque bloc, toutes les flèches possibles sont présentes (y compris celles reliant un point et lui-même) : il s'agit d'un sous-graphe complet.

Les différents blocs sont appelés *classes d'équivalence*.

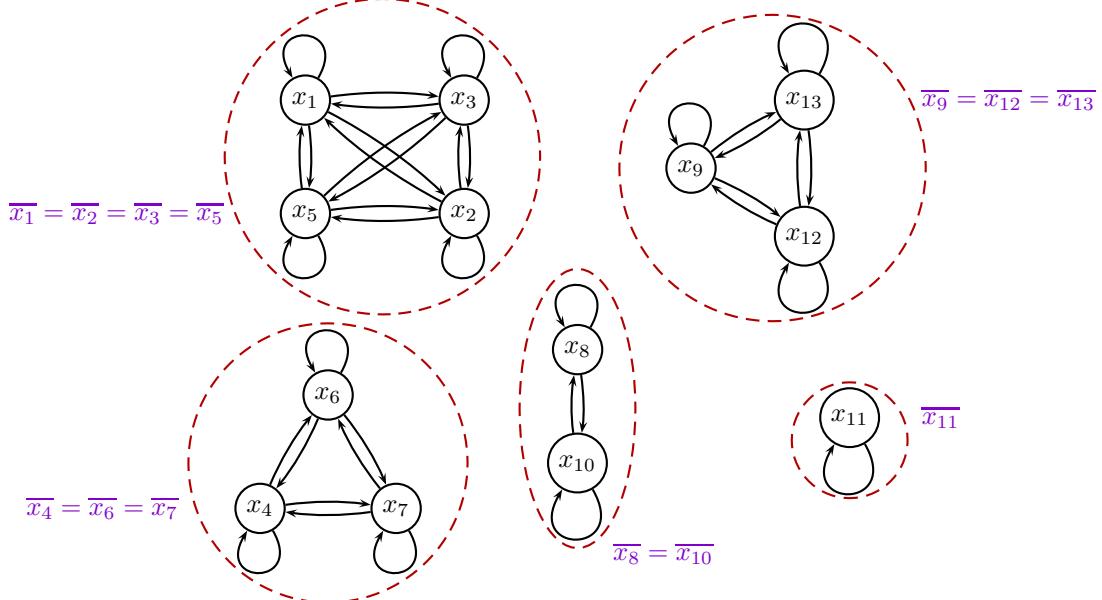


FIGURE 5.4 – Graphe d'une relation d'équivalence sur $E = \{x_1, x_2, \dots, x_{13}\}$.

Cette situation se généralise. C'est ce que nous étudions dans le paragraphe suivant.

II.2 Classes d'équivalence, ensembles quotients

Définition 5.2.3 (classes d'équivalence)

Soit \mathcal{R} une relation d'équivalence sur E , et $x \in E$. La *classe d'équivalence de x sous la relation \mathcal{R}* est le sous-ensemble C_x de E constitué des éléments en relation avec x :

$$C_x = \{y \in E \mid x\mathcal{R}y\}.$$

Lemme 5.2.4

Si y et z sont dans une même classe d'équivalence, alors $y\mathcal{R}z$.

▫ Éléments de preuve.

C'est juste la transitivité, en passant par x , définissant la classe commune. ▷

Théorème 5.2.5 (Partition formée par les classes d'équivalence)

Soit E un ensemble, et \mathcal{R} une relation d'équivalence sur E . L'ensemble des classes d'équivalence sous \mathcal{R} forme une partition de E .

▫ Éléments de preuve.

Il suffit de montrer que si $C_x \cap C_y \neq \text{vide}$, alors $C_x = C_y$, et par symétrie, $C_x \subset C_y$ suffit. Là encore, c'est la transitivité, en passant par un point appartenant aux 2 classes. ▷

En particulier, si $y \in C_x$, alors $C_x = C_y$.

Pour les propriétés « stables » par la relation d'équivalence, les points d'une même classe d'équivalence jouent des rôles similaires, et n'ont pas lieu d'être distingués. On formalise cela en introduisant un ensemble dont les éléments sont les classes d'équivalences (ainsi, tous les points d'une même classe représentent la même classe, et sont considérés comme égaux dans ce nouvel ensemble) :

Définition 5.2.6 (Ensemble quotient, HP)

L'ensemble des classes sous la relation \mathcal{R} s'appelle *l'ensemble quotient de E par R*, et est noté E/\mathcal{R} . C'est un sous-ensemble de $\mathcal{P}(E)$.

Ainsi, en notant \bar{x} la classe d'équivalence d'un élément x de E , l'ensemble E/\mathcal{R} est l'ensemble formé des éléments \bar{x} , où l'on impose $\bar{x} = \bar{y}$ dès que $x\mathcal{R}y$.

Exemples 5.2.7

1. On définit \mathbb{Q} comme étant le quotient $(\mathbb{Z} \times \mathbb{N}^*)/\equiv_{\mathbb{Q}}$.
2. On définit $\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}/\equiv_{[n]}$. C'est un ensemble à n éléments.

Définition 5.2.8 (projection canonique)

On appelle projection canonique de E sur E/\mathcal{R} l'application $\pi_{\mathcal{R}}$ qui à x associe sa classe \bar{x} . Par définition, $\pi_{\mathcal{R}}$ est surjective, et vérifie :

$$\forall(x, y) \in E^2 \quad x\mathcal{R}y \implies \pi_{\mathcal{R}}(x) = \pi_{\mathcal{R}}(y).$$

Cette projection canonique vérifie la propriété importante suivante (une telle propriété est appelée « propriété universelle ») :

Théorème 5.2.9 (Factorisation d'une application constante sur les classes d'équivalence)

Soit $f : E \rightarrow F$. Alors les deux propriétés suivantes sont équivalentes :

- (i) Pour tout $(x, y) \in E^2$, $x\mathcal{R}y \implies f(x) = f(y)$;
- (ii) il existe $g : E/\mathcal{R} \rightarrow F$ tel que $f = g \circ \pi_{\mathcal{R}}$.

▫ Éléments de preuve.

Définir $g(\bar{x})$ comme étant la valeur commune de tous les $f(y)$, pour $y \in \bar{x}$. ▷

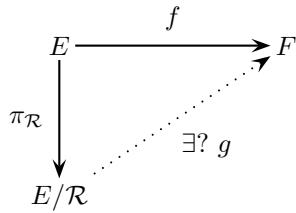


FIGURE 5.5 – Factorisation d'une application constante sur les classes d'équivalence

On dit dans la situation ci-dessus que la fonction f « passe au quotient ».

Ce résultat se généralise sans difficulté pour des fonctions de plusieurs variables, $f : E_1 \times \cdots \times E_n \rightarrow F$, les E_i étant chacun muni d'une relation d'équivalence \mathcal{R}_i . Dès lors que f « respecte » les relations

d'équivalences (donc la valeur de f ne dépend pas du choix des représentants des classes d'équivalence), on a la possibilité de factoriser f au travers de l'espace produit des espaces quotients $(E_1/\mathcal{R}_1) \times \cdots \times (E_n/\mathcal{R}_n)$. Cela permet notamment de définir assez facilement des lois (addition, multiplication) sur des ensembles quotients. C'est ce que nous étudions ci-dessous.

II.3 Congruences

L'ensemble \mathbb{Z} étant muni d'une loi d'addition et d'une loi de multiplication, on aimerait savoir si ces lois « passent au quotient » autrement dit, si elles permettent de définir une somme et un produit sur $\mathbb{Z}/n\mathbb{Z}$. La notion de congruence est adaptée à cette situation.

Définition 5.2.10 (Congruence)

Soit E un ensemble, muni d'un certain nombre d'opérations $\times_1, \times_2, \dots, \times_n$. On dit qu'une relation d'équivalence \mathcal{R} est une congruence sur $(E, \times_1, \dots, \times_n)$ si

$$\forall (x, y, x', y') \in E^4, \forall i \in \llbracket 1, n \rrbracket, (x \mathcal{R} x') \wedge (y \mathcal{R} y') \implies (x \times_i y) \mathcal{R} (x' \times_i y').$$

Proposition 5.2.11 (Congruence des entiers)

La relation de congruence des entiers $\equiv_{[n]}$ est une congruence sur $(\mathbb{Z}, +, \times)$.

▫ Éléments de preuve.

Vérifications de divisibilité immédiates.

▷

Proposition 5.2.12 (Passage au quotient des opérations, HP)

Soit $(E, \times_1, \dots, \times_n)$ un ensemble muni de n lois d'opérations, et \mathcal{R} une congruence sur $(E, \times_1, \dots, \times_n)$. Alors on peut définir sur E/\mathcal{R} des lois $\dot{\times}_1, \dots, \dot{\times}_n$ telles que pour tout $i \in \llbracket 1, n \rrbracket$, et tout $(x, y) \in E^2$:

$$\overline{x} \dot{\times}_i \overline{y} = \overline{x \times_i y}.$$

▫ Éléments de preuve.

La valeur de $x \times y$ ne dépendant que de la classe de x et celle de y , on peut définir $\overline{x} \times \overline{y}$ comme étant la valeur commune des $x' \times y'$, pour $x' \in \overline{x}$ et $y' \in \overline{y}$. ▷

Corollaire 5.2.13 (Addition et multiplication de $\mathbb{Z}/n\mathbb{Z}$)

On peut munir $\mathbb{Z}/n\mathbb{Z}$ d'une addition $\dot{+}$ et d'une multiplication $\dot{\times}$, notées plus simplement $+$ et \times (la multiplication est parfois aussi simplement notée par un point \cdot , ou même simplement omise), telles que :

$$\forall (x, y) \in \mathbb{Z}^2, \overline{x} + \overline{y} = \overline{x + y} \text{ et } \overline{x} \times \overline{y} = \overline{x \times y}$$

Exemple 5.2.14

Table des lois de $\mathbb{Z}/2\mathbb{Z}$, $\mathbb{Z}/3\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z}$.

III Relations d'ordre

Une autre famille de relation est celle qui permet de définir des inégalités.

III.1 Définitions générales

Une relation d'ordre, sans autre précision, est toujours une relation d'ordre au sens large :

Définition 5.3.1 (relation d'ordre large)

Une relation d'ordre sur E est une relation réflexive, antisymétrique et transitive. On note souvent $x \leq y$ pour indiquer que y est en relation avec x . Les écritures $x \leq y$ et $y \geq x$ sont équivalentes.

Exemples 5.3.2

1. L'inégalité usuelle \leq sur \mathbb{N} , \mathbb{Z} , \mathbb{Q} ou \mathbb{R} définit une relation d'ordre sur ces ensembles.
2. L'inégalité opposée \geq définit une autre relation d'ordre ; il s'agit de la relation réciproque de \leq .
3. De façon générale, étant donné une relation d'ordre \mathcal{R} sur un ensemble, la relation \mathcal{R}^{-1} définie par $x\mathcal{R}^{-1}y$ ssi $y\mathcal{R}x$ est une relation d'ordre (duale de \mathcal{R})
4. La relation de divisibilité $a | b$ est une relation d'ordre sur \mathbb{N}^* mais pas sur \mathbb{Z} .
5. L'inclusion dans $\mathcal{P}(E)$.
6. Le raffinement des partitions d'un ensemble.
7. L'ordre produit sur \mathbb{N}^n
8. L'ordre lexicographique sur $E \times F$, et plus généralement sur $E_1 \times \cdots \times E_n$, tous ces ensembles étant ordonnés

La figure 5.6 donne le graphe orienté associé à une relation d'ordre sur un ensemble fini. Certaines flèches sont nécessairement présentes (les flèches d'un élément vers lui-même, par reflexivité), ou déduites des autres (par transitivité). On se limite alors souvent au diagramme constitué par les flèches élémentaires (engendrant les autres), c'est-à-dire les flèches entre deux éléments consécutifs. La restriction à ces flèches définit ce qu'on appelle la relation de couverture (on dit qu'un élément x couvre y si $x \geq y$, $x \neq y$, et s'il n'existe pas z tel que $x > z > y$). Le graphe associé à cette relation de couverture est appelé diagramme de couverture de la relation \leq .

Pour l'exemple donné dans la figure 5.6, on obtient alors le diagramme de la question 5.7. Par convention, dans ce diagramme, les flèches vont en montant : plus un élément est placé haut, plus il est « grand » pour la relation d'ordre (à condition de pouvoir être comparé).

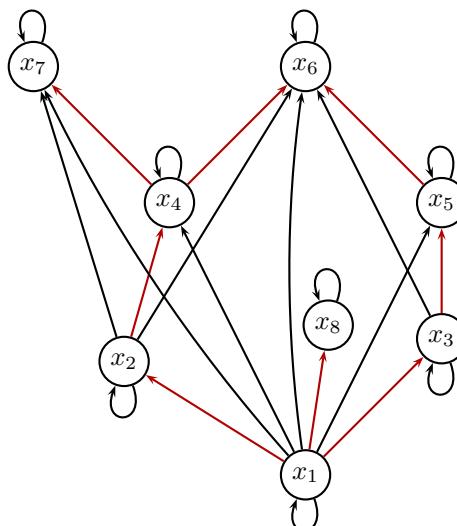


FIGURE 5.6 – Graphe d'une relation d'ordre sur $E = \{x_1, x_2, \dots, x_8\}$.

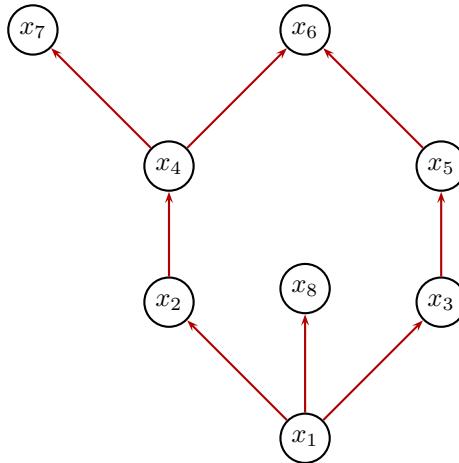


FIGURE 5.7 – Diagramme de couverture de la relation de la figure 5.6

Dans \mathbb{R} , vous avez l'habitude d'utiliser la relation d'ordre stricte. Elle peut se définir de façon générale :

Définition 5.3.3 (Relation d'ordre stricte)

Une relation d'ordre stricte est une relation antiréflexive et transitive.

Proposition 5.3.4 (Antisymétrie d'une relation d'ordre stricte)

Une relation d'ordre stricte est antisymétrique.

▫ Éléments de preuve.

Montrer qu'on ne peut pas avoir $x\mathcal{R}y$ et $y\mathcal{R}x$. En quoi cela entraîne-t-il l'antisymétrie ? ▷

Proposition 5.3.5 (D'une relation d'ordre large à une relation d'ordre stricte)

- Toute relation d'ordre large \leqslant définit une relation d'ordre stricte par $x < y$ si et seulement si $x \leqslant y$ et $x \neq y$.
- Réciproquement, toute relation d'ordre strict $<$ définit une relation d'ordre large \leqslant par $x \leqslant y$ si et seulement si $x < y$ ou $x = y$.

Remarque 5.3.6

Lorsqu'on parle de relation d'ordre, sans autre précision, il est en général sous-entendu qu'il s'agit d'une relation d'ordre large.

Dans \mathbb{R} muni de l'ordre usuel, on peut comparer deux à deux tous les éléments, mais ce n'est pas toujours le cas (cas de la divisibilité par exemple). Cette remarque motive la définition suivante :

Définition 5.3.7 (ordre total, ordre partiel)

- Soit \mathcal{R} une relation d'ordre sur un ensemble E . On dit que \mathcal{R} est une relation d'ordre total si pour tout $(x, y) \in E$, soit $x \leqslant y$ soit $y \leqslant x$.
- Dans le cas contraire, on dit que \mathcal{R} est une relation d'ordre partiel.

Exemples 5.3.8

1. L'ordre défini par le diagramme de la figure 5.7 n'est pas total : par exemple x_8 et x_7 ne sont pas comparables. En fait, le diagramme associé à une relation d'ordre total est linéaire.
2. L'ordre usuel sur \mathbb{N} , \mathbb{Z} , \mathbb{Q} ou \mathbb{R} , ainsi que l'ordre lexicographique sur \mathbb{N}^n , ou sur l'ensemble des mots définis à partir d'un alphabet sont des ordres totaux.
3. L'ordre produit sur \mathbb{N}^n (pour $n \geq 2$), la relation de divisibilité dans \mathbb{N}^* , la relation d'inclusion dans $\mathcal{P}(E)$ (lorsque E possède au moins 2 éléments), le raffinement des partitions de E (lorsque E possède au moins 3 éléments) sont des ordres partiels.
4. Si \leqslant est un ordre total, alors \geqslant est un ordre total.
5. Si E et F sont munis d'ordres totaux, l'ordre lexicographique sur $E \times F$ est total.

Proposition/Définition 5.3.9 (restriction d'une relation d'ordre)

Soit E un ensemble muni d'une relation d'ordre \mathcal{R} , et A un sous-ensemble de E . Alors \mathcal{R} définit sur A une relation d'ordre \mathcal{R}' par :

$$\forall (x, y) \in A^2, \quad x\mathcal{R}'y \iff x\mathcal{R}y.$$

Il s'agit de la restriction à A de la relation \mathcal{R} , ou de la relation induite par \mathcal{R} sur A . Elle est généralement notée \mathcal{R} également.

III.2 Minimalité, maximalité**Définition 5.3.10 (minimum, maximum)**

Soit (E, \leqslant) un ensemble muni d'une relation d'ordre.

1. Un élément m de E est appelé *plus petit élément de E* (ou *élément minimum*) si : $\forall m' \in E, m \leqslant m'$.
2. Un élément M de E est appelé *plus grand élément de E* (ou *élément maximum*) si : $\forall m' \in E, M \geqslant m'$.
3. Étant donné un sous-ensemble A de E , un élément minimum (*resp.* maximum) de A est un élément minimum (*resp.* maximum) pour la relation d'ordre \mathcal{R}' induite par \mathcal{R} sur A .

Proposition 5.3.11 (unicité du minimum)

*S'il existe, le plus petit élément de E (*resp.* de $A \subset E$) est unique. De même pour le plus grand élément.*

▫ Éléments de preuve.

S'il en existe 2, x et y , alors $x \leqslant y$ et $y \leqslant x$.

▷

Un ensemble n'a pas nécessairement de minimum ou de maximum (\mathbb{Z} par exemple, ou $\mathcal{P}(E) \setminus \{E, \emptyset\}$, pour la relation d'inclusion).

Exemple 5.3.12

Dans l'exemple de la figure 5.7, E admet-il un minimum ? un maximum ?

Définition 5.3.13 (élément minimal, maximal ; HP)

Soit (E, \leqslant) un ensemble muni d'une relation d'ordre.

1. Un élément m de E est appelé *élément minimal* de E s'il n'existe pas d'élément x de E tel que $x < m$.
2. Un élément M de E est appelé *élément maximal* s'il n'existe pas d'élément x de E tel que $x > M$.

Remarque 5.3.14 (distinction élément minimum/élément minimal)

Si l'ordre défini sur E est total, la notion d'élément minimal coïncide avec la notion d'élément minimum. Mais ce n'est plus vrai si la relation n'est que partielle, car $x < m$ n'est dans ce cas pas la négation de $x \geq m$.

Ainsi, m est minimum si et seulement si pour tout $x \in E$, on a $x \geq m$, alors que m est minimal si et seulement si pour tout $x \in E$, soit $x \geq m$, soit x est non comparable à E .

Exemples 5.3.15

1. Dans l'exemple de la figure 5.7, quels sont les éléments maximaux ? Sont-ils maximum ? Quels sont les éléments minimaux ?
Dans un diagramme de ce type, comment décririez-vous un élément maximal ?
2. Soit E ayant au moins deux éléments. Dans $\mathcal{P}(E) \setminus \{E, \emptyset\}$, décrire les éléments minimaux et maximaux.
A-t-on unicité de l'élément minimal, maximal ?
3. Décrire les éléments minimaux et maximaux dans $\mathbb{N} \setminus \{0, 1\}$ muni de la divisibilité. Y a-t-il un minimum ? un maximum ?
4. Trouver un ensemble ordonné ayant un unique élément minimal, mais pas d'élément minimum.

Proposition 5.3.16 (existence d'un élément minimal, HP)

Soit E un ensemble ordonné fini et non vide. Montrer que E admet un élément minimal.

▫ Éléments de preuve.

Considérer la chaîne de longueur maximale, et son plus petit élément.

▷

Corollaire 5.3.17 (minimum d'un ensemble fini, HP)

Soit E un ensemble fini. Alors E admet un minimum si et seulement si E admet un unique élément minimal, qui est alors aussi l'élément minimum.

▫ Éléments de preuve.

Dans le sens direct, remarquer que le minimum est comparable avec tout autre élément, donc aussi les éléments minimaux. Réciproquement, si l'élément minimal n'est pas minimum, il existe un élément x qui ne lui est pas comparable. Construire un minorant de x contredisant l'unicité de l'élément minimal.

▷

Définition 5.3.18 (minorant, majorant)

Soit (E, \leq) un ensemble muni d'une relation d'ordre. Soit $A \subset E$.

1. Un minorant m de A est un élément $m \in E$ tel que : $\forall a \in A, a \geq m$
2. Un majorant M de A est un élément $M \in E$ tel que : $\forall a \in A, a \leq M$

Définition 5.3.19 (borne supérieure, borne inférieure)

Soit (E, \leqslant) , et soit $A \subset E$.

1. La borne inférieure de A (ou *infimum* de A) dans E , notée $\inf_{x \in A} x$ ou $\inf_{x \in A} A$ est le plus grand des minorants de A , **s'il existe**.
2. La borne supérieure de A (ou *supremum* de A) dans E , notée $\sup_{x \in A} x$ ou $\sup_{x \in A} A$ est le plus petit des majorants de A , **s'il existe**.
3. Étant donnés x_1, \dots, x_n dans E , la borne inférieure (*resp.* la borne supérieure) des éléments x_1, \dots, x_n , notée $\inf(x_1, \dots, x_n)$ (*resp.* $\sup(x_1, \dots, x_n)$) est la borne inférieure (*resp.* supérieure) de l'ensemble $\{x_1, \dots, x_n\}$.

Exemple 5.3.20 (Propriété fondamentale de \mathbb{R})

Tout sous-ensemble non vide majoré de \mathbb{R} admet une borne supérieure. Cette propriété doit être prise comme axiome pour la construction de \mathbb{R} , soit découle immédiatement d'axiomes équivalents (il y a plusieurs façons équivalentes de construire \mathbb{R} , en imposant dans le cahier des charges des propriétés différentes).

Avertissement 5.3.21

Attention à l'ensemble dans lequel on considère la borne supérieure. Tout sous-ensemble borné de \mathbb{Q} admet une borne supérieure dans \mathbb{R} , mais pas nécessairement dans \mathbb{Q} .

Remarque 5.3.22

Dire qu'un entier s est le plus petit des majorants de A consiste à dire qu'il est un majorant de A et que tout autre majorant de A **lui est comparable** et est plus grand. S'il existe un majorant non comparable à s , s ne peut pas être la borne supérieure.

Ainsi, pour montrer que s est la borne supérieure de A :

- on montre que s est un majorant de A ;
- on montre que si M est un majorant de A , alors $s \leqslant M$.

Exemples 5.3.23 (bornes inférieures, supérieures)

1. Dans \mathbb{N}^* muni de la divisibilité, $\inf(a, b) = \text{pgcd}(a, b)$, et $\sup(a, b) = \text{ppcm}(a, b)$.
2. Dans $\mathcal{P}(E)$, muni de l'inclusion, $\inf(A, B) = A \cap B$, $\sup(A, B) = A \cup B$
3. Dans \mathbb{N}^2 muni de l'ordre produit, $\inf((x_1, y_1), (x_2, y_2)) = (\min(x_1, x_2), \min(y_1, y_2))$
4. Soit E muni d'un ordre total. Toute partie majorée de E admet-elle une borne supérieure ?

Proposition 5.3.24

Soit (E, \leqslant) , et $A \subset E$. A admet un maximum M (plus grand élément) si et seulement si A admet une borne supérieure b et si $b \in A$. Dans ce cas $M = b$. Énoncé similaire pour le minimum.

▫ Éléments de preuve.

Le maximum M , s'il existe, est nécessairement le plus petit des majorants de A , puisqu'il majore A , et que tout majorant de A majore en particulier M . Réciproque évidente, la borne supérieure étant un majorant. ▷

Définition 5.3.25 (Application croissante)

Soit E et F deux ensembles, munis chacun d'une relation d'ordre \leq_E et \leq_F respectivement. Une application $f : E \rightarrow F$ est dite :

- *croissante* si

$$\forall (x, y) \in E^2, \quad x \leq_E y \implies f(x) \leq_F f(y),$$

- *décroissante* si

$$\forall (x, y) \in E^2, \quad x \leq_E y \implies f(x) \geq_F f(y),$$

Avertissement 5.3.26

Prenez garde au fait que les propriétés usuelles des fonctions réelles croissantes ne sont pas toutes vraies dans une situation plus générale. Par exemple, étant donné E un ensemble fini, l'application de $\mathcal{P}(E)$ dans \mathbb{N} définie par $X \mapsto \text{Card}(X)$ est strictement croissante mais non injective !

III.3 Le lemme de Zorn (HP)

Pour terminer ce chapitre, nous donnons un résultat équivalent à l'axiome du choix (donc tout aussi indécidable), qui est la version sous laquelle l'axiome du choix est le plus fréquemment utilisé. Pour cela, nous commençons par donner une définition :

Définition 5.3.27 (ensemble inductif, HP)

Soit (E, \leq) un ensemble ordonné. On dit que E est un ensemble inductif si pour tout sous-ensemble $F \subset E$ totalement ordonné, F admet un majorant dans E .

Exemples 5.3.28

- Tout ensemble ordonné fini est inductif.
- L'ensemble (\mathbb{Z}, \leq) est-il inductif ?
- $(\mathcal{P}(E), \subset)$ est-il inductif ?

Théorème 5.3.29 (lemme de Zorn, ou de Kuratowski-Zorn, HP, admis)

(Avec AC) *Tout ensemble inductif admet un élément maximal.*

Le lemme de Zorn est en fait équivalent à l'axiome du choix.

Exemple 5.3.30

Tout ensemble ordonné fini est inductif. Mais le recours lemme de Zorn est-il pertinent dans ce cas ?

6

Les nombres réels

Au fond, en mathématiques, tout sort de l'itération obstinée et stupide de l'opérateur $n \mapsto n+1$ (addition de 1)

(René Thom)

Pour moi, la mathématique, c'est la conquête du continu par le discret.

(René Thom)

Introduction

Nous étudions dans ce chapitre les ensembles de nombres que nous utilisons usuellement : l'ensemble des nombres rationnels, puis l'ensemble des nombres réels. Nous étudierons dans le chapitre suivant l'ensemble des nombres complexes. La base de toute cette construction étant \mathbb{N} , ses deux opérations usuelles et sa relation d'ordre, nous essayons tout de même de faire une étude assez complète dans cette situation.

Le but n'est pas de faire une construction exhaustive des réels. Pour montrer la complexité de la chose, et donner une idée d'une construction possible, on donne tout de même quelques pistes, mais il est inenvisageable de faire toutes les preuves et toutes les vérifications dans le temps dont nous disposons.

I Un mot sur \mathbb{N} et \mathbb{Z}

I.1 Les entiers naturels

On admet l'existence d'un ensemble vérifiant les propriétés suivantes :

Définition 6.1.1 (L'ensemble \mathbb{N})

L'ensemble \mathbb{N} est un ensemble muni d'un élément initial 0 tel que :

- tout élément n admet un unique successeur noté $n + 1$,
- tout élément différent de 0 soit le successeur d'un unique élément de \mathbb{N} .
- 0 n'est le successeur d'aucun élément
- \mathbb{N} est minimal pour cette construction.

Les éléments de \mathbb{N} sont appelés entiers naturels.

Les unicités imposées impliquent la progression linéaire de \mathbb{N} .

Il s'agit donc d'une construction par induction structurelle, avec un élément initial 0 et une opération de construction, définissant $n + 1$ à partir de n .

Remarque 6.1.2

- Construire un ensemble ne vérifiant pas l'unicité du prédécesseur, et distinct de \mathbb{N} .
- Construire un ensemble distinct de \mathbb{N} vérifiant toutes les propriétés sauf la minimalité.

La minimalité se traduit par le fait que tout élément est accessible depuis 0 en itérant la prise de successeur. Cela se traduit par la proposition suivante :

Proposition 6.1.3 (Axiome de récurrence)

L'axiome de récurrence est équivalent à la propriété de minimalité imposée.

C'est pour cette raison que l'axiome de récurrence fait souvent partie de la définition axiomatique de \mathbb{N} .

▫ Éléments de preuve.

Considérer E minimal contenant 0 et stable par prise de successeur (c'est une définition par induction structurelle). Alors tout élément de E hormis 0 a un prédécesseur dans E , sinon on peut l'enlever.

- Considérer la propriété $\mathcal{P}(n) : n \in E$. Si l'axiome de récurrence est vrai, alors $E = \mathbb{N}$, d'où la minimalité de \mathbb{N}
- Réciproquement, si la propriété de minimalité est satisfaite, $E = \mathbb{N}$ ce qui implique l'axiome de récurrence (l'ensemble F des n tels que $\mathcal{P}(n)$ vérifie $E \subset F$)

▷

Cela ne signifie rien d'autre que le fait qu'en itérant la prise du successeur, on atteint tout élément de \mathbb{N} en partant de 0.

Une propriété importante de \mathbb{N} , qui pourrait également être prise en axiome (du fait qu'elle est équivalente à l'axiome de récurrence) est très justement appelée propriété fondamentale de \mathbb{N} :

Théorème 6.1.4 (propriété fondamentale de \mathbb{N})

Tout sous-ensemble non vide et majoré de \mathbb{N} admet un plus grand élément.

▫ Éléments de preuve.

On le montre par récurrence sur n , pour les sous-ensembles majorés par n , en discutant suivant que E , majoré par n , contient n ou non.

▷

Corollaire 6.1.5

Tout sous-ensemble non vide de \mathbb{N} admet un plus petit élément.

▫ Éléments de preuve.

Appliquer la propriété fondamentale à l'ensemble des minorants.

▷

En fait, la terminologie « propriété fondamentale » n'est pas anodine. On aurait pu construire une axiomatique de \mathbb{N} basée sur ce théorème (qui fonderait donc l'ensemble \mathbb{N}) plutôt que sur l'axiome de récurrence. C'est ce que montre le théorème suivant :

Théorème 6.1.6

La propriété fondamentale de \mathbb{N} est équivalente à l'axiome de récurrence.

▫ Éléments de preuve.

Pour démontrer que la propriété fondamentale entraîne l'axiome de récurrence, considérer le minimum de $E = \{n \mid \neg\mathcal{P}(n)\}$ et son prédécesseur.

▷

Les différentes opérations sur \mathbb{N} se définissent par récurrence (ce qui équivaut à les définir avec les successeurs) :

- $a + b$ est définie par récurrence sur b , par $a + 0 = a$ et $a + (b + 1) = (a + b) + 1$ (+1 désignant le successeur)
- $a \times b$ est définie par récurrence sur b par $a \times 0 = 0$ et $a \times (b + 1) = (a \times b) + a$

Vous connaissez les propriétés de ces opérations depuis longtemps :

Proposition 6.1.7 (propriétés de l'addition et du produit d'entiers)

Soit a, b et c des éléments de \mathbb{N} :

- $a + 0 = a$ (0 est élément neutre pour $+$)
- $0 + a = a$ (0 est élément neutre pour $+$)
- $a \times 0 = 0 \times a = 0$ (0 est absorbant pour \times)
- $a + 1 = 1 + a$
- $a \times 1 = 1 \times a = a$ (1 est neutre pour \times)
- $(a + b) + c = a + (b + c)$ ($+$ est associative)
- $a + b = b + a$ ($+$ est commutative)
- $a \times (b + c) = a \times b + a \times c$ (\times est distributive sur $+$)
- $a \times b = b \times a$ (\times est commutative)
- $(a \times b) \times c = a \times (b \times c)$ (\times est associative)
- $ab = 0 \implies a = 0$ ou $b = 0$ (intégrité de (\mathbb{N}, \times))
- $a + b = 0 \implies a = 0$ et $b = 0$
- $a + b = a + c \implies b = c$ (régularité pour $+$)
- Si $a \neq 0$, $ab = ac \implies b = c$ (régularité pour \times)

▫ Éléments de preuve.

Les 3 premières propriétés découlent de l'initialisation de la définition des lois, les autres se montrent par récurrence sur l'une des variables. ▷

Lemme 6.1.8 (Définition de la relation d'ordre)

$a \leq b$ si et seulement s'il existe $n \in \mathbb{N}$ tel que $a + n = b$

Proposition 6.1.9

\leq est une relation d'ordre total sur \mathbb{N} .

▫ Éléments de preuve.

Pour montrer que la relation est total : récurrence sur a , et distinguer suivant que $a = b$, $b - 1$ ou non. Utiliser l'existence d'un prédécesseur. ▷

On vérifie alors les différentes propriétés de la relation d'ordre vis-à-vis des opérations, en utilisant le lemme suivant :

Proposition 6.1.10 (Relation d'ordre, somme et produit)

Soit a, b, c, d des entiers naturels. Alors

- si $a \leq c$ et $b \leq d$ alors $a + b \leq c + d$ avec égalité si et seulement si $a = c$ et $b = d$
- si $0 < a \leq c$ et $0 < b \leq d$ alors $ab \leq cd$ avec égalité si et seulement si $a = c$ et $b = d$

I.2 Les entiers relatifs

Grossièrement, il s'agit de symétriser les éléments de \mathbb{N} : tout élément non nul n de \mathbb{N} définit un élément $-n$ appelé opposé de n . On étend la notion de successeur en disant que 0 est successeur de -1 et si n est non nul $-n$ est successeur de $-(n+1)$. Autrement dit, $-(n+1)+1 = -n$.

On a vu en TD qu'on peut définir \mathbb{Z} plus formellement comme quotient de $\mathbb{N} \times \mathbb{N}$ par la relation $(x+n, y+n) \sim (x, y)$. Le couple (x, y) définit alors l'entier $x-y$. Par exemple, $n \in \mathbb{N}$ est représenté par $(n, 0)$ et $-n$ est représenté par $(0, n)$.

On définit alors la somme et le produit sur les couples d'entiers naturels, et on vérifie qu'il s'agit de congruences :

- $(x, y) + (x', y') = (x+x', y+y')$
- $(x, y) \times (x', y') = (xx'+yy', xy'+x'y)$

On remarquera que les couples $(n, 0)$ et $(0, m)$ forment un système de représentant des classes d'équivalence. Autrement dit, tout entier relatif est représenté de façon unique soit sous la forme $(n, 0)$ (cela correspond à l'entier positif n) soit sous la forme $(0, m)$ (cela correspond à l'entier négatif $-m$). On vérifie sans peine que :

- Les deux lois sur \mathbb{Z} prolongent celles de \mathbb{N}
- $0_{\mathbb{Z}} = \overline{(0, 0)}$ est neutre pour l'addition et que $-m$ est l'opposé de m (0 dans le couple $(0, 0)$ désignant le $0_{\mathbb{N}}$)
- $1_{\mathbb{Z}} = \overline{(1, 0)}$ est neutre pour le produit.
- l'addition est associative, commutative
- la multiplication est associative, commutative et distributive sur l'addition.

On résume toutes ces propriétés en disant que $(\mathbb{Z}, +, \times)$ est un anneau commutatif. On étudiera plus précisément la structure d'anneau dans un chapitre ultérieur.

Proposition 6.1.11 (\mathbb{Z} peut être vu comme un prolongement de \mathbb{N})

L'application :

$$\begin{aligned} i : \mathbb{N} &\leftrightarrow \mathbb{Z} \\ n &\longmapsto \overline{(n, 0)} \end{aligned}$$

est une injection compatible avec les lois + et ×.

▫ Éléments de preuve.

C'est l'unicité de la représentation d'un élément de \mathbb{Z} sous la forme $(n, 0)$ ou $(0, m)$. ▷

Ainsi, \mathbb{N} peut être identifiée à son image dans \mathbb{Z} , et via cette identification, \mathbb{N} peut être considéré comme un sous-ensemble de \mathbb{Z} , les opérations $+$ et \times vérifiant les mêmes propriétés que la version initiale de \mathbb{N} . La relation d'ordre définie sur \mathbb{N} se prolonge à \mathbb{Z} en la définissant d'abord sur les couples :

$$(x, y) \leqslant (x', y') \iff (x+y' \leqslant x'+y).$$

Cette relation respecte la relation d'équivalence définissant \mathbb{Z} et passe donc au quotient, définissant une relation d'ordre total sur \mathbb{Z} . On vérifie facilement que cette relation prolonge celle de \mathbb{N} .

bte

II De \mathbb{Q} à \mathbb{R}

II.1 Construction de \mathbb{Q}

Nous avons déjà évoqué dans un chapitre antérieur une façon de construire \mathbb{Q} comme ensemble des quotients $\frac{a}{b}$ de deux entiers relatifs. Plus précisément, pour définir correctement les cas d'égalités entre fraction, il convient de définir \mathbb{Q} comme l'ensemble des classes d'équivalence de $\mathbb{Z} \times \mathbb{N}^*$ muni de la relation $\equiv_{\mathbb{Q}}$ définie par :

$$(a, b) \equiv (c, d) \iff ad - bc = 0.$$

En d'autres termes, \mathbb{Q} est l'espace quotient de $\mathbb{Z} \times \mathbb{N}^*$ par la relation $\equiv_{\mathbb{Q}}$.

Définition 6.2.1 (Notation usuelle pour un rationnel)

La classe $\overline{(a, b)}$ du couple (a, b) est notée $\frac{a}{b}$.

La définition de la relation donne les conditions d'égalité de deux fractions $\frac{a}{b}$ et $\frac{c}{d}$. C'est ce passage au quotient qui permet de gérer de façon rigoureuse la non unicité de l'écriture d'un rationnel sous forme d'un couple (numérateur, dénominateur).

Théorème 6.2.2 (Définition de l'addition et du produit de rationnels)

Les lois définies sur $\mathbb{Z} \times \mathbb{N}^$ par $(a, b) + (c, d) = (ad + bc, bd)$ et $(a, b) \times (c, d) = (ac, bd)$ passent au quotient, définissant sur \mathbb{Q} les lois pouvant être décrites avec les notations usuelles par :*

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} \quad \text{et} \quad \frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}.$$

▫ Éléments de preuve.

En vertu des résultats du chapitre précédent, il suffit de vérifier que les lois définies sur $\mathbb{Z} \times \mathbb{Z}^*$ sont des congruences, ce qui ne pose pas de problème majeur. ▷

Théorème 6.2.3 (Propriété des lois de \mathbb{Q})

- Les lois $+$ et \times sont associatives (i.e. $(x + y) + z = x + (y + z)$ et de même pour \times)
- Les lois $+$ et \times sont commutatives (i.e. $x + y = y + x$ et de même pour \times)
- La loi \times est distributive sur $+$ (i.e. $x \times (y + z) = x \times y + x \times z$)
- L'élément $0 = \frac{0}{1}$ est neutre pour $+$, et tout élément $\frac{a}{b}$ admet un opposé $\frac{-a}{b}$
- Le rationnel $\frac{a}{b}$ est égal à 0 si et seulement si $a = 0$.
- L'élément $1 = \frac{1}{1}$ est neutre pour \times , et tout élément $\frac{a}{b}$ non nul est inversible, d'inverse $\frac{b}{a}$.

Ainsi, \mathbb{Q} est un corps.

Remarque 6.2.4 (Inclusion de \mathbb{Z} dans \mathbb{Q})

Les entiers $a \in \mathbb{Z}$ peuvent être identifiés aux rationnels $\frac{a}{1}$ (dans le sens où $a \mapsto \frac{a}{1}$ définit une injection de \mathbb{Z} dans \mathbb{Q}). Via cette identification, on peut considérer que $\mathbb{Z} \subset \mathbb{Q}$.

On dit que \mathbb{Q} est le corps des fractions de l'anneau (intègre) \mathbb{Z} . L'intégrité de \mathbb{Z} est une propriété stipulant que pour tout $(a, b) \in (\mathbb{Z}^*)^2$, $ab \neq 0$. Cette propriété a été utilisée pour justifier l'inversibilité des rationnels non nuls.

La construction que nous venons de faire peut être faite à partir de n'importe quel anneau intègre, et permet de définir le corps des fractions de cet anneau. Nous retrouverons cette construction lorsque nous construirons les fractions rationnelles formelles à partir des polynômes formels. Nous n'expliciterons plus à ce moment les détails de la construction, du fait que tout se passe très précisément comme dans la situation ci-dessus.

II.2 Relation d'ordre dans \mathbb{Q}

Lemme 6.2.5

Soit q un rationnel. Alors le signe de l'entier naturel ab ne dépend pas de la représentation de q sous la forme $q = \frac{a}{b}$.

▫ Éléments de preuve.

Considérer $q = \frac{a}{b} = \frac{c}{d}$, et multiplier la relation définissant l'égalité de ces fractions par bd . ▷

Ce lemme nous assure que la définition suivante ne dépend pas de la représentation choisie pour q :

Définition 6.2.6 (Signe d'un rationnel)

Un rationnel $q = \frac{a}{b}$ est dit positif ou nul si $ab \geq 0$, et négatif ou nul si $ab \leq 0$. Le cas d'égalité n'est obtenu dans les deux cas que pour $q = 0$.

On note \mathbb{Q}^- l'ensemble des rationnels négatifs ou nuls.

Définition 6.2.7 (Inégalité dans \mathbb{Q})

Soit $q, q' \in \mathbb{Q}$, on dit que q est inférieur ou égal à q' et on note $q \leq q'$, si $q - q' \in \mathbb{Q}^-$.

De façon fort heureuse et assez immédiate, nous obtenons :

Proposition 6.2.8 (Cohérence de la notion de négativité)

Soit $q \in \mathbb{Q}$. Alors q est négatif ou nul si et seulement si $q \leq 0$, et q est positif ou nul si et seulement si $q \geq 0$.

Des vérifications élémentaires montrent que :

Théorème 6.2.9

La relation \leq ainsi définie sur \mathbb{Q} est une relation d'ordre total.

II.3 De l'existence de nombres non rationnels

Note Historique 6.2.10

On connaît l'existence des nombres irrationnels depuis l'antiquité. Ce sont les mathématiciens grecs qui, en premier, en ont été conscients, via l'étude de la diagonale du carré, et peut-être même avant cela, de la diagonale du pentagone.

Les nombres étant défini comme des longueurs, ils parlent de longueurs *incommensurables* (elles ne peuvent pas se mesurer à l'aide d'une unité commune).

Dans ce qui suit, on adopte la compréhension intuitive de \mathbb{R}^* comme mesure de longueurs.

Définition 6.2.11 (Nombres incommensurables)

Soit $(x, y) \in (\mathbb{R}^*)^2$. On dit que x et y sont incommensurables si $\frac{x}{y}$ est irrationnel.

Proposition 6.2.12 (Existence de nombres irrationnels)

Si n n'est pas un carré parfait (donc si n ne s'écrit pas sous la forme $n = m^2$ pour un certain entier m), alors \sqrt{n} est irrationnel.

▫ Éléments de preuve.

C'est une adaptation de la démonstration de l'irrationnalité de $\sqrt{2}$, en remarquant que par hypothèse, l'un des facteurs premiers de la décomposition primaire de n est de multiplicité impaire. ▷

Remarque 6.2.13

Cela nécessite d'avoir déjà prouvé l'existence d'un réel positif \sqrt{n} tel que $(\sqrt{n})^2 = n$, ce que nous ferons un peu plus loin.

II.4 L'ensemble \mathbb{R}

L'ensemble \mathbb{R} est alors obtenu en « bouchant les trous » laissés par les éléments de \mathbb{Q} , un peu comme on coulerait du mortier pour celler un ensemble de petites pierres, ou comme le bitume autour des gravillons. La façon de percevoir les trous de \mathbb{Q} est d'étudier l'exemple suivant :

Exemple 6.2.14 (un sous-ensemble borné de \mathbb{Q} n'admettant pas de borne supérieure)

Soit $E = \{x \in \mathbb{Q}_+ \mid x^2 \leq 2\}$. Alors E est borné, et n'admet pas dans \mathbb{Q} de borne supérieure.

Si on se rapproche de plus en plus du bord de cet intervalle, on tombe dans un trou... il n'y a rien au bord !

C'est ce vide que l'on comble en construisant \mathbb{R} comme l'ensemble \mathbb{Q} , complété des bornes supérieures de tous les sous-ensembles non vides bornés. Là encore, on a besoin de le faire sous forme d'un quotient pour une certaine relation d'équivalence, donnant une condition pour que deux bornes supérieures soit égales. Cette construction, qui n'est pas au programme, se résume par la propriété fondamentale de \mathbb{R} , *fondamentale* dans le sens où elle est intrinsèque à la définition de \mathbb{R} . Cette propriété ne pouvant être justifiée que par la manière rigoureuse de construire \mathbb{R} , nous l'admettrons.

Ainsi, on prolonge l'ensemble ordonné $(\mathbb{Q}, +)$ en un ensemble ordonné $(\mathbb{R}, +)$ le plus petit possible, assurant l'existence des bornes supérieures. Remarquez que cette « définition » très vague englobe aussi la définition de la relation d'ordre sur \mathbb{R} , prolongeant celle de \mathbb{Q} . Cette relation est totale. On admet également la possibilité de prolonger l'addition et le produit à \mathbb{R} , ainsi que la différence.

Ainsi, par construction même (c'est notre cahier des charges) :

Axiome 6.2.15 (Propriété fondamentale de \mathbb{R})

Soit E un sous-ensemble non vide et majoré de \mathbb{R} . Alors E admet une borne supérieure dans \mathbb{R} .

Évidemment, on en a une version équivalente pour la borne inférieure :

Théorème 6.2.16 (Propriété fondamentale de \mathbb{R} , exprimée avec la borne inférieure)

Soit E un sous-ensemble non vide et minoré de \mathbb{R} . Alors E admet une borne inférieure dans \mathbb{R} .

◊ Éléments de preuve.

Appliquer la propriété fondamentale à l'ensemble des minorants, puis montrer que la borne supérieure de cet ensemble est encore un minorant. ▷

Une variante consiste à appliquer la propriété fondamentale à $-E$.

Note Historique 6.2.17

La propriété de la borne supérieure a été énoncée (et démontrée, mais avec une erreur due à une absence de définition correcte de \mathbb{R}) en 1817 par le mathématicien tchèque d'origine italienne Bernhard Bolzano, en vue de donner une démonstration rigoureuse du théorème des valeurs intermédiaires (aussi appelé théorème de Bolzano), jusque-là démontré par un dessin (Cauchy, auteur de ce théorème se contente d'un dessin comme preuve)

III Nombres réels

Notre axiome de la construction de \mathbb{R} est la propriété fondamentale. Nous admettons la possibilité de prolonger les opérations $+$ et \times de \mathbb{Q} à \mathbb{R} . Nous admettons que ces lois munissent \mathbb{R} d'une structure d'anneau commutatif (cf I-2 : les deux lois sont associatives, commutatives, admettent un élément neutre, on a l'existence des opposés pour $+$ et \times est distributive sur $+$). On admet également l'intégrité, à savoir le fait que $ab = 0$ implique $a = 0$ ou $b = 0$.

III.1 Rappels sur les opérations et les inégalités

Il est indispensable de bien savoir manipuler les inégalités. En effet, l'analyse peut se définir comme l'étude d'approximations infinitésimales, ces approximations s'obtenant souvent par majorations et minorations. Le caractère infinitésimal se traduit par le fait qu'on s'autorise une marge d'erreur ε , mais que ε peut devenir aussi petit qu'on veut.

Remarquons que la définition de \mathbb{R} par la propriété fondamentale nécessite implicitement la définition de la relation d'ordre \leqslant , venant donc de la construction choisie de \mathbb{R} . Nous aurons l'occasion d'évoquer différentes constructions possibles plus tard. Pour le moment, nous admettons l'existence de cette relation d'ordre. La donnée de la relation d'ordre nous permet de définir \mathbb{R}_- et \mathbb{R}_+ par comparaison à 0, ainsi que \mathbb{R}_* et \mathbb{R}_+^* . Nous admettons les points suivants :

Propriété 6.3.1

La relation d'ordre sur \mathbb{R} vérifie les 4 propriétés élémentaires suivantes :

- (i) *C'est une relation d'ordre totale*
- (ii) $\forall (x, y) \in \mathbb{R}^2, x \leqslant y \iff y - x \in \mathbb{R}_+$
- (iii) $\forall (x, y) \in (\mathbb{R}_+)^2, x + y \geqslant 0$ avec égalité ssi $x = y = 0$
- (iv) $\forall (x, y) \in (\mathbb{R}_+)^2, xy \geqslant 0$.

Une conséquence facile de ces propriétés est la règle des signes pour le produit :

Corollaire 6.3.2

Soit x, y deux réels.

- (i) *Si $x \leqslant 0$ alors $-x \geqslant 0$*
- (ii) *Si $x \geqslant 0$ alors $-x \leqslant 0$*
- (iii) *Si $x \geqslant 0$ et $y \leqslant 0$ ou si $x \leqslant 0$ et $y \geqslant 0$, alors $xy \leqslant 0$*
- (iv) *Si $x \leqslant 0$ et $y \leqslant 0$, alors $xy \geqslant 0$*

▫ Éléments de preuve.

Par (ii) des propriétés précédentes

Par contraposée en utilisant (i) et en se ramenant au point précédent.

En utilisant l'un des deux points précédents pour se ramener à (iii). Le signe se gère bien en remarquant que $-x = (-1) \times x$. On peut alors le déplacer en utilisant l'associativité et la commutativité de \times

De même. ▷

On en déduit alors les propriétés usuelles des inégalités dans \mathbb{R} :

Proposition 6.3.3 (Manipulations élémentaires d'inégalités)

Soit a, b, c et d des réels.

- (i) Si $a \leq b$ et $c \leq d$, alors $a + c \leq b + d$.
- (ii) Si $a \leq b$ alors $-b \leq -a$.
- (iii) Si $a \leq b$ et $c \leq d$, alors $a - d \leq b - c$
- (iv) Si $0 < a \leq b$ et $0 < c \leq d$, alors $0 < ac \leq bd$, avec égalité si et seulement si $a = b$ et $c = d$.
L'inégalité reste vraie pour des valeurs positives ou nulles, mais on perd alors le cas d'égalité.
- (v) Si $a \geq 0$ et $c \leq d$, alors $ac \leq ad$.
- (vi) Si $a \leq 0$ et $c \leq d$, alors $ac \geq ad$.
- (vii) Pour toutes les autres situations de produit d'inégalité, raisonner d'abord sur la valeur absolue, puis ajouter le signe.

△ Éléments de preuve.

- (i) Utiliser (ii) et (iii) des propriétés
- (ii) Par (ii) des propriétés et le corollaire
- (iii) Combinaison des deux points précédents
- (iv) Utiliser (ii) et (iii) des propriétés et le corollaire, en introduisant bc . La stricte positivité se gère par l'intégrité.
- (v) Étudier les 3 cas suivants : $c \leq d \leq 0$, $c \leq 0 < d$, $0 < c \leq d$.
- (vi) Se ramener à la situation précédente.

▷

Évidemment, certaines de ces propriétés se généralisent de façon immédiate à un plus grand nombre de termes. Par exemple, si $(a_i)_{i \in \llbracket 1, n \rrbracket}$ et $(b_i)_{i \in \llbracket 1, n \rrbracket}$ vérifie $a_i \leq b_i$ pour tout $i \in \llbracket 1, n \rrbracket$, on a alors $\sum_{i=1}^n a_i \leq \sum_{i=1}^n b_i$, avec égalité si et seulement si $a_i = b_i$ pour tout $i \in \llbracket 1, n \rrbracket$.

L'obtention d'inégalités (par exemple la majoration d'une quantité pour contrôler son ordre de grandeur, souvent dans des études de convergence) est essentiel en analyse. Nous indiquons quelques démarches possibles pour obtenir de telles inégalités. Ces méthodes seront largement développées ultérieurement.

Terminologie 6.3.4

On rappelle que « majorer » une quantité A , c'est trouver B tel que $A \leq B$, « minorer » une quantité A , c'est trouver B tel que $A \geq B$. Souvent, A dépend d'un certain nombre de paramètres ou variables, et on cherche un majorant B ne dépendant plus de certains de ces paramètres (il arrivera fréquemment que B continue de dépendre de certains paramètres, mais pas de tous).

Méthode 6.3.5 (Majorer, minorer)

Pour obtenir des inégalités, on peut :

- tout passer du même côté (si l'inégalité est donnée) et essayer de factoriser pour déterminer le signe.
Exemple : $xy \leq \frac{1}{2}(x^2 + y^2)$;
- procéder par étude de fonction, si l'inégalité est donnée : on passe tout du même côté, et on étudie le signe de la fonction obtenue, grâce à une étude de variations.
Exemple : pour tout $x \geq 0$, $\ln(1 + x) \geq x - \frac{x^2}{2}$;

- utiliser une propriété de convexité ou de concavité : une fonction dérivable f est convexe si f' est croissante (donc la pente est de plus en plus forte). Intuitivement, la convexité se traduit par le fait que les tangentes sont sous la courbe, et les cordes sont au-dessus de la courbe.
Exemples : $e^x \geqslant 1 + x$, $\ln(1 + x) \leqslant x$, $\sin(x) \leqslant x$ pour $x \geqslant 0$, $\sin(x) \geqslant \frac{2}{\pi}x$ pour $x \in [0, \frac{\pi}{2}]$;
- Utiliser des inégalités classiques (en premier lieu l'inégalité triangulaire, l'inégalité de Cauchy-Schwarz, l'inégalité arithmético-géométrique...). L'inégalité triangulaire est à utiliser dès lors qu'on cherche à majorer la valeur absolue (ou le module) d'une somme dont on sait majorer la valeur absolue de chaque terme : il faut d'abord sortir ces termes de la valeur absolue globale, et ce grâce à l'inégalité triangulaire.

Le but de la fin de ce paragraphe est justement l'étude des plus importantes de ces inégalités classiques.

Définition 6.3.6 (Valeur absolue)

Soit $x \in \mathbb{R}$. La valeur absolue de x , notée $|x|$, est le réel obtenu de x en changeant si besoin son signe de sorte à obtenir une quantité positive :

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

Remarque 6.3.7

La valeur absolue est souvent utilisée lorsqu'on veut montrer qu'une quantité reste bornée, c'est-à-dire peut être à la fois majorée et minorée. En effet, majorer $|A|$ revient à majorer et minorer A , puisque $|A| \leqslant B$ équivaut à $-B \leqslant A \leqslant B$.

Pour montrer l'inégalité triangulaire, nous décomposons un réel en sa partie positive et sa partie négative (pour tout réel, l'une de ces deux parties sera nulle, assez logiquement).

Définition 6.3.8 (Partie positive, partie négative d'un réel)

Soit $x \in \mathbb{R}$.

- On appelle *partie positive* de x , et on note x^+ , le réel défini par :

$$x^+ = \max(0, x) = \begin{cases} x & \text{si } x \geqslant 0 \\ 0 & \text{si } x < 0 \end{cases}$$

- On appelle *partie négative* de x , et on note x^- , le réel défini par :

$$x^- = -\min(0, x) = \max(0, -x) = \begin{cases} -x & \text{si } x < 0 \\ 0 & \text{si } x \geqslant 0 \end{cases}$$

Propriétés 6.3.9 (Propriétés des parties positives et négatives)

Soit x un réel. Alors :

- $x^+ \geqslant 0$ et $x^- \geqslant 0$;
- $x^+ = 0$ ou $x^- = 0$;
- $x = x^+ - x^-$;
- l'égalité précédente est minimale dans le sens suivant : pour tout $(y, z) \in \mathbb{R}_+$, si $x = y - z$ alors $y \geqslant x^+$ et $z \geqslant x^-$;

- (v) $|x| = x^+ + x^- = \max(0, x) - \min(0, -x)$.
(vi) $(-x)^+ = x^-$ et $(-x)^- = x^+$.

▫ Éléments de preuve.

Les points (i), (ii) découlent de la définition, les autres s'obtiennent par discussion sur le signe de x .

▷

Corollaire 6.3.10 (Inégalités triangulaires)

Soit x et y deux réels. Alors :

- (i) $(x+y)^+ \leq x^+ + y^+$
- (ii) $(x+y)^- \leq x^- + y^-$
- (iii) $|x+y| \leq |x| + |y|$

Chacune de ces inégalités est une égalité si et seulement si x et y sont de même signe.

▫ Éléments de preuve.

Les points (i) et (ii) s'obtiennent ensemble en utilisant la propriété de minimalité (iv) de la proposition précédente, appliquée à $x+y$ exprimé avec x^+ , x^- , y^+ et y^- .

Le point (iii) en découle par somme. ▷

On obtient alors, par une récurrence immédiate sur le cardinal de I :

Corollaire 6.3.11 (Inégalité triangulaire pour les sommes)

Soit $(a_i)_{i \in I}$ une famille finie de réels. Alors

$$\left| \sum_{i \in I} a_i \right| \leq \sum_{i \in I} |a_i|.$$

Théorème 6.3.12 (Inégalité de Cauchy-Schwarz numérique)

Soient $x_1, \dots, x_n, y_1, \dots, y_n$ des réels. On a alors :

$$\left| \sum_{i=1}^n x_i y_i \right|^2 \leq \left(\sum_{i=1}^n x_i^2 \right) \left(\sum_{i=1}^n y_i^2 \right),$$

avec égalité si et seulement si les vecteurs (x_1, \dots, x_n) et (y_1, \dots, y_n) sont colinéaires.

En notant, pour $X = (x_1, \dots, x_n)$ et $Y = (y_1, \dots, y_n)$

$$\langle X, Y \rangle = \sum_{i=1}^n x_i y_i$$

le produit scalaire canonique de \mathbb{R}^n , et

$$\|X\| = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{\langle X, X \rangle}$$

la norme euclidienne canonique, l'inégalité de Cauchy-Schwarz se réexprime de la sorte :

$$|\langle X, Y \rangle| \leq \|X\| \cdot \|Y\|.$$

▫ Éléments de preuve.

Utiliser le fait que le polynôme $\|\lambda X + Y\|^2$ en la variable λ est de signe constant : que peut-on en déduire sur son discriminant ? (Attention, ne pas oublier de traiter le cas où le polynôme n'est pas de degré 2)

Pour le cas d'égalité, cela correspond à $\Delta = 0$: qu'est-ce que ça signifie pour le polynôme ci-dessus ?
▷

Remarque 6.3.13

La démonstration n'utilise que le fait que pour tout X , $\langle X, X \rangle \geq 0$, avec égalité ssi $X = 0$, et la symétrie du produit scalaire. On définira plus tard dans l'année une notion générale de produit scalaire, vérifiant ces propriétés. Ainsi, la formule de Cauchy-Schwarz restera valable pour tout produit scalaire, la norme associée à ce produit scalaire se définissant comme on l'a fait pour la norme euclidienne à partir du produit scalaire canonique.

Théorème 6.3.14 (Inégalité arithmético-géométrique, HP)

Pour tout $X = (x_1, \dots, x_n) \in (\mathbb{R}_+^*)^n$,

$$\frac{1}{n}(x_1 + \dots + x_n) \geq \sqrt[n]{x_1 \dots x_n}.$$

Cette inégalité dit qu'on est gentil avec vous en calculant votre moyenne scolaire avec la moyenne arithmétique plutôt que la moyenne géométrique.

▫ Éléments de preuve.

Récurrence un peu atypique avec des retours en arrière : on montre d'abord par récurrence le cas où n est une puissance de 2, puis en regroupant judicieusement des termes, on montre la propriété pour n quelconque en redescendant à partir des puissances de 2. C'est la démonstration de Cauchy. Il existe beaucoup d'autres démonstrations, notamment une démonstration très rapide par un argument de convexité, mais cela nécessite quelques connaissances supplémentaires que vous verrez plus tard. ▷

III.2 Division euclidienne dans \mathbb{R}

Le principe de la division euclidienne dans \mathbb{R} repose sur le résultat suivant, bien utile par ailleurs, notamment pour prouver des propriétés de densité :

Proposition 6.3.15 (Propriété d'Archimède)

Soit x et y deux réels strictement positifs. Il existe un entier $n \in \mathbb{N}$ tel que $x < ny$.

▫ Éléments de preuve.

Par l'absurde, appliquer la propriété fondamentale de \mathbb{R} à $\{ny, n \in \mathbb{N}\}$, et considérer n_0 tel que $n_0 y$ soit proche de la borne supérieure. Trouver une contradiction en considérant $(n_0 + 1)y$. ▷

Corollaire 6.3.16

Pour tout $x > 0$ et tout $y > 0$, il existe un rationnel r tel que $0 < rx < y$.

▫ Éléments de preuve.

Appliquer la propriété d'Archimède, puis multiplier par $\frac{1}{n}$. ▷

En particulier, en prenant $x = 1$, on peut toujours trouver un rationnel strictement inférieur à un réel strictement positif y arbitraire.

Une conséquence importante de ce résultat est l'inversibilité de tout réel non nul :

Théorème 6.3.17 (Le corps \mathbb{R})

Tout réel non nul est inversible. Combiné aux propriétés usuelles admises des opérations de \mathbb{R} , ce résultat affirme que \mathbb{R} est un anneau commutatif dans lequel tout élément est inversible, c'est-à-dire un corps.

▫ Éléments de preuve.

Se ramener au cas $x > 0$. Considérer $y_0 = \sup\{y \mid yx \leq 1\}$. Si $y_0x < 1$, considérer r tel que $0 < rx < 1 - y_0x$, et contredire la définition de y_0 . Même type d'argument si $y_0x > 1$. ▷

Une variante de la propriété d'Archimède est :

Corollaire 6.3.18

Soit x et y strictement positifs. Il existe un unique entier $n \in \mathbb{N}$ tel que $ny \leq x < (n+1)y$. Il existe également un unique entier n' tel que $n'y < x \leq (n'+1)y$. Sauf lorsque $\frac{x}{y}$ est entier, on a $n = n'$. Le résultat se généralise à x négatif.

▫ Éléments de preuve.

Considérer n minimal tel que $x < (n+1)y$. Pour le deuxième point, discuter suivant que x est un multiple de y ou non. Pour x négatif, appliquer ce qui précède à $-x$. L'unicité provient du fait que les intervalles définis par les ny sont disjoints. ▷

Ce résultat se réexprime sous la forme familière suivante :

Théorème 6.3.19 (Division euclidienne, Euclide)

Soit $x \in \mathbb{R}$ et $y \in \mathbb{R}_+^$. Il existe un unique entier n et un unique réel $r \in [0, y[$ tels que $x = ny + r$.*

C'est un résultat très concret : si un menuisier doit couper des planches de longueur donnée y dans une grande planche de longueur x , n est le nombre de planches de la bonne longueur qu'il peut obtenir, et r est la longueur du bout inutile qu'il lui reste à la fin, donc la chute.

Note Historique 6.3.20

Archimède est légèrement postérieur à Euclide (3^e siècle avant J.-C.). La propriété d'Archimède figure en fait déjà dans les *Éléments* d'Euclide. Archimède utilise largement cette propriété, sans pour autant prétendre à sa paternité.

III.3 Densité de \mathbb{Q} et $\mathbb{R} \setminus \mathbb{Q}$ dans \mathbb{R}

On peut alors se demander s'il y a beaucoup de nombres irrationnels, et comment ils se répartissent entre les nombres rationnels pour former \mathbb{R} . Un élément de réponse est apporté par la propriété de densité, affirmant qu'il y a des rationnels et des irrationnels un peu partout dans \mathbb{R} : il n'existe pas dans \mathbb{R} d'intervalle non vide ou non réduit à un singleton, aussi petit soit-il, ne possédant ni rationnel ni irrationnel. Intuitivement, si ce n'était pas le cas, tous les réels à l'intérieur d'un intervalle ne contenant pas de rationnels ne seraient d'aucune utilité pour exprimer des bornes supérieures d'ensembles de rationnels, car trop éloignés des rationnels les plus proches. Notre construction de \mathbb{R} ne serait alors pas minimale. Commençons par définir rigoureusement la propriété de densité dans \mathbb{R} .

Définition 6.3.21 (Densité dans \mathbb{R})

Un sous-ensemble E de \mathbb{R} est dense dans \mathbb{R} si pour tout $(x, y) \in \mathbb{R}^2$ tel que $x < y$, il existe $z \in E$ tel que $x < z < y$.

Autrement dit, entre deux éléments quelconques de \mathbb{R} (aussi proches soient-ils), il existe toujours un élément de E : les éléments de E vont s'infiltre un peu partout.

Cette notion de densité se généralise à d'autres ensembles que \mathbb{R} , mais d'un point de vue topologique.

Théorème 6.3.22 (Densité des rationnels et des irrationnels dans \mathbb{R})

Les ensembles \mathbb{Q} et $\mathbb{R} \setminus \mathbb{Q}$ sont denses dans \mathbb{R} .

▫ Éléments de preuve.

- \mathbb{Q} : Appliquer la propriété d'Archimède avec un rationnel r inférieur à $y - x$.
- $\mathbb{R} \setminus \mathbb{Q}$: Se ramener à la densité de \mathbb{Q} en translatant d'un irrationnel (par exemple $\sqrt{2}$)

▷

Bref, il existe beaucoup de nombres irrationnels. Et en fait beaucoup plus que des rationnels. En effet \mathbb{Q} est dénombrable, et comme \mathbb{R} ne l'est pas, $\mathbb{R} \setminus \mathbb{Q}$ non plus. Il n'est pas très compliqué de montrer que $\mathbb{R} \setminus \mathbb{Q}$ a puissance du continu (*i.e.* a le même cardinal que \mathbb{R}).

III.4 Nombres transcendants

Les nombres irrationnels peuvent eux-même être classés en nombres « peu irrationnels » et nombres « très irrationnels », ou plutôt, en adoptant une terminologie plus correcte, en nombres algébriques (solutions d'une équation polynomiale en \mathbb{Q}) et en nombres transcendants (les autres).

Définition 6.3.23 (Nombres algébriques, transcendants sur \mathbb{Q} , HP)

Soit $x \in \mathbb{R}$.

- On dit que x est algébrique sur \mathbb{Q} s'il existe un polynôme P à coefficients dans \mathbb{Q} tel que $P(x) = 0$.
- On dit que x est transcendant sur \mathbb{Q} s'il n'est pas algébrique.

En admettant provisoirement que tout polynôme a un nombre fini de racines, on obtient le joli résultat suivant :

Proposition 6.3.24 (cardinal de l'ensemble des nombres algébriques, HP)

L'ensemble des nombres algébriques est dénombrable.

▫ Éléments de preuve.

L'ensemble des polynômes à coefficients rationnels est dénombrable, car peut s'écrire comme union dénombrable d'ensembles dénombrables. L'union sur ces polynômes des racines de chaque polynôme est donc aussi dénombrable.

▷

Corollaire 6.3.25 (existence de nombres transcendants, HP)

Il existe des nombres transcendants.

Note Historique 6.3.26

On connaît l'existence de nombres irrationnels depuis la Grèce classique. En revanche, la notion de nombres transcendants est beaucoup plus récente :

- Leibniz est le premier, en 1682, à envisager la possibilité de l'existence de nombres transcendants.
- Ce n'est qu'en 1844 que Liouville justifie l'existence de nombres transcendants, en construisant pour l'occasion et en montrant la transcendance du réel $c = \sum_{j=1}^{+\infty} 10^{-j!}$, appelé depuis « constante de Liouville ».
- Hermite prouve en 1873 la transcendance de e . C'est la première fois qu'on montre la transcendance d'un réel qui avait une existence antérieure.
- Lindemann montre en 1882 la transcendance de π , mettant ainsi fin à 3 millénaires de recherches infructueuses pour réaliser la quadrature du cercle.
- Les travaux de Cantor permettent de justifier l'existence de nombres transcendants sans avoir à en contruire, par simple considération des cardinaux.

III.5 Partie entière, partie décimale

Nous étudions maintenant les représentations des réels dans la vie pratique. Pour cela, nous commençons par séparer la partie entière et la partie décimale, en définissant rigoureusement ces notions.

Proposition/Définition 6.3.27 (Partie entière)

Soit $x \in \mathbb{R}$. Les cinq définitions suivantes sont équivalentes :

- $[x] = \max\{n \in \mathbb{Z} \mid n \leq x\}$;
- $[x] = \min\{n \in \mathbb{Z} \mid n > x\} - 1$;
- $[x]$ est l'unique entier tel que $[x] \leq x < [x] + 1$;
- $[x]$ est l'unique entier tel que $x - 1 < [x] \leq x$.
- $[x]$ est le quotient de la division euclidienne de x par 1.

L'entier $[x]$ défini par l'une de ces propriétés (donc toutes) est appelé *partie entière de x* . On trouve aussi parfois les notations $[x]$ et $E(x)$.

On donne le graphe de la fonction partie entière : $x \mapsto [x]$ en figure 6.1

On définit parfois aussi la *partie entière par excès*, notée $\lceil x \rceil$, comme étant la plus petit entier supérieur ou égal à x :

$$\lceil x \rceil = \min\{n \in \mathbb{N} \mid n \geq x\}.$$

La partie entière par excès est alors caractérisée par son appartenance à n et l'un ou l'autre des encadrements suivants :

$$x \leq \lceil x \rceil < x + 1 \quad \text{et} \quad \lceil x \rceil - 1 < x \leq \lceil x \rceil.$$

Proposition 6.3.28 (Relation entre partie entière et partie entière par excès)

Soit $x \in \mathbb{R}$. On a alors

- $\lceil x \rceil = \begin{cases} [x] + 1 & \text{si } x \notin \mathbb{Z} \\ [x] & \text{si } x \in \mathbb{Z} \end{cases}$
- $\lceil -x \rceil = -\lceil x \rceil$.

On reconnaît dans l'encadrement définissant la partie entière et la partie entière par excès le corollaire de la propriété d'Archimède.

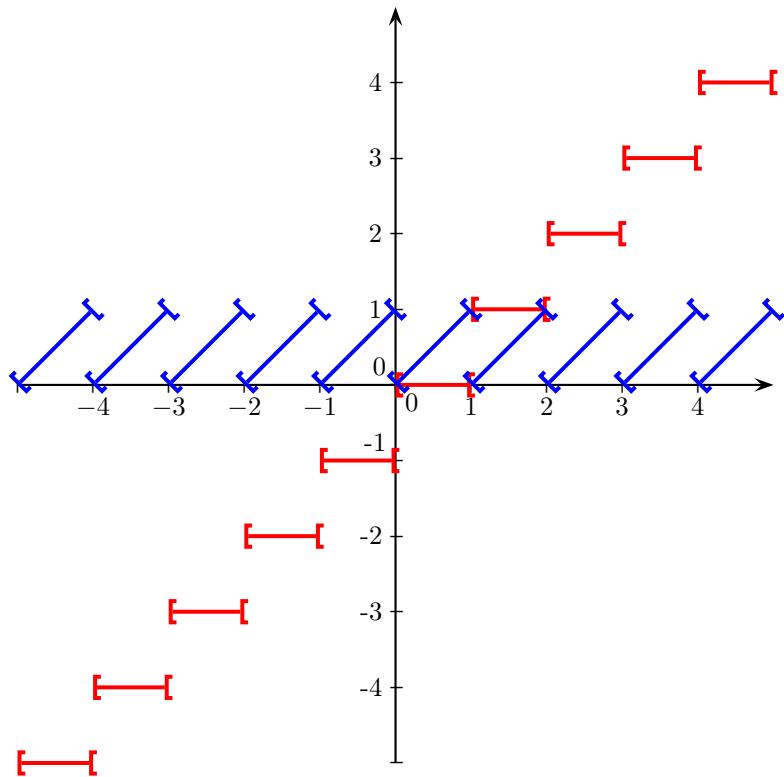


FIGURE 6.1 – Graphe de la partie entière et de la partie décimale

Définition 6.3.29 (Partie décimale, voir graphe 6.1)

La partie décimale de x est le réel $x - \lfloor x \rfloor$. Il s'agit du reste de la division euclidienne de x par 1. Intuitivement, c'est le réel obtenu en ne gardant que les chiffres après la virgule, mais cette interprétation n'est valable que pour les réels positifs.

La partie décimale de x est parfois notée $\{x\}$.

On a alors par définition même : $x = \lfloor x \rfloor + \{x\}$.

Propriétés 6.3.30 (propriétés de la partie entière)

1. $\forall x, y \in \mathbb{R}, \lfloor x \rfloor + \lfloor y \rfloor + 1 \geq \lfloor x + y \rfloor \geq \lfloor x \rfloor + \lfloor y \rfloor$;
2. $\forall x, y \in \mathbb{R}_+, \lfloor xy \rfloor \geq \lfloor x \rfloor \cdot \lfloor y \rfloor$;
3. $\forall x \in \mathbb{R}, \forall n \in \mathbb{Z}, \lfloor x + n \rfloor = \lfloor x \rfloor + n$.

▫ Éléments de preuve.

Le premier point s'obtient en écrivant $x = \lfloor x \rfloor + \{x\}$ et $y = \lfloor y \rfloor + \{y\}$ et en remarquant que $\{x\} + \{y\} \in [0, 2[$. Les deux autres points sont évidents. ▷

III.6 Représentation décimale**Note Historique 6.3.31 (petit aperçu historique de la représentation des nombres)**

- Dès l'antiquité, les mathématiciens se rendent compte de la nécessité de rendre la notion de nombre indépendante de toute unité ou toute échelle. Ainsi, les Grecs définissent un nombre comme un rapport entre deux grandeurs de même type, par exemple entre deux longueurs. Ainsi, le théorème de Thalès

(premier mathématicien grec important, début du 1er millénaire avant JC) s'exprime sous forme de rapports de longueur.

- Les premières représentations des entiers sont des numérations additives (on ajoute des symboles pour faire des nombres plus gros). Le premier système de numération connu est simplement une succession d'encoches dans un bout de bois ou un os. D'ailleurs, le mot « calcul » dérive du mot « calculus » signifiant caillou (pensez aux calculs rénaux!), car les bergers utilisaient des cailloux pour compter leurs moutons.
- Beaucoup plus tard, il reste encore des vestiges de ce système de numération chez les romains, même s'ils disposent d'un système un peu plus sophistiqué (plusieurs symboles pour désigner différentes grandeurs, et notation soustractive).
- La numération de position (qui correspond à une numération dans une base donnée, la position d'un chiffre déterminant le coefficient multiplicatif qui lui sera appliquée) apparaît vraisemblablement à Babylone. Les nombres y sont représentés en base 60 (y compris pour les décimales) ; les « chiffres » de 1 à 60 sont représentés grâce à une numération additive, à l'aide de deux symboles de valeur 1 et 10. Le zéro n'existant pas encore, la position des chiffres n'est pas toujours très claire.
- Il reste d'ailleurs dans notre civilisation des vestiges de cette numération en base 60. Lesquels ?
- Les fractions apparaissent dès le 2e millénaire avant JC, à Babylone, où un calcul très complexe des fractions est mis en place. Imaginez-vous faire des opérations sur des fractions en base 60... Les règles calculatoires (sommes, produits) ne sont pas encore bien établies, et restent intuitive. La plupart des calculs sur les fractions sont faits à l'aide de tables.
- La découverte des nombres irrationnels date probablement de Pythagore (diagonale du carré), mais le secret est gardé. Hippase de Métaponte dévoile aux non initiés l'existence de grandeurs incommensurables (*i.e.* dont le rapport est irrationnel). Selon la légende, il est jeté du haut d'une falaise pour avoir révélé le secret pythagoricien. Quelle est la part de vérité dans cette histoire ? C'est dur à dire. Il est possible aussi que la découverte des irrationnels provienne de propriétés géométriques du pentagone : une construction fractale d'une suite de pentagones permet de montrer géométriquement que l'algorithme d'Euclide pour le rapport entre la diagonale et le côté ne se termine pas, ce qui équivaut à l'incommensurabilité de ces deux grandeurs.
- La numération actuelle est une numération en base 10, et une numération de position (l'un n'entraînant pas l'autre, la notation romaine n'est pas une numération de position, mais est bien une numération en base 10). La numération de position s'est imposée suite à la diffusion des ouvrage de Al Khwarizmi diffusant le système de numération indien. L'intermédiaire arabe de cette diffusion a eu pour conséquence la terminologie de « chiffres arabes », mais il s'agit bien de « chiffres indiens », même si la graphologie de ces chiffres a beaucoup changé. L'importance de l'apport est bien plus le système de numération par position (avec un symbole pour représenter 0) que la graphie précise des chiffres.

Notation 6.3.32 (nombres décimaux)

- Nous notons \mathbb{D} l'ensemble des nombres décimaux, c'est à dire des réels x tels qu'il existe $n \in \mathbb{N}$ tel que $10^n x$ est entier.
- Étant donné $n \in \mathbb{N}$, nous notons \mathbb{D}_n l'ensemble des nombres décimaux tels que $10^n x \in \mathbb{Z}$; Par exemple $\mathbb{D}_0 = \mathbb{Z}$, et \mathbb{D}_1 sont les décimaux s'écrivant avec au plus un chiffre après la virgule.

Proposition 6.3.33 (Approximations décimales d'un réel x)

Soit x un réel et $n \in \mathbb{N}^*$. Il existe un unique élément y de \mathbb{D}_n tel que

$$y_n \leq x < y_n + 10^{-n}.$$

- Le décimal y_n est appelé valeur approchée décimale à la précision 10^{-n} par défaut
- Le décimal $y_n + 10^{-n}$ est appelé valeur approchée décimale à la précision 10^{-n} par excès.

▫ Éléments de preuve.

Définir y_n à l'aide d'une partie entière.

▷

Les inégalités satisfaites par les y_n et leur unicité montrent facilement que :

Lemme 6.3.34

Pour tout $n \in \mathbb{N}^*$, il existe $a_n \in \llbracket 0, 9 \rrbracket$ tel que $y_n - y_{n-1} = \frac{a_n}{10^n}$.

Théorème 6.3.35 (Existence du développement décimal de x)

Soit $x \in \mathbb{R}$. Il existe un entier relatif y_0 et pour tout $n \in \mathbb{N}^*$, des entiers $a_n \in \llbracket 0, 9 \rrbracket$ tels que

$$x = y_0 + \sum_{n=1}^{+\infty} a_n 10^{-n} = y_0 + \lim_{N \rightarrow +\infty} \sum_{n=1}^N a_n 10^{-n}.$$

▫ Éléments de preuve.

Les a_n sont ceux du lemme précédent. ▷

Théorème 6.3.36 (Unicité du développement décimal de x)

Soit $x \in \mathbb{R}$.

1. Si x n'est pas décimal, x admet un unique développement décimal.
2. Si x est décimal, x admet deux développements décimaux exactement, l'un terminant uniquement par des 9, l'autre terminant uniquement par des 0

▫ Éléments de preuve.

Considérer la première différence dans les développements décimaux, et majorer les restes suivant cette différence : ces restes ne peuvent compenser la différence que si l'un est nul et l'autre obtenu pour les valeurs maximales des a_n . ▷

Définition 6.3.37 (Développement propre)

On appelle développement décimal propre de x l'unique développement de x si x n'est pas décimal, ou l'unique développement de x terminant par des 0 si x est décimal. Ainsi, tout réel admet un unique développement décimal propre

Proposition 6.3.38 (Caractérisation des rationnels par leur développement décimal, HP)

Un nombre x est rationnel si et seulement si son développement décimal est périodique à partir d'un certain rang.

▫ Éléments de preuve.

Le sens direct s'obtient par l'algorithme de la division posée, qui tourne en boucle. La réciproque est obtenue en exprimant $10^n x - x$, où n est la longueur de la période. ▷

Tout ce qui a été fait ci-dessus en base 10 se généralise immédiatement en base $b \in \llbracket 2, +\infty \rrbracket$. Un développement en base b est une décomposition de x sous la forme

$$x = y_0 + \sum_{n=1}^{+\infty} a_n b^{-n}$$

où $y_0 \in \mathbb{Z}$ et les a_n sont dans $\llbracket 0, b-1 \rrbracket$. On obtient par les mêmes procédés :

Théorème 6.3.39 (Existence et unicité d'un développement en base b)

Tout réel admet un développement en base b . Ce développement est unique si $b^n x$ n'est entier pour aucune valeur entière de n . Dans le cas contraire, x admet exactement deux développements en base b , l'un terminant par des 0, l'autre par des $b - 1$.

Proposition 6.3.40

L'entier y_0 lui-même peut se décomposer en base b (de façon unique) sous la forme

$$y_0 = \sum_{m=0}^{m_0} a'_m b^m = \sum_{n=-m_0}^0 a_n b^{-n},$$

en posant $a_n = a'_{-n}$. Ainsi, le développement en base b de x peut se réécrire :

$$x = \sum_{n=-m_0}^{+\infty} a_n b^{-n}.$$

▫ Éléments de preuve.

Décomposer y_0 par récurrence forte, en effectuant une division euclidienne par b . On peut aussi se ramener directement au théorème précédent en divisant au préalable par b^{m_0} , où m_0 est tel que $y_0 < b^{m_0}$. ▷

IV Intervalles

IV.1 Description des intervalles

Nous définissons les intervalles par leur propriété de convexité :

Définition 6.4.1 (ensemble convexe, figure 6.2)

Soit E un sous-ensemble de \mathbb{R}^n . On dit que E est convexe si et seulement si pour tout couple de points A et B de E , le segment $[AB]$ est entièrement inclus dans E .

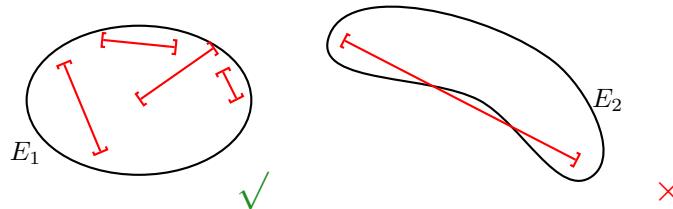


FIGURE 6.2 – Un sous-ensemble convexe E_1 et un sous-ensemble non convexe E_2 de \mathbb{R}^2

Définition 6.4.2 (Intervalle)

Un intervalle I de \mathbb{R} est un sous-ensemble convexe I de \mathbb{R} , c'est-à-dire tel que :

$$\forall (a, b) \in I^2, \forall x \in \mathbb{R}, a \leq x \leq b \implies x \in I.$$

Intuitivement, « il n'y a pas de trou dans un intervalle ».

Théorème 6.4.3 (Inventaire des intervalles réels)

Tout intervalle I de \mathbb{R} est d'une des formes suivantes, pour certaines valeurs réelles a et b :

- $[a, b] = \{x \in \mathbb{R}, a \leq x \leq b\}, a \leq b ;$
- $]a, b[= \{x \in \mathbb{R}, a < x < b\}, a < b ;$
- $[a, b[= \{x \in \mathbb{R}, a \leq x < b\}, a < b ;$
- $]a, b] = \{x \in \mathbb{R}, a < x \leq b\}, a < b ;$
- $[a, +\infty[= \{x \in \mathbb{R}, x \geq a\} ;$
- $]a, +\infty[= \{x \in \mathbb{R}, x > a\} ;$
- $] - \infty, b] = \{x \in \mathbb{R}, x \leq b\} ;$
- $] - \infty, b[= \{x \in \mathbb{R}, x < b\} ;$
- $] - \infty, +\infty[= \mathbb{R} ;$
- $\emptyset.$

▫ Éléments de preuve.

Si I est non vide, sa borne inférieure a et sa borne supérieure b existent (éventuellement infinies).

Montrer, par la propriété de convexité, que $]a, b[\subset I \subset [a, b]$. ▷

Remarquez que le premier cas pour $a = b$ dit que tous les singlettes $\{a\}$ sont des intervalles.

On définit de la même manière les intervalles de \mathbb{Q} comme sous-ensemble convexes de \mathbb{Q} (comme on reste dans \mathbb{Q} , les trous ne se voient pas dans la propriété de convexité)

Proposition 6.4.4 (description des intervalles de \mathbb{Q})

Un sous-ensemble I de \mathbb{Q} est un intervalle de \mathbb{Q} si et seulement s'il existe un intervalle $I_{\mathbb{R}}$ de \mathbb{R} tel que $I = I_{\mathbb{R}} \cap \mathbb{Q}$.

▫ Éléments de preuve.

À peu près le même principe ; les bornes inférieure et supérieure sont réelles, tout le reste est rationnel.

▷

Les intervalles de \mathbb{Q} n'admettent pas de description de la forme $]a, b[$ intrinsèque à \mathbb{Q} : on ne peut pas exprimer les bornes des intervalles sans sortir de \mathbb{Q} . Par exemple, $]0, \sqrt{2}[\cap \mathbb{Q}$ n'admet pas de borne supérieure dans \mathbb{Q} .

Remarque 6.4.5

L'inventaire des intervalles de \mathbb{R} est strictement équivalente à la propriété de la borne supérieure, et pourrait être considérée comme propriété fondamentale de \mathbb{R} en lieu et place de la propriété de la borne supérieure : un réel est alors une classe d'équivalence d'intervalles de \mathbb{Q} de même borne supérieure dans \mathbb{R} (cette notion d'égalité étant à définir intrinsèquement à \mathbb{Q} , pour que la définition soit valide).

Un intervalle est donc délimité par deux réels (ou les infinis), et chacune des deux bornes, si elle est finie, peut être ou ne pas être dans l'intervalle (une borne infinie est toujours exclue de l'intervalle, l'infini n'étant pas un réel). L'appartenance ou non des bornes à l'intervalle nous incite à donner un classement des intervalles :

Définition 6.4.6 (intervalles ouverts, fermés, semi-ouverts)

- On dit qu'un intervalle est ouvert s'il est de la forme $]a, b[$, $]a, +\infty[$, $] - \infty, b[$, \mathbb{R} ou \emptyset .
- On dit qu'un intervalle est fermé s'il est de la forme $[a, b]$, $[a, +\infty[$, $] - \infty, b]$, \mathbb{R} ou \emptyset .
- On dit qu'un intervalle est semi-ouvert s'il est de la forme $[a, b[$ ou $]a, b]$.

Remarquez qu'il existe des intervalles à la fois ouverts et fermés (\mathbb{R} et \emptyset)

IV.2 Intervalles et topologie

La notion d'intervalle est en fait liée à des notions de « topologie » plus générales (la topologie étant l'étude des sous-ensembles « ouverts » et « fermés » d'un ensemble). Nous nous limitons à une brève introduction de ces notions dans \mathbb{R}^n , la distance que nous utilisons étant la distance euclidienne canonique : si $X = (x_1, \dots, x_n)$ et $Y = (y_1, \dots, y_n)$, la distance entre X et Y est :

$$d(X, Y) = \sqrt{\sum_{i=1}^n (y_i - x_i)^2}.$$

En particulier, si $x, y \in \mathbb{R}^1 = \mathbb{R}$, $d(x, y) = |y - x|$.

Ces notions se généralisent dans des espaces muni de distances plus générales (espaces métriques).

Définition 6.4.7 (Boule dans \mathbb{R}^n , figure 6.3)

Soit $x \in \mathbb{R}^n$ et $r \in \mathbb{R}_+$.

1. La boule ouverte de centre x et de rayon r est : $B(x, r) = \{y \in \mathbb{R}^n \mid d(y, x) < r\}$
2. La boule fermée de centre x et de rayon r est : $\overline{B}(x, r) = \{y \in \mathbb{R}^n \mid d(y, x) \leq r\}$

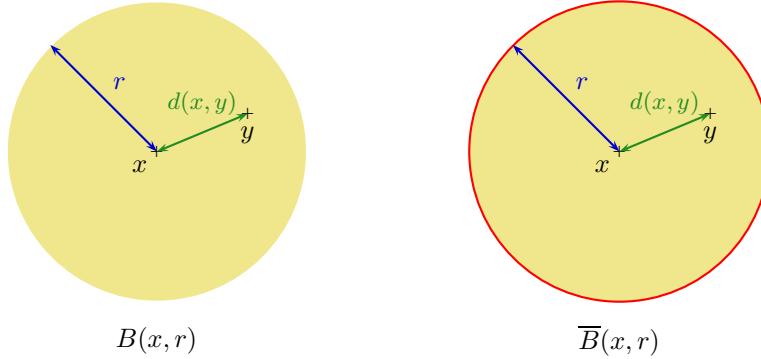


FIGURE 6.3 – Boules dans \mathbb{R}^n

Exemple 6.4.8

Dans \mathbb{R} , les boules sont des intervalles (voir figure 6.4) :

- $B(x, r) =]x - r, x + r[$
- $\overline{B}(x, r) = [x - r, x + r]$

En fait, tout intervalle borné ouvert est une boule ouverte, tout intervalle borné fermé est une boule fermée :

- $]a, b[= B\left(\frac{a+b}{2}, \frac{b-a}{2}\right)$,
- $[a, b] = \overline{B}\left(\frac{a+b}{2}, \frac{b-a}{2}\right)$

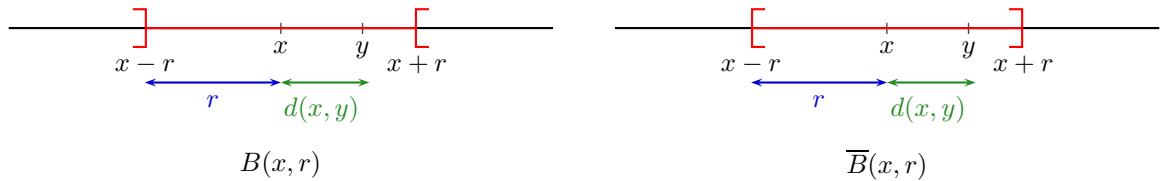


FIGURE 6.4 – Boule ouverte, boule fermée dans \mathbb{R}

Remarque 6.4.9

Il est important de retenir qu'une majoration de certaines valeur absolue se traduit par l'appartenance à une boule :

- une majoration du type $|x - a| \leq r$ traduit l'appartenance de x à la boule fermée $\overline{B}(a, r)$ de centre a de rayon r , donc à l'intervalle $[a - r, a + r]$;
- une majoration du type $|x - a| < r$ traduit l'appartenance de x à la boule ouverte $B(a, r)$ de centre a de rayon r , donc à l'intervalle $(a - r, a + r)$;

Définition 6.4.10 (Voisinage, figure 6.5)

Soit $x \in \mathbb{R}^n$. Un *voisinage* V de x est un sous-ensemble V de \mathbb{R}^n tel qu'il existe une boule ouverte centrée en x entièrement contenue dans V :

$$\exists \varepsilon > 0, B(x, \varepsilon) \subset V, \quad \text{i.e.} \quad \exists \varepsilon > 0, \forall y \in E, d(y, x) < \varepsilon \implies y \in V.$$

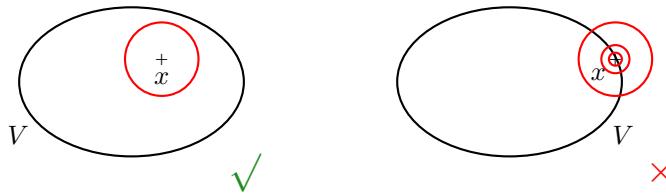


FIGURE 6.5 – Voisinage de x

En gros, V est un voisinage de x si x est « à l'intérieur de V », et non sur un bord. En s'éloignant un peu de x , on ne sort pas de V .

Exemple 6.4.11 (Voisinages dans \mathbb{R})

- Dans \mathbb{R} , un voisinage de x est un ensemble contenant un intervalle (a, b) tel que $x \in (a, b)$.
- Par extension et commodité, on dit parfois qu'un ensemble contenant un intervalle $(a, +\infty)$ est un voisinage de $+\infty$. Version symétrique pour $-\infty$.

Définition 6.4.12 (sous-ensemble ouvert)

- Un *ouvert* U de \mathbb{R}^n est un sous-ensemble U de \mathbb{R}^n qui est voisinage de tous ses points
- De manière équivalente, $U \subset \mathbb{R}^n$ est ouvertssi :

$$\forall x \in U, \exists \varepsilon > 0, B(x, \varepsilon) \subset U.$$

Intuitivement, un ouvert est un ensemble dont le « bord » est flou : on peut s'en approcher, mais jamais l'atteindre en restant dans U . Ainsi, l'image qu'il faut en garder est qu'un ouvert est un ensemble ne contenant pas son bord. Évidemment, n'ayant pas défini la notion de bord, ceci reste une image.

Définition 6.4.13 (sous-ensemble fermé)

Un sous-ensemble F de \mathbb{R}^n est *fermé* si son complémentaire $\complement_E F$ est ouvert.

Cette fois, intuitivement, c'est le complémentaire qui ne contient pas son bord, donc F , lui contient tout son bord.

Exemples 6.4.14

1. Les intervalles ouverts sont des sous-ensembles ouverts de \mathbb{R} .
2. Les intervalles fermés sont des sous-ensembles fermés de \mathbb{R} .
3. Les intervalles semi-ouverts ne sont ni ouverts ni fermés.
4. Plus généralement, une boule ouverte de \mathbb{R}^n est un ouvert ; une boule fermée de \mathbb{R}^n est un fermé (ce qui justifie la terminologie adoptée lors des définitions de ces objets)
5. \mathbb{R} et \emptyset sont des sous-ensembles à la fois fermés et ouverts de \mathbb{R} .
6. On peut montrer que les sous-ensembles ouverts de \mathbb{R} sont les unions disjointes d'intervalles ouverts.

Proposition 6.4.15 (union, intersection d'ouverts et de fermés)

1. Toute union quelconque d'ouverts est un ouvert ;
2. Toute intersection d'un nombre fini d'ouverts est un ouvert ;
3. Toute intersection quelconque de fermés est un fermé ;
4. Toute union d'un nombre fini de fermés est un fermé.

△ Éléments de preuve.

1. Si x est dans l'union, il existe une boule centrée en x restant dans l'un des ouverts, donc dans leur union.
2. Prendre le minimum des rayons des boules centrées en x restant dans chaque ouvert. Pourquoi doit-on se limiter au cas fini ?
3. Par complémentation
4. De même.

▷

Exemples 6.4.16

Voici deux contre-exemples à bien garder en tête :

1. Contre-exemple pour une intersection infinie d'ouverts : $\bigcap_{n=1}^{+\infty} \left[-\frac{1}{n}, 1 \right] = [0, 1[$.
2. Contre-exemple pour une union infinie de fermés : $\bigcup_{n=1}^{+\infty} \left[\frac{1}{n}, 1 \right] =]0, 1]$.

V Droite achevée $\bar{\mathbb{R}}$

Par commodité, il est parfois intéressant de pouvoir considérer les deux infinis comme des éléments comme les autres. Cela permet en particulier d'unifier certains énoncés et certaines démonstrations, qui sinon, nécessiteraient une disjonction de cas.

Définition 6.5.1 (droite achevée réelle)

La droite achevée réelle, notée $\bar{\mathbb{R}}$, est l'ensemble $\mathbb{R} \cup \{-\infty, +\infty\}$.

Définition 6.5.2 (relation d'ordre sur $\overline{\mathbb{R}}$)

On peut prolonger l'ordre de \mathbb{R} en un ordre sur $\overline{\mathbb{R}}$ en posant :

$$\forall x \in \overline{\mathbb{R}}, -\infty \leq x \leq +\infty.$$

Définition 6.5.3 (règles calculatoires dans $\overline{\mathbb{R}}$)

On peut prolonger partiellement les opérations de \mathbb{R} sur $\overline{\mathbb{R}}$, en posant :

- $-(+\infty) = -\infty$
- $\forall x \in \overline{\mathbb{R}} \setminus \{-\infty\}, x + (+\infty) = +\infty$
- $\forall x \in \overline{\mathbb{R}} \setminus \{+\infty\}, x + (-\infty) = -\infty$
- $\frac{1}{+\infty} = \frac{1}{-\infty} = 0$
- $\forall x \in \overline{\mathbb{R}}_+, x \times (+\infty) = +\infty, x \times (-\infty) = -\infty,$
- $\forall x \in \overline{\mathbb{R}}_-, x \times (+\infty) = -\infty, x \times (-\infty) = +\infty.$

En revanche, certaines opérations ne peuvent pas être définies de façon cohérente, comme le montre l'étude des formes indéterminées dans le calcul des limites.

Définition 6.5.4 (Formes indéterminées)

Les opérations suivantes ne sont pas définies, et définissent les formes indéterminées de la somme et du produit dans $\overline{\mathbb{R}}$:

- $-\infty + (+\infty)$
- $0 \times (+\infty)$
- $0 \times (-\infty).$

Pour l'étude des limites, ces formes donnent également des formes indéterminées pour les puissances, par passage à l'exponentielle (voir le chapitre sur les suites).

Proposition 6.5.5 (Bornes supérieures)

Tout sous-ensemble E de $\overline{\mathbb{R}}$ admet une borne supérieure dans $\overline{\mathbb{R}}$

▫ Éléments de preuve.

Éliminer les cas $+\infty \in \overline{\mathbb{R}}$, $E = \emptyset$ et $E = \{-\infty\}$, pour se ramener au cas d'un sous-ensemble non vide $E' = E \cap \mathbb{R}$. Discuter suivant qu'il est majoré ou non. ▷

Remarque 6.5.6

Le résultat est bien sûr aussi valable pour les bornes inférieures.

Les intervalles de $\overline{\mathbb{R}}$ étant définis comme ceux de \mathbb{R} par la propriété de convexité, on obtient la description suivante :

Proposition 6.5.7 (Inventaire des intervalles de $\overline{\mathbb{R}}$)

Les intervalles de $\overline{\mathbb{R}}$ sont les intervalles de \mathbb{R} , et les intervalles de la forme $]a, +\infty]$, $[a, +\infty]$, $[-\infty, a[$, $[-\infty, a]$, $]-\infty, +\infty]$, $[-\infty, +\infty[, et $[-\infty, +\infty] = \overline{\mathbb{R}}$, pour $a \in \mathbb{R}$.$

▫ Éléments de preuve.

Comme dans le cas réel, si a et b désignent les bornes inférieure et supérieure de I , on a la double inclusion $]a, b[\subset [a, b]$. ▷

Le corps \mathbb{C} des complexes

Au reste, tant les vraies racines que les fausses ne sont pas toujours réelles, mais quelquefois seulement imaginaires, c'est à dire qu'on peut bien toujours en imaginer autant que j'ai dit en chaque équation, mais qu'il n'y a quelquefois aucune quantité qui corresponde à celles qu'on imagine ; comme encore qu'on puisse en imaginer trois en celle-ci :

$$x^3 - 6x^2 + 13x - 10 = 0,$$

il n'y en a toutefois qu'une réelle qui est 2, et pour les deux autres, quoiqu'on les augmente ou diminue, ou multiplie en la façon que je viens d'expliquer, on ne saurait les rendre autres qu'imaginaires.

(René Descartes)

Ce que nous nommons temps imaginaire est en réalité le temps réel, et ce que nous nommons temps réel n'est qu'une figure de notre imagination.

(Stephen Hawking)

Les nombres complexes sont nés de l'étude des solutions des équations du troisième degré. Dans un premier temps, ils ont été utilisés sans se préoccuper de leur donner un sens précis : ils étaient des outils abstraits, imaginaire pourrait-on dire, pour accéder aux solutions, pouvant elles être bien réelles, des équations considérées.

I Les nombres complexes : définition et manipulations

I.1 Définition, forme algébrique

Note Historique 7.1.1

- Les nombres complexes ont été introduits par Cardan et Bombelli au 16-ième siècle, comme moyen d'exprimer certaines racines de polynômes de degrés 3 ou 4. A cette époque, l'introduction des nombres imaginaires (*via* des racines de réels négatifs) est un pur artifice.
- Ainsi, dès leur origine, les nombres complexes sont introduits pour pallier au fait que certains polynômes à coefficients réels n'ont pas de racines dans \mathbb{R} , comme par exemple $X^2 + 1$.
- La notation i est introduite par Euler en 1777 pour remplacer la notation $\sqrt{-1}$.

D'un point de vue formel, \mathbb{C} est défini comme le plus petit sur-corps de \mathbb{R} dans lequel le polynôme $X^2 + 1$ admet une racine (c'est ce qu'on appelle un corps de rupture du polynôme $X^2 + 1$, correspondant dans ce cas au corps de décomposition, le plus petit corps dans lequel le polynôme peut se factoriser en polynômes de degré 1).

Ainsi, il s'agit d'un ensemble contenant un élément i , racine de $X^2 + 1$, vérifiant donc $i^2 = -1$, et muni d'une addition et d'un produit prolongeant celles de \mathbb{R} , avec les mêmes propriétés. En fait, la relation $i^2 = -1$ et les propriétés de commutativité, associativité et distributivité déterminent entièrement les opérations sur \mathbb{C} . Nous donnons la définition suivante :

Définition 7.1.2 (ensemble \mathbb{C} des nombres complexes)

L'ensemble des nombres complexes \mathbb{C} est l'ensemble \mathbb{R}^2 , muni des opérations suivantes :

- $(a, b) + (a', b') = (a + a', b + b')$;
- $(a, b) \times (a', b') = (aa' - bb', ab' + a'b)$.

Pour tout réel λ et tout complexe $z = (a, b)$, on peut définir λz par $\lambda z = (\lambda a, \lambda b)$.

Remarque 7.1.3

On identifie un réel λ au complexe $(\lambda, 0)$ (possible du fait de l'injectivité de $\lambda \mapsto (\lambda, 0)$, et du respect des lois $+$ et \times). Via cette identification, on peut considérer que $\mathbb{R} \subset \mathbb{C}$, et on vérifie facilement que la somme et le produit définis ci-dessus sur \mathbb{C} prolongent les lois de \mathbb{R} , et que la multiplication d'un complexe z par un réel λ définie ci-dessus correspond au produit dans \mathbb{C} , lorsqu'on voit le réel λ comme un élément de \mathbb{C} .

Définition 7.1.4 (définition de la forme algébrique ; partie réelle, partie imaginaire)

- On note $1 = (1, 0)$, et $i = (0, 1)$
- On a alors, pour tout $z = (a, b) \in \mathbb{C}$,

$$z = a \cdot 1 + b \cdot i = a + i b.$$

(du fait de l'identification de la remarque précédente). C'est la *forme algébrique* du nombre complexe z .

- Soit $z = a + i b$, avec $a, b \in \mathbb{R}$.
 - * Le réel a est appelé *partie réelle de z* , et est noté $\text{Re}(z)$;
 - * Le réel b est appelé *partie imaginaire de z* , et est noté $\text{Im}(z)$
- Un nombre $z \in \mathbb{C}$ tel que $\text{Re}(z) = 0$ est appelé *nombre imaginaire pur*.
- Un nombre $z \in \mathbb{C}$ vérifie $\text{Im}(z) = 0$ ssi $z \in \mathbb{R}$.

Proposition 7.1.5 (propriétés liées au produit)

1. $i^2 = -1$
2. Le produit $(a + i b)(a' + i b')$ est simplement obtenu par utilisation des règles de distributivité et par la relation $i^2 = -1$.
3. Si $z \neq 0$, alors z est inversible, et, si $z = a + i b$ avec $(a, b) \in \mathbb{R}^2$, on a l'expression de l'inverse :

$$z^{-1} = \frac{a - i b}{a^2 + b^2}.$$

▫ Éléments de preuve.

Ce sont des vérifications immédiates. ▷

Théorème 7.1.6 (structure de \mathbb{C})

L'ensemble \mathbb{C} muni des opérations ci-dessus est un corps.

▫ Éléments de preuve.

Il faut vérifier la commutativité des opérations, leur associativité, la distributivité du produit sur la somme, l'existence du neutre additif 0 et du neutre multiplicatif 1, l'existence des opposés $-x$ et l'inversibilité de tout complexe non nul. Fastidieux mais pas difficile. ▷

Dans la pratique, un nombre complexe est représenté sous sa forme algébrique $a + ib$, ou sa forme trigonométrique que nous rappellerons plus loin. On perd un peu de vue le point de vue initial du couple (d'ailleurs, on peut introduire \mathbb{C} de façon différente). Nous donnons dans la définition suivante la démarche inverse, permettant de revenir de \mathbb{C} à \mathbb{R}^2 . Cette interprétation est fructueuse pour la géométrie du plan, pouvant ainsi être étudiée sous l'angle des nombres complexes.

Définition 7.1.7 (affixe d'un point du plan)

Soit $A = (a, b)$ un point de \mathbb{R}^2 . L'*affixe* du point A est le nombre complexe $z_A = a + ib$.

Désormais, nous abandonnons la notation d'un complexe sous forme d'un couple, et nous représenterons un nombre complexe sous la forme $a + ib$.

Remarquons que la construction de \mathbb{C} à partir de \mathbb{R} est un cas particulier d'une construction plus générale d'« extensions monogènes d'un corps \mathbb{K} », à la base de la théorie de Galois : étant donné une racine x d'une équation polynomiale P à coefficients dans \mathbb{K} , $\mathbb{K}[x]$ est l'ensemble de toutes les sommes, produits, quotients qu'on peut former à partir des éléments de \mathbb{K} et de x . Par exemple $\mathbb{Q}[\sqrt{2}] = \{(a + b\sqrt{2}), (a, b) \in \mathbb{Q}^2\}$, et $\mathbb{Q}[\sqrt[3]{2}] = \{a + b\sqrt[3]{2} + c(\sqrt[3]{2})^2, (a, b, c) \in \mathbb{Q}^3\}$ (on peut montrer que tout quotient peut s'exprimer ainsi aussi).

De ce point de vue, $\mathbb{C} = \mathbb{R}[i]$.

Nous avons donc défini \mathbb{C} comme un corps de rupture sur \mathbb{R} du polynôme $X^2 + 1$, et même un corps de décomposition, puisqu'on a alors la factorisation suivante : $X^2 + 1 = (X + i)(X - i)$. En fait, cette propriété est beaucoup plus générale, ainsi que le prouve le théorème suivant, d'une importance capitale :

Théorème 7.1.8 (d'Alembert-Gauss, admis)

Tout polynôme non constant à coefficients complexes admet au moins une racine dans \mathbb{C} .

On démontrera plus tard que ceci implique que tout polynôme à coefficients complexes se factorise en polynômes de degré 1.

Le théorème de d'Alembert-Gauss se réexprime ainsi : \mathbb{C} est algébriquement clos, ce qui signifie qu'il n'existe pas d'autre nombre algébrique sur \mathbb{C} que les nombres complexes eux-mêmes.

Le théorème de d'Alembert-Gauss, restreint aux polynômes à coefficients réels, allié au fait que \mathbb{C} est par définition le plus petit corps dans lequel $X^2 + 1$ admet une racine, s'exprime en disant que \mathbb{C} est la clôture algébrique de \mathbb{R} .

Note Historique 7.1.9

Le théorème de d'Alembert-Gauss est d'une importance capitale, puisque c'est ce résultat qui motive la construction de \mathbb{C} .

- Il est conjecturé depuis longtemps déjà lorsque d'Alembert en propose une preuve en 1743. Cette preuve n'est pas satisfaisante, Gauss va jusqu'à la qualifier de *petitio principii*, puisqu'elle part de l'hypothèse de l'existence de racines « fictives ».
- La première preuve complète et rigoureuse revient à Gauss, au 19-ième siècle.

Nous voyons maintenant quelques notions directement liées à la forme algébrique des nombres complexes

Définition 7.1.10 (conjugué d'un nombre complexe)

Soit $z = a + \mathrm{i}b$ (avec $(a, b) \in \mathbb{R}^2$) un nombre complexe. Le *conjugué* de z est le nombre complexe

$$\overline{z} = a - \mathrm{i}b.$$

Propriétés 7.1.11 (propriétés de la conjugaison dans \mathbb{C})

Soit z et z' deux nombres complexes. Alors :

1. $\overline{\overline{z}} = z$ (autrement dit, la conjugaison est une involution) ;
2. $z = \overline{z} \iff z \in \mathbb{R}$;
3. $z = -\overline{z} \iff z$ imaginaire pur ;
4. $\operatorname{Re}(z) = \frac{z + \overline{z}}{2}$ et $\operatorname{Im}(z) = \frac{z - \overline{z}}{2\mathrm{i}}$;
5. $\overline{z + z'} = \overline{z} + \overline{z'}$, $\overline{zz'} = \overline{z} \cdot \overline{z'}$, $\overline{z^{-1}} = \overline{z}^{-1}$, $\overline{\left(\frac{z}{z'}\right)} = \frac{\overline{z}}{\overline{z'}}$.

▫ Éléments de preuve.

Vérifications faciles. Pour l'inverse, utiliser la formule obtenue tantôt, ou conjuguer la relation $zz^{-1} = 1$. ▷

I.2 Module

Nous définissons maintenant quelques notions liées à la structure euclidienne de \mathbb{R}^2 , donc aux propriétés métriques.

Définition 7.1.12 (module d'un nombre complexe)

Soit $(a, b) \in \mathbb{R}^2$, et $z = a + \mathrm{i}b$. Le *module* de z est le réel positif défini par

$$|z| = \sqrt{a^2 + b^2}.$$

Remarque 7.1.13

Via la correspondance entre \mathbb{C} et \mathbb{R}^2 le module d'un nombre complexe correspond à la norme des vecteurs. Ainsi, si A est le point d'affixe z et O l'origine, alors $|z| = \|\overrightarrow{OA}\|$.

Exemples 7.1.14

1. Décrire l'ensemble des nombres complexes z tels que $|z - a| = r$, où $a \in \mathbb{C}$ et $r \in \mathbb{R}_+^*$.
2. Même question avec l'inéquation $|z - a| \leq r$.

Propriétés 7.1.15 (propriétés du module)

Soit z et z' deux nombres complexes. Alors :

1. $z = 0 \iff |z| = 0$;
2. $|\operatorname{Re}(z)| \leq |z|$ et $|\operatorname{Im}(z)| \leq |z|$;
3. $|z|^2 = z\overline{z}$ (expression du module à l'aide du conjugué)
4. $|zz'| = |z| \cdot |z'|$ et si $z' \neq 0$, $\left|\frac{z}{z'}\right| = \frac{|z|}{|z'|}$ (multiplicativité du module)

5. $|z| = |\bar{z}|$ (*invariance du module par conjugaison*)
 6. $|z + z'| \leq |z| + |z'|$ (*inégalité triangulaire, ou sous-additivité du module*).
L'égalité est vérifiée si et seulement si $z = 0$ ou s'il existe $\lambda \in \mathbb{R}_+$ tel que $z' = \lambda z$.

▫ Éléments de preuve.

Le seul point non trivial est l'inégalité triangulaire. Comparer les carrés, en exprimant les modules à l'aide des conjugués. Se ramener à une comparaison de la partie réelle et du module de $z \cdot \bar{z}'$. ▷

On en déduit notamment une méthode pour exprimer sous forme algébrique un quotient de deux complexes donnés sous forme algébrique (pour les autres opérations, cela ne pose aucune difficulté).

Méthode 7.1.16 (Expression algébrique d'un quotient)

Soit z_1 et z_2 deux nombres complexes donnés sous forme algébrique, avec $z_2 \neq 0$. Pour trouver la forme algébrique du quotient $\frac{z_1}{z_2}$, multipliez le dénominateur et le numérateur par \bar{z}_2 , c'est-à-dire considérez l'expression $\frac{z_1 \cdot \bar{z}_2}{z_2 \cdot \bar{z}_2}$.

De la sorte, le dénominateur est maintenant un réel.

II Trigonométrie

II.1 Cercle trigonométrique, formules de trigonométrie

Littéralement, « trigonométrie » signifie « mesure des trois angles », donc se rapporte aux propriétés des angles d'un triangle. On fait donc ainsi référence aux interprétations géométriques usuelles des fonctions trigonométriques.

Note Historique 7.2.1

- La trigonométrie (l'étude des mesures dans le triangle) existe depuis l'antiquité (Égypte, Babylone, Grèce), et est développée en rapport avec l'astronomie.
- Le sinus, sous sa forme actuelle, a été introduit par les indiens aux alentours de 500 ap JC, pour l'étude des angles célestes. La première table connue date de 499, et est attribuée au mathématicien indien Aryabhata. En 628, Brahmagupta construit une approximation de la fonction sinus par interpolation.
- Ces notions nous sont parvenus grâce aux travaux de synthèse des mathématiciens arabes des 9^e et 10^e siècles (essentiellement basés dans les actuelles Irak, Iran et Khazakstan)
- Auparavant, les grecs utilisaient plutôt la mesure de la corde, ce qui est moins commode, mais assez équivalent.
- Le nom de « sinus » provient d'un mot sanscrit signifiant « arc », apparaissant dans l'ouvrage de Aryabhata, et transcrit phonétiquement en arabe, puis déformé en un mot proche signifiant « repli de vêtement ». Il a été traduit en latin au 12^e siècle par le mot « sinus » signifiant « pli ».

Définition 7.2.2 (cercle trigonométrique)

Le cercle trigonométrique (ou cercle unité) est le sous-ensemble de \mathbb{C} , noté \mathbb{U} (comme « unité »), constitué des nombres complexes de module 1 :

$$\mathbb{U} = \{z \in \mathbb{C} \mid |z| = 1\}.$$

Le cercle trigonométrique correspond dans l'interprétation géométrique des complexes au cercle de centre $(0, 0)$ et de rayon 1 de \mathbb{R}^2 .

Définition 7.2.3 (fonctions trigonométriques, figure 7.1)

Soit $x \in \mathbb{R}$. On considère le cercle trigonométrique dans le plan euclidien canonique. Soit z le point du cercle trigonométrique tel que le rayon correspondant du cercle trigonométrique forme avec l'axe des réels un angle (orienté dans le sens direct) de x . On définit alors les fonctions cosinus, sinus et tangente par :

$$\cos(x) = \operatorname{Re}(z), \quad \sin(x) = \operatorname{Im}(z) \quad \text{et} \quad \tan(x) = \frac{\sin(x)}{\cos(x)} \text{ si } \cos(x) \neq 0.$$

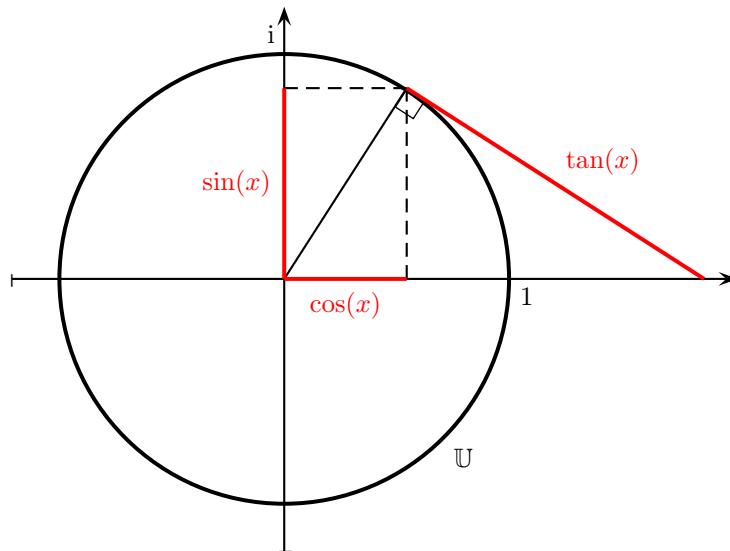


FIGURE 7.1 – Fonctions trigonométriques

Remarque 7.2.4

La bonne définition des fonctions trigonométriques nécessite de justifier l'existence et l'unicité d'un point de \mathbb{U} sur une demi-droite issue de 0. Elle nécessite aussi d'avoir défini la notion d'angle, ce qu'on fait ici de façon élémentaire en prenant la proportion convenable de l'angle total de 2π .

On retrouve facilement l'interprétation usuelle sur les triangles (définitions données au collège), pour un angle $\alpha \in]0, \frac{\pi}{2}[$: si ABC est un triangle rectangle en A , et d'angle en B égal à α , alors :

$$\sin(\alpha) = \frac{\text{côté opposé}}{\text{hypoténuse}}, \quad \cos(\alpha) = \frac{\text{côté adjacent}}{\text{hypoténuse}}, \quad \tan(\alpha) = \frac{\text{côté opposé}}{\text{côté adjacent}}.$$

On utilise parfois une autre fonction trigonométrique, version symétrique de la tangente :

Définition 7.2.5 (cotangente)

Soit $x \in \mathbb{R}$. La cotangente est définie, pour tout x tel que $\sin(x) \neq 0$, par :

$$\operatorname{cotan}(x) = \frac{\cos(x)}{\sin(x)}.$$

Ainsi, en tout point en lequel à la fois $\sin(x) \neq 0$ et $\cos(x) \neq 0$, on a $\operatorname{cotan}(x) = \frac{1}{\tan(x)}$.

Proposition 7.2.6 (domaines de définition des fonctions trigonométriques)

1. *Les fonctions sin et cos sont définies sur \mathbb{R} .*

2. La fonction \tan est définie sur $\bigcup_{n \in \mathbb{Z}} \left[-\frac{\pi}{2} + n\pi, \frac{\pi}{2} + n\pi \right] = \mathbb{R} \setminus \left\{ \frac{\pi}{2} + n\pi, n \in \mathbb{Z} \right\}$.

3. La fonction \cotan est définie sur $\bigcup_{n \in \mathbb{Z}}]n\pi, (n+1)\pi[= \mathbb{R} \setminus \{n\pi, n \in \mathbb{Z}\}$.

Les propriétés suivantes sont à bien comprendre et visualiser sur le cercle trigonométrique.

Proposition 7.2.7 (Symétries de sin et cos)

1. \sin et \cos sont 2π -périodiques ;
2. \sin est impaire et \cos est paire ;
3. $\forall x \in \mathbb{R}, \cos(\pi + x) = -\cos(x)$, et $\sin(\pi + x) = -\sin(x)$.
4. $\forall x \in \mathbb{R}, \cos(\pi - x) = -\cos(x)$, et $\sin(\pi - x) = \sin(x)$.
5. $\forall x \in \mathbb{R}, \cos\left(\frac{\pi}{2} - x\right) = \sin(x)$, et $\sin\left(\frac{\pi}{2} - x\right) = \cos(x)$.
6. $\forall x \in \mathbb{R}, \cos\left(\frac{\pi}{2} + x\right) = -\sin(x)$, et $\sin\left(\frac{\pi}{2} + x\right) = \cos(x)$.

▫ Éléments de preuve.

Du fait de la définition géométrique donnée, la preuve l'est aussi : il s'agit d'étudier les symétries du cercle trigonométrique. ▷

Proposition 7.2.8 (symétries de tan et cotan)

1. \tan et \cotan sont π -périodiques ;
2. \tan et \cotan sont impaires ;
3. pour tout x dans le domaine de \tan , $\tan(\pi - x) = -\tan(x)$;
4. pour tout x dans le domaine de \cotan , $\cotan(\pi - x) = -\cotan(x)$;
5. pour tout $x \in \mathbb{R} \setminus \left\{ \frac{n\pi}{2}, n \in \mathbb{Z} \right\}$, $\tan\left(\frac{\pi}{2} - x\right) = \cotan(x)$ et $\cotan\left(\frac{\pi}{2} - x\right) = \tan(x)$;
6. pour tout $x \in \mathbb{R} \setminus \left\{ \frac{n\pi}{2}, n \in \mathbb{Z} \right\}$, $\tan\left(\frac{\pi}{2} + x\right) = -\cotan(x)$ et $\cotan\left(\frac{\pi}{2} + x\right) = -\tan(x)$;

Nous rappelons :

Proposition 7.2.9 (Valeurs particulières des fonctions trigonométriques)

Voici un tableau des valeurs particulières à bien connaître, entre 0 et $\frac{\pi}{2}$ (les autres s'obtiennent par les symétries) :

	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
sin	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1
cos	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
tan	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	—
cotan	—	$\sqrt{3}$	1	$\frac{1}{\sqrt{3}}$	0

▫ Éléments de preuve.

Amusez-vous à retrouver ces valeurs géométriquement, en vous servant du théorème de Pythagore. ▷

Nous obtenons les graphes des fonctions trigonométriques, les variations pouvant être obtenues par des considérations purement géométriques (voir figure 7.2).

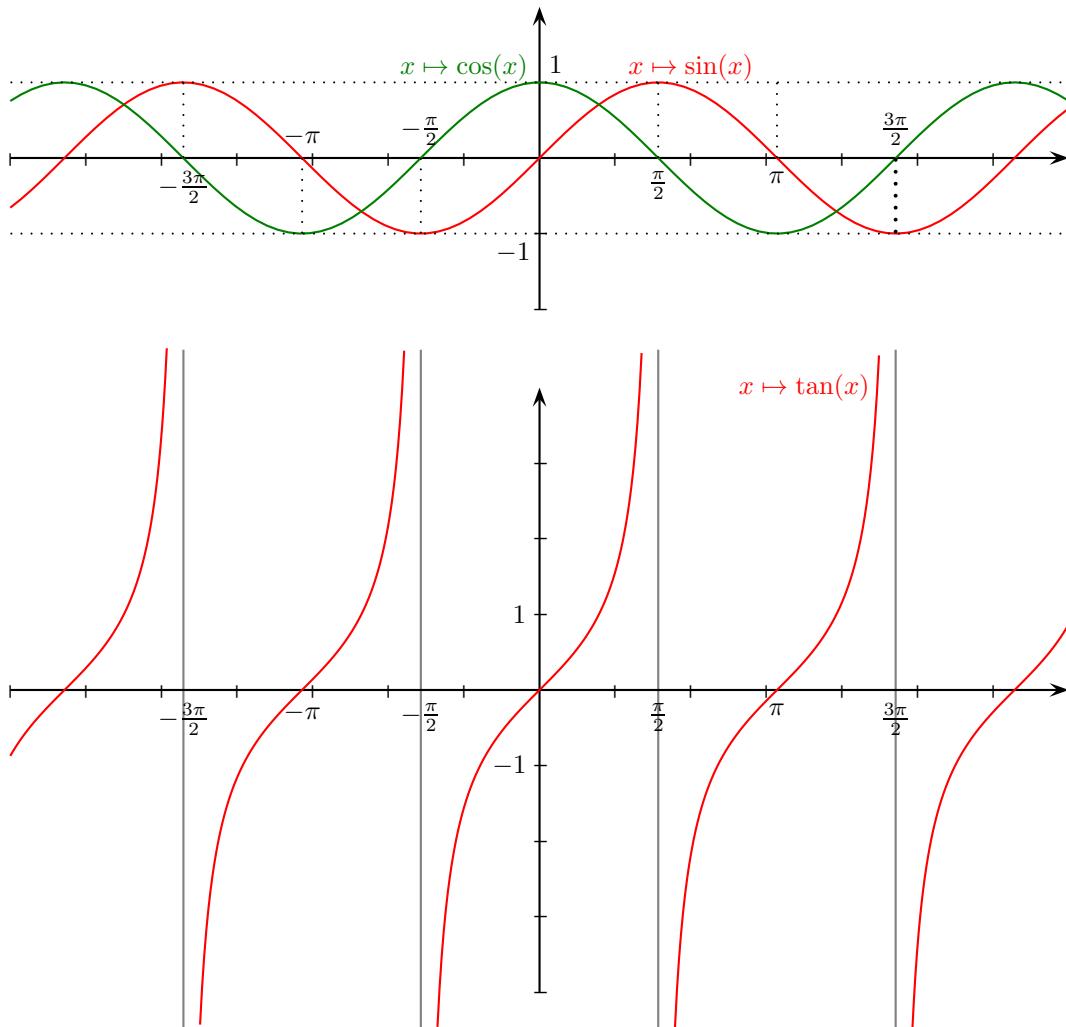


FIGURE 7.2 – Graphe des fonctions trigonométriques

Et voici les inévitables formules à retenir...

Proposition 7.2.10 (identité remarquable, ou théorème de Pythagore)

Pour tout $x \in \mathbb{R}$, $\sin^2(x) + \cos^2(x) = 1$.

▫ Éléments de preuve.

C'est juste dire que $\cos(x) + i\sin(x)$ est sur le cercle... ▷

Proposition 7.2.11 (formules d'addition)

Soit a et b deux réels. Alors :

- (i) $\sin(a + b) = \sin(a)\cos(b) + \sin(b)\cos(a)$
- (ii) $\sin(a - b) = \sin(a)\cos(b) - \sin(b)\cos(a)$

$$(iii) \cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$$

$$(iv) \cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b)$$

(v) Pour tout $(a, b) \in \mathbb{R}^2$ tel que $\tan(a)$, $\tan(b)$ et $\tan(a+b)$ soient définis :

$$\tan(a+b) = \frac{\tan(a) + \tan(b)}{1 - \tan(a)\tan(b)}$$

(vi) Pour tout $(a, b) \in \mathbb{R}^2$ tel que $\tan(a)$, $\tan(b)$ et $\tan(a-b)$ soient définis :

$$\tan(a-b) = \frac{\tan(a) - \tan(b)}{1 + \tan(a)\tan(b)}$$

(vii) Pour tout $(a, b) \in \mathbb{R}^2$ tel que $\cotan(a)$, $\cotan(b)$ et $\cotan(a+b)$ soient définis :

$$\cotan(a+b) = \frac{\cotan(a)\cotan(b) - 1}{\cotan(a) + \cotan(b)}$$

△ Éléments de preuve.

(ii) découle de (i) en considérant $-b$, (iii) découle de (i) en considérant $\sin(\frac{\pi}{2} + a + b)$, et (iv) découle de (iii). (v), (vi) et (vii) en découlent par quotient et simplifications.

On peut montrer (i) géométriquement par projections signées. Cela revient à considérer des produits scalaires, mais ce point de vue anticiperait un peu trop les derniers chapitres de l'année. ▷

Proposition 7.2.12 (formules de duplication des angles)

Soit a un réel. Alors :

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

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

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

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

△ Éléments de preuve.

Appliquer ce qui précède avec $a = b$. ▷

Proposition 7.2.13 (formules de linéarisation des carrés, ou formules de Carnot)

Soit $a \in \mathbb{R}$:

$$(i) \cos^2(a) = \frac{1 + \cos(2a)}{2}$$

$$(ii) \sin^2(a) = \frac{1 - \cos(2a)}{2}$$

△ Éléments de preuve.

Utiliser (ii) de la proposition précédente. ▷

Proposition 7.2.14 (formules de développement, ou transformation de produit en somme)

Soit a et b deux réels.

$$(i) \sin(a)\sin(b) = \frac{1}{2}[\cos(a-b) - \cos(a+b)]$$

$$(ii) \cos(a)\cos(b) = \frac{1}{2}[\cos(a-b) + \cos(a+b)]$$

$$(iii) \sin(a)\cos(b) = \frac{1}{2}[\sin(a+b) + \sin(a-b)]$$

▫ Éléments de preuve.

Sommer ou soustraire 2 par 2 les formules d'addition.

▷

Proposition 7.2.15 (formules de factorisation, ou formules de Simpson)

Soit p et q deux réels.

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

▫ Éléments de preuve.

Changement de variable dans les formules de développement.

▷

Proposition 7.2.16 (formules de l'arc moitié)

Soit x un réel tel que $t = \tan\left(\frac{x}{2}\right)$ soit défini. Alors :

- (i) $\sin(x) = \frac{2t}{1+t^2}$
- (ii) $\cos(x) = \frac{1-t^2}{1+t^2}$
- (iii) $\tan(x) = \frac{2t}{1-t^2}$ si $t \neq \pm 1$.

▫ Éléments de preuve.

Utiliser la formule de duplication de l'angle et le fait que $\frac{1}{\cos^2(y)} = 1 + \tan^2(y)$.

▷

Proposition 7.2.17 (formule de factorisation de $a \cos x + b \sin x$)

Soit a , b et x trois réels, $a \neq 0$. Alors

$$a \cos(x) + b \sin(x) = \frac{a}{\cos \varphi} \cos(x - \varphi), \text{ où } \tan(\varphi) = \frac{b}{a}.$$

▫ Éléments de preuve.

Utilisation facile de la formule d'addition.

▷

En physique, $\frac{a}{\cos(\varphi)}$ est appelé *amplitude* et φ est appelé la *phase*.

II.2 L'exponentielle complexe et applications à la trigonométrie

Définition 7.2.18 (exponentielle complexe)

On définit l'exponentielle complexe sur les nombres imaginaires purs par :

$$\forall \theta \in \mathbb{R}, \quad e^{i\theta} = \cos \theta + i \sin \theta.$$

Proposition 7.2.19

La fonction $\theta \mapsto e^{i\theta}$ est surjective de \mathbb{R} sur \mathbb{U} . Plus précisément, c'est une bijection de tout intervalle $[\alpha, \alpha + 2\pi]$ sur \mathbb{U} , ainsi que de tout intervalle $[\alpha, \alpha + 2\pi[$ sur \mathbb{U} .

▫ Éléments de preuve.

La surjectivité provient de la définition-même de cosinus et sinus comme projetés d'un point du cercle. L'injectivité modulo 2π provient de la résolution des équations $\cos(\theta) = a, \sin(\theta) = b$, provenant des variations de ces fonctions. ▷

Corollaire 7.2.20

La fonction de $\mathbb{R}_+^* \times]-\pi, \pi]$ sur \mathbb{C}^* définie par $(r, \theta) \mapsto r e^{i\theta}$ est bijective.

Définition 7.2.21 (forme trigonométrique)

- Ainsi, tout nombre complexe non nul z s'écrit sous la forme $z = r e^{i\theta}$ appelée forme trigonométrique de z , avec $r > 0$;
- r est unique, égal au module de z ;
- θ est unique modulo 2π , appelé argument de z .
- L'unique argument θ de l'intervalle $]-\pi, \pi]$ est appelé *argument principal* de z et est noté $\arg(z)$.

Proposition 7.2.22 (formules d'Euler)

Soit $\theta \in \mathbb{R}$. Alors :

$$\cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad \text{et} \quad \sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

Théorème 7.2.23 (formules trigonométriques d'addition)

Pour tout $(\theta, \theta') \in \mathbb{R}^2$, $e^{i(\theta+\theta')} = e^{i\theta} e^{i\theta'}$.

▫ Éléments de preuve.

Le rapport de l'un à l'autre se fait par les formules d'addition. ▷

Remarque 7.2.24

Cette formule synthétise les formules d'addition. Si vous les avez oubliées, vous les retrouvez facilement en considérant la partie réelle ou la partie imaginaire de l'égalité ci-dessus.

Corollaire 7.2.25 (Formule de (De) Moivre, 1707)

Pour tout $\theta \in \mathbb{R}$ et $n \in \mathbb{N}$:

$$e^{in\theta} = (e^{i\theta})^n \quad \text{soit:} \quad \cos(n\theta) + i \sin(n\theta) = (\cos \theta + i \sin \theta)^n.$$

▫ Éléments de preuve.

Réurrence triviale à partir du théorème précédent.

▷

Note Historique 7.2.26

- Abraham de Moivre était un mathématicien ami des physiciens et astronomes Newton et Halley. Il faudrait théoriquement dire « formule de De Moivre », (selon la règle de conservation de la particule onomastique pour les noms d'une syllabe, comme de Gaulle), mais on dit plus souvent « formule de Moivre ».
- La version démontrée par Moivre est la version donnée avec les fonctions sin et cos, le lien avec les propriétés de l'exponentielle n'ayant été découvertes que plus tard par Euler (19^e siècle), qui est à l'origine de la notation exponentielle $e^{i\theta}$.

L'utilisation des exponentielles complexes permet de simplifier un certain nombre de calculs liés à la trigonométrie. Nous présentons ci-dessous quelques méthodes à connaître.

Méthode 7.2.27 (Principe de symétrisation des arguments)

Cette méthode permet d'exprimer une somme ou une différence de deux exponentielles à l'aide des fonctions trigonométriques. C'est notamment intéressant pour obtenir la partie réelle et la partie imaginaire sous forme factorisée. Soit a et b deux réels. Alors :

- $e^{ia} + e^{ib} = e^{i\frac{a+b}{2}} \left(e^{i\frac{a-b}{2}} + e^{-i\frac{a-b}{2}} \right) = 2 \cos\left(\frac{a-b}{2}\right) e^{i\frac{a+b}{2}}$.
- $e^{ia} - e^{ib} = e^{i\frac{a+b}{2}} \left(e^{i\frac{a-b}{2}} - e^{-i\frac{a-b}{2}} \right) = 2i \sin\left(\frac{a-b}{2}\right) e^{i\frac{a+b}{2}}$.

Remarquez que c'est une façon commode de retenir ou retrouver les formules de factorisation des fonctions trigonométriques (transformation d'une somme en produit).

Exemple 7.2.28

Factoriser $1 + e^{ia}$.

Méthode 7.2.29 (Linéarisation)

Le but est d'exprimer $\cos^n \theta$ ou $\sin^n \theta$ en fonction de $\cos(k\theta)$ et $\sin(k\theta)$, $k \in \mathbb{N}$. Principe du calcul :

1. Exprimer $\cos \theta$ (ou $\sin \theta$) à l'aide des formules d'Euler ;
2. Développer à l'aide de la formule du binôme de Newton ;
3. Regrouper dans le développement les exponentielles conjuguées et les réexprimer à l'aide des fonctions sin et cos en utilisant la formule d'Euler dans l'autre sens.

Exemple 7.2.30

1. Linéariser $\cos^4(x)$
2. Linéariser $\sin^6(x)$. En déduire $\int_0^\pi \sin^6(x) dx$.
3. Proposer plus efficace pour le calcul de $\int_0^\pi \sin^5(x) dx$

Méthode 7.2.31 (« délinéarisation », ou les polynômes de Tchébychev)

Il s'agit de la méthode inverse, consistant à écrire $\cos(n\theta)$ ou $\sin(n\theta)$ en fonction des puissances de $\cos(x)$ et/ou $\sin(x)$. Le principe du calcul :

1. On utilise la formule de Moivre pour exprimer $\cos(n\theta)$ ou $\sin(n\theta)$ comme partie réelle ou imaginaire de $(\cos(\theta) + i \sin(\theta))^n$.
2. On développe cette expression à l'aide de la formule du binôme de Newton
3. On utilise l'identité remarquable $\sin^2 x + \cos^2 x = 1$ pour exprimer la partie réelle (ou imaginaire) sous forme d'un polynôme en $\cos(x)$ (pour $\cos(n\theta)$) ou le produit de $\cos(x)$ par un polynôme en $\sin(x)$ (pour $\sin(n\theta)$)

Remarque 7.2.32

Les polynômes obtenus ainsi s'appelle polynômes de Tchébychev, de première espèce pour les cosinus, et de seconde espèce pour les sinus. On peut définir ces polynômes par récurrence et redémontrer directement à l'aide de ces relations de récurrence et des formules de trigonométrie le fait qu'ils assurent la délinéarisation de $\cos(n\theta)$ et $\sin(n\theta)$. La méthode ci-dessus permet alors d'obtenir une expression explicite de ces polynômes.

Méthode 7.2.33 (Sommes de sin et cos)

Le principe général est d'écrire une somme de sin (ou de cos) sous forme de partie imaginaire (ou réelle) d'une somme d'exponentielles. On peut alors souvent exploiter le caractère géométrique du terme $e^{in\theta}$, par utilisation des propriétés des sommes géométriques, ou de la formule du binôme etc.

Exemple 7.2.34 (noyau de Dirichlet)

Soit a et b deux réels, vérifiant $b \neq 0$ $[2\pi]$, et

$$C = \cos a + \cos(a + b) + \cos(a + 2b) + \cdots + \cos(a + nb) = \sum_{k=0}^n \cos(a + kb).$$

$$\text{Alors } C = \cos\left(a + \frac{bn}{2}\right) \cdot \frac{\sin\left(\frac{n+1}{2} \cdot b\right)}{\sin\frac{b}{2}}.$$

Nous terminons ce paragraphe par une présentation rapide de l'exponentielle complexe générale.

Définition 7.2.35 (Exponentielle complexe)

Soit z un nombre complexe. On définit alors $e^z = e^{\operatorname{Re}(z)} \times e^{i \operatorname{Im}(z)}$.

Si z est réel ou imaginaire pur, on retrouve respectivement l'exponentielle réelle et l'exponentielle définie sur les imaginaires purs.

De façon immédiate, on a les résultats suivants :

Proposition 7.2.36 (Parties réelle, imaginaire, module et argument de e^z)

- (i) $\operatorname{Re}(e^z) = e^{\operatorname{Re}(z)} \cos(\operatorname{Im}(z))$
- (ii) $\operatorname{Im}(e^z) = e^{\operatorname{Re}(z)} \sin(\operatorname{Im}(z))$
- (iii) $|e^z| = e^{\operatorname{Re}(z)}$
- (iv) $\arg(e^z) \equiv \operatorname{Im}(z) [2\pi]$

Théorème 7.2.37 (propriété d'addition)

Soit $(z, z') \in \mathbb{C}^2$. Alors $e^{z+z'} = e^z e^{z'}$.

Proposition 7.2.38 (cas d'égalité)

Soit $(z, z') \in \mathbb{C}^2$. On a $e^z = e^{z'}$ si et seulement si $\operatorname{Re}(z) = \operatorname{Re}(z')$ et $\operatorname{Im}(z) \equiv \operatorname{Im}(z') \pmod{2\pi}$.

Proposition 7.2.39 (recherche de l'image réciproque)

Soit $a \in \mathbb{C}^2$. Alors :

- si $a = 0$, l'équation $e^z = a$ n'a pas de solution ;
- si $a \neq 0$, l'équation $e^z = a$ a une infinité de solutions, décrites par :

$$\operatorname{Re}(z) = \ln|a| \quad \text{et} \quad \operatorname{Im}(z) \equiv \arg(a) \pmod{2\pi}.$$

III Racines d'un nombre complexe

III.1 Racines n -ièmes

Définition 7.3.1 (racines n -ièmes, groupe \mathbb{U}_n)

- Soit $n \in \mathbb{N}^*$ et $z \in \mathbb{C}$. Une racine n -ième de z est une racine (complexe) du polynôme $X^n - z$, donc un nombre complexe ω tel que $\omega^n = z$
- Une racine n -ième de l'unité est une racine n -ième de 1.
- L'ensemble des racines n -ièmes de l'unité est noté \mathbb{U}_n .

Nous verrons plus tard que cet ensemble \mathbb{U}_n possède une structure de *groupe* (notion définie plus tard).

Proposition 7.3.2 (Explication des racines de l'unité)

Le groupe \mathbb{U}_n des racines n -ièmes de l'unité est constitué de n éléments deux à deux distincts et donnés par :

$$\mathbb{U}_n = \{\omega_k = e^{i \frac{2\pi k}{n}}, k \in \llbracket 0, n-1 \rrbracket\}.$$

▫ Éléments de preuve.

Les racines de 1 sont de module 1. On peut donc les rechercher sous la forme $e^{i\theta}$. Résoudre $e^{in\theta} = 1$.

▷

Ces racines se répartissent de façon régulière sur le cercle trigonométrique, de façon à former les sommets d'un polygone régulier à n côtés (figure 7.3)

Proposition 7.3.3 (racines n -ièmes de z , figure 7.4)

Soit $z = re^{i\theta}$ un nombre complexe. Alors :

- (i) Une racine n -ième particulière de z est $z_0 = \sqrt[n]{r} \cdot e^{i \frac{\theta}{n}}$
- (ii) z possède exactement n racines n ièmes, données par :

$$\xi_k = z_0 \omega, \quad \omega \in \mathbb{U}_n,$$

où $\omega_0, \dots, \omega_{n-1}$ sont les racines n -ièmes de l'unité.

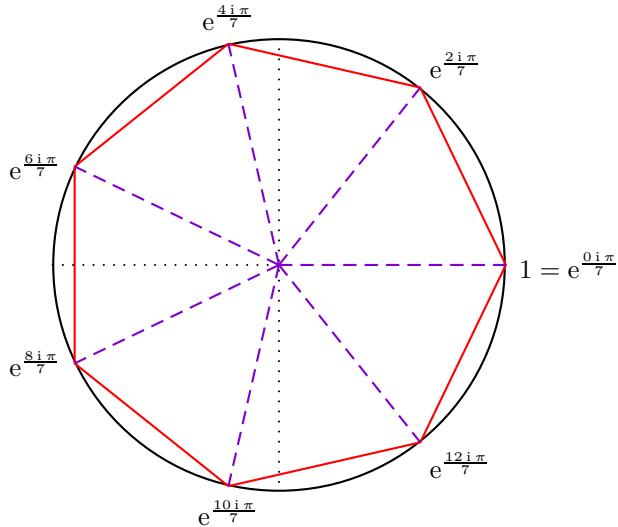


FIGURE 7.3 – Répartition des racines de l'unité sur le cercle trigonométrique

(iii) Ainsi, pour $z = re^{i\theta}$, on obtient la description explicite des racines n -ièmes :

$$\xi_k = \sqrt[n]{r} \cdot e^{i\left(\frac{\theta+2k\pi}{n}\right)}, \quad k \in \llbracket 0, n-1 \rrbracket.$$

▫ Éléments de preuve.

De même, rechercher les racines sous la forme $\rho e^{i\varphi}$.

▷

Proposition 7.3.4

Soit $\omega \in \mathbb{U}_n \setminus \{1\}$. Alors $\sum_{i=0}^{n-1} \omega^i = 0$.

▫ Éléments de preuve.

Formule de sommation des suites géométriques.

▷

En particulier, puisque pour tout $k \in \llbracket 0, n \rrbracket$, $\omega_k = \omega_1^k$, on obtient :

Corollaire 7.3.5 (somme des racines n -ièmes de l'unité)

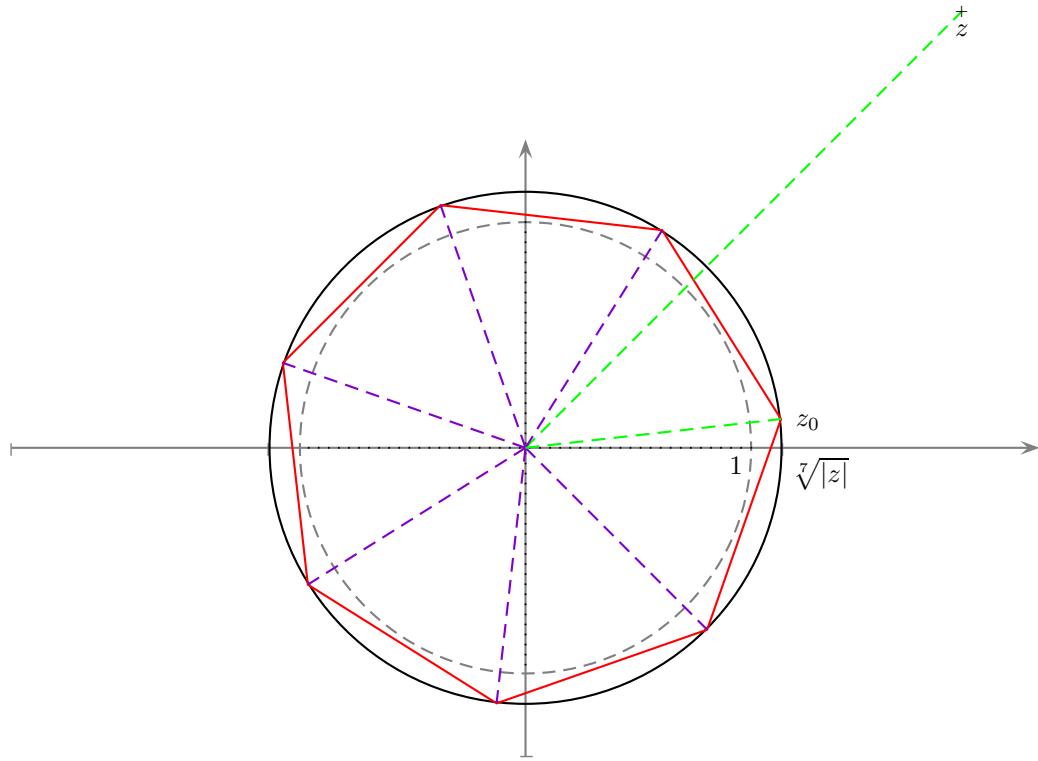
Soit $n \in \mathbb{N}^* \setminus \{1\}$. Alors $\sum_{\omega \in \mathbb{U}_n} \omega = 0$

Corollaire 7.3.6 (somme des racines n -ièmes de z)

Soit $n \geq 2$, et $\{\xi_0, \dots, \xi_{n-1}\}$ l'ensemble des racines n -ièmes de z , alors $\sum_{k=0}^{n-1} \xi_k = 0$.

Remarque 7.3.7

Ce résultat est un cas particulier des relations entre coefficients et racines d'un polynôme.

FIGURE 7.4 – Répartition des racines de z dans le plan complexe**Notation 7.3.8 (Le complexe j)**

On note j la racine cubique de l'unité $j = e^{i\frac{2\pi}{3}} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$.

Les racines cubiques de 1 sont donc $1, j$ et j^2 .

Proposition 7.3.9 (Propriétés de j)

- (i) $j^3 = 1$, donc $j^n = j^r$, où r est le reste de la division de n par 3
- (ii) $\bar{j} = j^2$
- (iii) $j^2 + j + 1 = 0$

▫ Éléments de preuve.

Le point (ii) provient de $j\bar{j} = 1$

▷

On peut également décrire les racines 6-ièmes de 1 à l'aide de j

Proposition 7.3.10 (racines 6-ièmes de 1)

Les racines 6-ièmes de 1 sont, par ordre croissant d'arguments positifs : $1, -j^2, j, -1, j^2, -j$.

III.2 Cas des racines carrées : expression sous forme algébrique

Les racines carrées z peuvent être obtenues par la méthode général décrite ci-dessus, lorsqu'on connaît z sous forme trigonométrique : Si $z = re^{i\theta}$, z possède deux racines carrées :

$$z_1 = \sqrt{r}e^{i\frac{\theta}{2}} \quad \text{et} \quad z_2 = -\sqrt{r}e^{i\frac{\theta}{2}}.$$

Dans le cas particulier des racines carrées, on dispose également d'une méthode permettant d'obtenir facilement l'expression algébrique des racines carrées de z lorsque z est donné lui-même sous forme algébrique.

Méthode 7.3.11 (recherche des racines carrées sous forme algébrique)

Soit $z = a + i b$ un nombre complexe sous forme algébrique ($a, b \in \mathbb{R}$). Pour trouver les racines carrées sous forme algébrique :

1. Considérer une racine $z' = c + i d$
2. Identifier les parties imaginaires et réelles dans l'égalité $(z')^2 = z$: en retenir essentiellement la valeur de $c^2 - d^2$ et le signe de cd .
3. Donner l'égalité des modules de $(z')^4$ et de z^2 . Cela donne la valeur de $c^2 + d^2$.
4. Résoudre le système en c^2 et d^2 donné par les équations ci-dessus.
5. Des quatre solutions pour le couple (c, d) , garder les deux seules qui donnent le bon signe de cd .

On peut aussi s'en sortir sans utiliser l'égalité des modules, en constatant que l'étape 2 donne la valeur de la somme et du produit de c^2 et $-d^2$, donc une équation du second degré dont ces réels sont les solutions.

Exemple 7.3.12

1. Rechercher les racines carrées de $3 + 5i$.
2. Trouver les solutions de l'équation du second degré : $(2 + i)z^2 - iz + 1 = 0$

La méthode de résolution ci-dessus montre que les racines (réelles ou complexes) d'un polynôme de degré 2 peuvent être exprimées à l'aide de radicaux (*i.e.* les parties réelles et imaginaires peuvent être exprimées à l'aide des 4 opérations usuelles à partir des nombres rationnels, des coefficients de l'équation, et des fonctions « racine » définies sur \mathbb{R}_+ , ici la racine carrée).

Note Historique 7.3.13

- La résolution d'équations polynomiales par radicaux a motivé une part importante de la recherche mathématique, jusqu'à ce que Niels Abel prouve l'impossibilité de résoudre l'équation général du 5-ième degré par radicaux. Peu de temps après, Évariste Galois élucide complètement le problème, dans un mémoire rédigé peu avant sa mort prématurée en 1832, et dans une lettre rédigée à la hâte à un ami, la veille du duel qui devait lui être fatal (il avait alors 20 ans). Dans ce mémoire, on y trouve en particulier les balbutiements de la théorie des groupes.
- Carl Friedrich Gauss montre que les racines de $X^n - 1$ (donc les racines n -ièmes de l'unité), peuvent s'exprimer par radicaux si n est premier.
- Il va plus loin, en montrant que si n est un entier premier de la forme $2^{2^k} + 1$, alors les solutions peuvent s'exprimer sous forme de radicaux carrés. Ce résultat amène la constructibilité à la règle et au compas du pentagone (déjà connu depuis bien longtemps), de l'heptadécagone, *i.e.* le polygone à 17 côtés (Gauss en donne une construction) puis des polygones à 257 et 65537 côtés. On ne connaît pas à ce jour d'autre nombre premier de la forme $2^{2^k} + 1$ (nombres de Fermat). On ne sait pas s'il y en a d'autres.
- Pierre-Laurent Wantzel montre la réciproque en 1837 : les seuls polygones constructibles sont les polygones dont le nombre de côtés est un nombre premier de la forme $2^{2^k} + 1$, ou des nombres ayant comme uniques facteurs (qui doivent être simples) ces nombres premiers ou 2 (en multiplicité quelconque). Ce théorème est connu sous le nom de théorème de Gauss-Wantzel.

IV Nombres complexes et géométrie

IV.1 Affixes

Nous terminons cette étude des nombres complexes par un bref aperçu de l'efficacité de l'utilisation des nombres complexes pour l'étude de la géométrie du plan. Nous rappelons que, par la construction que nous avons donnée, \mathbb{C} s'identifie au plan \mathbb{R}^2 . Nous rappelons :

Définition 7.4.1 (affixe)

1. L'affixe d'un point $(a, b) \in \mathbb{R}^2$ est le complexe $z_A = a + i b$.
2. L'affixe d'un vecteur $\vec{u} = \begin{pmatrix} a \\ b \end{pmatrix}$ de \mathbb{R}^2 est le complexe $z_{\vec{u}} = a + i b$.

On commence par traduire sous forme complexe certaines propriétés géométriques :

Proposition 7.4.2 (affixe d'un vecteur défini par un bipoint)

Soit A et B deux points d'affixe z_A et z_B . Alors $z_{\overrightarrow{AB}} = z_B - z_A$

Proposition 7.4.3 (norme d'un vecteur)

Soit \vec{u} un vecteur de \mathbb{R}^2 d'affixe $z_{\vec{u}}$. Alors $\|\vec{u}\| = |z_{\vec{u}}|$.

IV.2 Alignement, orthogonalité, angles

Proposition 7.4.4 (Expression complexe du déterminant et du produit scalaire)

Soit $z_u = a + i b$ et $z_v = c + i d$. Alors

- (i) $\text{Re}(z_u \overline{z_v}) = \langle \vec{u}, \vec{v} \rangle$ (produit scalaire canonique)
- (ii) $\text{Im}(z_u \overline{z_v}) = \det(\vec{u}, \vec{v}) = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - bc$ (déterminant)

▫ Éléments de preuve.

Expliquer $z_u \overline{z_v}$.

▷

Proposition 7.4.5 (traduction de l'alignement, de l'orthogonalité)

Avec les notations sous-entendues évidentes :

1. \vec{u} et \vec{v} sont colinéaires si et seulement si $\text{Im}(z_u \overline{z_v}) = 0$.
2. A, B et C sont alignés si et seulement si $\text{Im}((z_C - z_A)(\overline{z_C} - \overline{z_B})) = 0$.
3. \vec{u} et \vec{v} sont orthogonaux si et seulement si $\text{Re}(z_u \overline{z_v}) = 0$.

▫ Éléments de preuve.

1. Sauf dans le cas trivial où $v = 0$, cela équivaut à dire que $\frac{z_u}{z_v}$ est réel.
2. A, B et C sont alignés si et seulement si \overrightarrow{AC} et \overrightarrow{BC} sont colinéaires.
3. C'est la définition (via le produit scalaire) de l'orthogonalité.

▷

Remarque 7.4.6

Si \vec{v} est non nul, il est parfois plus commode de caractériser la colinéarité par le fait que $\frac{z_u}{z_v}$ est réel, et l'orthogonalité par le fait que $\frac{z_u}{z_v}$ est imaginaire.

Proposition 7.4.7 (interprétation géométrique de $\frac{b-a}{c-a}$.)

Soit a , b et c trois complexes, et A , B et C les points de \mathbb{R}^2 d'affixe a , b et c . Alors

$$\arg\left(\frac{b-a}{c-a}\right) = (\overrightarrow{AC}, \overrightarrow{AB})$$

△ Éléments de preuve.

L'écrire en notation trigonométrique.

▷

IV.3 Transformations du plan

Enfin, les transformations usuelles du plan peuvent être traduites par des fonctions simples de \mathbb{C} dans \mathbb{C} , ce qui simplifie souvent leur étude ou leur utilisation.

Proposition 7.4.8 (Interprétation complexe des transformations usuelles du plan)

1. Soit \vec{u} un vecteur du plan, d'affixe z_u . La translation de vecteur \vec{u} correspond dans \mathbb{C} à la fonction $z \mapsto z + z_u$.
2. Soit A un point du plan, d'affixe z_A , et θ un réel. La rotation de centre A et d'angle θ (dans le sens trigonométrique, ou direct) correspond dans \mathbb{C} à la fonction $z \mapsto z_A + e^{i\theta}(z - z_A)$.
3. Soit A un point du plan, d'affixe z_A , et λ un réel. L'homothétie de centre A et de rapport λ correspond dans \mathbb{C} à la fonction $z \mapsto z_A + \lambda(z - z_A)$.
4. Soit D une droite passant par le point A d'affixe z_A , et de vecteur directeur unitaire \vec{u} , d'affixe z_u . La symétrie orthogonale d'axe D (ou reflexion d'axe D) est donnée par la fonction $z \mapsto z_u^2(\bar{z} - \bar{z}_A) + z_A$

△ Éléments de preuve.

Si on note M le point d'affixe z et M' le point d'affixe $\varphi(z)$, en vertu des résultats précédents :

1. $\overrightarrow{MM'} = \vec{u}$;
2. $\|AM'\| = \|AM\|$, et l'angle entre les deux vecteurs est l'argument du quotient des affixes ;
3. $\overrightarrow{AM'} = \lambda \overrightarrow{AM}$;
4. Si θ est l'argument de z_u , faire une rotation d'angle $-\theta$ nous ramène à une symétrie d'axe horizontal, et si on translate encore de $-\overrightarrow{OA}$ cette symétrie correspond à la conjugaison. Quelle relation cela donne-t-il entre $\frac{z - z_A}{z_u}$ et $\frac{z' - z_A}{z_u}$?

▷

Exemple 7.4.9 (Cas particuliers importants)

1. La rotation d'angle $\frac{\pi}{2}$ se traduit par une multiplication par i . S'en souvenir dans des contextes d'orthogonalité !
2. La rotation d'angle $\frac{\pi}{3}$ se traduit par la multiplication par j . La rotation d'angle $\frac{\pi}{6}$ se traduit par la multiplication par $-\bar{j} = -j^2$. S'en souvenir dans tous les contextes faisant intervenir une symétrie d'ordre 3, notamment pour l'étude des triangles équilatéraux.

Remarquez dans le 4 qu'on a supposé que \overrightarrow{u} est unitaire, donc que z_u est de la forme $e^{i\theta}$. Ainsi, la multiplication par $\overline{z_u}$ correspond à une rotation qui nous ramène à un axe horizontal. Si cet axe passe par 0, la symétrie correspond alors à la conjugaison (ce qui fait partir la barre du coefficient multiplicatif z_u).

Attention, le caractère unitaire de \overrightarrow{u} est important ici.

Remarque 7.4.10 (Forme des applications associées aux transformations usuelles)

Les transformations usuelles étudiées ci-dessus s'écrivent toutes sous la forme $z \mapsto az + b$ ou $z \mapsto a\bar{z} + b$.

On montre que réciproquement une application de ce type correspond à une transformation usuelle, ou une composée de transformations usuelles du plan :

Théorème 7.4.11 (Interprétation des transformations affines de \mathbb{C})

- (i) Soit $\varphi : z \mapsto az + b$, $a \neq 0$. Alors :
 - Si $a = 1$, φ représente une translation ;
 - Si $a = \lambda e^{i\theta} \neq 1$, il existe un point C tel que φ représente la composée d'une rotation d'angle θ de centre C et d'une homothétie de rapport λ de même centre C .
- (ii) Soit $\varphi : z \mapsto a\bar{z} + b$, $a = \lambda e^{i\theta} \neq 0$ et $u = e^{i\frac{\theta}{2}}$. Alors, il existe un point A tel que φ représente la composée d'une symétrie d'axe D porté par u et passant par A , d'une translation de vecteur αu , $\alpha \in \mathbb{R}$ (un glissement le long de la droite D) et d'une homothétie de centre O de rapport λ .

▫ Éléments de preuve.

- (i) si $a \neq 1$, mettre sous la forme $\varphi(z) - z_A = \lambda e^{i\theta}(z - z_A)$. Cette équation détermine z_A en fonction de a et b .
- (ii) L'homothétie étant appliquée en dernier, on est ramené à la description de $z \mapsto z' = u^2\bar{z} + \frac{b}{\lambda}$. Chercher à la mettre sous la forme $z' = u^2(\bar{z} - \overline{z_A}) + z_A + tu$, où $t \in \mathbb{R}$ est colinéaire à u . Se ramener à la forme algébrique de $\frac{b}{\lambda u}$ pour déterminer t et $\text{Im}(\overline{u}z_A)$. La partie réelle de cette quantité peut être choisie quelconque (c'est dire que le choix du point A sur la droite D est quelconque)

▷

Méthode 7.4.12

Il faut savoir retrouver, sur des formes explicites, les points A , vecteurs u etc., en utilisant la démarche suggérée dans la démonstration.

En particulier, cette représentation des transformations usuelles du plan est efficace pour exprimer des composées :

Lemme 7.4.13 (Composée d'applications affines)

- Soit $\varphi_k : z \mapsto a_k z + b_k$, $k \in \llbracket 1, n \rrbracket$. Alors la composée $\varphi = \varphi_n \circ \dots \circ \varphi_1$ est de la forme $z \mapsto az + b$, où $a = a_1 \dots a_n$
- Si un nombre N des φ_k sont de la forme $z \mapsto a_k \bar{z} + b_k$, on obtiendra une composée $z \mapsto az + b$ si le nombre N est pair, et $z \mapsto a\bar{z} + b$ s'il est impair. Mais le coefficient a est un peu plus délicat à exprimer de façon générale en fonction des a_i (mais sans difficulté dans l'étude de cas particuliers).

▫ Éléments de preuve.

Récurrence sur n .

▷

En particulier, en combinant les deux résultats précédents, on obtient :

Proposition 7.4.14 (composition de rotations, translations et homothéties)

Toute transformation obtenue par composition de translations, de rotations (d'angle total θ), et d'homothéties (dont le produit des rapports est λ) est

- soit la composition d'une rotation d'angle θ et d'une homothétie de même centre, de rapport λ si $\lambda e^{i\theta} \neq 1$,
- soit une translation si $\lambda e^{i\theta} = 1$.

IV.4 Isométries et similitudes

Définition 7.4.15 (isométries et similitudes)

Soit $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ une application.

1. On dit que F est une isométrie affine si F conserve les longueurs, donc si pour tout $(A, B) \in (\mathbb{R}^2)^2$, $\|F(A)F(B)\| = \|AB\|$. Cela se traduit par une application $\varphi : \mathbb{C} \rightarrow \mathbb{C}$ vérifiant $|\varphi(z_2) - \varphi(z_1)| = |z_2 - z_1|$ pour tous z_1 et z_2 .
2. On dit que F est une isométrie vectorielle si F est une isométrie affine telle que $F(O) = O$, où $O = (0, 0)$.
3. On dit que F est une similitude affine s'il existe $\lambda \in \mathbb{R}_+^*$ tel que pour tout $(A, B) \in (\mathbb{R}^2)^2$, $\|f(A)f(B)\| = \lambda \|AB\|$, ce qui se traduit par $|\varphi(z_2) - \varphi(z_1)| = \lambda |z_2 - z_1|$.
4. On dit que F est une similitude vectorielle si F est une similitude affine telle que $F(O) = O$.

Lemme 7.4.16 (Composée de similitudes ou d'isométries)

La composée de deux similitudes affines de rapports λ et λ' est une similitude affine de rapport $\lambda\lambda'$, vectorielle si les deux le sont. En particulier, la composée de deux isométries est une isométrie.

Les rotations, translations et symétries étant des isométries, et les homothéties des similitudes, il vient de l'étude précédente et des propriétés de composition que :

Théorème 7.4.17 (Les $z \mapsto az + b$ ou $a\bar{z} + b$ sont des similitudes)

Les applications $z \mapsto az + b$ et $z \mapsto a\bar{z} + b$ correspondent à des similitudes, qui sont des isométries si de plus $|a| = 1$

▫ Éléments de preuve.

On peut aussi facilement en faire une preuve directe.

▷

On étudie maintenant la réciproque, en commençant par le cas vectoriel :

Théorème 7.4.18 (Les similitudes vectorielles sont des $z \mapsto az$ ou $a\bar{z}$)

Soit φ l'application associée à une similitude vectorielle, et $a = \varphi(1)$. Alors $\varphi : z \mapsto az$ ou $\varphi : z \mapsto a\bar{z}$. De plus, φ est une isométrie si et seulement si $|a| = 1$.

▫ Éléments de preuve.

Fixer un z_0 non réel. En utilisant le fait que les triangles définis par $0, 1, z_0$ et $0, a, \varphi(z_0)$, sont semblables, en déduire que $\varphi(z_0) = az_0$ ou $\varphi(z_0) = a\bar{z}_0$. Conclure en faisant de même pour z quelconque et en considérant aussi la distance de $\varphi(z)$ à $\varphi(z_0)$. ▷

Ainsi, on peut de façon naturelle classer les isométries et les similitudes en deux catégories :

Définition 7.4.19 (isométrie, similitude directe, indirecte)

- Une isométrie (*resp.* similitude) vectorielle de type $z \mapsto az$ est appelée isométrie (*resp.* similitude) directe. Il s'agit des rotations (*resp.* homothéties-rotations).
- Une isométrie (*resp.* similitude) vectorielle de type $z \mapsto a\bar{z}$ est appelée isométrie (*resp.* similitude) indirecte.

En remarquant que φ est une similitude affine ssi $\varphi - \varphi(0)$ est une similitude vectorielle, on obtient alors :

Proposition 7.4.20 (Caractérisation des isométries et des similitudes affines)

Les similitudes affines de \mathbb{R}^2 correspondent exactement aux applications complexes $f : z \mapsto az + b$ ou $f : z \mapsto a\bar{z} + b$, $a \in \mathbb{C}^*, b \in \mathbb{C}$; ce sont des isométries ssi $|a| = 1$.

On définit de même que dans le cas vectoriel les isométries et similitudes affines directes et indirectes.

Exemple 7.4.21

À quelles isométries correspondent les applications $z \mapsto \bar{z}$ et $z \mapsto i\bar{z}$?

IV.5 Caractérisation de certains objets géométriques

Proposition 7.4.22 (Caractérisation des droites)

La droite passant par A d'affixe a et B d'affixe b est l'ensemble constitué des points M d'affixe z tels que l'une des propriétés équivalentes suivantes soit vérifiée :

- (i) il existe $t \in \mathbb{R}$ tel que $z = (1 - t)a + tb$
- (ii) $z = a$, ou $z \neq a$ et $\arg(z - a) \equiv \arg(b - a) [\pi]$
- (iii) $\frac{z - a}{b - a} \in \mathbb{R}$.

▫ Éléments de preuve.

Exprimer la colinéarité de \overrightarrow{AM} et \overrightarrow{AB} . ▷

Proposition 7.4.23 (Caractérisation des cercles)

Un sous-ensemble C de \mathbb{C} est un cercle éventuellement vide si et seulement si il existe un complexe α et un réel β tels que C soit l'ensemble des points d'affixe z vérifiant :

$$z \cdot \bar{z} + \alpha z + \bar{\alpha} \cdot \bar{z} + \beta = 0.$$

L'ensemble C est dans ce cas non vide si et seulement si $\beta \leq \alpha\bar{\alpha}$ et dans ce cas, son centre est le point d'affixe $-\bar{\alpha}$ et son rayon est $r = \sqrt{\alpha\bar{\alpha} - \beta}$.

▫ Éléments de preuve.

Écrire cette équation sous la forme $(z + \bar{\alpha})(\bar{z} + \alpha) = \rho$. ▷

Méthode 7.4.24

Il est inutile de retenir l'expression de l'affixe du centre et du rayon. Il faut en revanche être capable de les retrouver rapidement en faisant la factorisation précédente.